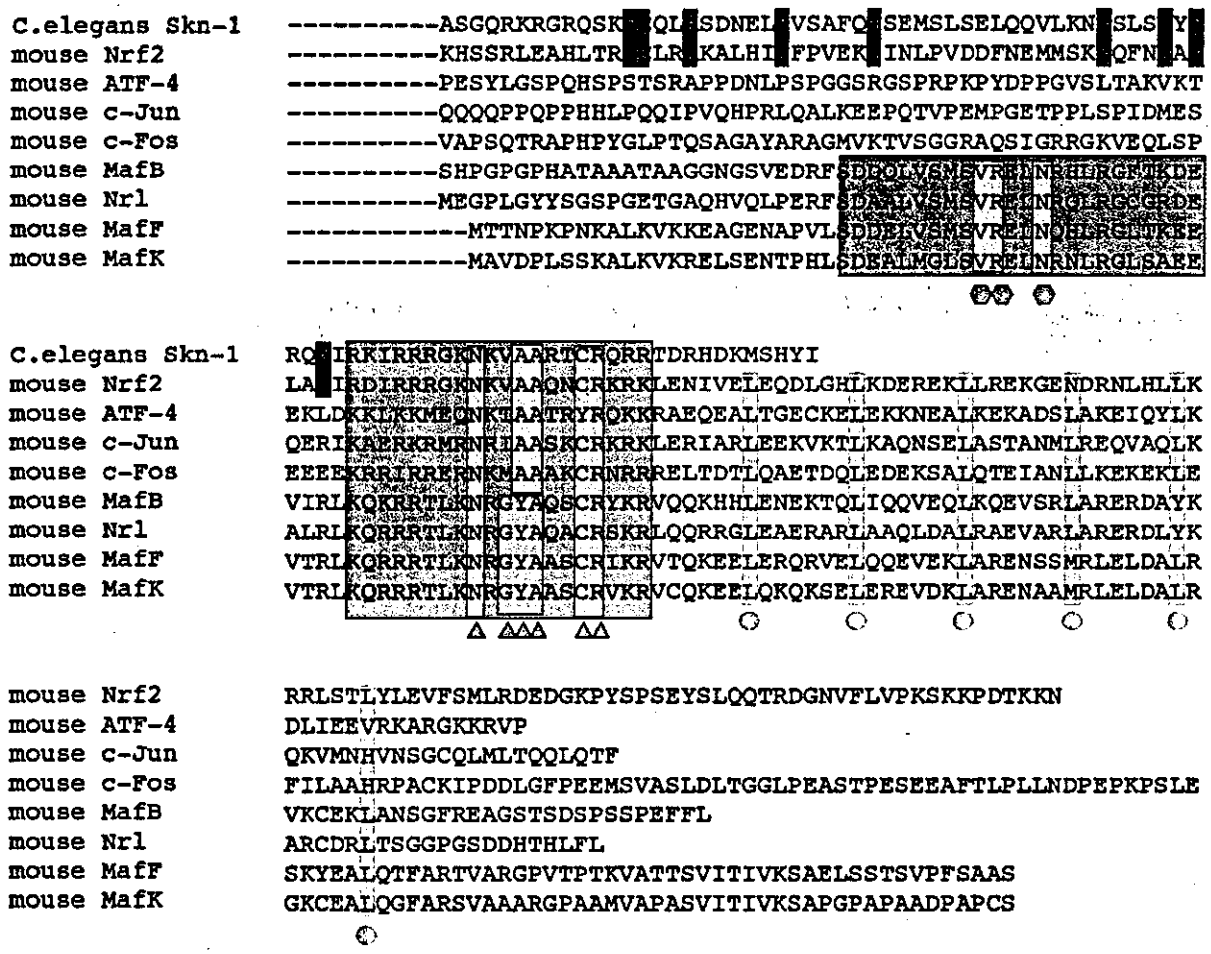

Fig. 1. Sequence alignment of various bZip transcription factors. Basic regions are highlighted in blue. Blue triangles indicate amino acid residues that are important for contacting DNA. Heptad repeats of the leucine zippers are highlighted in yellow and indicated by yellow circles. The extended homology region (EHR), a well-conserved motif in Maf family proteins, is highlighted in pink. Red hexagons indicate the critical residues of the EHR required for recognition of the flanking sequence of the MARE. Conserved residues in the CNC domains that are specific for the CNC family proteins are highlighted in green.

nents of erythroid-specific gene transcription (Martin et al., 1996).

In particular, the erythroid-specific transactivator NF-E2 was revealed by functional analysis to transduce signals through the AP-1-like motif (Mignotte et al., 1989b). Subsequently, NF-E2 was purified from mouse erythroleukemia (MEL) cells as a novel bZip transcription factor composed of a 45 kDa subunit and an 18 kDa subunit (Andrews et al., 1993a). Expressed specifically in hematopoietic cells, the p45 subunit is most closely related to the Drosophila 'cap-'n'-collar' (thus CNC) transcription factor (Mohler et al., 1991). The p18 subunit was later identified as MafK, one of the small Maf proteins (Andrews et al., 1993b). The observation that recombinant p45 fails to bind to DNA by itself showed that MafK is necessary for NF-E2 transcriptional activity.

The consensus NF-E2 binding sequence is an asymmetric 11-bp sequence 'TGCTGACTCAT' that differs from a T-MARE by the presence of a T at the position of the G residue directly 5′ of the TRE in T-MARE (Fig. 2). The strong resemblance that the T-MARE bears to the NF-E2 binding site prompted our initial investigations of the binding properties of p45 and small Mafs to the NF-E2 site using EMSA, thus establishing the discovery that small Mafs constitute the heterodimeric partner molecules of p45 in mediating NF-E2 activity (Igarashi et al., 1994).

2. Expansion of the CNC transcription factor family

2.1. Nrf1, Nrf2, and Nrf3

Since the isolation of p45, the search for other NF-E2 (or NF-E2-type MARE) binding factors led to the identification of Nrf1, Nrf2, and Nrf3. All of these factors share the conserved CNC-like bZip motif and exploit small Maf proteins as obligatory heterodimeric partner molecules for binding to the MARE (Toki et al., 1997; Kobayashi et al., 1999). A probe containing the tandem MARE of the β-globin LCR was used to isolate Nrf1/ICR-F1 from a cDNA expression library (Chan et al., 1993; Caterina et

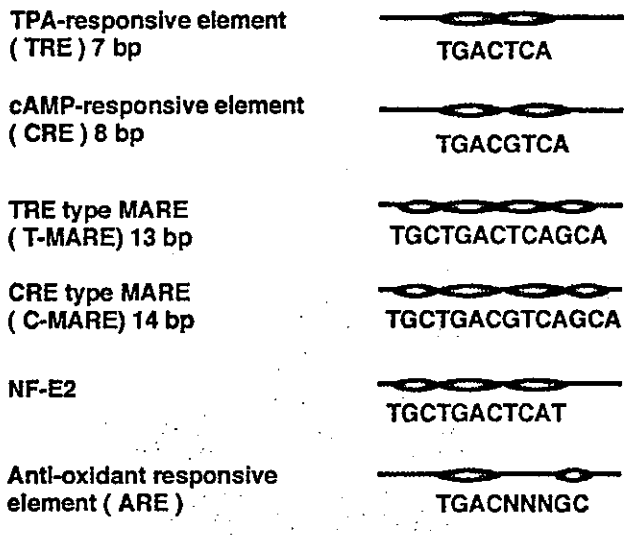

Fig. 2. Consensus sequences recognized by bZip transcription factors. TRE and CRE are included within T-MARE and C-MARE, respectively, and establish the core sequence of the MARE (in green). Three bases on each side of the core form the flanking sequence (in red), which is critical for recognition of the MARE by Maf family proteins. NF-E2 and the ARE are asymmetrical sequences composed of a half site of TRE and a half site of the MARE (or one flanking sequence).

al., 1994). Nrf2 was identified while screening for a molecule that can interact with the NF-E2 binding motif (Moi et al., 1994). The chicken orthologue of Nrf2, ECH, was independently isolated in the search for CNC homologues (Itoh et al., 1995). The fourth member of the CNC family of proteins, Nrf3, was isolated as an EST clone encoding an amino acid sequence homologous to that of Nrf1 (Kobayashi et al., 1999).

Since p45 is expressed in erythroid, megakaryocyte, and mast cell lineages (Andrews et al., 1993a), it was suggested that p45 is the major transcriptional activator in these lineages. However, in contrast to expectation that mice lacking p45 would probably die of anemia during early development, p45-null mutant mice exhibited normal erythropoiesis. Instead, the mutant mouse showed high mortality due to severe hemorrhage resulting from the absence of circulating platelets (Shivdasani et al., 1995). Large megakaryocytes with high ploidy accumulated in the bone marrow and spleen of the p45 mutants, and ultrastructural analysis revealed abundant formation of demarcation membranes. Proplatelet formation, a pathway of platelet production from megakaryocytes, was remarkably inhibited in p45-null megakaryocytes (Lecine et al., 1998), indicating that disruption of p45 causes arrest in a late stage of megakaryopoiesis.

Nrf1, Nrf2, and Nrf3 are expressed in a wide range of tissues, with each factor showing a distinct, but overlapping expression profile. Nrf1 germline mutant mice were found to die at late gestation of anemia. However, Nrf1-deficient ES cells can contribute to the development of adult blood cells in chimeric mice, indicating that the defective erythropoiesis in *nrf1*-null mutant animals is non-cell autonomous (Chan et al., 1998). On the other hand, Nrf2-deficient mice did not display any abnormality in hematopoiesis (Chan et al., 1996; Itoh et al., 1997).

In order to test for possibly compensating activities among members of the CNC family in the erythroid lineage, mice double mutant for p45 and Nrf2 were generated (Kuroha et al., 1998; Martin et al., 1998). Examination of these animals showed that a lack of both p45 and Nrf2 did not exacerbate erythroid or megakaryocytic dysfunctions observed in p45-null mice, suggesting that Nrf2 is dispensable for normal hematopoiesis.

Many reports have suggested that the NF-E2-type MARE is important for erythroid-specific gene regulation. However, no conclusive evidence has been obtained from direct genetic analyses to conclude that the contribution of the NF-E2 binding sequence is indispensable for hematopoiesis in vivo. Since four candidate molecules exist for the large subunit and three exist for the small subunit of NF-E2, we surmise that the resulting complexity might make it difficult to uncover the 'loss of function' effect.

2.2. *Bach1 and Bach2*

In the search for proteins that interact with MafK, the bZip transcription factors Bach1 and Bach2 were discovered by yeast two hybrid screening. Bach1 is a ubiquitous factor and is highly expressed in hematopoietic organs, such as the bone marrow and fetal liver (Igarashi et al., 1998), while Bach2 is most abundant in the brain and B lymphocytes (Oyake et al., 1996; Muto et al., 1998). The Bach factors do not contain a canonical transcriptional activation domain and repress transcription when expressed transiently in fibroblasts, although Bach1 was found to function as a transcriptional activator in cultured erythroid cells (Oyake et al., 1996). Furthermore, the corepressor SMRT has been shown to interact with Bach2, but not with Bach1 (Muto et al., 1998).

Bach1 and Bach2 also possess a BTB domain N-terminal to the CNC-bZip domain (Oyake et al., 1996). In addition to enabling heterodimer formation with MafK through their respective zipper motifs, Bach1 can also interact with fellow Bach-MafK dimers through their BTB domains to generate multimeric and multivalent DNA binding complexes (Igarashi et al., 1998; Yoshida et al., 1999). Through this ability to form multimers, Bach1 may be able to mediate interactions between distant, multiple MAREs, and thus may serve as an architectural transcription factor. Another interesting feature of Bach1 is that its C-terminus contains four dipeptide Cysteine-Proline motifs which serve as a heme-binding domain (Ogawa et al., 2001). In the presence of heme, the DNA binding activity of Bach1 is inhibited and, consequently, its function as a transcriptional repressor in reporter transfection assays is abrogated. With the identification of Bach family proteins, it was revealed that small Maf proteins have a wider choice of partner molecules for bind-

ing to the MARE. Through the choice of dimeric partner, small Mafs can switch transcriptional activity from repression to activation.

3. 'Nrf2 era' of detoxifying and antioxidant gene regulation

3.1. Nrf2 is a key regulator in the electrophilic counter-attack response

The study of transcriptional regulation through the MARE brought about a new paradigm in the field of molecular toxicology. Xenobiotic exposure provokes induction of detoxicating enzymes responsible for converting xenobiotics into their less harmful, more hydrophilic forms. Xenobiotics are detoxicated in two-steps: a phase I reaction followed by a phase II reaction. During phase I detoxication, the cytochrome P450 mono-oxygenase system catalyses the oxidation of xenobiotics to their metabolically activated forms. The activated substrates are then catalysed to non-toxic metabolites by phase II enzymes such as glutathione *S*-transferase (GST) and UDP-glucuronosyl-transferase (UGT).

Interestingly, phase II detoxicating enzymes are induced by phase I metabolites as well as by antioxidants, which are very often electrophilic. Pioneering work in this field has shown that many of the phase II enzyme genes are induced via the antioxidant responsive element (ARE) (Rushmore et al., 1991). This critical regulatory element was also called the electrophile responsive element (EpRE) (Friling et al., 1990). The minimum ARE sequence required for transcriptional induction by electrophiles was reported to be TGACNNNGC. The striking similarity between the MARE and ARE sequences (Fig. 2) led us to hypothesize that some of the CNC-Maf factors may also act through the ARE to regulate the transcription of cytoprotective enzyme genes. This was indeed found to be the case for the CNC-bZip factor Nrf2.

Nrf2 is highly expressed in detoxication organs, such as liver and kidney, and organs exposed to the external environment, such as skin, lung, and digestive tract. In this regard, Nrf2 was chosen as the best candidate with which to study the possible interaction of CNC members with the ARE. We therefore generated *nrf2* germline mutant mice and examined their responsiveness to electrophiles (Itoh et al., 1997). Whereas the expression of phase II enzymes was remarkably induced in the liver and intestine of both wild-type and *nrf2*-heterozygous knockout mice, induction was markedly diminished in *nrf2*-homozygous knockout mice. This result, thus, unequivocally demonstrates the critical role that Nrf2 plays in resistance to xenobiotic toxicity. Without Nrf2, insufficient induction of cytoprotective enzymes brings about an increased susceptibility of cells to toxic xenobiotics including acetaminophen, butylated hydroxytoluene, and diesel exhaust (McMahon et al., 2001; Enomoto et al., 2001; Chan and Kan, 1999; Aoki et al., 2001). In addition to this disrupted induction profile, cancer chemoprevention mechanisms are also abolished in mice deficient in Nrf2 (Ramos-Gomez et al., 2001). For example, Oltipraz is known to act as a chemoprotective reagent against various classes of carcinogens and prevents benzo[*a*]pyrene-induced carcinogenesis. However, Oltipraz failed to provide chemoprotection in *nrf2* gene knockout mice, indicating that Nrf2 plays a critical role in the action of this drug.

Nrf2 was found to be a general regulator of oxidative stress-inducible enzymes/proteins, such as heme oxygenase-1 and peroxiredoxin 1 (Ishii et al., 2000). Through in vivo analysis, Kleeberger and colleagues recently demonstrated that susceptibility to hyperoxia is linked to the *nrf2* locus. A point mutation was detected in the promoter region of the *nrf2* gene in C57Bl/6J mice, a strain of mice known to be sensitive to hyperoxia (Cho et al., 2002). This study demonstrated that Nrf2 is fundamental to defense against reactive oxygen species (ROS) and implied that Nrf2 may be involved in the pathogenesis of various lung and other chronic diseases. Thus, protection against electrophiles and ROS share a common mechanism, with Nrf2 functioning as the central mediator.

3.2. Regulation of Nrf2 activity

Elucidation of the relationship between the structural domains of Nrf2 and their function showed that deletion of the N-terminal Neh2 domain enhances the transcriptional activity of Nrf2. This suggested that the N-terminus is necessary for recruiting a negative regulator of Nrf2. This repressor, by the name of Keap1, was identified in a yeast two hybrid screen using the Neh2 domain as bait (Itoh et al., 1999). Keap1 localizes to the actin cytoskeleton in the cytoplasm and possesses a BTB domain and a double glycine repeat (DGR) domain within its N-terminus and C-terminus, respectively. Simultaneous expression of Nrf2 and Keap1 in cultured cells revealed that Nrf2 associates with Keap1 in the cytoplasm, but upon addition of electrophiles to the culture medium, Nrf2 translocates into nuclei and concludes in activated reporter gene transcription.

In certain cases, the level of Nrf2 mRNA is increased by a positive feedback mechanism originating at the ARE in the promoter region of the *nrf2* gene (Kwak et al., 2002). Post-translational modification of Nrf2 may also represent an important regulatory step in the activation of Nrf2. For instance, phosphorylation of Nrf2 by protein kinase C has been reported (Huang et al., 2000). Nrf2 activation is also a downstream consequence of ERK and p38 MAP kinase signal transduction cascades (Zipper and Mulcahy, 2000). While these observations provide important clues regarding the mechanism of Nrf2 activation, our understanding of this process is incomplete. We envisage that determination of the mechanisms involved will eventually lead to the identification of sensor molecules for electrophiles and oxidative stress.

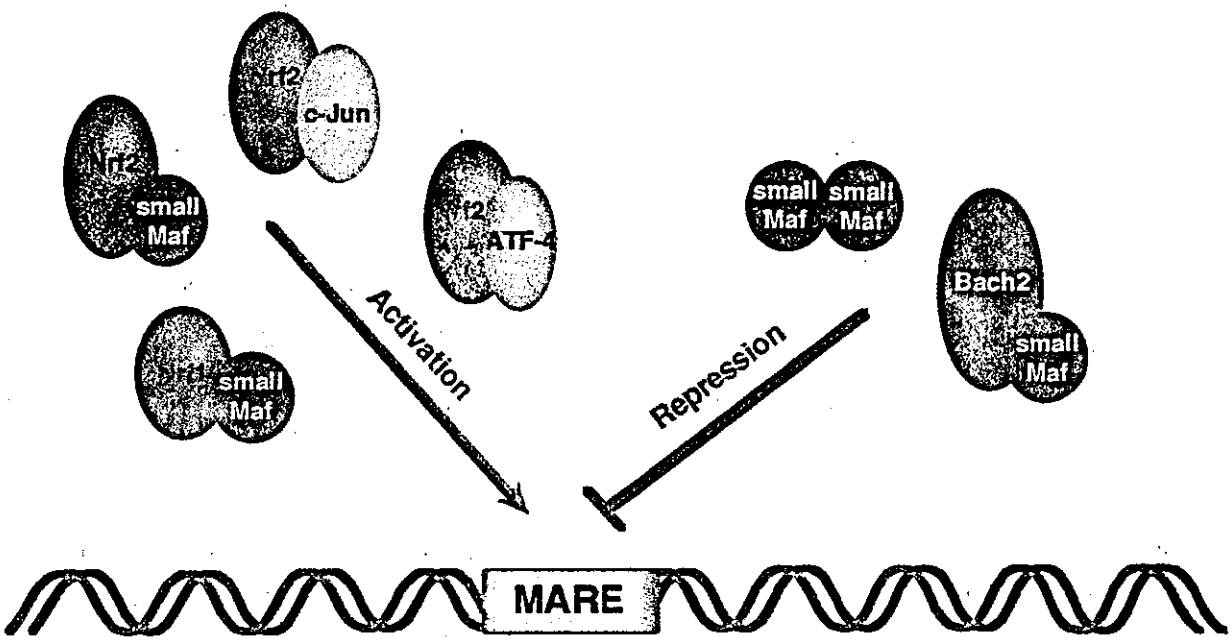

Fig. 3. Various dimer formations interacting with the MARE. Some of these dimers are transcriptional activators, while others repress transcription. Heterodimers containing Nrf1 or Nrf2 can activate transcription. Nrf2 possesses two separate activation domains that act cooperatively. Since small Maf proteins lack activation domains, their homodimers serve as repressors. Bach2 recruits the corepressor SMRT, resulting in transcriptional inhibition.

3.3. Regulatory network surrounding the ARE

Of the CNC family, Nrf2 was found to be the most potent transcriptional activator and usually activates reporter gene transcription by nearly 100-fold (Katoh et al., 2001). This strong activation potential is generated by the synergistic interaction between the coactivator CBP and the two transactivation domains of Nrf2 (Katoh et al., 2001). The requirement for a 'GC' motif in the ARE consensus sequence strongly supports the contention that small Maf proteins serve as the heterodimeric partner molecules of Nrf2 (Kusunoki et al., 2002). Indeed, Nrf2 cannot bind to the ARE or MARE as a monomer, but requires dimerization with one of the small Maf proteins in order to bring about transactivation (Itoh et al., 1997).

The factors c-Jun and ATF4 were also reported to form heterodimers with Nrf2 (Venugopal and Jaiswal, 1998; Dhakshinamoorthy and Jaiswal, 2000; He et al., 2001; Bloom et al., 2002). Since there is a wide variety of possible dimer combinations among bZip transcription factors, it is critically important to establish the major partner molecule of Nrf2 in vivo. Studies using genetically engineered mice seem to provide the most powerful approach to this end. Examination of the antioxidant response in mice lacking ATF-4 or each small Maf is now ongoing in our laboratory.

It has been suggested that CNC transcription factors in addition to Nrf2 may contribute to the cellular defense mechanism against xenobiotic and oxidative stress. Embryonic fibroblasts (MEF) from Nrf1-deficient mouse embryos displayed enhanced sensitivity to oxidative stress and an increased accumulation of free radicals (Kwong et al., 1999). Red blood cells deficient in p45 accumulated ROS and showed a shortened life span due to decreased expression of GST and catalase (Chan et al., 2001). Thus, Nrf1 and p45 are likely participants in ARE-mediated gene regulation in certain cell lineages. While the contribution of these CNC factors to ARE-dependent gene regulation is now clear, we surmise that Nrf2 plays the central role since mice deficient in Nrf2 show a striking loss-of-function phenotype in terms of cytoprotective enzyme induction.

A high level of transcriptional complexity arises from the involvement of several bZip transcription factors acting as a gene regulatory network through the ARE (Fig. 3). Of these factors, the roles of the Bach proteins are the most intriguing since it has been suggested that they antagonize the function of Nrf2. Bach2 normally localizes in the cytoplasm, but translocates to the nucleus in response to an electrophilic stimulus where, upon binding to the MARE, carries out its function as a transcriptional repressor (Hoshino et al., 2000). The C-terminal region of Bach2 contains a unique structure, referred to as the cytoplasmic targeting signal, which is critical for the nuclear-cytoplasmic translocation mediated by exportin/CRM1.

4. Small Maf proteins serve as a determinant of transcriptional activity

4.1. Abundance of small Maf proteins serves as a transcriptional switch

The CNC and Bach family transcription factors require

small Maf partners in order to bind to the MARE sequence. Since small Mafs do not possess any canonical transactivation domain, small Maf homodimers function as transcriptional repressors. Thus, a deficiency in small Mafs would impair the function of CNC and Bach factors yet an excess of small Mafs is also predicted to increase small Maf homodimer formation and lead to repressed transcription. Indeed, when an increasing amount of MafK was co-expressed in cultured cells with a constant level of p45, transcriptional activity from the MARE reporter gene changed dynamically from positive to negative (Igarashi et al., 1994; Nagai et al., 1998). Even a two-fold increase in small Maf abundance was found to convert transcriptional activity from a maximal level to a suppressed level. This result suggested that the quantitative balance between the small Mafs and their heterodimeric partner molecules serves as a molecular switch of gene expression (Igarashi et al., 1994; Nagai et al., 1998). Importantly, small Maf expression levels change dynamically during embryonic development and cell lineage differentiation (Igarashi et al., 1995; Motohashi et al., 1996, 1998; Katsuoka et al., 2000).

We decided that we should confirm our hypothesis that small Mafs serve as a molecular switch in vivo. The small Maf factors MafG, MafK and MafF have similar biochemical activities (Kataoka et al., 1995) and their expression patterns overlap considerably (Onodera et al., 1999). As expected, no apparent abnormalities were detected in either *mafK*-null or *mafF*-null mutant mice and only mild thrombocytopenia and subtle motor defects were observed in mice lacking MafG (Kotkow and Orkin, 1996; Shavit et al., 1998; Onodera et al., 1999). However, the phenotype of *mafG*::*mafK* compound mutant mice was significantly greater in severity than that displayed by the *mafG* mutation alone; thrombocytopenia was exacerbated, phenocopying the *p45* mutation, and the neurological phenotype was conspicuous at a much earlier stage in development (Onodera et al., 2000). These results indicate that megakaryocytes and neurons are sensitive to a reduction in the abundance of small Maf proteins and, therefore, that the total amount of small Mafs must be maintained at a quite specific abundance for proper cellular function.

For this reason, we chose to use megakaryocytes as the target tissue in further examination of small Maf function. From this analysis, we discovered that both *mafG*-null megakaryocytes and transgenic MafG-overexpressing megakaryocytes were defective in proplatelet formation (Motohashi et al., 2000). We concluded that carefully titrated expression of small Maf proteins is vital for achieving maximum transcriptional activity through the MARE, with the absolute quantity of small Mafs determining the fine balance that is normally achieved for MARE-dependent gene regulation in vivo. To further examine the CNC-small Maf regulatory system beyond this model, it is necessary to measure the affinity among the bZip molecules to form heterodimers or homodimers. The difference in DNA-binding preference of each dimeric molecule must also be considered.

4.2. New concept for the roles of small Maf proteins in transcriptional regulation

Treatment of MEL cells with dimethylsulfoxide (DMSO) markedly increases the number of hemoglobinized cells, thus serving as a paradigmatic model system in which to analyse transcriptional gene regulation and the reorganization of eukaryotic nuclei during terminal cell differentiation. Mice deficient in p45 display normal β-globin gene expression, even though p45/small Maf heterodimers bind to MAREs present in several regulatory regions in the β-globin gene locus (Forsberg et al., 2000). A chromatin immunoprecipitation (CHIP) experiment was performed using mouse fetal liver cells and anti-p45 antibody and showed that MAREs within the hypersensitive site 2 (HS2) region of the β-globin LCR were significantly enriched. This indicates that, in vivo, p45 directly and specifically binds to MAREs within the HS2 site. The CHIP assay also showed that small Maf proteins bind predominantly to the β-globin locus in uninduced MEL cells. When MEL cells were induced by DMSO, the β-major globin gene was activated, with recruitment of p45 and small Maf to the β-globin LCR and gene promoter (Sawado et al., 2001). These results indicate that the composition of transcription factors bound to the MARE changes from before and after induction by DMSO.

Interestingly, it has been suggested that relocation of small Maf proteins within the nucleus may be responsible for the change in transcription factor composition at the MARE (Francastel et al., 2001). p45 is concentrated in distinct foci dispersed throughout the nuclei of both undifferentiated and differentiated MEL cells. In comparison, small Maf proteins co-localize with heterochromatin prior to induction by DMSO, but with p45 after induction (Francastel et al., 2001). Upon induction, relocation of small Maf proteins coincides with movement of the β-globin locus from heterochromatin to foci containing p45. Thus, it is tempting to speculate that the cue for differentiation is first bestowed upon small Maf proteins, which initially sequester the β-globin locus within heterochromatin. Subsequently, the small Mafs bring the β-globin locus to the p45 foci where genes are actively transcribed.

These lines of evidence have introduced a new concept to the understanding of regulatory mechanisms directed by bZip transcription factors (Fig. 4). Small Mafs, as a homodimer or as part of a repressive heterodimer, repress transcription by competing with transcriptional activators for the MARE. In addition, small Mafs act as transcriptional repressors by sequestering gene loci within heterochromatin. Alteration in subnuclear localization has thus emerged as a potentially important mechanism in the regulatory system directed by small Maf and CNC proteins.

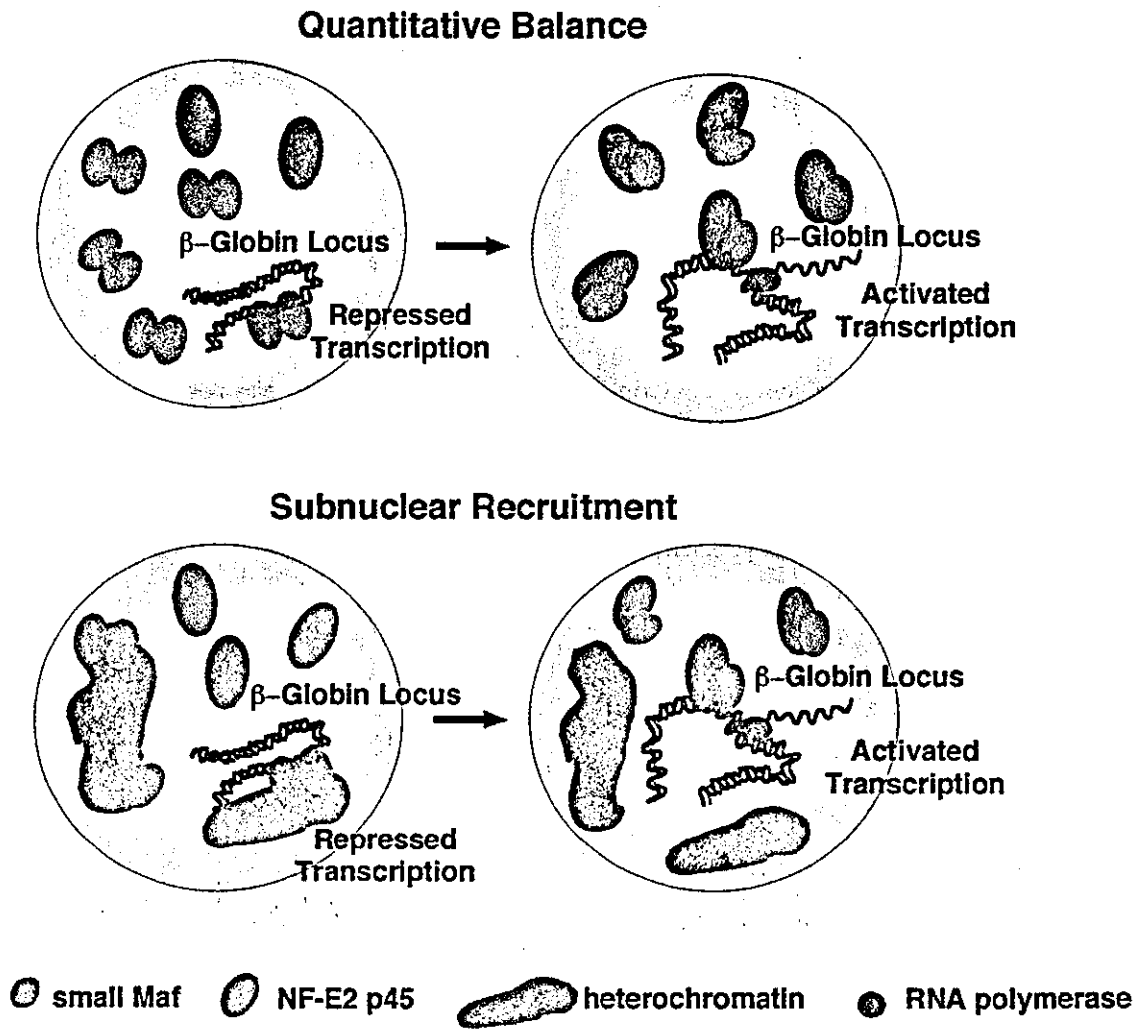

Fig. 4. Two models for the switch from transcriptional inactivation to transcriptional activation. When MEL cells are treated with DMSO, erythroid-specific genes are activated. A shift in the abundance of small Maf and p45 leads to an increase in p45/small Maf heterodimers and subsequent gene activation (upper scheme). Small Maf proteins relocate from heterochromatin to p45-containing foci, bringing the gene loci to the transcriptionally active subnuclear compartments (lower scheme).

5. Deciphering the network comprised of Maf and CNC proteins

5.1. Structural basis of the unique DNA recognition mode of Maf family proteins

Transcription factors of the bZip family possess a basic region and a leucine zipper domain and include proteins such as Skn-1 (contains only a basic region), Nrf2, c-Jun, c-Fos, ATF-4, and Maf (Fig. 1). Maf proteins recognize 13 bp T-MAREs containing the TRE consensus (TGACTCA) or 14 bp C-MAREs containing the CRE consensus (TGACGTCA). In these MAREs, the core consensus motif is flanked on each side by three conserved residues: 'TGC' and 'GCA' at the 5′-end and 3′-end, respectively. Similarly, the 'GC' bases on the outside of the TRE core in the NF-E2 and ARE sequences are conserved. The DNA-binding specificity unique to Maf proteins is achieved through their inherent recognition of these flanking sequences (Kerppola and Curran, 1994). Other bZip factors, in contrast, recognize primarily the core consensus sequences of 7–8 bp in length, such as the TRE and CRE. Importantly, then, members of the Maf family differ from other bZip factors in their DNA binding specificity, suggesting that they carry out a different mode of DNA binding.

The EHR, located on the N-terminal side of the basic region, is conserved only within the Maf family. We recently determined the solution structure of the DNA-binding domain of MafG that contains the EHR and basic region. The structure consists of three α-helices and resembles the fold of the DNA-binding domain of Skn-1. We deduced from determination of the solution structure of the MafG DNA-binding domain that the valine, arginine, and asparagine side chains of the extended homology region (EHR; Fig. 1) are critical for recognition of the flanking region (Kusunoki et al., 2002). Importantly, these three amino

acids are conserved only among the Maf family proteins of various species, from Drosophila to human (Fig. 1). Furthermore, within the conserved basic regions of the bZip factors, Maf proteins differ from the others by two amino acids; Mafs contain the sequence NXGYAXXCR, whereas the others contain the sequence NXXAAXXCR. Thus, the alanine is substituted by tyrosine in Mafs and the glycine residue is unique to Maf family members (Fig. 1). The tyrosine and glycine residues in the basic domain of the Maf proteins facilitate the conformational change required for Maf recognition of the GC flanking sequence (Dlakic et al., 2001). In summary, three amino acids in the EHR, valine, arginine, and asparagine, and two amino acids in the basic region, tyrosine and glycine, form the structural basis for the binding specificity of Maf family proteins.

5.2. Cross-talk among CNC-small Maf heterodimers, Maf homodimers and AP-1

What is the biological significance of the unique DNA-binding specificity of Maf proteins? Gene targeting/loss-of-function studies have shed light on the function of large Maf proteins and identified target genes in vivo. There was a severe impairment in the differentiation of lens fiber cells in mice lacking c-Maf and the expression of a subset of crystallin genes was abolished (Kawauchi et al., 1999). Cytological analysis also suggested that c-Maf plays critical roles in T-cell differentiation (Kim et al., 1999a). Conditional gene disruption must be carried out to evaluate the contribution of c-Maf to lymphopoiesis, since c-*maf*-null mutants experience lethality just after birth. The gene affected in mice with the *kreisler* (*kr*) mutation was identified as *mafB*. This mutant mouse exhibited abnormal behavior as characterized by head tossing and running in circles because of abnormal segmentation of the hindbrain and defective differentiation of the inner ear (Cordes and Barsh, 1994). MafB regulates the expression of the hoxb3 gene in rhombomere 5 in the hindbrain of mouse embryos (Manzanares et al., 1997).

Nrl is expressed exclusively in retina and neuronal cells (Swaroop et al., 1992). Mutations in human Nrl were found to be associated with autosomal dominant retinitis pigmentosa (Bessant et al., 1999). Nrl-deficient mutant mice exhibited abnormal differentiation of photoreceptor cells (Mears et al., 2001). As a result, rod cell activity was completely lost and the functional transformation of rod cells to a subtype of cone cells was observed. Chicken L-Maf, whose counterpart has not been identified in the mouse, is also a key regulator of lens differentiation (Ogino and Yasuda, 1998). Small Maf-deficient mice display defective thrombopoiesis, neuronal dysfunction, and altered protein composition in red cell membranes (Shavit et al., 1998; Onodera et al., 2000). Combined, these findings imply that Maf proteins are involved in the processes of cell differentiation and maintenance of homeostasis.

Various MAREs can be subdivided into several classes (Kerppola and Curran, 1994). One type of MARE is bound by both AP-1 and Maf dimers, whereas another is bound exclusively by Maf dimers. Conservation of the flanking region is essential for proper recognition by Mafs and the difference in affinity may be attributed to a variation of the TRE core sequence within MAREs. A well-conserved TRE core sequence enables AP-1 to bind to the MARE, while a MARE containing an atypical TRE core is bound exclusively by Maf family proteins. Indeed, Kataoka et al. showed that the common target genes of Maf and AP-1 are responsible for cell transformation, while Maf-specific target genes are not (Kataoka et al., 2001). The latter group of genes might be involved in cell differentiation and maturation, which has been suggested by the analysis of maf gene knockout mice. Loss-of-function analysis has not succeeded in interpreting the contribution of MARE-dependent gene regulation to cell proliferation or oncogenesis, probably because of the redundant regulation by the many bZip transcription factors. Deciphering the complex gene regulatory network composed of Maf and CNC families of bZip transcription factors represents a challenging project. This would indeed form the prototype for investigation of the transcriptional mechanisms cooperatively regulated by multiple transcription factors.

Acknowledgements

This work was supported in part by JSPS, CREST, PROBRAIN and the Ministry of Education, Science, Sports and Culture of Japan.

References

Aoki, Y., Sato, H., Nishimura, N., Takahashi, S., Itoh, K., Yamamoto, M., 2001. Accelerated DNA adduct formation in the lung of the Nrf2 knockout mouse exposed to diesel exhaust. Toxicol. Appl. Pharmacol. 173, 154–160.

Andrews, N.C., Erdjument-Bromage, H., Davidson, M.B., Tempst, P., Orkin, S.H., 1993a. Erythroid transcription factor NF-E2 is a haematopoietic-specific basic-leucine zipper protein. Nature 362, 722–728.

Andrews, N.C., Kotkow, K.J., Ney, P.A., Erdjument-Bromage, H., Tempst, P., Orkin, S.H., 1993b. The ubiquitous subunit of erythroid transcription factor NF-E2 is a small basic-leucine zipper protein related to the v-maf oncogene. Proc. Natl. Acad. Sci. USA 90, 11488–11492.

Benkhelifa, S., Provot, S., Lecoq, O., Pouponnot, C., Calothy, G., Felder-Schmittbuhl, M.P., 1998. mafA, a novel member of the maf proto-oncogene family, displays developmental regulation and mitogenic capacity in avian neuroretina cells. Oncogene 17, 247–254.

Bessant, D.A., Payne, A.M., Mitton, K.P., Wang, Q.L., Swain, P.K., Plant, C., Bird, A.C., Zack, D.J., Swaroop, A., Bhattacharya, S.S., 1999. A mutation in NRL is associated with autosomal dominant retinitis pigmentosa. Nat. Genet. 4, 355–356.

Bloom, D., Dhakshinamoorthy, S., Jaiswal, A.K., 2002. Site-directed mutagenesis of cysteine to serine in the DNA binding region of Nrf2 decreases its capacity to upregulate antioxidant response element-mediated expression and antioxidant induction of NAD(P)H:quinone oxidoreductase gene. Oncogene 21, 2191–2200.

Caterina, J.J., Donze, D., Sun, C.W., Ciavatta, D.J., Townes, T.M., 1994. Cloning and functional characterization of LCR-F1: a bZIP transcrip-

tion factor that activates erythroid-specific, human globin gene expression. Nucleic Acids Res. 22, 2383–9231.

Chan, K., Kan, Y.W., 1999. Nrf2 is essential for protection against acute pulmonary injury in mice. Proc. Natl. Acad. Sci. USA 96, 12731–12736.

Chan, J.Y., Han, X.-L., Kan, Y.W., 1993. Cloning of Nrf1, an NF-E2-related transcription factor, by genetic selection in yeast. Proc. Natl. Acad. Sci. USA 90, 11371–11375.

Chan, K., Lu, R., Chang, J.C., Kan, Y.W., 1996. NRF2, a member of the NFE2 family of transcription factors, is not essential for murine erythropoiesis, growth, and development. Proc. Natl. Acad. Sci. USA 93, 13943–13948.

Chan, J.Y., Kwong, M., Lu, R., Chang, J., Wang, B., Yen, T.S.B., Kan, Y.W., 1998. Targeted disruption of the ubiquitous CNC-bZIP transcription factor, Nrf-1, results in anemia and embryonic lethality in mice. EMBO J. 17, 1779–1787.

Chan, J.Y., Kwong, M., Lo, M., Emerson, R., Kuypers, F.A., 2001. Reduced oxidative-stress response in red blood cells from p45NFE2-deficient mice. Blood 97, 2151–2158.

Cho, H.-Y., Jedlicka, A.E., Reddy, S.P.M., Zhang, L.-Y., Kensler, T.W., Kleeberger, S.R., 2002. Linkage analysis of susceptibility to hyperoxia. Nrf2 is a candidate gene. Am. J. Respir. Cell Mol. Biol. 26, 21–51.

Cordes, S.P., Barsh, G.S., 1994. The mouse segmentation gene kr encodes a novel basic domain-leucine zipper transcription factor. Cell 79, 1025–1034.

Cox, T.C., Bawden, M.J., Martin, A., May, B.K., 1991. Human erythroid 5-aminolevulinate synthase: promoter analysis and identification of an iron-responsive element in the mRNA. EMBO J. 10, 1891–1902.

Dhakshinamoorthy, S., Jaiswal, A.K., 2000. Small Maf (MafG and MafK) proteins negatively regulate antioxidant response element-mediated expression and antioxidant induction of the NAD(P)H:Quinone oxidoreductase1 gene. J. Biol. Chem. 275, 40134–40141.

Dlakic, M., Grinberg, A.V., Leonard, D.A., Kerppola, T.K., 2001. DNA sequence-dependent folding determines the divergence in binding specificities between Maf and other bZip proteins. EMBO J. 20, 828–840.

Ellenberger, T.E., Brandl, C.J., Struhl, K., Harrison, S.C., 1992. The GCN4 basic region leucine zipper binds DNA as a dimer of uninterrupted alpha helices: crystal structure of the protein-DNA complex. Cell 71, 1223–1237.

Enomoto, A., Itoh, K., Nagayoshi, E., Haruta, J., Kimura, T., O'Connor, T., Harada, T., Yamamoto, M., 2001. High sensitivity of Nrf2 knockout mice to acetaminophen hepatotoxicity associated with decreased expression of ARE-regulated drug metabolizing enzymes and antioxidant genes. Toxicol. Sci. 59, 169–177.

Forsberg, E.C., Downs, K.M., Bresnick, E.H., 2000. Direct interaction of NF-E2 with hypersensitive site 2 of the β-globin locus control region in living cells. Blood 96, 334–339.

Francastel, C., Magis, W., Groudine, M., 2001. Nuclear relocation of a transactivator subunit precedes target gene activation. Proc. Natl. Acad. Sci. USA 98, 12120–12125.

Friling, R.S., Bensimon, A., Tichauer, Y., Daniel, V., 1990. Xenobiotic-inducible expression of murine glutathione S-transferase Ya subunit gene is controlled by an electrophile-responsive element. Proc. Natl. Acad. Sci. USA 87, 6258–6262.

Fujiwara, K.T., Kataoka, K., Nishizawa, M., 1993. Two new members of the maf oncogene family, mafK and mafF, encode nuclear b-Zip proteins lacking putative trans-activator domain. Oncogene 8, 2371–2380.

Glover, J.N., Harrison, S.C., 1995. Crystal structure of the heterodimeric bZIP transcription factor c-Fos-c-Jun bound to DNA. Nature 373, 257–261.

He, C.H., Gong, P., Hu, B., Stewart, D., Choi, M.E., Choi, A.M., Alam, J., 2001. Identification of activating transcription factor 4 (ATF4) as an Nrf2-interacting protein. Implication for heme oxygenase-1 gene regulation. J. Biol. Chem. 276, 20858–20865.

Hoshino, H., Kobayashi, A., Yoshida, M., Kudo, N., Oyake, T., Motohashi, H., Hayashi, N., Yamamoto, M., Igarashi, K., 2000. Oxidative stress abolishes leptomycin B-sensitive nuclear export of transcription repressor Bach2 that counteracts activation of Maf recognition element. J. Biol. Chem. 275, 15370–15376.

Huang, H.C., Nguyen, T., Pickett, C.B., 2000. Regulation of the antioxidant response element by protein kinase C-mediated phosphorylation of NF-E2-related factor 2. Proc. Natl. Acad. Sci. USA 97, 12475–12480.

Igarashi, K., Kataoka, K., Itoh, K., Hayashi, N., Nishizawa, M., Yamamoto, M., 1994. Regulation of transcription by dimerization of erythroid factor p45 with small Maf proteins. Nature 367, 568–572.

Igarashi, K., Itoh, K., Motohashi, H., Hayashi, N., Matuzaki, Y., Nakauchi, H., Nishizawa, M., Yamamoto, M., 1995. Activity and expression of murine small Maf family protein MafK. J. Biol. Chem. 270, 7615–7624.

Igarashi, K., Hoshino, H., Muto, A., Suwabe, N., Nishikawa, S., Nakauchi, H., Yamamoto, M., 1998. Multivalent DNA binding complex generated by small Maf and Bach1 as a possible biochemical basis for beta-globin locus control region complex. J. Biol. Chem. 273, 11783–11790.

Ishii, T., Itoh, K., Takahashi, S., Sato, H., Yanagawa, T., Katoh, Y., Bannai, S., Yamamoto, M., 2000. Transcription factor Nrf2 coordinately regulates a group of oxidative stress-inducible genes in macrophages. J. Biol. Chem. 275, 16023–16029.

Itoh, K., Igarashi, K., Hayashi, N., Nishizawa, M., Yamamoto, M., 1995. Cloning and characterization of a novel erythroid cell-derived CNC family transcription factor heterodimerizing with the small Maf family proteins. Mol. Cell. Biol. 15, 4184–4193.

Itoh, K., Chiba, T., Takahashi, S., Ishii, T., Igarashi, K., Katoh, Y., Oyake, T., Hayashi, N., Satoh, K., Hatayama, I., Yamamoto, M., Nabeshima, Y., 1997. An Nrf2/small Maf heterodimer mediates the induction of phase II detoxifying enzyme genes through antioxidant response elements. Biochem. Biophys. Res. Commun. 236, 313–322.

Itoh, K., Wakabayashi, N., Katoh, Y., Ishii, T., Igarashi, K., Engel, J.D., Yamamoto, M., 1999. Keap1 represses nuclear activation of antioxidant responsive elements by Nrf2 through binding to the amino-terminal Neh2 domain. Genes Dev. 13, 76–86.

Kajihara, M., Kawauchi, S., Kobayashi, M., Ogino, H., Takahashi, S., Yasuda, K., 2001. Isolation, characterization, and expression analysis of zebrafish large Mafs. J. Biochem. (Tokyo) 129, 139–146.

Kataoka, K., Nishizawa, M., Kawai, S., 1993. Structure–function analysis of the *maf* oncogene product, a member of the b-zip protein family. J. Virol. 67, 2133–2141.

Kataoka, K., Fujiwara, K.T., Noda, M., Nishizawa, M., 1994a. MafB, a new Maf family transcription activator that can associate with Maf and Fos but not with Jun. Mol. Cell. Biol. 14, 7581–7591.

Kataoka, K., Noda, M., Nishizawa, M., 1994b. Maf nuclear oncoprotein recognizes sequences related to an AP-1 site and forms heterodimers with both Fos and Jun. Mol. Cell. Biol. 14, 700–712.

Kataoka, K., Igarashi, K., Itoh, K., Fujiwara, K.T., Noda, M., Yamamoto, M., Nishizawa, M., 1995. Small Maf proteins heterodimerize with Fos and potentially act as competitive repressors of NF-E2 transcription factor. Mol. Cell. Biol. 15, 2180–2190.

Kataoka, K., Shioda, S., Yoshitomo-Nakagawa, K., Handa, H., Nishizawa, M., 2001. Maf and Jun nuclear oncoproteins share downstream target genes for inducing cell transformation. J. Biol. Chem. 276, 36849–36856.

Katsuoka, F., Motohashi, H., Onodera, K., Suwabe, N., Engel, D.J., Yamamoto, M., 2000. One enhancer mediates *mafK* transcriptional activation in both hematopoietic and cardiac muscle cells. EMBO J. 19, 2980–2991.

Katoh, Y., Itoh, K., Yoshida, E., Miyagishi, M., Fukamizu, A., Yamamoto, M., 2001. Two domains of Nrf2 cooperatively bind CBP, a CREB binding proteins, and synergistically activate transcription. Genes Cells 6, 857–868.

Kawauchi, S., Takahashi, S., Nakajima, O., Ogino, H., Morita, M., Nishizawa, M., Yasuda, K., Yamamoto, M., 1999. Regulation of lens fiber cell differentiation by transcription factor c-Maf. J. Biol. Chem. 274, 19254–19260.

Kerppola, T.K., Curran, T., 1994. A conserved region adjacent to the basic domain is required for recognition of an extended DNA binding site by Maf/Nrl family proteins. Oncogene 9, 3149–3158.

Kim, J.I., Ho, I.-C., Grusby, M.J., Glimcher, L.H., 1999a. The transcription factor c-Maf controls the production of interleukin-4 but not other Th2 cytokines. Immunity 10, 745–751.

Kim, J.I., Li, T., Ho, I.C., Grusby, M.J., Glimcher, L.H., 1999b. Requirement for the c-Maf transcription factor in crystallin gene regulation and lens development. Proc. Natl. Acad. Sci. USA 96, 3781–3785.

Kobayashi, A., Ito, E., Toki, T., Kogame, K., Takahashi, S., Igarashi, K., Hayashi, N., Yamamoto, M., 1999. Molecular cloning and functional characterization of a new Cap'n' collar family transcription factor Nrf3. J. Biol. Chem. 274, 6443–6452.

Kotkow, K.J., Orkin, S.H., 1996. Complexity of the erythroid transcription factor NF-E2 as revealed by gene targeting of the mouse p18 NF-E2 locus. Proc. Natl. Acad. Sci. USA 93, 3514–3518.

Kuroha, T., Takahashi, S., Komeno, T., Itoh, K., Nagasawa, T., Yamamoto, M., 1998. Ablation of Nrf2 function does not increase the erythroid or megakaryocytic cell lineage dysfunction caused by p45 NF-E2 gene disruption. J. Biochem. (Tokyo) 123, 376–379.

Kusunoki, H., Motohashi, H., Katsuoka, F., Morohashi, A., Yamamoto, M., Tanaka, T., 2002. Solution structure of the DNA-binding domain of MafG. Nat. Struct. Biol. 4, 252–256.

Kwak, M.-K., Itoh, K., Yamamoto, M., Kensler, T.W., 2002. Enhanced expression of the transcription factor Nrf2 by Cancer chemopreventive agents: role of antioxidant response element-like sequences in the nrf2 promoter. Mol. Cell. Biol. 22, 2883–2892.

Kwong, M., Kan, Y.W., Chan, J.Y., 1999. The CNC basic leucine zipper factor, Nrf1, is essential for cell survival in response to oxidative stress-inducing agents. J. Biol. Chem. 274, 37491–37498.

Lecine, P., Villeval, J.L., Vyas, P., Swencki, B., Xu, Y., Shivdasani, R.A., 1998. Mice lacking transcription factor NF-E2 provide in vivo validation of the proplatelet model of thrombocytopoiesis and show a platelet production defect that is intrinsic to megakaryocytes. Blood 92, 1608–1616.

Lim, K.C., Ishihara, H., Riddle, R.D., Yang, Z., Andrews, N., Yamamoto, M., Engel, J.D., 1994. Structure and regulation of the chicken erythroid delta-aminolevulinate synthase gene. Nucleic Acids Res. 22, 1226–1233.

Manzanares, M., Cordes, S., Kwan, C.T., Sham, M.H., Barsh, G.S., Krumlauf, R., 1997. Segmental regulation of Hoxb-3 by kreisler. Nature 387, 191–195.

Martin, D.I., Fiering, S., Groudine, M., 1996. Regulation of beta-globin gene expression: straightening out the locus. Curr. Opin. Genet. Dev. 6, 488–495.

Martin, F., van Deursen, J.M., Shivdasani, R.A., Jackson, C.W., Troutman, A.G., Ney, P.A., 1998. Erythroid maturation and globin gene expression in mice with combined deficiency of NF-E2 and Nrf-2. Blood 91, 3459–3466.

McMahon, M., Itoh, K., Yamamoto, M., Chanas, S.A., Henderson, C.J., McLellan, L.I., Wolf, C.R., Cavin, C., Hayes, J.D., 2001. The Cap'n'Collar basic leucine zipper transcription factor Nrf2 (NF-E2 p45-related factor 2) controls both constitutive and inducible expression of intestinal detoxification and glutathione biosynthetic enzymes. Cancer Res. 61, 3299–3307.

Mears, A.J., Kondo, M., Swain, P.K., Takada, Y., Bush, R.A., Saunders, T.L., Sieving, P.A., Swaroop, A., 2001. Nrl is required for rod photoreceptor development. Nat. Genet. 29, 447–452.

Mignotte, V., Eleouet, J.F., Raich, N., Romeo, P.-H., 1989a. Cis- and trans-acting elements involved in the regulation of the erythroid promoter of the human porphobilinogen deaminase gene. Proc. Natl. Acad. Sci. USA 86, 6548–6552.

Mignotte, V., Wall, L., deBoer, E., Grosveld, F., Romeo, P.-H., 1989b. Two tissue-specific factors bind the erythroid promoter of the human porphobilinogen deaminase gene. Nucleic Acids Res. 17, 37–54.

Mohler, J., Vani, K., Leung, S., Epstein, A., 1991. Segmentally restricted, cephalic expression of a leucine zipper gene during Drosophila embryogenesis. Mech. Dev. 34, 3–9.

Moi, P., Chan, K., Asunis, I., Cao, A., Kan, Y.W., 1994. Isolation of NF-E2-related factor 2 (Nrf2), a NF-E2-like basic leucine zipper transcriptional activator that binds to the tandem NF-E2/AP1 repeat of the β-globin locus control region. Proc. Natl. Acad. Sci. USA 91, 9926–9930.

Motohashi, H., Igarashi, K., Onodera, K., Takahashi, S., Ohtani, H., Nakafuku, M., Nishizawa, M., Engel, J.D., Yamamoto, M., 1996. Mesodermal- vs. neuronal-specific expression of MafK is elicited by different promoters. Genes Cells 1, 223–238.

Motohashi, H., Ohta, J., Engel, J.D., Yamamoto, M., 1998. A core region of the mafK gene IN promoter directs neurone-specific transcription in vivo. Genes Cells 3, 671–684.

Motohashi, H., Katsuoka, F., Shavit, J., Engel, J.D., Yamamoto, M., 2000. Positive or negative MARE-dependent transcriptional regulation is determined by the abundance of small Maf proteins. Cell 103, 865–875.

Muto, A., Hoshino, H., Madisen, L., Yanai, N., Obinata, M., Karasuyama, H., Hayashi, N., Nakauchi, H., Yamamoto, M., Groudine, M., Igarashi, K., 1998. Identification of Bach2 as a B-cell-specific partner for small Maf proteins that negatively regulate the immunoglobulin heavy chain gene 3′ enhancer. EMBO J. 17, 5734–5743.

Nagai, T., Igarashi, K., Akasaka, J., Furuyama, K., Fujita, H., Hayashi, N., Yamamoto, M., Sassa, S., 1998. Regulation of NF-E2 activity in erythroleukemia cell differentiation. J. Biol. Chem. 273, 5358–5365.

Nishizawa, M., Kataoka, K., Goto, N., Fujiwara, K.T., Kawai, S., 1989. v-maf, a viral oncogene that encodes a leucine zipper motif. Proc. Natl. Acad. Sci. USA 86, 7711–7715.

Ogawa, K., Sun, J., Taketani, S., Nakajima, O., Nishitani, C., Sassa, S., Hayashi, N., Yamamoto, M., Shibahara, S., Fujita, H., Igarashi, K., 2001. Heme mediates derepression of Maf recognition element through direct binding to transcription repressor Bach1. EMBO J. 20, 2835–2843.

Ogino, H., Yasuda, K., 1998. Induction of lens differentiation by activation of a bZip transcription factor, L-Maf. Science 280, 115–118.

Onodera, K., Shavit, J.A., Motohashi, H., Kastuoka, F., Akasaka, J., Engel, J.D., Yamamoto, M., 1999. Characterization of the murine *mafF* gene. J. Biol. Chem. 274, 21162–21169.

Onodera, K., Shavit, J.A., Motohashi, H., Yamamoto, M., Engel, D.J., 2000. Perinatal synthetic lethality and hematopoietic defects in compound *mafG::mafK* mutant mice. EMBO J. 19, 1335–1345.

Oyake, T., Itoh, K., Motohashi, H., Hayashi, N., Hoshino, H., Nishizawa, M., Yamamoto, M., Igarashi, K., 1996. Bach proteins belong to a novel family of BTB-basic leucine zipper transcription factors that interact with MafK and regulate transcription through the NF-E2 site. Mol. Cell. Biol. 16, 6083–6095.

Ramos-Gomez, M., Kwak, M.-K., Dolan, P.M., Itoh, K., Yamamoto, M., Talalay, P., Kensler, T.W., 2001. Sensitivity to carcinogenesis is increased and chemoprotective efficacy of enzyme inducers is lost in *nrf2* transcription factor-deficient mice. Proc. Natl. Acad. Sci. USA 98, 3410–3415.

Rushmore, T.H., Morton, M.R., Pickett, C.B., 1991. The antioxidant responsive element. Activation by oxidative stress and identification of the DNA consensus sequence required for functional activity. J. Biol. Chem. 266, 11632–11639.

Sakai, M., Imaki, J., Yoshida, K., Ogata, A., Matsushima-Hibaya, Y., Kuboki, Y., Nishizawa, M., Nishi, S., 1997. Rat maf related genes: specific expression in chondrocytes, lens and spinal cord. Oncogene 14, 745–750.

Sawado, T., Igarashi, K., Groudine, M., 2001. Activation of beta-major globin gene transcription is associated with recruitment of NF-E2 to the beta-globin LCR and gene promoter. Proc. Natl. Acad. Sci. USA 98, 10226–10231.

Shavit, J.A., Motohashi, H., Onodera, K., Akasaka, J., Yamamoto, M., Engel, J.D., 1998. Impaired megakaryopoiesis and behavioral defects in mafG-null mutant mice. Genes Dev. 12, 2164–2174.

Shivdasani, R.A., Rosenblatt, M.F., Zucker-Franklin, D., Jackson, C.W.,

Hunt, P., Saris, C.J.M., Orkin, S.H., 1995. Transcription factor NF-E2 is required for platelet formation independent of the actions of thrombopoietin/MGDF in megakaryocyte development. Cell 81, 695–704.

Swaroop, A., Xu, J., Pawar, H., Jackson, A., Skolnick, C., Agarwal, N., 1992. A conserved retina-specific gene encoded a basic motif/leucine zipper domain. Proc. Natl. Acad. Sci. USA 89, 266–270.

Taketani, S., Inazawa, J., Nakahashi, Y., Abe, T., Tokunaga, R., 1992. Structure of the human ferrochelatase gene. Exon/intron gene organization and location of the gene to chromosome 18. Eur. J. Biochem. 205, 217–222.

Talbot, D., Grosveld, F., 1991. The 5′HS2 of the globin locus control region enhances transcription through the interaction of a multimeric complex binding at two functionally distinct NF-E2 binding sites. EMBO J. 10, 1391–1398.

Toki, T., Itoh, J., Kitazawa, J., Arai, K., Hatakeyama, K., Akasaka, J., Igarashi, K., Nomura, N., Yokoyama, M., Yamamoto, M., Ito, E., 1997. Human small Maf proteins form heterodimers with CNC family transcription factors and recognize the NF-E2 motif. Oncogene 14, 1901–1910.

Venugopal, R., Jaiswal, A.K., 1998. Nrf2 and Nrf1 in association with Jun proteins regulate antioxidant response element-mediated expression and coordinated induction of genes encoding detoxifying enzymes. Oncogene 17, 3145–3156.

Yoshida, C., Tokumasu, F., Hohmura, K.I., Bungert, J., Hayashi, N., Nagasawa, T., Engel, J.D., Yamamoto, M., Takeyasu, K., Igarashi, K., 1999. Long range interaction of cis-DNA elements mediated by architectural transcription factor Bach1. Genes Cells. 4, 643–655.

Zipper, L.M., Mulcahy, R.T., 2000. Inhibition of ERK and p38 MAP kinases inhibits binding of Nrf2 and induction of GCS genes. Biochem. Biophys. Res. Commun. 278, 484–492.

Solution structure of the DNA-binding domain of MafG

Hideki Kusunoki[1,2], Hozumi Motohashi[1,2], Fumiki Katsuoka[1], Akio Morohashi[3,4], Masayuki Yamamoto[1,2] and Toshiyuki Tanaka[2,3]

[1]*Institute of Basic Medical Science, University of Tsukuba, Tsukuba, Ibaraki 305-8575, Japan.* [2]*Center for Tsukuba Advanced Research Alliance, University of Tsukuba, Tsukuba, Ibaraki 305-8577, Japan.* [3]*Institute of Applied Biochemistry, University of Tsukuba, Tsukuba, Ibaraki 305-8572, Japan.* [4]*Banyu Tsukuba Research Institute, Tsukuba, Ibaraki 300-2611, Japan.*

Published online: 4 March 2002, DOI: 10.1038/nsb771

The Maf family proteins, which constitute a subgroup of basic region-leucine zipper (bZIP) proteins, function as transcriptional regulators of cellular differentiation. Together with the basic region, the Maf extended homology region (EHR), conserved only within the Maf family, defines the DNA binding specific to Mafs. Here we present the first NMR-derived structure of the DNA-binding domain (residues 1–76) of MafG, which contains the EHR and the basic region. The structure consists of three α-helices and resembles the fold of the DNA-binding domain of Skn-1, a developmental transcription factor of *Caenorhabditis elegans*. The structural similarity between MafG and Skn-1 enables us to propose a possible mechanism by which Maf family proteins recognize their consensus DNA sequences.

The Maf family of transcription factors are basic region-leucine zipper (bZIP) proteins, characterized by the presence of a specific amino acid sequence, the Maf extended homology region (EHR), located on the N-terminal side of the basic region[1,2]. So far, seven family members have been reported and subdivided into two groups: large Mafs (26–39 kDa), which have a putative transactivation domain at their N-termini, and small Mafs (17–18 kDa), which lack such a domain[1–3] (Fig. 1*a*). Maf family proteins play important roles in cellular differentiation and morphogenesis, and their unique functions seem to be determined by their selective dimerization and DNA-binding specificity[1–3]. Mafs can form homodimers or heterodimers with other bZIP transcription factors, such as AP-1 and Cap-N-Collar (CNC) family members, through their leucine zipper domains. For example, small Mafs are known to dimerize with p45 (the large subunit of the nuclear factor-erythroid 2, NF-E2) to activate erythroid cell-specific transcription[1–3].

As for DNA-binding specificity, Maf family proteins recognize a relatively long palindromic DNA sequence, TGC TGACTCAGCA or TGCTGACGTCAGCA, known as the Maf recognition element (MARE). The MARE sequence contains the well-known AP-1 recognition site, the 12-*O*-tetradecanoylphorbol-13-acetate (TPA)-responsive element (TRE; TGACTCA) or the cyclic AMP-responsive element (CRE; TGACGTCA) (Fig. 1*b*)[1–5]. Therefore, Mafs may recognize the TRE or CRE site in a similar way as AP-1 proteins[6]. However, the recognition of extended sequence elements (flanking regions) on both sides of the AP-1 core site distinguishes Mafs from AP-1 proteins in the DNA-binding mode. The DNA-binding specificity of Mafs is believed to be achieved through recognition of the flanking region by the EHR, which is conserved only within the Maf family and not found in other bZIP proteins[2,6]. To understand the unique DNA-binding mode of Mafs, we determined the solution structure of the MafG (a small Maf) DNA-binding domain (residues 1–76), which shares no sequence homology with any protein whose structure has been determined. This structure provides the first opportunity to understand how the EHR recognizes the flanking region.

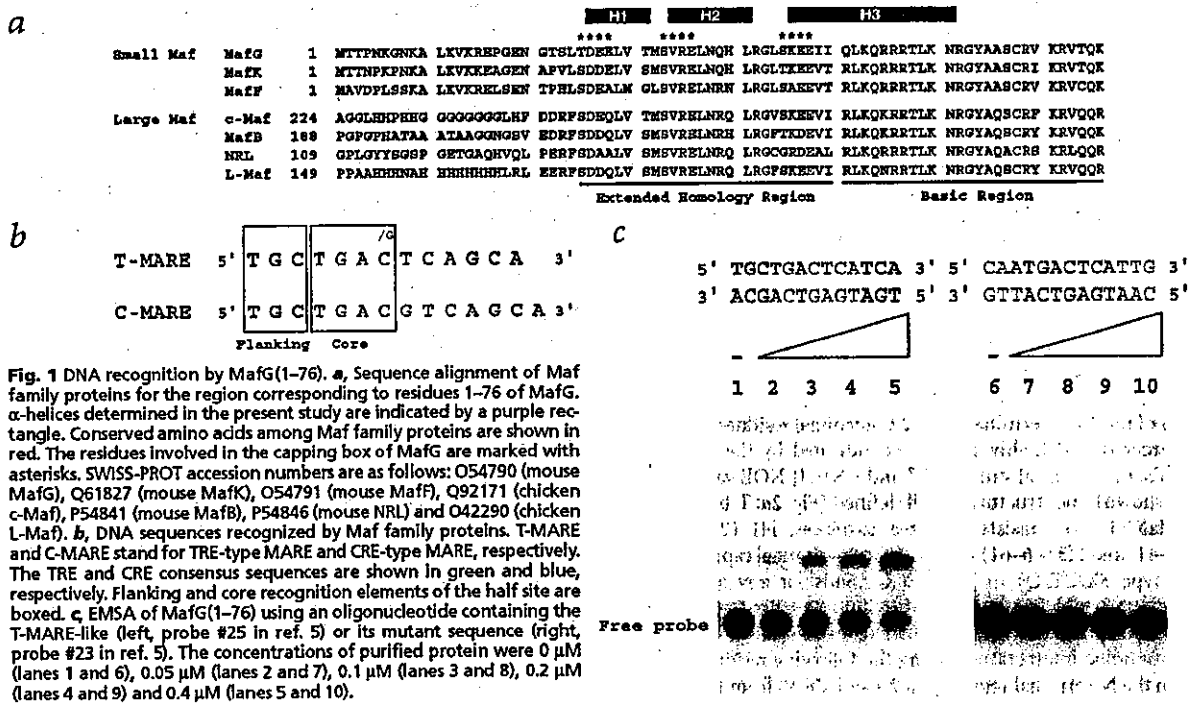

Fig. 1 DNA recognition by MafG(1–76). ***a***, Sequence alignment of Maf family proteins for the region corresponding to residues 1–76 of MafG. α-helices determined in the present study are indicated by a purple rectangle. Conserved amino acids among Maf family proteins are shown in red. The residues involved in the capping box of MafG are marked with asterisks. SWISS-PROT accession numbers are as follows: O54790 (mouse MafG), Q61827 (mouse MafK), O54791 (mouse MafF), Q92171 (chicken c-Maf), P54841 (mouse MafB), P54846 (mouse NRL) and O42290 (chicken L-Maf). ***b***, DNA sequences recognized by Maf family proteins. T-MARE and C-MARE stand for TRE-type MARE and CRE-type MARE, respectively. The TRE and CRE consensus sequences are shown in green and blue, respectively. Flanking and core recognition elements of the half site are boxed. ***c***, EMSA of MafG(1–76) using an oligonucleotide containing the T-MARE-like (left, probe #25 in ref. 5) or its mutant sequence (right, probe #23 in ref. 5). The concentrations of purified protein were 0 μM (lanes 1 and 6), 0.05 μM (lanes 2 and 7), 0.1 μM (lanes 3 and 8), 0.2 μM (lanes 4 and 9) and 0.4 μM (lanes 5 and 10).

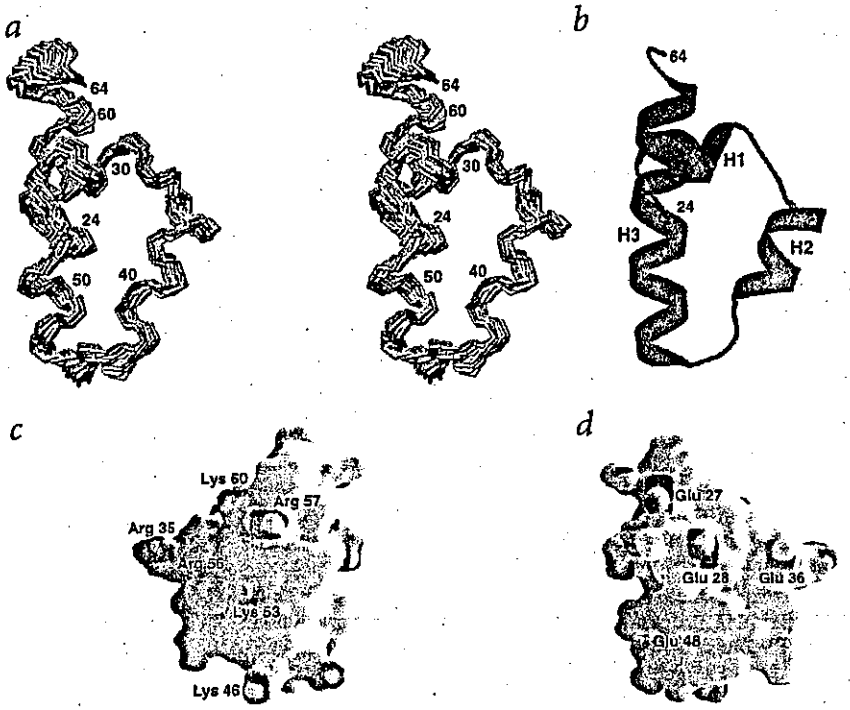

Fig. 2 3D structure of MafG(1–76). *a*, Stereo view of a best-fit superposition of the backbone atoms (N, Cα and C′) of the 20 NMR-derived structures of MafG(1–76). The main chain atoms of the 20 structures are superimposed against the energy-minimized average structure using residues 24–64. The 23 N-terminal and 12 C-terminal residues, which are not well defined because they lack many experimental restraints, are omitted throughout panels (*a–d*). *b*, Ribbon diagram of the energy-minimized average structure of MafG(1–76). The α-helices are shown in purple and labeled. *c,d*, Electrostatic potential surfaces of MafG(1–76). Positive and negative potentials are in blue and red, respectively. The orientation of the image (*d*) is the same as that in panel (*b*). The image in (*c*) is related to that in (*d*) by a 180° rotation along the vertical axis.

DNA-binding ability of MafG

MafG(1–76) contains both the EHR and the basic region, which are necessary for DNA recognition and binding, but lacks the leucine zipper domain required for dimerization. First, we examined the functional properties of MafG(1–76). Electrophoretic mobility shift assays (EMSA) showed that this protein binds to the TRE-type MARE (T-MARE)-like sequence TGCTGACTCATCA with a K_d of 0.3 μM, but not to CAATGACTCATTG, which has mutations introduced in the flanking region (Fig. 1*c*). In addition, a sedimentation equilibrium analysis revealed that MafG(1–76) exists as a monomer in solution, with calculated and measured molecular weights of 9,000 and 9,700 ± 800, respectively. These results indicate that MafG(1–76) alone functions as the DNA-binding domain that retains the DNA-binding specificity of the natural protein. The relatively low affinity of MafG(1–76) for the DNA sequence results from its lack of a leucine zipper domain, which is responsible for stabilizing the protein–DNA complex through dimerization.

Structure description

The 3D structure of MafG(1–76) has been determined on the basis of a total of 843 NMR-derived restraints. With the exception of the 23 N-terminal and 12 C-terminal residues, which are disordered and highly mobile as indicated by the paucity of NOEs, the chemical shift index[7] and $^{15}N\{^1H\}$ NOE values[8] (data not shown), the structure is well defined (Fig. 2*a*; Table 1).

MafG(1–76) consists of three α-helices, H1 (26–31), H2 (34–41) and H3 (46–61) (Fig. 2*b*). An N-terminal capping box of the type SXX(E/Q) or TXXE (X stands for any amino acid residue)[9], which is known to function as a helix stop signal, is found in all helices (Fig. 1*a*). The helices are stabilized by hydrophobic interactions among the following residues: Leu 24 from the N-terminal region; Leu 29 and Val 30 from H1; Met 32 from a turn between H1 and H2; Leu 37 and Leu 41 from H2; Leu 44 from a turn between H2 and H3; and Ile 49, Leu 52 and Leu 59 from H3. Conservation of these hydrophobic residues and the N-terminal capping boxes among Mafs (Fig. 1*a*) indicates that there is a similarity in the 3D structures of their DNA-binding domains. The EHR constitutes the region from helix H1 to the first one-third of helix H3, whereas the basic region comprises the rest of helix H3 and the C-terminal flexible part.

Most of the basic amino acid residues reside on one side of the protein surface (Fig. 2*c,d*): Arg 35 from H2 and Lys 46, Lys 53, Arg 56, Arg 57 and Lys 60 from H3. These basic residues, except for Lys 46, are completely conserved within the Maf family (Fig. 1*a*). In addition, loss of DNA binding was reported for the R57E mutant of MafG and the corresponding mutants of MafK, MafF and MafB[4,5]. Therefore, the basic cluster on the protein surface is probably involved in the interaction with target DNA sequences. We will refer to this surface as the DNA-binding surface.

DNA-binding site of MafG(1–76)

To identify the DNA-binding site of MafG(1–76), we performed an NMR titration, in which the 2D ^1H-^{15}N HSQC spectra of MafG(1–76) were recorded with successive additions of DNA containing the T-MARE-like sequence. Large chemical shift changes were observed upon DNA binding, mainly in the region from helix H2 to the C-terminus (Fig. 3). This region contains both the EHR and the basic region, indicating their involvement in the interaction with DNA. The changes continued on the addition of up to half the equivalent of DNA, and only a single set of protein signals was observed throughout the titration experiment. These results indicate that MafG(1–76) binds to DNA as a 2:1 complex with two-fold symmetry and that the exchange between free and bound protein conformations is fast on the NMR time scale.

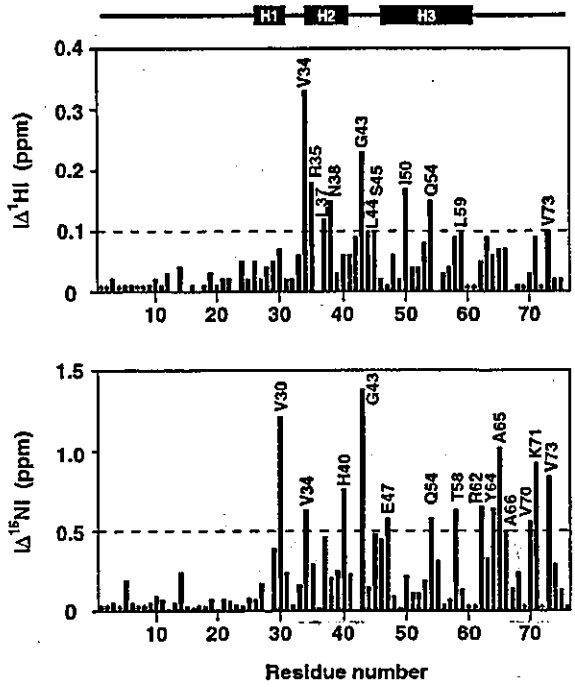

Fig. 3 Chemical shift changes of amide protons (upper panel) and nitrogens (lower panel) of MafG(1–76) upon binding to the DNA containing the T-MARE-like sequence. Absolute chemical shift differences between DNA-free and DNA-bound states are plotted *versus* residue number. Residues showing large chemical shift changes (≥0.1 p.p.m. for ^{1}H or ≥0.5 p.p.m. for ^{15}N) are marked with a one-letter code and a residue number. Unassigned residues in either state and Pro residues are indicated using asterisks. A representation of the secondary structure of MafG(1–76) is shown at the top.

Comparison of MafG(1–76) with other DNA-binding domains

A helix-turn-helix (HTH) motif, one of the most common DNA-binding motifs, is found in many eukaryotic and prokaryotic transcriptional regulatory proteins. The second helix of the motif usually binds to the major groove of DNA and, therefore, is mainly involved in recognition and interaction with specific DNA sequences[10]. MafG(1–76) contains a similar HTH motif (H2 and H3), but the interhelical angle between the two helices (149°) is considerably larger than the typical value for the common HTH motif (~110°)[10]. Thus, MafG(1–76) may have a protein–DNA interaction mode distinct from that of the ordinary HTH-type DNA-binding domains.

Recently, the 3D structure of the DNA-binding domain of Skn-1, a developmental transcription factor from *Caenorhabditis elegans*, was determined in both DNA-free[11] and DNA-bound states[12]. This domain shares 25% sequence homology with MafG(1–76), but their sequence similarity has not been identified by a database search. The structured region of MafG resembles the fold of the Skn-1 DNA-binding domain (backbone root mean square (r.m.s.) deviation of 1.2 Å for helices H1, H2 and H3), although MafG lacks an N-terminal helix (H0) and has a shorter C-terminal helix (H3) (Fig. 4*a*). Their structural similarity is supported by the following features of the three helices: similar interhelical angles, a capping box sequence at the N-terminus and conservation of the hydrophobic residues involved in their stabilization (Fig. 4*a*,*c*). In addition, a basic cluster is present on the surface of Skn-1 in a position similar to that of MafG(1–76) (Fig. 4*b*).

DNA-binding specificity of Maf proteins

The structural similarity between Skn-1 and MafG provides insight into the mechanism by which Maf family proteins might recognize the consensus sequences (Fig. 1*b*). Skn-1 achieves selectivity for the consensus sequence (G/A)TCAT, an AP-1-like half site[11,12], by direct interaction through the five residues in the NXXAAXXCR sequence of its basic region (Fig. 4*d*)[12]. In addition, many of the basic residues in the basic region (closed circles, Fig. 4*c*) make contact with the phosphate backbones of the DNA[12]. A similar amino acid sequence, NXXYAXXCR, and most of the basic residues are conserved in the basic region of Mafs (Figs 1*a*, 4*c*) and form a DNA-binding surface analogous to that of Skn-1 (Fig. 4*b*). Furthermore, a chemical shift perturbation experiment showed that these residues are involved in the interaction with the MARE sequence (Fig. 3). Considering the above, the Maf family probably uses a DNA-binding mode common to Skn-1 in recognition of the AP-1 site. A similar mode of DNA-binding is also probably used by AP-1 bZIP proteins; they possess the NXXAAXXC(/S)R sequence, and almost identical interactions between the protein and the DNA were found in their complex structures[13,14]. However, complete conservation of a Tyr instead of an Ala residue at the fourth position in the NXXYAXXCR sequence (Fig. 1*a*) indicates the possibility that the basic region of Mafs has a slightly different DNA-binding mode.

How then does the EHR, essential for the unique DNA-binding specificity of Mafs, recognize the flanking region? In the structure of the Skn-1–DNA complex, the hydrophobic side chain of Leu 482 in helix H2 was found to interact with the ribose group of the guanine nucleotide that resides outside the consensus sequence (red, Fig. 4*d*)[12]. Therefore, the corresponding region on the DNA-binding surface of MafG may interact with the nucleotide bases in the flanking region that lie within three base pairs from the AP-1 core region. Side chains of Val 34, Arg 35 and Asn 38 are present in this region (Fig. 4*b*), and the six residues of helix H2 that contain them are completely conserved within the EHR (Fig. 1*a*). In addition, these residues displayed perturbed chemical shifts upon DNA binding (Fig. 3). To assess the roles of these residues in recognition of the flanking region, EMSA was performed using a series of mutants, in which one of the three residues was replaced by an Ala residue. A significant reduction in DNA-binding affinity was observed for all three mutants (V34A, R35A and N38A) (Fig. 4*e*), indicating that Val 34, Arg 35 and Asn 38 on the DNA-binding surface are the most probable residues involved in the binding of the flanking region and the contribution of helix H2 to DNA recognition is quantitatively different between MafG and Skn-1. Truncation of the N-terminus of c-Maf to Glu 259 (corresponding to Glu 36 of MafG) was reported to result in a loss of DNA-binding specificity[6]. This truncated c-Maf was able to bind to both the T-MARE sequence and the sequence whose flanking region was mutated. This suggests that Val 34 and Arg 35 play an important role in the specific recognition of the flanking region.

Conclusions

The present structure determination of MafG(1–76), which shows no sequence homology to any protein of known structure, has revealed a structural feature relevant to its DNA recognition mechanism. Within Maf family proteins, the structure formed by the three helices of the EHR (the EHR fold) is present just

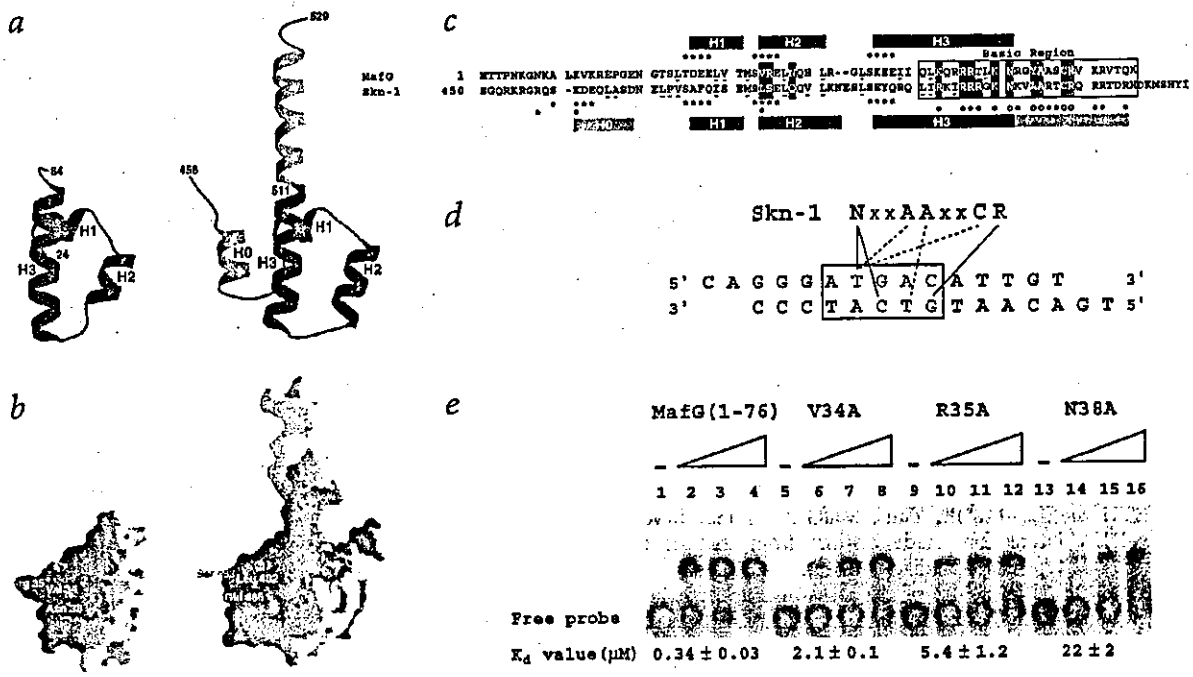

Fig. 4 Comparison of MafG(1–76) with the Skn-1 DNA-binding domain. *a*, DNA-binding domains of MafG (left) and Skn-1 (right) (PDB accession code 1SKN). The helices of MafG and the corresponding regions of Skn-1 are in purple; the other helical regions of Skn-1 are in cyan. *b*, DNA-binding surfaces of MafG (left) and Skn-1 (right). The basic residues of MafG (Lys 53, Arg 56, Arg 57 and Lys 60) and Skn-1 (Arg 503, Arg 506, Arg 507, Arg 508 and Lys 510) are in blue. The five residues of Skn-1 that interact with the DNA bases and the corresponding residues of MafG (only Asn 61 and Tyr 64 in this panel) are in green. Val 34, Arg 35 and Asn 38 of MafG and the corresponding residues of Skn-1 are shown in red. *c*, Sequence alignment of the MafG and Skn-1 DNA-binding domains (SWISS-PROT accession number P34707). Secondary structure elements of MafG (upper) and Skn-1 (lower) are shown. The residues that form the hydrophobic cores of MafG and Skn-1 are underlined. The residues corresponding to the colored residues in panel (*b*) are highlighted in a matching color. The capping box sequences are indicated with asterisks. Open and closed circles show the residues of Skn-1 that interact with the bases and phosphate backbones of DNA, respectively. *d*, DNA sequence used in the structure determination of the Skn-1–DNA complex and its interaction with the NXXAAXXCR sequence of Skn-1. The consensus recognition site of Skn-1 is boxed. The AP-1 core region is in green. Hydrogen bonds and van der Waals contacts are indicated by solid and dotted lines, respectively. *e*, EMSA of MafG(1–76) and its mutants using an oligonucleotide containing the T-MARE-like sequence (probe #25 in ref. 5). The concentrations of purified protein were 0 μM (lanes 1, 5, 9 and 13), 0.2 μM (lanes 2, 6, 10 and 14), 0.8 μM (lanes 3, 7, 11 and 15), and 3 μM (lanes 4, 8, 12 and 16).

before the basic region. This structure enables Mafs to make a broader area of contact with DNA and to recognize longer DNA sequences. In particular, the two residues at the beginning of helix H2 (Val 34 and Arg 35) are positioned to recognize the flanking region. Therefore, disruption of the EHR fold will lead to a loss of DNA-binding affinity and/or specificity of Mafs. This EHR fold is not found in AP-1 family members such as GCN4, c-Fos and c-Jun. The EHR fold is designed for specific recognition of the flanking region and, thus, distinguishes Mafs from other proteins by virtue of its DNA-binding specificity.

Methods

Sample preparation. MafG(1–76) or its mutant protein was expressed in *Escherichia coli* as a His_6-tagged protein. M9 minimal medium containing $^{15}NH_4Cl$ or $^{15}NH_4Cl/^{13}C_6$-D-glucose was used to prepare ^{15}N- or $^{15}N/^{13}C$-enriched protein, respectively. The His_6-tagged protein was purified by a Ni-NTA superflow column (Qiagen) and then cleaved with thrombin. The recombinant protein that contained three extra vector-derived amino acids (Gly-Ser-His) at the N-terminus was further purified using S2 (BioRad), HiLoad 16/60 Superdex 75 pg (Pharmacia) and Benzamidine Sepharose 6B (Pharmacia) columns. NMR samples contained 1.5–2.0 mM protein in 20 mM sodium phosphate buffer, pH 6.7, 10 mM dithiothreitol (DTT)-d_{10} and 10% (v/v) ^2H_2O. Comparison of the proton NMR spectra indicated that none of the mutations disrupted the three-dimensional structure of MafG(1–76).

NMR spectroscopy. NMR spectra were obtained at 25 °C using a Bruker AVANCE DRX 800 or a Varian UNITY INOVA 500 spectrometer. All data were processed using nmrPipe and nmrDraw[15], and analyzed with PIPP[16]. Backbone and side chain ^1H, ^{13}C and ^{15}N resonances were assigned using standard triple resonance experiments[17]. Stereo-specific assignments of Val and Leu methyl groups were obtained by analyzing a 2D ^1H-^{13}C CT-HSQC spectrum on a 10% ^{13}C-enriched sample[18]. Three sets of $^{15}N\{^1H\}$ NOE spectra with and without the NOE effect were recorded and analyzed as described[8].

Structure calculation. Distance restraints were obtained from 2D NOESY and 3D ^{15}N-edited and ^{13}C-edited NOESY experiments (mixing time of 100 ms) as described[19]. Dihedral Φ angle restraints were deduced from $^3J_{NH\alpha}$ coupling constants measured by a 3D HNHA experiment[20]: Φ was restrained to –60 ± 30° for residues with $^3J_{NH\alpha} < 5.5$ Hz. Structure calculations were performed using a simulated annealing protocol[21] within X-PLOR v.3.851 (ref. 22). From the 50 calculated structures, 20 structures with the lowest total restraint energy were selected. Structure figures were generated using MOLMOL[23], MOLSCRIPT[24] and RASTER3D[25], or GRASP[26].

EMSA. Proteins were incubated with 0.25 nM of ^{32}P-labeled DNA probe at 37 °C for 30 min in 10 μl of EMSA buffer (20 mM sodium phosphate, pH 6.5, 10 mM DTT, 4 mM $MgCl_2$, 50 mM NaCl, 0.1 mg ml^{-1} BSA and 0.1 mg ml^{-1} poly(dI-dC)). The resulting mixture was subjected to native polyacrylamide gel electrophoresis and

Table 1 Structural statistics of the 20 structures of MafG(1–76)

R.m.s. deviations[1]	
Experimental distance restraints (Å)	
All (817)	0.031 ± 0.002
Interresidue sequential (|i – j| = 1) (237)	0.028 ± 0.011
Interresidue short range (1 < |i – j| < 5) (200)	0.036 ± 0.005
Interresidue long range (|i – j| ≥ 5) (94)	0.029 ± 0.002
Intraresidue (286)	0.029 ± 0.008
Experimental dihedral restraints (°) (26)	0.27 ± 0.02
Idealized covalent geometry	
Bonds (Å)	0.006 ± 0.0001
Angles (°)	1.73 ± 0.002
Impropers (°)	1.02 ± 0.01
Energies (kcal mol^{-1})	
F_{NOE}[2]	39.1 ± 7.0
F_{cdih}[2]	0.19 ± 0.28
F_{repel}[3]	8.4 ± 2.1
E_{L-J}[4]	–94.5 ± 17.0
Average r.m.s. difference (Å)[5]	
Residues 24–64	0.69 ± 0.17 (1.38 ± 0.13)
Residues in α-helices	0.56 ± 0.13 (1.34 ± 0.13)
Ramachandran map analyses (%)[6]	
Allowed regions	99.2 ± 1.0
Disallowed regions	0.8 ± 1.0

[1]The number of each type of restraint used in the structure calculation is given in parentheses. None of the calculated structures show violations >0.50 Å for the distance restraints or 5.0° for the dihedral restraints.
[2]F_{NOE} and F_{cdih} were calculated using force constants of 50 kcal mol^{-1} Å^{-2} and 200 kcal mol^{-1} rad^{-2}, respectively.
[3]F_{repel} was calculated using a final value of 4.0 kcal mol^{-1} Å^{-4} with the van der Waals hard sphere radii set to 0.75× those in the parameter set PARALLHSA supplied with X-PLOR[22].
[4]E_{L-J} is the Lennard-Jones van der Waals energy calculated using the CHARMM[30] empirical energy function and is not included in the target function for the simulated annealing calculation.
[5]The average r.m.s. differences from the energy-minimized average structure are given for selected residues. The value for backbone atoms (N, Cα and C′) is followed by that for all heavy atoms in parentheses.
[6]The values were obtained by PROCHECK-NMR[31].

visualized by autoradiography. The K_d value was determined as described[27] on the basis of the results obtained using protein concentrations from 0 to 50 μM.

DNA titration. The 15-bp oligonucleotides 5′-GTGCTGACTCATCAG-3′ and 5′-CTGATGAGTCAGCAC-3′, which contain the T-MARE-like sequence, were used. The 2D ^{1}H-^{15}N HSQC spectra of MafG(1–76) with 0, 0.13, 0.26, 0.39, 0.52, 0.65, 0.78, 0.91 and 1.04 equivalents of DNA were recorded at 37 °C. In this experiment, sodium chloride was added to a final concentration of 50 mM.

Sedimentation equilibrium analysis. Sedimentation equilibrium experiments were carried out at 20 °C using a Beckman XL-I analytical ultracentrifuge. Data sets were collected at three rotor speeds of 20,000, 25,000 and 30,000 rpm and three protein concentrations of 150, 220 and 370 μM. The nine data sets obtained at an absorbance of 280 nm were subjected to global fitting, and the molecular weight was calculated using a nonlinear least-squares technique[28]. Partial specific volume was calculated from the amino acid composition of the recombinant protein by using SEDNTERP[29].

Coordinates. Coordinates have been deposited in the Protein Data Bank (accession code 1K1V).

Acknowledgments
We thank E. Arai and F. Arisaka for ultracentrifuge analysis, T. Maeda for useful discussion, K. Yap for providing a program to calculate interhelical angles and T. O'Connor for critical reading of the manuscript. This work was supported by grants from JSPS and TARA (T.T.); the Ministry of Education, Science, Sports and Culture of Japan (H.M. and M.Y.); JSPS and CREST (M.Y.); and PROBRAIN (H.M.).

Competing interests statement
The authors declare that they have no competing financial interests.

Correspondence should be addressed to T.T. *email: ttanaka@tara.tsukuba.ac.jp*

Received 2 October, 2001; accepted 24 January, 2002.

1. Motohashi, H., Shavit, J.A., Igarashi, K., Yamamoto, M. & Engel, J.D. *Nucleic Acids Res.* **25**, 2953–2959 (1997).
2. Blank, V. & Andrews, N.C. *Trends Biochem. Sci.* **22**, 437–441 (1997).
3. Motohashi, H., Katsuoka, F., Shavit, J.A., Engel, J.D. & Yamamoto, M. *Cell* **103**, 865–875 (2000).
4. Kataoka, K. *et al. Mol. Cell. Biol.* **15**, 2180–2190 (1995).
5. Kataoka, K., Fujiwara, K.T., Noda, M. & Nishizawa, M. *Mol. Cell. Biol.* **14**, 7581–7591 (1994).
6. Kerppola, T.K. & Curran, T. *Oncogene* **9**, 3149–3158 (1994).
7. Wishart, D.S. & Sykes, B.D. *Methods Enzymol.* **239**, 36–81 (1994).
8. Farrow, N.A. *et al. Biochemistry* **33**, 5984–6003 (1994).
9. Harper, E.T. & Rose, G.D. *Biochemistry* **32**, 7605–7609 (1993).
10. Wintjens, R. & Rooman, M. *J. Mol. Biol.* **262**, 294–313 (1996).
11. Lo, M.C., He, S., Pelczer, I., Pal, S. & Walker, S. *Proc. Natl. Acad. Sci. USA* **95**, 8455–8460 (1998).
12. Rupert, P.B., Daughdrill, G.W., Bowerman, B. & Matthews, B.W. *Nature Struct. Biol.* **5**, 484–491 (1998).
13. Ellenberger, T.E., Brandl, C.J., Struhl, K. & Harrison, S.C. *Cell* **71**, 1223–1237 (1992).
14. Glover, J.N. & Harrison, S.C. *Nature* **373**, 257–261 (1995).
15. Delaglio, F. *et al. J. Biomol. NMR* **6**, 277–293 (1995).
16. Garrett, D.S., Powers, R., Gronenborn, A.M. & Clore, G.M. *J. Magn. Reson.* **95**, 214–220 (1991).
17. Sattler, M., Schleucher, J. & Griesinger, C. *Prog. NMR Spect.* 34, 93–158 (1999).
18. Neri, D., Szyperski, T., Otting, G., Senn, H. & Wüthrich, K. *Biochemistry* **28**, 7510–7516 (1989).
19. Tomomori, C. *et al. Nature Struct. Biol.* **6**, 729–734 (1999).
20. Vuister, G.W. & Bax, A. *J. Am. Chem. Soc.* **115**, 7772–7777 (1993).
21. Nilges, M., Clore, G.M. & Gronenborn, A.M. *FEBS Lett.* **229**, 317–324 (1988).
22. Brünger, A.T. X-PLOR version 3.1: A system for X-ray crystallography and NMR (Yale University Press, New Haven; 1993).
23. Koradi, R., Billeter, M. & Wüthrich, K. *J. Mol. Graph.* **14**, 51–55 (1996).
24. Kraulis, P.J. *J. Appl. Crystallogr.* **24**, 946–950 (1991).
25. Merritt, E.A. & Bacon, D.J. *Methods Enzymol.* **277**, 505–524 (1997).
26. Nicholls, A., Sharp, K.A. & Honig, B. *Proteins* **11**, 281–296 (1991).
27. Azam, T.A. & Ishihama, A. *J. Biol. Chem.* **274**, 33105–33113 (1999).
28. Johnson, M.L., Correia, J.J., Yphantis, D.A. & Halvorson, H.R. *Biophys. J.* **36**, 575–588 (1981).
29. Laue, T.M., Bhairavi, D.S., Ridgeway, T.M. & Pelletier, S.L. In *Analytical ultracentrifugation in biochemistry and polymer science* (eds Harding, S.E., Rowe, A.J. & Horton, J.C) 90–125 (Royal Society of Chemistry, London; 1992).
30. Brooks, B.R. *et al. J. Comp. Chem.* **4**, 187–217 (1983).
31. Laskowski, R.A., Rullmann, J.A.C., MacArthur, M.W., Kaptein, R. & Thornton, J.M. *J. Biomol. NMR* **8**, 477–486 (1996).

THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 277, No. 5, Issue of February 1, pp. 3456–3463, 2002
Printed in U.S.A.

A Sulforaphane Analogue That Potently Activates the Nrf2-dependent Detoxification Pathway*

Received for publication, October 24, 2001
Published, JBC Papers in Press, November 12, 2001, DOI 10.1074/jbc.M110244200

Yasujiro Morimitsu‡, Yoko Nakagawa‡, Kazuhiro Hayashi‡, Hiroyuki Fujii‡, Takeshi Kumagai‡, Yoshimasa Nakamura‡, Toshihiko Osawa‡, Fumihiko Horio§, Ken Itoh¶, Katsuyuki Iida¶, Masayuki Yamamoto¶, and Koji Uchida‡‖

From the ‡Laboratory of Food and Biodynamics and the §Division of Biomodeling, Graduate School of Bioagricultural Sciences, Nagoya University, Nagoya 464-8601, Japan and the ¶Center for Tsukuba Advanced Research Alliance and Institute of Basic Medical Sciences, University of Tsukuba, Tsukuba 305-8577, Japan

Exposure of cells to a wide variety of chemoprotective compounds confers resistance to a broad set of carcinogens. For a subset of the chemoprotective compounds, protection is generated by an increase in the abundance of the protective phase II detoxification enzymes, such as glutathione *S*-transferase (GST). We have recently developed a cell culture system, using rat liver epithelial RL 34 cells, that potently responds to the phenolic antioxidants resulting in the induction of GST activity (Kawamoto, Y., Nakamura, Y., Naito, Y., Torii, Y., Kumagai, T., Osawa, T., Ohigashi, H., Satoh, K., Imagawa, M., and Uchida, K. (2000) *J. Biol. Chem.* 275, 11291–11299.) In the present study, we investigated the phase II-inducing potency of an isothiocyanate compound *in vitro* and *in vivo* and examined a possible induction mechanism. Based on an extensive screening of vegetable extracts for GST inducer activity in RL34 cells, we found Japanese horseradish, wasabi (*Wasabia japonica*, syn. Eutrema wasabi), as the richest source and identified 6-methylsulfinylhexyl isothiocyanate (6-HITC), an analogue of sulforaphane (4-methylsulfinylbutyl isothiocyanate) isolated from broccoli, as the major GST inducer in wasabi. 6-HITC potently induced both class α GSTA1 and class π GSTP1 isozymes in RL34 cells. In animal experiments, we found that 6-MSHI was rapidly absorbed into the body and induced hepatic phase II detoxification enzymes more potently than sulforaphane. The observations that (i) 6-HITC activated the antioxidant response element (ARE), (ii) 6-HITC induced nuclear localization of the transcription factor Nrf2 that binds to ARE, and (iii) the induction of phase II enzyme genes by 6-HITC was completely abrogated in the *nrf2*-deficient mice, suggest that 6-HITC is a potential activator of the Nrf2/ARE-dependent detoxification pathway.

Xenobiotic metabolizing enzymes play a major role in regulating the toxic, oxidative damaging, mutagenic, and neoplastic effects of chemical carcinogens. Mounting evidence has indicated that the induction of phase II detoxification enzymes, such as glutathione *S*-transferases (GSTs),[1] results in protection against toxicity and chemical carcinogenesis, especially during the initiation phase. The GSTs are a family of enzymes that catalyze the nucleophilic addition of the thiol of reduced glutathione (GSH) to a variety of electrophiles (for a review, see Ref. 1). In addition, the GSTs bind with varying affinities to a variety of aromatic hydrophobic compounds. It is now generally accepted that the GSTs are encoded by at least five different gene families. Four (Classes α, μ, π, and θ) of the gene families encode the cytosolic GSTs, whereas the fifth encodes a microsomal form of the enzyme. It has been shown that the induction of GST is associated with the reduced incidence and multiplicity of tumors (2, 3). Recently, two transgenic rodent studies clearly demonstrated that one of the GST isozymes can profoundly alter the susceptibility to chemical carcinogenesis in mouse skin (4) and rat liver (5). Thus, the induction of GSTs is regarded as one of the most important determinants in cancer susceptibility and that its elevated synthesis is required to prevent toxic compounds from accumulating in the cells. The induction of phase II enzymes, such as GSTs, is reported to be evoked by an extraordinary variety of chemical agents, including Michael reaction acceptors, diphenols, quinones, isothiocyanates, peroxides, vicinal dimercaptans, and others (6–8). With few exceptions, these agents are electrophiles, and accordingly, many of these inducers are substrates for phase II detoxification enzymes.

Epidemiologic studies have found that persons who consume a high proportion of green and yellow vegetables in their diet have a decreased risk of developing some types of cancer (9, 10). Subsequent laboratory work has led to the isolation of various phase II inducers from fruits and vegetables that reduce the incidence of experimental carcinogenesis in animal models. Among them are included β-carotene from a variety of vegetables and fruits (11) and the monoterpenes D-limonene and D-carvone from various food plants including Citrus species (12). Later, as an approach for the detection of novel phase II inducing cancer chemoprotective agents, Talalay and his colleagues developed an *in vitro* assay system using cultured Hepa 1c1c7 murine hepatoma cells (13). They then used this assay to demonstrate that Brassica vegetables are particularly rich sources of phase II inducers and to identify sulforaphane (4-methylsulfinylbutyl isothiocyanate) as the principal phase II

* This research project was mainly supported by the Program for Promotion of Basic Research Activities for Innovative Biosciences (PROBRAIN). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‖ To whom correspondence should be addressed: Laboratory of Food and Biodynamics, Graduate School of Bioagricultural Sciences, Nagoya Univ., Nagoya 464-8601, Japan. Tel.: 81-52-789-4127; Fax: 81-52-789-5741; E-mail: uchidak@agr.nagoya-u.ac.jp.

[1] The abbreviations used are: GST, glutathione *S*-transferase; 6-HITC, 6-methylsulfinylhexyl isothiocyanate; HPLC, high-performance liquid chromatography; MS, mass spectrometry; TMS, tetramethylsilane; PBS, phosphate-buffered saline; γ-GCS, γ-glutamylcysteine synthetase; ARE, antioxidant response element; NQO1, quinone reductase (NAD(P)H:(quinone-acceptor) oxidoreductase 1, EC 1.6.99.2).

inducer in broccoli extracts (14). They also have demonstrated that sulforaphane is a dose-related inhibitor of carcinogen-induced mammary tumorigenesis in rats (15).

We have recently developed a cell culture system, using rat liver epithelial RL34 cells, that potently responds to the already-known phase II inducers, such as phenolic antioxidants and α, β-unsaturated aldehydes, resulting in the induction of the GST activity (16). In the present study, using the RL34 cells, we determined the GST induction potencies of food plants and found that the wasabi extracts induce GST activity with great potency. We provided an analysis of the wasabi extracts that demonstrate an isothiocyanate compound as a principal inducer of phase II enzymes. Moreover, we have investigated the phase II inducing potency of this compound *in vitro* and *in vivo* and examined a possible induction mechanism.

EXPERIMENTAL PROCEDURES

Materials

Authentic 6-methylsulfinylhexyl isothiocyanate (6-HITC) and sulforaphane were synthesized by the oxidation of 6-methylthiohexyl isothiocyanate and 4-methylthiobutyl isothiocyanate (kind gifts of Kinjirushi Wasabi, Co., Ltd., Nagoya, Japan), respectively. The 6-HITC-related compounds were synthesized in our laboratory. All other chemicals were purchased from Wako Pure Chemical Industries (Osaka, Japan). Wasabi (*Wasabia japonica*, syn. Eutrema wasabi) cultivated in Shizuoka, Japan, was also obtained from Kinjirushi Wasabi Co., Ltd. Low- and high-resolution fast atom bombardment-mass spectrometry was measured using a JEOL JMS-700 (MStation) instrument. NMR spectra were recorded with a Bruker AMX600 (600 MHz) instrument. Ultraviolet absorption spectra were measured with a Hitachi U-Best-50 spectrophotometer, and fluorescence spectra were recorded with a Hitachi F-2000 spectrometer. Liquid chromatography-MS) was measured with a JASCO PlatformII-LC instrument.

Cell Culture

RL34 cells were obtained from the Japanese Cancer Research Resources Bank. The cells were grown as monolayer cultures in Dulbecco's modified Eagle's medium supplemented with 5% heat-inactivated fetal bovine serum, penicillin (100 units/ml), streptomycin (100 μg/ml), L-glutamine (588 μg/ml), and 0.16% $NaHCO_3$ at 37 °C in an atmosphere of 95% air and 5% CO_2.

Extraction and Isolation Procedures

Wasabi roots (water wasabi) harvested in Shizuoka, Japan, were kind gifts from Kinjirushi Wasabi Co, Ltd. The wasabi roots (1.3 kg) were smashed with a grater, and the homogenates were stored at room temperature for 10 min. The homogenates were then sequentially extracted with ethyl acetate, 2 liters), *n*-butanol (2 liters) and water (1 liters). The ethyl acetate extract was further separated by silica gel column chromatography (silica gel BM-300, Fuji-Silicia Chem), and the acetone fraction, which showed the most potent activity, was analyzed by reversed-phase HPLC (yield, 70 mg), using a Develosil ODS-HG-5 (8 × 250 mm) column. The flow solvent was methanol/water = 3/2 (v/v) at a flow rate of 2.0 ml/min. Detection was carried out at 254 nm. Identification of the active compound was done by spectroscopic analyses (17). The NMR and MS data were measured using a Bruker ARX 400 and a JEOL MStation MS-700, respectively. The IR, UV, and optical rotation data were measured by a JASCO FT/IR-8300, a Beckman DU7500 and a JASCO DIP-370, respectively. The spectral data of 6-HITC were as follows: ^{13}C-NMR ($CDCl_3$, TMS), d (ppm): 22.4 (C-4), 26.2 (C-3), 27.9 (C-5), 29.6 (C-2), 38.6 (sulfinyl *methyl*), 44.9 (C-6), 54.3 (C-1), 130.0 (-NCS); ^{1}H-NMR ($CDCl_3$, TMS), d (ppm): 1.4–1.6 (4H, m, H-3 and H-4), 1.73 (2H, tt, J = 6.7, 7.1 Hz, H-5), 1.81 (2H, tt, J = 6.4, 7.4 Hz, H-2), 2.58 (^{3}H, s, sulfinyl *methyl*), 2.72 (2H, dt, J = 7.1, 13.1 Hz, H-6), 3.54 (2H, t, J = 6.4 Hz); IR (liquid film, $CHCl_3$), νmax (cm^{-1}): 3002 (CH strech), 2108 (-NCS), 1032 (sulfinyl); UV (MeOH), λmax (nm): 243 (ϵ 1210); $[\alpha]_D$: −65.0 (c 0.600, $CHCl_3$, 23 °C).

Gas-Liquid Chromatography Analysis of 6-HITC in Wasabi Extracts

The wasabi extract for the gas-liquid chromatography analyses was prepared using dichloromethane (CH_2Cl_2) as the extracting solvent. The CH_2Cl_2 was carefully removed under atmospheric pressure in a N_2 stream on the ice bath. The residue was redissolved with an adequate amount of CH_2Cl_2. A 1 μl-aliquot of the CH_2Cl_2 solution of the wasabi extract was injected into the gas-liquid chromatography equipment. The quantitative analyses of 6-HITC and 6-methylthiohexyl isothiocyanate were carried out with a Hitachi G-3500 gas chromatograph (column, DB-1, 0.25 mm × 30 m; carrier gas, N_2 at 1 ml/min; temp. program, 60–230 °C at 5 °C/min). Each peak was identified by the retention time of each authentic or synthetic sample and the gas-liquid chromatography-mass spectrometry analysis using a JEOL MStation MS-700 mass spectrometer linked to a Hewlett-Packard 6890 (column, DB-1, 0.25 mm × 30 m; carrier gas, He at 1 ml/min; temp. program, 60–300 °C at 8 °C/min; column inlet split rate, 1/100). These analyses were repeated three times.

Enzyme Activity Assays

The total GST activity was measured in cytosolic fractions (105,000 × *g*) in the presence of 0.1% bovine serum albumin with 1-chloro-2,4-dinitrobenzene as a substrate (18), whereas quinone reductase (NAD(P)H:(quinone-acceptor) oxidoreductase 1 (NQO1)) activity was determined using menadione as the substrate (19). Cytochrome P4501A1-mediated ethoxyresorufin *O*-deethylase activity was measured using the procedures of Kennedy *et al.* (20). The protein concentration was determined using the bicinchoninic acid protein assay (Pierce).

Western Blot Analysis

The homogenates prepared from the cells or animal tissues were treated with SDS-sample buffer (without dye or 2-mercaptoethanol) and immediately boiled for 5 min. The protein concentrations were determined using the BCA protein assay kit (Pierce). One hundred μg of the proteins were separated by SDS-polyacrylamide gel electrophoresis in the presence of 2-mercaptoethanol and electro-transferred onto an Immobilon membrane (Millipore, Bedford, MA). To detect the immunoreactive proteins, we used horseradish peroxidase-conjugated anti-rabbit IgG and ECL blotting reagents (Amersham Biosciences, Inc.). The polyclonal antibodies against GSTA1 were the kind gifts from Dr. K. Satoh of Hirosaki University School of Medicine. The polyclonal antibodies against GSTP1 and GSTM1 were obtained from Medical and Biological Laboratories Co., Ltd. (Nagoya, Japan) and Oxford Biomedical Research, Inc. (Oxford, MI), respectively. Polyclonal rabbit antisera raised against mouse Nrf2 was used as previously described (21). The anti-nuclear Lamin B antiserum was purchased from Santa Cruz Biotechnology (Palo Alto, CA).

Nuclear Translocation of Nrf2

The cells treated with Me_2SO or 6-HITC were fixed overnight in PBS containing 2% paraformaldehyde and 0.2% picric acid at 4 °C. The membranes were permeabilized by exposing the fixed cells to PBS containing 0.3% Triton X-100. The cells were then sequentially incubated in PBS solutions containing blocking serum (5% normal rabbit serum) and immunostained with the anti-Nrf2 polyclonal antibody (Santa Cruz, Santa Cruz, CA). The cells were then incubated for 1 h in the presence of fluorescein isothiocyanate-labeled rabbit anti-goat (DAKO A/S, Glostrup, Denmark), rinsed with PBS containing 0.3% Triton X-100, and covered with anti-fade solution. Images of the cellular immunofluorescence were acquired using a confocal laser microscope (Bio-Rad, Hercules, CA) with a 40× objective (488-nm excitation and 518-nm emission).

Plasmid Preparation

The annealed oligonucleotide of GSTA1 ARE (top strand: TCGAGTAGCTTGGAAATGACATTGCTAATGGTGACAAAGCAACTTTG; bottom strand: TCGACAAAGTTGCTTTGTCACCATTAGCAATGTCATTTCCAAGCTAC) was ligated to the *Xho*I and *Sal*I sites of the Bluescript SK(−) plasmid, and the plasmid with three GSTA1 ARE inserts in tandem was selected. The *Kpn*I and *Hinc*II fragment of the plasmid was blunted and then subcloned into the *Sma*I site of the pRBGP3 plasmid (22). The annealed oligonucleotide of the human NQO1 ARE (top strand: CGCGTTCAGAGATTTCAGTCTAGAGTCACAGTGACTTGGCAAAATCG; bottom strand: CTAGCGATTTTGCCAAGTCACTGTGACTCTAGACTGAAATCTCTGAA) was ligated to the *Mlu*I and *Nhe*I site of the pRBGP3.

RNA Blot Hybridization

RL34 cells were maintained in Iscoves's modified Dulbecco's medium and 10% fetal bovine serum. Fifteen μg of the total cellular RNAs extracted by ISOGEN (NipponGene) was electrophoresed and transferred to Zeta-Probe GT membranes (Bio-Rad Japan, Tokyo). The mem-

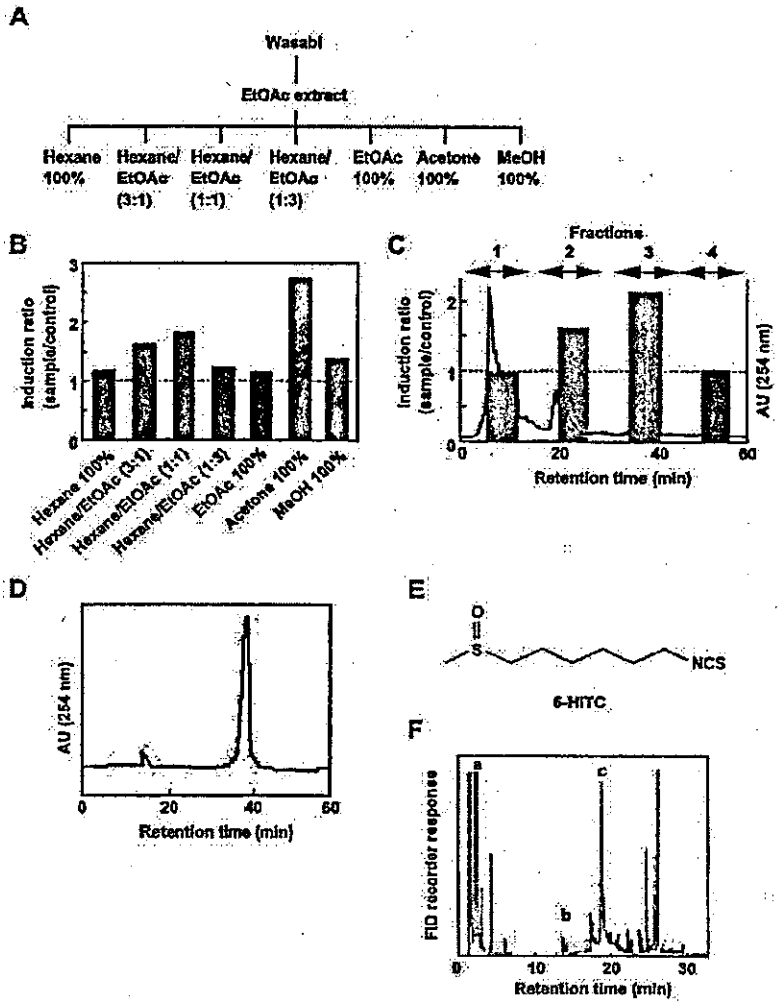

FIG. 1. **Isolation and identification of a major GST inducer from wasabi.** *A*, scheme for fractionation of ethyl acetate extract of wasabi by silica gel chromatography. *B*, GST inducer potency of fractions separated by silica gel chromatography. *C*, fractionation of the active fraction (the acetone fraction in *panel B*) by reverse-phase HPLC and GST inducer potency of each fraction (fractions 1-4). *D*, purification of active compound from fraction 3 in *panel C*. *E*, chemical structure of 6-HITC. *F*, gas-liquid chromatography analysis of wasabi extract. Peaks: *a*, allylisothiocyanate; *b*, 6-methylthiohexyl isothiocyanate; *c*, 6-HITC.

branes were probed with ^{32}P-labeled cDNA probes as indicated in the figures. 18S RNA cDNA was used as the positive control.

Transient Transfection Assay

RL34 cells were maintained in Iscoves's modified Dulbecco's medium supplemented with 10% fetal bovine serum and seeded in 5×10^4/well in 12-well dishes 24 h before transfection. The cells were transfected with plasmids using FuGENE (Roche Molecular Biochemicals) according to the manufacturer's instructions. Nine hours after the transfection, the medium was changed to the fresh medium, and the cells were treated with Me_2SO or 6-HITC (5 μM). After 36 h, the luciferase assay was performed by utilizing the Luciferase Assay System (Promega, Madison, WI) following the supplier's protocol and measured in a Biolumat Luminometer (Berthold, Bad Wildbad, Germany). Transfection efficiencies were routinely normalized by the activity of a co-transfected *Renilla* luciferase. Normally, three independent experiments, each carried out in duplicate, were performed, and the mean values were presented with the standard error of means (S.E.).

Animal Study

Determination of 6-HITC and Its Dithiocarbamate in the Plasma—Male Wistar rats (Japan SLC Inc., Hamamatsu, Japan) were obtained at 7 weeks of age and individually housed in stainless wire-mesh cages at 23 ± 0.3 °C with a 12-h light cycle. They were fed unrestricted amounts of water, and the control diet was as follows: 20% casein, 3.5% mineral (93G-MX), 5.0% vitamin (93-VX), 0.2% choline chloride, 5.0% corn oil, 4.0% cellulose powder, 22.1% sucrose, and 44.2% starch. After 7 days of feeding the control diet, food was withheld for 24 h, and then 6-HITC dissolved in sesame oil was orally administered to four rats by direct stomach intubation. Blood samples were collected by heart puncture using heparinized needles and syringes under anesthesia with diethyl ether. The plasma was immediately obtained from the collected blood by centrifugation at 1600 × *g* for 15 min at 4 °C. The plasma separation was finished within 30 min. An aliquot of the plasma was acidified with one-tenth volume of phosphoric acid and stored at −80 °C until used. The levels of 6-MSHI in plasma were measured by the cyclocondensation assay as previously reported (23).

6-HITC-administered Mice—Female ICR mice (Japan SLC Inc., Hamamatsu, Japan) were obtained at 4 weeks of age and individually housed in plastic cages (five/cage) at 23 ± 0.3 °C with a 12-h light cycle. They were fed for 12 days unrestricted amounts of water, and the control diet was as follows: 20% casein, 3.5% mineral (93G-MX), 5.0% vitamin (93-VX), 0.2% choline chloride, 5.0% corn oil, 4.0% cellulose powder, 22.1% sucrose, and 44.2% starch. From the thirteenth day, 6-HITC or sulforaphane was administered to these mice by gavage in daily doses of 15 μmol for 5 days (14).

Nrf2 Knockout Mice—A single dose of 15 μmol of 6-HITC was suspended in olive oil and administered to adult female Nrf2 knockout mice (24) or ICR control animals (Nrf2 +/+) by gavage.

RESULTS

Isolation and Identification of a Major GST Inducer from Wasabi—We have previously shown that the ethyl acetate extracts of wasabi exhibited the most significant enhancement of GST activity at the concentration of 2.5 μg/ml (25). Hence, we performed the activity-guiding separation of a principal inducer from wasabi. Wasabi roots (1.3 kg) were extracted with 1.5 liters of ethyl acetate at 4 °C, and the extract was fractionated into seven fractions by silica gel chromatography (Fig. 1*A*). The acetone fraction exhibited the most potent GST-inducing activity (Fig. 1*B*) was further fractionated into four fractions by

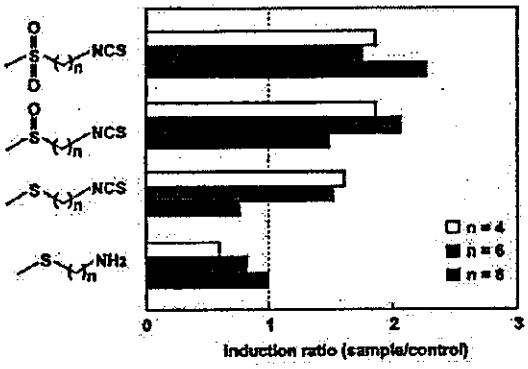

FIG. 2. **Induction of GST in RL34 cells by 6-HITC and analogues.** The cells were treated with the compounds (25 μM) for 24 h. After incubation, the cytosols were prepared and assayed for GST activity.

preparative HPLC (Fig. 1*C*). Final purification of the inducer was carried out by preparative HPLC under the same conditions (Fig. 1*D*); evaporation of pooled active fraction gave 70 mg of a colorless liquid. The spectroscopic data, including ^{1}H- and ^{13}C-NMR, and EI-MS (molecular ion at *m/z* of 149 [M^+], C_8H_7NS), of this compound were completely identical to those of 6-HITC (Fig. 1*E*). As shown in Fig. 1*F*, 6-HITC was one of the major components in the wasabi extract. The concentration of 6-HITC in the wasabi was ~550–556 μg/g wet body weight of wasabi root.

Structure-Activity Relationship—To define the structural features of 6-HITC in the RL34 cells, we synthesized the following analogues of 6-HITC and measured their GST inducer potencies according to Equations 1–4.

$$CH_3\text{-}S\text{-}(CH_2)_n\text{-}NH_2\ (n = 4, 6, \text{or } 8) \qquad \text{(Eq. 1)}$$

$$CH_3\text{-}S\text{-}(CH_2)_n\text{-}N = C = S\ (n = 4, 6, \text{or } 8) \qquad \text{(Eq. 2)}$$

CH_3-SO-$(CH_2)_n$-N = C = S (n = 4, 6, or 8;

n = 4, sulforaphane; n = 6, 6-HITC) (Eq. 3)

$$CH_3\text{-}SO_2\text{-}(CH_2)_n\text{-}N = C = S\ (n = 4, 6, \text{or } 8) \qquad \text{(Eq. 4)}$$

As shown in Fig. 2, the methylthioalkylamines (Eq. 1) had no induction potency, whereas the methylthioalkyl isothiocyanate (Eq. 2) significantly induced GST activity, indicating that the isothiocyanate moiety was essential for the induction of GST activity. It was also found that the inducer potency was influenced by the oxidation state of sulfur and the number of methylene groups in the bridge linking the thiomethyl and isothiocyanate moieties. These data indicate that, among the isothiocyanates, 6-HITC is one of the most potent phase II inducers.

In Vitro Induction of Phase II Enzymes by 6-HITC—As shown in Figs. 3, *A* and *B*, 6-HITC induced GST activities in time- and dose-dependent manners. To examine the GST isozyme responsible for the increase in the GST activity of the 6-HITC-treated RL34 cells, an immunoblot analysis was carried out using the GST class-specific antibodies to confirm the apparent induction of GST proteins. The immunoblot analysis demonstrated a significant increase in the levels of the class α GSTA1 and class π GSTP1 by treatment with 6-HITC (Fig. 3*C*), while the amount of the class μ isozyme (GSTM1) was nearly unchanged (data not shown). These results indicated that the induction of GST activity by 6-HITC resulted, at least, from the enhanced expression of GSTA1 and GSTP1. The treatment with 6-HITC also resulted in a significant increase in the intracellular GSH levels (data not shown), suggesting the induction of a rate-limiting enzyme of GSH biosynthesis, γ-glutamylcysteine synthetase (γ-GCS), by 6-HITC.

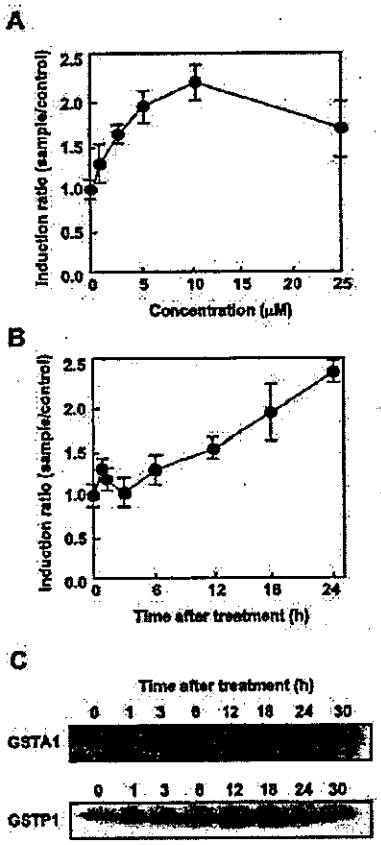

FIG. 3. **Induction of GST by 6-HITC.** *A*, dose-dependent induction of GST activity by 6-HITC. *B*, time-dependent induction of GST activity by 6-HITC. *C*, immunoblot analysis of GST isozymes in the RL34 cells treated with 25 μM 6-HITC. *Upper*, GSTA1; *Lower*, GSTP1.

Absorption of 6-HITC—Prior to the *in vivo* assessment of induction of the phase II enzymes, we examined the absorption of 6-HITC into the body. For this purpose, the cyclocondensation assay was utilized (Fig. 4*A*) (23). To validate this assay for measurement of the 6-HITC, the GSH conjugate of 6-HITC (dithiocarbamate) was incubated with 1,2-benzenedithiol, and the reaction products were analyzed by reverse-phase HPLC. As shown in Fig. 4*B*, the reaction product (1,3-benzodithiole-2-thione) linearly increased with the dithiocarbamate concentration. Thus, this assay was demonstrated to provide a valid measurement of 6-HITC and its dithiocarbamate metabolite. To examine the absorption of 6-HITC into the blood stream, rats were orally administered 6-HITC, and the levels of 6-HITC in the plasma were measured by the cyclocondensation assay. As shown in Fig. 4*C*, upon treatment of plasma with 1,2-benzenedithiol, the cyclocondensation product was formed. The peak was identified as 1,3-benzodithiole-2-thione by comparison with the authentic compound based on the retention time in the HPLC analysis, UV-visible spectrum, and liquid chromatography-MS. The plasma 6-HITC concentration reached a maximum within 30 min after the administration and began to fall within 1 h (Fig. 4*D*). These results suggest that 6-HITC is absorbed and rapidly enters the circulatory system as is and/or as its GSH conjugate. Interestingly, the subsequent decrease in the level of 6-HITC was relatively slow, and it was detected in the plasma even 4 h after a single administration of 6-HITC.

In Vivo Induction of Phase II Enzymes by 6-HITC—6-HITC

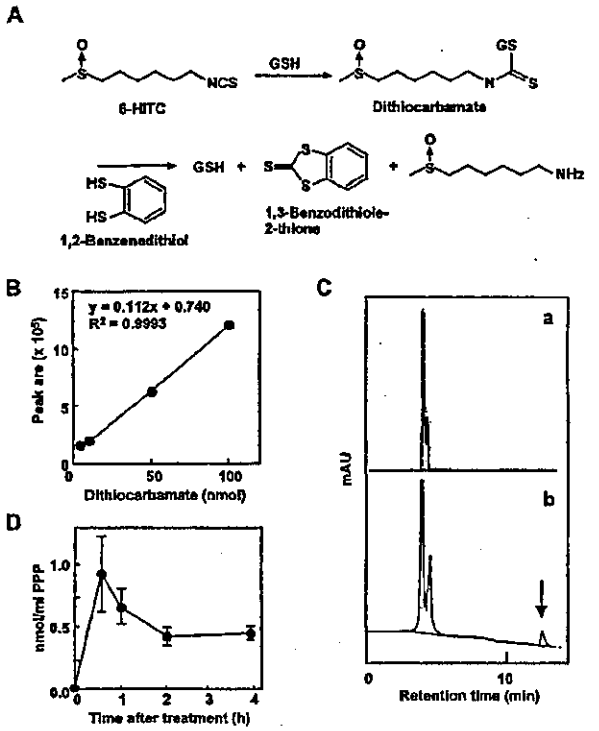

FIG. 4. **Absorption of 6-HITC into body.** *A*, the cyclocondensation reaction of 6-HITC with 1,2-benzenedithiol. *B*, standard curve for the dithiocarbamate derivative (6-HITC-GSH adduct). *C*, HPLC chromatogram of rat plasma extracts after oral administration of 6-HITC. *Panel A*, HPLC chromatogram of plasma extracts from control rats. *Panel B*, HPLC chromatogram of plasma extracts from 6-HITC-administered rats at 1 h. *D*, concentration of 6-HITC in rat plasma after oral administration of 6-HITC. Values are means + S.E., $n = 3$.

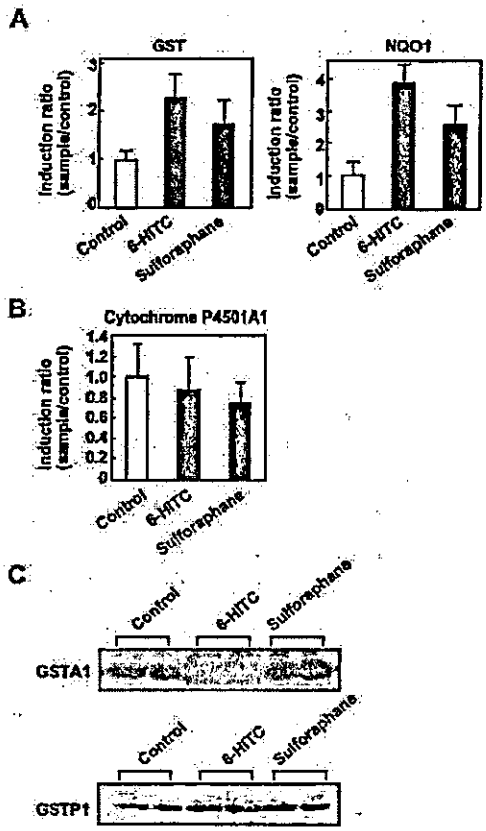

FIG. 5. **Effect of 6-HITC and sulforaphane administrations on mouse hepatic detoxification enzyme activities.** *A*, GST and NQO1 activities. *B*, cytochrome P4501A1 activity. *C*, immunoblot analysis of GSTA1 and GSTP1.

was then investigated in dietary studies performed with female ICR mice. To assess the potential of 6-HITC to induce phase II detoxification enzymes, animals were treated with 6-HITC or sulforaphane, a well known phase II inducer isolated from broccoli (14), by gavage in a daily dose of 15 μmol for 5 days, and changes in the hepatic GST, NQ01, and cytochrome P4501A1 activities were examined. There were no significant differences in the body and liver weights (data not shown). As shown in Fig. 5*A*, hepatic GST and NQ01 activities were more potently induced by 6-HITC than by sulforaphane. Whereas, only slight increases in the GST and NQ01 activities were observed in the lung and kidney (data not shown). It was also observed that both 6-HITC and sulforaphane moderately inhibited cytochrome P4501A1-mediated ethoxyresorufin *O*-deethylase activity (Fig. 5*B*); however, this inhibition was not significant in comparison with the control group. Taken together, these data suggest that induction by 6-HITC was monofunctional (induction of phase II and inhibition of phase I). To further determine the GST isozymes responsible for the increase in the hepatic GST activity of the 6-HITC- or sulforaphane-treated mice, an immunoblot analysis was carried out using the GST class-specific antibodies. As shown in Fig. 5*C*, 6-HITC induced hepatic GSTA1 and GSTP1 isozymes more prominently than sulforaphane. The difference in the GST activity levels between the 6-HITC- and sulforaphane-treated mice may therefore correspond to that in the levels of the GSTA1 isozyme.

Induction of Phase II Enzyme Gene Expression and Activation of ARE—6-HITC, as well as sulforaphane, induced the gene expression of the phase II enzymes, such as GSTP1, γ-GCS (heavy chain), and γ-GCS (light chain) (Fig. 6). The data also showed that, similar to the data in Fig. 5, the induction of the phase II enzyme gene expression by 6-HITC was significantly more prominent than by sulforaphane. It has been established that the induction of the phase II enzyme activity occurs at the transcriptional level and is regulated by a *cis*-acting element that is present in the promoters of the phase II enzyme genes defined as the antioxidant response element (ARE) (26) or electrophile response element (27). Hence, we examined the involvement of ARE in the 6-HITC-induced phase II enzyme gene expression and found that 6-HITC potently stimulated the activity of the ARE reporter genes (Fig. 7).

Activation of Transcription Factor Nrf2—Several lines of evidence indicate that a member of the basic leucine zipper transcription factor family, Nrf2 (NF-E2-related factor 2), is involved in the activation of ARE (24, 28). To determine whether this transcription factor indeed contributes to the 6-HITC-stimulated activation of ARE, we examined the nuclear localization of Nrf2 in the 6-HITC-treated RL34 cells. As shown in Fig. 8*A*, only cytoplasmic labeling of Nrf2 with no nuclear staining was observed in the non-stimulated cells (*panel A*), whereas an intense nuclear labeling was observed in the 6-HITC-stimulated cells (*panel B*). In addition, the 6-HITC treatment of the cells led to a dose-dependent increase in the Nrf2 levels (Fig. 8*B*). Sulforaphane also showed a similar effect, but its induction potency was significantly lower than that of 6-HITC.

Effect of nrf2 Genotype and 6-HITC Treatment on Phase II Enzyme Activities—Finally, to directly show that Nrf2 is involved in the 6-HITC-induced expression of the phase II en-