MOLECULAR AND CELLULAR BIOLOGY, Aug. 1995, p. 4184–4193 Vol. 15, No. 8
0270-7306/95/$04.00+0

Cloning and Characterization of a Novel Erythroid Cell-Derived CNC Family Transcription Factor Heterodimerizing with the Small Maf Family Proteins

KEN ITOH,[1] KAZUHIKO IGARASHI,[1]* NORIO HAYASHI,[1] MAKOTO NISHIZAWA,[2] AND MASAYUKI YAMAMOTO[1]

Department of Biochemistry, Tohoku University School of Medicine, 2-1 Seiryomachi, Aoba-ku, Sendai 980-77,[1] and Department of Molecular Oncology, Kyoto University Faculty of Medicine, Sakyo-ku, Kyoto 606,[2] Japan

Received 28 November 1994/Returned for modification 5 January 1995/Accepted 12 April 1995

The chicken β-globin enhancer is critical for the tissue- and developmental stage-specific expression of the β-globin genes. This enhancer contains two indispensable *cis* elements, one containing two GATA sites and the other containing an NF-E2 site. To identify the putative transcription factor acting through the NF-E2 motif in the chicken β-globin enhancer, we screened chicken cDNA libraries with a mouse p45 NF-E2 cDNA probe and isolated cDNA clones which encode a protein of 582 amino acid residues. This protein contains a region that includes the basic region-leucine zipper domain which is well conserved among members of the CNC family proteins (Cap 'n' collar, p45 NF-E2, LCR-F1, Nrf1, and Nrf2). Hence, we named this protein ECH (erythroid cell-derived protein with CNC homology). ECH is expressed abundantly in cultured erythroid cells undergoing terminal differentiation, peripheral erythrocytes, and some nonhematopoietic tissues. Since most of the cDNA clones obtained from the chicken erythrocyte cDNA library encoded ECH, ECH is likely the predominant CNC family protein present in avian peripheral erythrocytes. Like p45 NF-E2, ECH can heterodimerize with any of the small Maf family proteins and bind the NF-E2 site as a heterodimer in vitro. In a transfection assay, ECH transactivates transcription depending on the presence of NF-E2 sites on the reporter gene plasmid. These results indicate that ECH is likely a key regulator of avian erythropoiesis.

The chicken β-globin gene locus provides an excellent model system for unraveling the molecular mechanisms underlying cell lineage- and developmental stage-specific expression of globin genes in erythroid cells, since most of the *cis*-acting regulatory domains are confined within relatively small regions of the locus (31). Pioneering work had identified a strong erythroid cell-specific enhancer (β enhancer) located between β- and ε-globin genes (reviewed in reference 9). Within the β enhancer, several *cis*-acting functionally significant elements have been defined (30). One of these elements, termed proximal region of β-enhancer site F2 (site F2p or site IIA), contains a motif, AGCTGAGTCAT (12, 30), which resembles the consensus binding sequence of mammalian erythroid transcription factor NF-E2. Although NF-E2 has been characterized in mice and humans (2, 3, 15, 22, 26), little is known about the identity of the putative factor(s) that binds the β-enhancer site F2p in avian erythroid cells.

Transcription factor NF-E2 was originally identified as an erythroid cell-restricted DNA binding activity which recognizes the consensus sequence TGCTGA(G/C)TCA(T/C) (2, 22) containing a 12-*O*-tetradecanoylphorbol-13-acetate (TPA)-responsive element (TGA(G/C)TCA). This motif is frequently found within the regulatory regions of globin genes and several other erythroid genes, and the functional importance of the NF-E2 sites for expression of the associated genes has been demonstrated (21, 27, 28, 30, 33). The recent purification of the NF-E2 activity from murine and human erythroleukemia cells and the subsequent cDNA cloning have revealed that NF-E2 is a heterodimeric protein that consists of a large and a small subunit (p45 and p18, respectively), both of which contain basic region-leucine zipper (b-zip) motifs (2, 3, 26).

The chicken small Maf family proteins (MafF, MafG, and MafK) constitute a subgroup in the Maf proto-oncoprotein family (11, 16). These proteins possess the characteristic b-zip domain but lack any canonical *trans* activation domain. Various dimeric combinations of the small Maf family proteins can recognize a DNA sequence motif called T-MARE (TGCTGA(G/C)TCAGCA) (17) or its related motifs (16). Furthermore, all of the chicken small Maf family proteins can form heterodimers with the large subunit of mouse NF-E2 (p45), and the resultant heterodimers can bind the NF-E2 consensus motif with higher affinity than p45 (15). Recent analysis of *mafK* in mice unequivocally established that the small subunit of NF-E2 (p18) is encoded by the *mafK* gene (14).

The small Maf family proteins are expressed in various hematopoietic and nonhematopoietic lineages in chickens (11, 15), whereas p45 is specifically expressed in erythroid, megakaryocytic, and mast cells in mice (2). In the heterodimeric association of mouse p45 and MafK (or any of the chicken small Maf family proteins), the small Maf subunit protein confers DNA binding activity to p45, while p45 subunit directly activates transcription via its amino-terminal transactivation domain (14, 15). In the absence of p45, however, the small Maf family proteins repress NF-E2 site-mediated transcriptional activation by interfering with the binding of AP-1-like factors to the NF-E2 site (15). Thus, both positive and negative transcriptional regulation can be mediated through NF-E2 sites, depending on the relative amounts of p45, the small Maf family proteins, and AP-1-like factors.

Mysteries regarding effector molecules of NF-E2 sites still abound since there also exist three p45-related proteins, Nrf1 (7), LCR-F1 (6), and Nrf2 (25), in human erythroid cells. Since LCR-F1 and Nrf1 differ from each other exclusively at their

* Corresponding author. Mailing address: Department of Biochemistry, Tohoku University School of Medicine, 2-1 Seiryomachi, Aoba-ku, Sendai 980-77, Japan. Phone: 81-22-273-9065. Fax: 81-22-274-5698.

amino termini, they are thought to be transcribed from a single gene via alternative exon usage (6). Analysis of the human TCF11 gene, which encodes Nrf1, supports this hypothesis (20). These proteins show high structural similarity with p45 in the b-zip and surrounding region but are distinct otherwise. Because of the structural similarity to a domain of *Drosophila* transcription factor Cap 'n' collar (CNC) (24), the region conserved among these proteins is referred to as a CNC domain (6). LCR-F1/Nrf1 and Nrf2 are expressed in a wide range of tissues and cells, including erythroid cells (6, 7, 25). LCR-F1 has been shown to be a more potent transcriptional activator than p45 in erythroid cells (6). Thus, these proteins can potentially regulate erythroid cell-specific gene expression through NF-E2 sites.

In this paper, we report the cloning and characterization of chicken cDNAs encoding a novel member of the CNC (or p45 NF-E2)-related protein family. We named the protein ECH, for erythroid-derived protein with CNC homology. Our analysis shows that ECH may be the predominant CNC member expressed in chicken erythroid cells. Expression of ECH increases significantly during terminal differentiation of cultured erythroid cells. Like p45, ECH can bind the NF-E2 site within site F2p as a heterodimer with any of the small Maf family proteins in vitro. In vivo, ECH transactivates transcription depending on the presence of NF-E2 sites on the reporter gene plasmid. These results suggest that ECH is likely a key regulator acting on NF-E2 sites during avian erythropoiesis.

MATERIALS AND METHODS

Library screening. An anemic chicken peripheral erythrocyte (RBC) cDNA library in λgt11 (RBC4) (36) was screened with a cDNA fragment coding for the b-zip domain of mouse p45 NF-E2. The library was plated on 150-mm-diameter petri dishes at a density of 5×10^4 PFU per plate. Duplicate phage lift filters (BiodyneA; Pall) were made from each plate and processed for hybridization as described previously (32). Hybridization to the probe was carried out in the presence of 25% formamide, 1% sodium dodecyl sulfate (SDS), 5× SSC (1× SSC is 0.15 M NaCl plus 0.015 M sodium citrate), 5× Denhardt's solution, and 0.2 mg of sonicated salmon sperm DNA per ml at 42°C. The filters were washed in 2× SSC–0.1% SDS at 42°C for 1 h. Positive clones were purified by three additional rounds of plaque hybridization screening. Inserts of positive phage clones were then subcloned into plasmid pBluescript SK(−) (Stratagene) by using *Eco*RI and *Not*I sites in the vectors. A 5-day-old chicken whole embryo cDNA library (in λZapII; generous gift from Atsushi Kuroiwa) was screened with a cDNA fragment cloned as described above. cDNA inserts were recovered as plasmids by in vivo excision as instructed by the manufacturer (Stratagene). These screenings resulted in phage clones, including λRBC18 and λCE5 (from the RBC4 and whole embryo libraries, respectively).

5′-RACE analysis. 5′-RACE (rapid amplification of the cDNA end) analysis was carried out with the 5′-AmpliFinder RACE system (Clontech). Positions of the oligonucleotides used as primers are indicated in Fig. 1A. cDNA was synthesized by using cp103 (5′-TAGTGTTTGTTGTTCCTGTGTCACCGT-3′) as a primer and 2 μg of polyadenylated RNA isolated from HD3 cells as the template. The first PCR was carried out with cp106 (5′-CCTGTCTCCTCATCTAA CACCAGCT-3′) and the anchor primer. Nested PCR was carried out with cp108 (5′-AGGCTTTCTCCCGCTCTTTCTGGAG-3′) and the anchor primer. Specific amplification products were cloned into plasmid pBluescript.

Sequence analysis. DNA sequences were determined on both strands by the dideoxynucleotide method, using *Taq* polymerase cycle sequencing (ABI Japan) and deletion subclones of plasmids as templates. A composite cDNA sequence was generated from the sequence information for the inserts in λRBC18, λCE5, and RACE clones.

Plasmid constructions. The *Eco*RI-*Not*I fragment of λRBC18 was cloned between the *Eco*RI and *Not*I sites of pBluescript SK(−), resulting in pKI18. The 357-bp *Pst*I fragment of pKI18 was subcloned in the *Pst*I site of pBluescript SK(−), resulting in pKI121. pCE5 was generated by in vivo excision of λCE5. A prokaryotic expression plasmid for ECH (pMBPECH) was constructed as follows. pKI121 was first digested with *Not*I, blunted with T4 polymerase, and partially digested with *Pst*I. The 846-bp cDNA fragment containing the b-zip coding domain was subcloned into pMAL-c2 (New England Biolabs), using the *Pst*I and *Hin*dIII sites. Prokaryotic expression plasmids for MafF, MafG, and MafK have been described elsewhere (16). A eukaryotic expression plasmid for ECH (pEFECH) was constructed by inserting the 2,254-bp *Bam*HI (which cuts the vector sequence)-*Spe*I fragment of pCE5 into the *Bss*HII site of pEF*Bss*HII, a modified version of pEF-BOS (23). Luciferase reporter plasmids (pRBGP2 and pRBGP4) and a p45 expression plasmid (pEFp45) were described previously (15).

In vitro transcription-translation analysis. pCE5 was first linealized with *Kpn*I and then transcribed by T3 RNA polymerase in the presence of cap analog [m^7G(5′)ppp(5′)G; Pharmacia]. The in vitro translation reaction was carried out with rabbit reticulocyte lysates (Promega) in the presence of [^{35}S]methionine. Translation products were analyzed by SDS-polyacrylamide gel electrophoresis and autoradiography.

Genomic DNA blot hybridization analysis. High-molecular-weight DNA was extracted from the liver of a 6-week-old chicken as described previously (32). The DNA (25 μg) was digested with *Bam*HI, *Eco*RI, or *Hin*dIII, subjected to electrophoresis in a 0.8% agarose gel, and transferred onto a Zeta Probe nylon membrane (Bio-Rad). The membrane was hybridized with the pKI121 labeled with [α-^{32}P]dCTP (Amersham) by the random priming method. Hybridization and washing conditions were as described previously (32).

RNA blot hybridization analysis. Poly(A)$^+$ RNA was isolated from various cell lines or tissues of adult chickens by the acid guanidinium thiocyanate-phenol-chloroform extraction method (8) and with oligo(dT)-latex beads (Takara, Kyoto, Japan). Poly(A)$^+$ RNA (1.5 μg) was separated on agarose-formaldehyde gels and transferred onto Zeta Probe membranes. The membranes were hybridized with ^{32}P-labeled antisense RNA of ECH, washed, and analyzed by autoradiography as previously described (10). The same RNA samples were hybridized with chicken glyceraldehyde-3-phosphate dehydrogenase (GAPDH) cDNA.

Antibody preparation. pMBPECH was introduced into *Escherichia coli* SG12036, and the fusion protein was expressed and purified by amylose resin affinity column chromatography as described previously (17). Purified protein was used to immunize a Japanese White rabbit after collection of the preimmune serum. Antibody was purified by ammonium sulfate precipitation as described previously (13).

EGMSA. The maltose-binding protein (MBP) fusion proteins were expressed and purified as described above. Oligonucleotides containing site F2p and the NF-E2 site of the chicken β enhancer (5′-TCGCCCGAAAGGAGCTGACTCA TGCTAGCCC-3′) were labeled with [γ-^{32}P]ATP by T4 polynucleotide kinase. The electrophoretic gel mobility shift assay (EGMSA) was performed as described previously (17).

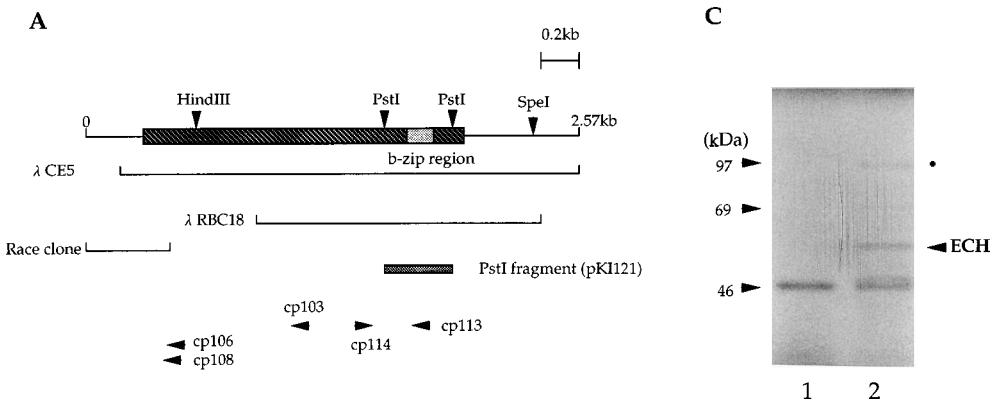
Cotransfection-transactivation assay. Quail fibroblast cell line QT6 was maintained in Dulbecco's modified Eagle's medium–10% fetal bovine serum. Transfection was carried out by the standard $CaPO_4$ precipitation method (32). In brief, QT6 cells were seeded in 24-well dishes (10^4 cells per well with 0.3 ml of medium) 24 h before transfection. Various amounts of pEFECH were transfected in the presence of 10 ng of the reporter plasmid and 50 ng of pENL (β-galactosidase expression plasmid). The total amount of transfected plasmids was adjusted by reference to the amount of parental plasmid pEF-*Bss*HII. Cytoplasmic extracts were prepared 36 h after transfection, and luciferase activities were quantitated with an Autolumat luminometer (Berthold). Relative luciferase reporter activities were determined by normalizing the luciferase activity to β-galactosidase activity.

Nucleotide sequence accession number. The nucleotide sequence data reported in this paper will appear in GSDB, DDBJ, EMBL, and NCBI nucleotide sequence databases with accession number D49365.

RESULTS

Isolation of cDNAs encoding a protein with a CNC domain. Our aim was to identify a chicken transcription factor which regulates transcription through site F2p of the chicken β-globin enhancer. Since site F2p has a DNA sequence motif similar to that of the NF-E2 binding site, we expected that the structure of such a factor would have some structural similarity to that of p45 NF-E2. Thus, we screened an anemic chicken peripheral RBC cDNA library (RBC4) under low stringency by using a cDNA fragment encoding the b-zip region of mouse p45 NF-E2. As a result, 42 recombinant bacteriophage clones were isolated in the initial screening of 5×10^5 recombinant plaques, and partial structures of the cDNA inserts in 4 of the 42 bacteriophage clones were determined by sequencing the 5′ and 3′ termini of the cDNA inserts. All four inserts had the same 3′-terminal sequences, but they varied in length at their 5′ termini. We determined the entire nucleotide sequence of the longest cDNA insert in clone λRBC18 and found that the insert had a reading frame that encoded a b-zip structure that is remarkably similar to that of mouse p45 (Fig. 1A).

To obtain a complete open reading frame (ORF) structure, we screened another cDNA library, which was made from mRNA isolated from 5-day-old whole chicken embryos (a generous gift from Atsushi Kuroiwa), using a cDNA fragment that



B

```
gatgtttcaccctgcccttagagattagacgctttcttcaggggtagcaaggtatgaggaagaaggtgctttttcgcagcgccagtgaaacagaaggaac
aaatggaacgaattatgaattttgaacgttgttgctgaggttaatctactgttggaagagatctgccttgccccgtcatcacctcttctgcaccgaactg
ggaataacatggtggcagtgaagaggaacagagagcagagaaaaggacatgaacttgattgacatccttttggaggcaagatatagaccttggggcaaggc
                                                    M  N  L  I  D  I  L  W  R  Q  D  I  D  L  G  A  R  R
gtgaagtttttgattttagtcaacgacagaaggagtatgaactcgagaaacagaagaaacttgaaaaggaaagacaagagcagctccagaaagagcggga 400
  E  V  F  D  F  S  Q  R  Q  K  E  Y  E  L  E  K  Q  K  K  L  E  K  E  R  Q  E  Q  L  Q  K  E  R  E
gaaagccttgctggctcagctggtgttagatgaggagacaggtgaatttgttccagctcagccagctcagcgcgttcagtcagaaaacgctgaaccacca 500
 K  A  L  L  A  Q  L  V  L  D  E  E  T  G  E  F  V  P  A  Q  P  A  Q  R  V  Q  S  E  N  A  E  P  P
atcagtttttcgcagagcacagatacttcaaaaccagaagaggccttgtcctttgatgactgcatgcagctcttggcagaagctttcccgttcatagatg 600
I  S  F  S  Q  S  T  D  T  S  K  P  E  E  A  L  S  F  D  D  C  M  Q  L  L  A  E  A  F  P  F  I  D  D
acaatgaggcttctccagctgcatttcaatcactggttcctgatcagattgatagtgacccagtcttcatttctgctaatcaaactcagccacccagctc 700
  N  E  A  S  P  A  A  F  Q  S  L  V  P  D  Q  I  D  S  D  P  V  F  I  S  A  N  Q  T  Q  P  P  S  S
acctggtatagttccacttactgatgcagagaatatgcagaacatagagcaagtttgggaagagttattgtcccttccagagttacagtgtctaaacatt 800
 P  G  I  V  P  L  T  D  A  E  N  M  Q  N  I  E  Q  V  W  E  E  L  L  S  L  P  E  L  Q  C  L  N  I
gaaaacgacaacctggctgaagtaagcacaatcacgagccctgaaaccaagccagcggagatgcacaacagctatgattactacaactcgttaccgatca 900
E  N  D  N  L  A  E  V  S  T  I  T  S  P  E  T  K  P  A  E  M  H  N  S  Y  D  Y  Y  N  S  L  P  I  M
tgagaaaagatgttaactgtggtccagatttcctggagaatatcgagggccctttttccagcattttgcagccagacgacagccagctgaatgtgaa 1000
  R  K  D  V  N  C  G  P  D  F  L  E  N  I  E  G  P  F  S  S  I  L  Q  P  D  D  S  S  Q  L  N  V  N
ctctttaaataactcattgactttgagctctgatttctgtgaagatttctataccaatttttatttgtgcaaaggggacggtgacacaggaacaacaaac 1100
 S  L  N  N  S  L  T  L  S  S  D  F  C  E  D  F  Y  T  N  F  I  C  A  K  G  D  G  D  T  G  T  T  N
actatcagtcagtcacttgcagatattttaagtgaacctattgatctttctgatttcccgctgtggagagcttttaacgatgaccactcaggaactgtac 1200
T  I  S  Q  S  L  A  D  I  L  S  E  P  I  D  L  S  D  F  P  L  W  R  A  F  N  D  D  H  S  G  T  V  P
cagagtgtaatgattctgactctggtatttcactgaatgcaaattctagtatagcatcacctgaacactctgttgaatcatctacctgtggagataagac 1300
  E  C  N  D  S  D  S  G  I  S  L  N  A  N  S  S  I  A  S  P  E  H  S  V  E  S  S  T  C  G  D  K  T
ttttggttgtagtgattctgaaatggaagacatggacagttctcctgggagtgtgccgcagggcaatgctagtgtgtactcatcgcggttccctgatcag 1400
 F  G  C  S  D  S  E  M  E  D  M  D  S  S  P  G  S  V  P  Q  G  N  A  S  V  Y  S  S  R  F  P  D  Q
gtacttccttctgtggagccgggcactcaaacgcccagtttgcaacgtatgaacacaccaaagaaagaccctcctgctggcccaggccaccctaaagctc 1500
V  L  P  S  V  E  P  G  T  Q  T  P  S  L  Q  R  M  N  T  P  K  K  D  P  P  A  G  P  G  H  P  K  A  P
cattcacaaaagataaaccttcaggccgtcttgaagctcatctcacaagagatgagcaaagagcaaaagctctgcagatccctttttccagtagaaaaaat 1600
  F  T  K  D  K  P  S  G  R  L  E  A  H  L  T  R  D  E  Q  R  A  K  A  L  Q  I  P  F  P  V  E  K  I
catcaatctccctgttgatgacttcaatgaaatgatgtcgaaggagcagttcagtgaagcccagctcgcactcattcgagatatacgcaggagaggcaag 1700
 I  N  L  P  V  D  D  F  N  E  M  M  S  K  E  Q  F  S  E  A  Q  L  A  L  I  R  D  I  R  R  R  G  K
aataaagtggctgcccaaaactgccgtaagagaaaactggaaaacatagtggaactggagcaagatttgagtcacctaaaagatgagagagagaaattgc 1800
N  K  V  A  A  Q  N  C  R  K  R  K  L  E  N  I  V  E  L  E  Q  D  L  S  H  L  K  D  E  R  E  K  L  L
ttaaagaaaaaggggagaatgacaaaagccttcgtcaaatgaaaaagcagttaaccaccttatatattgaggtgttcagcatgctgcgtgatgaagatgg 1900
  K  E  K  G  E  N  D  K  S  L  R  Q  M  K  K  Q  L  T  T  L  Y  I  E  V  F  S  M  L  R  D  E  D  G
gaagtcttactctcctagtgaatactcactgcagcaaactagagatggcaatatcttccttgttcctaaaagcagaaaggcagagactaaactgtgaaga 2000
 K  S  Y  S  P  S  E  Y  S  L  Q  Q  T  R  D  G  N  I  F  L  V  P  K  S  R  K  A  E  T  K  L  *
gcagcacggctggctggtgttctccgagttactatttttgtatcgttatcctaatagcttttactgtgaggtggaatgcagaattaggtaatatttttca
agtaattctatgcaatgatgatttgaaaattcatagttaagtttatgaaaggtgcaagttcaaaactaatggtgtaatgcagacatacgtatgcaaaatt
cataatctcacttttctaacaggcatttccttcttttagcaccactctaactagtttccatacatgtgtaaatatttaaagataccgtatttatatactg
ttcttatctcctattcatagatttgatatataagaataattttagactcctaaatataatttcttgcaagacaaacagtatggcttcttaaccttttgtg
aaagaaaaaaaaatgcaatagcgtttgtttccaatttttcttatgcatttataaatgctctatatttaatatatgattttaatttagtgtcaggttgata
ctcattgtttgattttgattcatcttacccgttcattaaacttttttattattgttaaaaaaaaaa
```

FIG. 1. (A) Schematic representation of ECH cDNA. The coding region, b-zip coding region, and UTRs are shown by striped boxes, the dotted box, and thin bars, respectively. Restriction enzyme sites are shown above the line. Phage and RACE clones from which the composite cDNA sequence was derived are shown below the line. The *Pst*I fragment which encodes the b-zip domain is shown as a crosshatched box and was cloned in pKI121. Positions and directions of oligonucleotides are shown with arrowheads. (B) Nucleotide sequence and conceptual translation of the ORF of ECH cDNA. (C) In vitro transcription and translation of ECH mRNA. Translation reactions in rabbit reticulocyte lysate were carried out without or with the ECH mRNA synthesized in vitro (lanes 1 and 2, respectively). The specific translation product is shown by the arrowhead at the right. A product of unknown origin is indicated by a dot. Positions of molecular weight markers and their sizes are shown at the left.

spans the b-zip coding region of λRBC18 (pKI121; Fig. 1A) as a probe. Seven additional clones were isolated. The cDNA inserts in these clones showed similar restriction enzyme site patterns, indicating that all cDNA inserts were derived from the same mRNA. Subsequent sequence analysis of the longest cDNA insert in λCE5 verified that the insert overlaps, but is more 5′ extended than, the cDNA insert in λRBC18 (Fig. 1A).

We then performed a 5′-RACE analysis to obtain additional upstream sequences. Sequence information from the 5′ end of λCE5 was used to design a set of primers (primers cp103, cp106, and cp108; Fig. 1A). Poly(A)$^+$ RNA from chicken erythroid cell line HD3 (4) was reverse transcribed by using cp103 as a primer, and an adaptor was ligated to the resultant cDNAs. The cDNAs were amplified by PCR using cp106 and a primer complementary to the adaptor sequence. In a second PCR, cp108 was used instead of cp106. We identified amplified cDNAs of approximately 450 bp (data not shown) and cloned the cDNAs into a plasmid vector. Sequence analysis of several independent clones revealed that all clones were initiated from the proper priming site and contained sequences that overlap the 5′ end of the cDNA in λCE5 (Fig. 1A).

A composite cDNA sequence (2,565 bp) was generated with the sequence information obtained from the phage clones (λRBC18 and λCE5) and the RACE clones (Fig. 1B). The sequence contains a single ORF flanked by several in-frame termination codons in both the 5′ and 3′ untranslated regions (UTRs). The ORF encodes a protein of 582 amino acid residues with a predicted molecular mass of 65.5 kDa. Nucleotide sequence surrounding the first ATG codon matches the Kozak consensus for efficient initiation of translation (18), strongly suggesting that this ATG is utilized as the translation initiation codon. The absence of an authentic polyadenylation consensus sequence (AAUAAA) in the 3′ UTR of the cDNA suggests that cDNA synthesis was initiated from an A-rich region within the 3′ UTR.

To verify the coding potential of the ORF within the cDNA in pCE5, in vitro transcription-translation with the reticulocyte lysate was performed. A specific translation product of 63 kDa was detected in SDS-polyacrylamide gel electrophoresis (Fig. 1C). The size is consistent with the predicted size of the conceptual ORF product, supporting the assignment of the initiation codon and ORF. The nature of the larger translation product (97 kDa) is not clear at present (see below).

Structure of an erythroid cell-derived CNC-homologous protein. The ORF encodes a protein with a characteristic b-zip domain (Fig. 1B, 2A, and 2B). The basic region is almost identical to those of p45 NF-E2 and its related proteins, Nrf1/LCR-F1, Nrf2, and CNC (the CNC family proteins; Fig. 2B). The conservation of amino acid sequence extends amino terminally from the basic region, indicating that the cDNAs isolated here encode a new member of the CNC family proteins (see below). Because of the similarity with CNC, we named the protein product encoded by these cDNAs ECH.

A region in ECH located carboxy terminally to the leucine zipper displays significant similarity with the corresponding regions in other CNC family proteins (Fig. 2B). In contrast, amino acid sequences in the leucine zipper regions are less conserved among the CNC family proteins, suggesting that these proteins may have different specificities in dimer formation. We found a segment in the amino-terminal region of ECH which is conserved among Nrf1 and Nrf2 (domain 2; Fig. 2A and C). LCR-F1, which is transcribed from the same gene as Nrf1, and p45 NF-E2 lack this conserved segment. Among the CNC family proteins thus far identified, ECH showed the highest similarity to Nrf2 (Fig. 2D).

ECH is the major CNC family gene expressed in chicken erythroid cells and is encoded by a single-copy gene. The results described above showed that 4 of the 42 clones isolated from the anemic chicken RBC cDNA library encoded portions of ECH. cDNA inserts in the remaining 38 clones were analyzed by PCR to confirm identity and length of the cDNA insert. We prepared a primer set which can amplify the b-zip region of ECH (cp114 and cp113; Fig. 1A) and carried out PCR. We detected the expected 228-bp PCR product in all of the 38 clones except λRBC9, -20, and -26 (numbering corresponds to that of λ phage recombinant clones obtained from the RBC4 library; data not shown). We also analyzed these 38 clones by PCR with the λgt11 forward primer and cp113 to detect shorter cDNA clones and to estimate the lengths of the cDNAs. Of 38 clones, 32, including λRBC20, were positive in the PCR amplification, yielding DNA fragments of variable lengths which hybridized with the b-zip probe (pKI121). λRBC20 gave an amplification product of 160 bp. Thus, λRBC20 contained a shorter ECH cDNA fragment whose 5′ end was located downstream of the cp114 primer. λRBC9 and -26 were again negative in this assay. The four clones which were negative in this PCR but were positive in the first PCR may contain a long insert which could not be amplified efficiently.

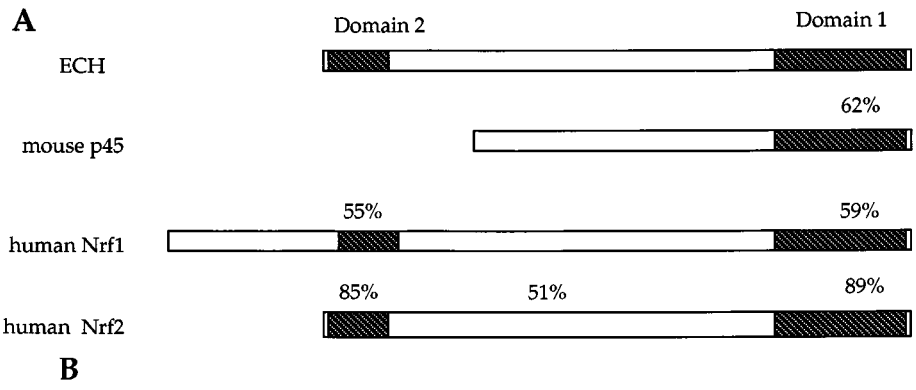
These results indicated that 40 of the 42 recombinant cDNA clones, most of which were judged to be independent clones, were originated from ECH mRNA. Thus, ECH is the predominant population among the phage clones isolated from the anemic chicken RBC cDNA library. Although frequencies of particular clones in a cDNA library do not precisely reflect the relative amounts of mRNAs within cells, the results nonetheless strongly suggest that ECH is the major CNC family gene expressed in chicken erythroid cells.

To determine whether ECH is encoded by a single-copy gene, we carried out a genomic DNA blot analysis. High-molecular-weight DNA isolated from chicken liver was digested with various restriction endonucleases, separated by gel electrophoresis through a 0.8% agarose gel, and transferred onto a nylon membrane. The latter was then hybridized with the *Pst*I DNA fragment containing the b-zip coding region of ECH (pKI121; Fig. 1A). We found the presence of a single band in each lane which specifically hybridized to the probe DNA (data not shown). The result strongly suggests that ECH is encoded by a single-copy gene.

Expression of ECH. To shed light on the functional roles of ECH in vivo, we examined the expression profile of ECH. Poly(A)$^+$ RNA samples were isolated from peripheral RBCs, various tissues, and cultured cell lines and were separated by electrophoresis through agarose gels, transferred onto nylon membranes, and hybridized with an ECH-specific probe (Fig. 3A). ECH transcript was detected as two distinct signals approximately 3.4 and 2.8 kb in size. Since ECH is encoded by a single-copy gene (see above), both of the ECH mRNA species are likely to be generated from the same gene. Consistent with this inference, there was heterogeneity at both 5′ and 3′ ends of the ECH cDNA clones (data not shown).

Of the tissues examined, ECH was expressed most abundantly in the peripheral RBCs (Fig. 3A, lane 1). The kidney and intestine expressed moderate levels of ECH mRNA, whereas skeletal muscle and the brain, liver, and heart expressed significantly less ECH mRNA (lanes 5 to 10). Considering the expression in the peripheral RBCs, these low levels of expression of ECH mRNA may be derived from contaminating blood cells. These results indicated that the expression levels of ECH differed substantially among various tissues.

Cell line HD3 is transformed by a strain of avian erythroblastosis virus that contains the v-*erbB* oncogene (4). These



B

Domain 1 (CNC homology domain)

basic region

```
ECH   RDEQRAKALQIPFPVEKIINLPVDDFNEMMSKEQFSEAQLALIRDIRRRGKNKVAAQNCRKR
Nrf2  RDELRAKALHIPFPVEKIINLPVVDFNENMSKEQFNEAQLALIRDIRRRGKNKVAAQNCRKR
p45   RDERRALAMKIPFPTDKIVNLPVDDFNELLAQYPLTESQLALVRDIRRRGKNKVAAQNCRKR
Nrf1  RDEHRARAMKIPFTNDKIINLPVEEFNELLSKYQLSEAQLSLIRDIRRRGKNKMAAQNCRKR
CNC   RDEKRARSLNIPISVPDIINLPMDEFNERLSKYDLSENQLSLIRDLRRRGKNKVAAQNCRKR

ECH   KLENIVELEQDLSHLKDEREKLLKEKGENDKSLRQMKKQLTTLYIEVFSMLRDEDGKSYSPS
Nrf2  KLENIVELEQDLDHLKDEKEKLLKEKGENDKSLHLLKKQLSTLYLEVFSMLRDEDGKPYSPS
p45   KLETIVQLERELERLSSERERLLRARGEADRTLEVMRQQLAELYHDIFQHLRDESGNSYSPE
Nrf1  KLDTILNLERDVEDLQRDKARLLREKVEFLRSLRQMKQKVQSLYQEVFGRLRDENGRPYSPS
CNC   KLDQILTLEDEVNAVVKRKTQLNQDRDHLESERKRISNKFAMLHRHVFQYLRDPEGNPCWPA

ECH   EYSLQQTRDGNIFLVPKSRKAETKL
Nrf2  EYSLQQTRDGNVFLVPKSKKPDVKKN
p45   EYVLQQAADGAIFLVPRGTKMEATD
Nrf1  QYALQYAGDGSVLLIPRTMADQQARRQERKPKDRRK
CNC   DYSLQQAADGSVYLLPREKSEGNNTATAASNAVSSA
```

C

```
Nrf2 MDLIDILWRQDIDLGVSREVFDFSQRRKEYELEKQKKLEKERQEQLQKEQEKAFFTQLQLDEETGEFLPIQPA
ECH  MNLIDILWRQDIDLGARREVFDFSQRQKEYELEKQKKLEKERQEQLQKEREKALLAQLVLDEETGEFVPAQPA
Nrf1 IDLIDILWRQDIDLGAGREVFDYSHRQKEQDVEKELR-DGGEQDTWAGEGAEALARNLLVDGETGESFPAQFP
```

FIG. 2. (A) Conserved domains (1 and 2) among chicken ECH, mouse p45, and human Nrf1 and Nrf2. The homology of each domain with the corresponding domain of ECH is shown between the bars. (B) Comparison of conserved domain 1 of ECH with CNC family proteins (mouse p45, ECH, human Nrf1, human Nrf2, and *Drosophila* CNC). Amino acids conserved among at least four proteins are shaded. Leucine zippers are indicated by dots above the sequences. (C) Comparison of the conserved domain 2 of ECH and human Nrf1 and Nrf2. Amino acids conserved among at least two proteins are shaded. (D) Comparison of amino acid sequences of ECH and human Nrf2. Identical amino acids are shaded.

cells can be induced to undergo terminal erythroid differentiation by shifting the culture to a nonpermissive temperature (42°C), thus inactivating the v-ErbB oncoprotein, and incubating cells in the presence of appropriate growth factors. During differentiation in vitro, the expression of erythroid cell-specific genes, such as globin genes, is known to increase significantly (5). HD3 cells which had been cultured under differentiation-inducing conditions expressed more ECH mRNA (Fig. 3A, lanes 3 and 4, and 3B). The upregulation of ECH mRNA during this period paralleled that of erythroid δ-aminolevulinate synthase mRNA (Fig. 3B), which is important for terminal differentiation of erythroid cells (10, 19, 37). Taken together, these results suggest that ECH is important in the terminal differentiation of erythroid cells. Lymphoid cell line MSB-1 (1) expressed less ECH mRNA than did HD3 (Fig. 3A, lane 2).

To identify the presence of ECH protein in erythroid cells, we prepared a rabbit antiserum directed against ECH by inoculating bacterially expressed ECH. Reactivity of the antibody was assessed in extracts prepared from QT6 cells transfected with the ECH expression plasmid. In an immunoblot analysis (Fig. 4), the antibody reacted with a protein whose

D

```
                                                                    60
ECH   MNLIDILWRQDIDLGARREVFDFSQRQKEYELEKQKKLEKERQEQLQKEREKALLAQLVL
Nrf2  MDLIDILWRQDIDLGVSREVFDFSQRRKEYELEKQKKLEKERQEQLQKEQEKAFFTQLQL
                                                                   120
      DEETGEFVPAQPAQRVQSENAEPPISFSQSTDTSKPEEALSFDDCMQLLAEAFPFIDDNE
      DEETGEFLPIQPAQHTQSETSGSA-NYSQVAHIPKSD-ALYFDDCMQLLAQTFPFVDDNE
                                                                   180
      ASPAAFQSLVPD---QIDSDPVFISANQTQPP--SSPGIVPLTDAENMQNIEQVWEELLS
      VSSATFQSLVPDIPGHIES-PVFIATNQAQSPETSVAQVAPVDLDGMQQDIEQVWEELLS
                                                                   240
      LPELQCLNIENDNLAEVSTITSPETKPAEMHNSYDYYNSLPIMRKDV-NCGPDFLENIEG
      IPELQCLNIENDKLVETTMVPSPEAKLTEV-DNYHFYSSIPSMEKEVGNCSPHFLNAFED
                                                                   300
      PFSSILQPDDSSQLNVNSLNNSLTLSSDFCEDFYTNFICAKGDGDT-GTTNTISQSLADI
      SFSSILSTEDPNQLTVNSLNSDATVNTDFGDEFYSAFIAEPSISNSMPSPATLSHSLSEL
                                                                   360
      LSEPIDLSDFPLWRAFNDDHSGTVPECNDSDSGISLNANSSIASPEHSVESSTCGDKTFG
      LNGPIDVSDLSLCKAFNQNHPESTAEFNDSDSGISLNTSPSVASPEHSVESSSYGDTLLG
                                                                   420
      CSDSEMEDMDSSPGSVPQ--GNASVYSS-RFPDQVLPSVEPGTQTPSLQRMNTPKKDPPA
      LSDSEVEELDSAPGSVKQNGPKTPVHSSGDMVQPLSPSQGQSTHVHDAQCENTPEKELPV
                                                                   480
      GPGHPKAPFTKDKPSGRLEAHLTRDEQRAKALQIPFPVEKIINLPVDDFNEMMSKEQFSE
      SPGHRKTPFTKDKHSSRLEAHLTRDELRAKALHIPFPVEKIINLPVVDFNEMMSKEQFNE
                                                                   540
      AQLALIRDIRRRGKNKVAAQNCRKRKLENIVELEQDLSHLKDEREKLLKEKGENDKSLRQ
      AQLALIRDIRRRGKNKVAAQNCRKRKLENIVELEQDLDHLKDEKEKLLKEKGENDKSLHL

      MKKQLTTLYIEVFSMLRDEDGKSYSPSEYSLQQTRDGNIFLVPKSRKAETKL
      LKKQLSTLYLEVFSMLRDEDGKPYSPSEYSLQQTRDGNVFLVPKSKKPDVKKN
```

FIG. 2—*Continued.*

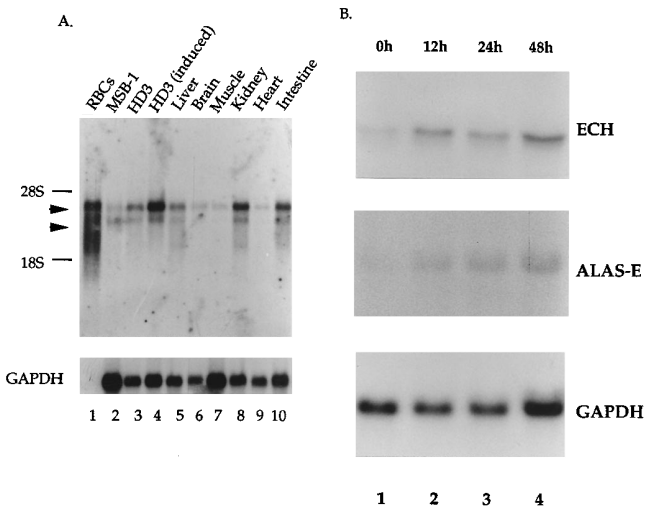


FIG. 3. Expression of ECH mRNA. (A) Poly(A)$^+$ RNA (1.5 μg) isolated from cultured cells (lymphoid cell line MSB-1, erythroid cell line HD3, and HD3 cultured in differentiation conditions for 2 days), anemic chicken peripheral RBCs induced with phenylhydrazine, and tissues of adult chickens as indicated above the lanes were separated on agarose-formaldehyde gels and transferred onto nylon membranes. The membranes were hybridized with either an ECH-specific probe (above) or a GAPDH-specific probe (below). The positions of ECH mRNAs are indicated by arrowheads; the positions of 18S and 28S rRNAs are shown by bars. (B) HD3 cells were cultured at the nonpermissive temperature (42°C) for various periods to induce erythroid differentiation, and RNA samples were harvested. Expression of erythroid δ-aminolevulinate synthase (ALAS-E), ECH, and GAPDH mRNAs was analyzed as described above.

molecular mass was 63 kDa and which was absent in cells transfected with the vector plasmid but present in cells transfected with the ECH expression plasmid (lanes 1 and 2). Thus, the antibody specifically recognized ECH. The antibody also reacted with a protein of 97 kDa which was present in cells transfected with the ECH expression plasmid. This band may represent ECH with some modification. However, the nature of this band is not clear at present.

This antibody reacted with a protein in the chicken RBC whole cell extract (basic extract) (12) whose size was the same as that of ECH expressed in the QT6 cells (lanes 3 and 4). Hence, ECH is present in chicken erythroid cells.

ECH binds DNA as heterodimers with small Maf family proteins. Mouse p45 NF-E2 acquires high DNA binding activity by forming heterodimers with the small Maf family proteins in vitro (3, 15). We showed that ECH binds to the consensus NF-E2 site as heterodimers with the small Maf family proteins (16). To test whether ECH can bind to the NF-E2 site within site F2p, the b-zip domains of ECH and the small Maf family proteins were expressed in *E. coli* as fusion proteins with MBP, and heterodimer formation was examined in an EGMSA.

As shown in Fig. 5A, we could not detect any DNA binding activity when a site F2p probe containing the NF-E2 site of the chicken β enhancer was incubated with bacterially expressed ECH alone (lane 2). Faint but specific DNA-protein complexes could be observed when the probe was incubated with the small Maf proteins (lanes 6 to 8). In contrast, the addition of ECH to any of the small Maf family proteins resulted in much more intense signals (lanes 3 to 5). The mobilities of the complexes formed in the presence of both ECH and MafF or MafK were similar to those of homodimeric complexes of MafF or MafK, respectively, since these fusion proteins were of similar sizes. On the other hand, the mobility of the complex observed in the presence of both ECH and MafG was faster than that of the MafG homodimeric complex (this is apparent in Fig. 5B). These results clearly demonstrated that ECH can interact with the small Maf family proteins similarly to p45, and the small Maf family proteins confer strong DNA binding activity to ECH.

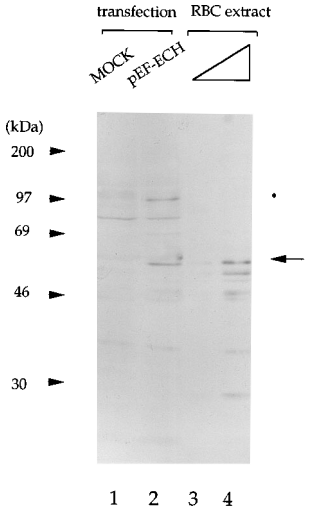


FIG. 4. Detection of ECH in erythroid extract with an anti-ECH antiserum. Cell extracts from QT6 cells transfected with the pEF-*Bss*HII vector or ECH expression plasmid (lanes 1 and 2, respectively) or increasing amounts of the basic extract from chicken RBCs (lane 3 and 4) were separated on an SDS-polyacrylamide gel, transferred onto a membrane, and reacted with the antibody against bacterially expressed ECH. The position of ECH is indicated by an arrow. An immunoreactive band of unknown origin is indicated with a dot. Molecular size markers are indicated by arrowheads.

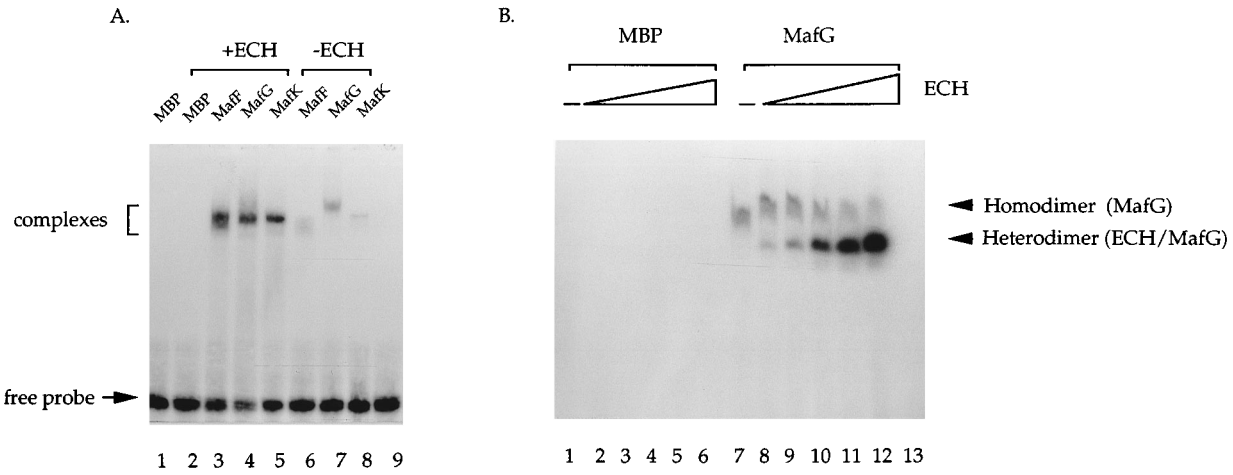


FIG. 5. Binding of ECH and small Maf heterodimers to site F2p in EGMSA. (A) MBP, MBP-MafF, MBP-MafG, or MBP-MafK fusion protein (100 ng) was incubated with the site F2p probe containing an NF-E2 site in the presence (lanes 2 to 5) or absence (lanes 6 to 9) of 20 ng of MBP-ECH fusion protein. (B) Binding buffer alone (lane 13) or twofold increments (2.5 to 40 ng) of MBP-ECH (lanes 2 to 6 and 8 to 12) were incubated with the site F2p probe in the absence (lanes 1 to 6) or presence (lanes 7 to 12) of 100 ng of MBP-MafG.

To compare the DNA binding activities of ECH for the NF-E2 site in the presence and absence of small Maf family proteins, increasing amounts of ECH were incubated with the site F2p probe in the presence or absence of a constant amount of MafG protein. Binding of the ECH homodimer was not detectable even with 40 ng of ECH, whereas binding of the heterodimer was detected with as little as 2.5 ng of ECH in the presence of excess MafG (Fig. 5B). These results suggest that the formation of heterodimers with the small Maf family proteins is an obligatory step for ECH to bind the NF-E2 site.

ECH is a strong transcriptional activator. We examined the function of ECH as a transcriptional regulator in a transient cotransfection assay using quail fibroblast QT6 cells as recipient cells. Two kinds of luciferase reporter gene plasmids were used in this analysis; pRBGP2 contains the NF-E2 site and surrounding region (site F2p) of the chicken β-globin enhancer upstream of a rabbit β-globin TATA box, whereas pRBGP4 contains a mutated NF-E2 site which cannot be bound by mouse NF-E2 or AP-1 (15). When these reporter plasmids were transfected into QT6 cells, luciferase reporter gene activity was significantly higher in cells transfected with the test plasmid pRBGP2 than in cells transfected with the control plasmid pRBGP4 (Fig. 6A). The results showed that the endogenous activity stimulating transcription from pRBGP2 was dependent on the presence of an intact NF-E2 site and suggested that the activity was probably due to the binding of AP-1 families to the NF-E2 site, as the NF-E2 site encompasses a TPA-responsive element.

Cotransfection of the ECH expression plasmid with the reporter pRBGP2 increased the level of luciferase expression in a dose-dependent manner (Fig. 6A). Contrary to this result, cotransfection of p45 NF-E2 barely affected the reporter gene expression from pRBGP2 by itself (data not shown; see reference 15). Cotransfection of the ECH expression plasmid with pRBGP4 did not result in an increase in the level of reporter gene transcription. These results indicate that ECH is a transcriptional activator that requires an NF-E2 site in target genes for its activity.

When we cotransfected a MafK expression plasmid with pRBGP2, MafK efficiently repressed the expression of the reporter gene from pRBGP2 (Fig. 6B) as described previously (15). However, this MafK-mediated repression was antagonized efficiently by concomitant transfection of the ECH expression plasmid. By increasing the amount of the ECH expression plasmid, more than 20-fold as much transactivation was attained compared with that in the absence of any effector plasmids. Transcription stimulation observed in the presence of exogenously expressed MafK and ECH was again significantly higher than that achieved by coexpressing MafK and p45 NF-E2, and similar results were obtained with MafG and MafF instead of MafK (data not shown). Thus, depending on the amounts of ECH and the small Maf family proteins within cells, NF-E2 site-dependent transcription can be regulated both positively and negatively within a dynamic range of 100-fold.

To compare directly the transcriptional activation potentials of p45 NF-E2 and ECH, we transfected various amounts of the p45 expression plasmid in the presence of fixed amounts of the ECH expression plasmid and the reporter plasmid. As shown in Fig. 6C, p45 interfered with ECH-mediated activated transcription in a dose-dependent manner. This *trans*-dominant inhibitory effect of p45 against ECH is consistent with the hypothesis that ECH is a more potent *trans* activator than p45 when tethered on the NF-E2 sites in a fibroblast environment.

DISCUSSION

In this report, we described the isolation and characterization of chicken cDNAs which encode a novel transcription factor, ECH. Salient features of the study include comparison of the structure, expression profile, and *trans* activation activity of ECH with those of p45 NF-E2 and other CNC family factors.

Since p45, Nrf1/LCR-F1, Nrf2, and ECH possess virtually identical basic regions, it is likely that these proteins can recognize related DNA motifs. In the case of NF-E2, p45 and the small Maf family proteins recognize the shorter and the longer halves of the NF-E2 site, respectively [TGCTGA(G/C)TCA(T/C); the underlined region is recognized by p45 (3, 28a)]. Thus, it is possible that basic regions of Nrf1/LCR-F1, Nrf2, and ECH also recognize the same four-base sequence motif as p45 does. In fact, ECH could bind the authentic NF-E2 site in the porphobilinogen deaminase (PBGD) promoter as a heterodimer with the small Maf family proteins (data not shown),

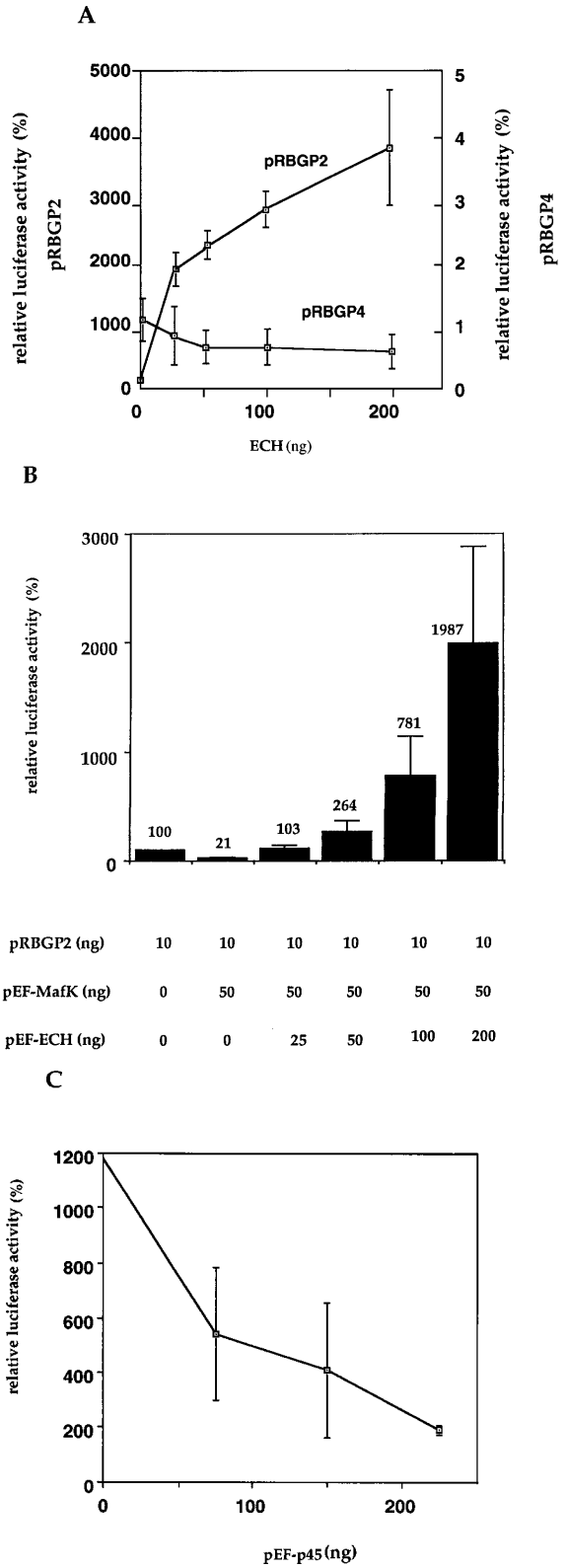


FIG. 6. Transcription activation by ECH. (A) Increasing amounts of an ECH expression plasmid were transfected into quail fibroblast QT6 cells together with the reporter plasmid pRBGP2 or pRBGP4. Luciferase activity of pRBGP2 in the absence of any effector plasmids was set at 100%, and the results of the three independent experiments each carried out in duplicate are shown. Bars indicate 1 standard deviation. Note that pRBGP4 expressed 100-fold less luciferase activity than pRBGP2 in the absence of the ECH expression plasmid. (B) Various amounts of ECH and MafK expression plasmids were cotransfected with pRBGP2. Results are presented as described above. (C) Various amounts of a mouse p45 NF-E2 expression plasmid were transfected with 25 ng of the ECH expression plasmid. Luciferase activity of pRBGP2 in the absence of any effector plasmids was set at 100%, and the results are presented as described above.

indicating that p45 and ECH recognize the similar sequence motifs as parts of complexes with the small Maf family proteins. Although LCR-F1 was shown to bind an NF-E2 site as a homodimer (6), it remains possible that LCR-F1 (and Nrf1) binds NF-E2 sites more strongly by forming heterodimers with the small Maf family proteins or a related protein.

The leucine zipper within the b-zip protein mediates dimer formation. In this regard, it is interesting that among the CNC family proteins, amino acid residues between the leucine repeats are not strictly conserved. Since amino acid residues between leucine repeats contribute to specificities in dimer formation (35), these proteins may choose distinct partners for dimer formation. However, ECH could form a heterodimer with any of the small Maf family proteins as efficiently as p45 NF-E2 does (Fig. 5A and data not shown). Thus, ECH has a dimerization specificity very similar to that of p45 NF-E2 at least in vitro.

Expression of ECH is most abundant in peripheral RBCs. Erythroid cell line HD3 expressed more ECH RNA than did lymphoid cell line MSB-1. Among the nonhematopoietic tissues, the intestine and kidney expressed significant levels of ECH mRNA. In contrast to the expression profile of ECH, both Nrf1/LCR-F1 and Nrf2 are expressed in a wide range of tissues and cells in addition to erythroid cells (6, 7, 25). On the other hand, expression of p45 NF-E2 is restricted to hematopoietic cells and the intestine (2, 26, 29). Thus, the expression profile of ECH is unique among the CNC family genes.

We observed strong transcriptional activation of reporter gene expression (more than 30-fold) in QT6 cells by ECH which is dependent on the presence of an intact NF-E2 site in the reporter plasmid. The difference between the basal reporter gene activities of pRBGP2 and pRBGP4 (Fig. 6A) indicates that the NF-E2 site is a target of transcriptional activation of more than 100-fold by an endogenous factor in QT6 cells. Hence, ECH can activate transcription through NF-E2 sites by several thousand-fold over basal transcription directed by a TATA box. Thus, ECH is a very potent transcriptional activator affecting the NF-E2 site. The *trans* activation ability of ECH is in clear contrast to that of p45 NF-E2, since mouse p45 alone could activate transcription only marginally through NF-E2 sites compared with endogenous activity (15). Even in the presence of exogenous small Maf family proteins, p45 can activate transcription far less significantly than ECH does (data not shown) (15). This finding was corroborated by the results in Fig. 6C. In accord with our observations, the transactivating potential of p45 NF-E2 was significantly weaker than that of LCR-F1 in an erythroid environment (6). The strong transactivating potential of ECH was also evident when erythroleukemia cells were used as recipient cells in the cotransfection analysis (unpublished observation). These observations, taken together, suggest that ECH is more like LCR-F1 in terms of transcription-stimulating activity.

The results of EGMSA (Fig. 5) demonstrated that ECH cannot bind the NF-E2 site as a homodimer but requires formation of a heterodimer with the small Maf family proteins for binding to the site. This prerequisite for the small Maf family proteins for DNA binding poses a question with respect to the results of the cotransfection assay (Fig. 6) as to how ECH bound NF-E2 sites and transactivated reporter gene expression

in the absence of the small Maf expression plasmids. As previously shown, QT6 cells express mRNAs for the small Maf family proteins to some extent (15). Thus, one explanation for the ECH function in QT6 cells would be that ECH formed a heterodimer with the endogenous small Maf family proteins and acquired the NF-E2 site-specific binding ability. Consistent with this hypothesis, NF-E2 site-binding activity that is generated in QT6 cells by the expression of ECH is reactive with an anti-MafK antibody (unpublished observation).

Although p45 NF-E2, Nrf1/LCR-F1, and Nrf2 are all expressed in mammalian erythroid cells, the situation in chicken erythroid cells appears different. Most of the clones (40 of the 42 clones) isolated from the anemic chicken RBC cDNA library by virtue of their similarity with the p45 b-zip domain were found to encode ECH. Furthermore, screening of the 5-day-old chicken embryo cDNA library with the b-zip coding region of ECH cDNA under low-stringency conditions yielded only ECH cDNA. Since the cloning procedures should allow isolation of other cDNA clones encoding CNC-related proteins, it appears that ECH is the predominant CNC family protein in chicken erythroid cells. Thus, most of the genes associated with NF-E2 sites are expected to be regulated by ECH (with the small Maf proteins) in chicken erythroid cells. This consideration raises an interesting question: why are the molecules affecting NF-E2 sites in erythroid cells so different in chickens (ECH) and mammals (p45 NF-E2, Nrf1/LCR-F1, and Nrf2)? An attractive hypothesis is that the differences among the molecules in the various organisms may reflect the divergent mechanisms of gene regulation through NF-E2 sites in these organisms. GATA-1, another erythroid transcription factor, was suggested to be such a case on the basis of the structural and functional differences between human/mouse and chicken GATA-1 molecules (34).

However, it is still possible that homologs of the three factors (p45 NF-E2, Nrf1/LCR-F1, and Nrf2) are expressed in chicken erythroid cells as well. Results of genomic Southern blotting experiments under low-stringency conditions indicated the presence of other ECH-related genes in chickens (unpublished observation). Reciprocally, an ECH homolog may be present in mammalian erythroid cells. Since ECH shows the highest sequence similarity with human Nrf2, these two proteins may be cross-species homologs. However, Nrf2 is expressed ubiquitously (25), in clear contrast to the expression profile of ECH. Thus, the relationship of these two factors remains to be established. In any case, alternative usage of the CNC family proteins with distinct activities may be important for developmental regulation of erythroid genes, such as switching processes in globin gene expression. The differences between these factors in terms of transcription activation potentials (strong activators like ECH and LCR-F1, and presumably Nrf1, and a weak activator like p45 NF-E2, as defined in transient transfection assays) or in other aspects may lead to differential utilization during development and maturation of erythroid cells.

ACKNOWLEDGMENTS

We thank A. Kuroiwa for the chicken whole embryo cDNA library, K. Kataoka for various suggestions, and K. C. Lim for critical reading of the manuscript.

This work was supported in part by grants-in-aid from the Ministry of Education, Science and Culture and grants from the Uehara Memorial Foundation and the Naito Foundation to K.I. and M.Y. and by a grant from the Yamanouchi Foundation for Research on Metabolic Disorders to N.H.

REFERENCES

1. **Akiyama, Y., and S. Kato.** 1974. Two cell lines from lymphomas of Marek's disease. Biken J. **17:**105–116.
2. **Andrews, N. C., H. Erdjument-Bromage, M. B. Davidson, P. Tempst, and S. H. Orkin.** 1993. Erythroid transcription factor NF-E2 is a haematopoietic-specific basic-leucine zipper protein. Nature (London) **362:**722–728.
3. **Andrews, N. C., K. J. Kotkow, P. A. Ney, H. Erdjument-Bromage, P. Tempst, and S. H. Orkin.** 1993. 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.
4. **Beug, H., G. Doederlien, C. Freudenstein, and T. Graf.** 1982. Erythroblast cell lines transformed by a ts mutant of avian erythroblastosis virus: a model system to study erythroid differentiation in vitro. J. Cell. Physiol. Suppl. **1:**195–207.
5. **Beug, H., S. Palmieri, C. Freudenstein, H. Zentgraf, and T. Graf.** 1982. Hormone-dependent terminal differentiation in vitro of chicken erythroleukemia cells transformed by ts mutants of avian erythroblastosis virus. Cell **28:**907–919.
6. **Caterina, J. J., D. Donze, C. W. Sun, D. J. Ciavatta, and T. M. Towns.** 1994. Cloning and functional characterization of LCR-F1: a bZIP transcription factor that activates erythroid-specific, human globin gene expression. Nucleic Acids Res. **12:**2383–2391.
7. **Chan, J. Y., X.-L. Han, and Y. W. Kan.** 1993. Cloning of Nrf1, an NF-E2-related transcription factor, by genetic selection in yeast. Proc. Natl. Acad. Sci. USA **90:**11371–11375.
8. **Chomczynski, I., and N. Sacchi.** 1987. Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. Anal. Biochem. **162:**156–159.
9. **Engel, J. D., H. Beug, J. H. LaVail, M. W. Zenke, K. Mayo, M. W. Leonard, K. P. Foley, Z. Yang, J. M. Kornhauser, L. J. Ko, K. C. Lim, K. M. George, and K. Breigel.** 1992. Cis and trans regulation of tissue-specific transcription. J. Cell Sci. Suppl. **16:**21–31.
10. **Fujita, H., M. Yamamoto, T. Yamagami, N. Hayashi, and S. Sassa.** 1991. Erythroleukemia differentiation. Distinctive responses of the erythroid-specific and the nonspecific delta-aminolevulinate synthase mRNA. J. Biol. Chem. **266:**17494–17502.
11. **Fujiwara, K. T., K. Kataoka, and M. Nishizawa.** 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.
12. **Gallarda, J. L., K. P. Foley, Z. Yang, and J. D. Engel.** 1989. The β-globin stage selector factor is erythroid-specific promoter/enhancer binding protein NF-E4. Genes Dev. **3:**1845–1859.
13. **Harlow, E., and D. Lane.** 1988. Antibodies, a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
14. **Igarashi, K., K. Itoh, H. Motohashi, N. Hayashi, Y. Matuzaki, H. Nakauchi, M. Nishizawa, and M. Yamamoto.** 1995. Activity and expression of murine small Maf family protein MafK. J. Biol. Chem. **270:**7615–7624.
15. **Igarashi, K., K. Kataoka, K. Itoh, N. Hayashi, M. Nishizawa, and M. Yamamoto.** 1994. Regulation of transcription by dimerization of erythroid factor NF-E2 p45 with small Maf proteins. Nature (London) **367:** 568–572.
16. **Kataoka, K., K. Igarashi, K. Itoh, K. T. Fujiwara, M. Noda, M. Yamamoto, and M. Nishizawa.** 1995. Small Maf proteins heterodimerize with Fos and may act as competitive repressors of the NF-E2 transcription factor. Mol. Cell. Biol. **15:**2180–2190.
17. **Kataoka, K., M. Noda, and M. Nishizawa.** 1994. 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.
18. **Kozak, M.** 1989. The scanning model for translation: an update. J. Cell Biol. **108:**229.
19. **Lake-Bullock, H., and H. A. Dailey.** 1993. Biphasic ordered induction of heme synthesis in differentiating murine erythroleukemia cells: role of erythroid 5-aminolevulinate synthase. Mol. Cell. Biol. **13:**7122–7132.
20. **Luna, L., O. Johnsen, A. H. Skartlien, F. Pedeutour, C. Turc-Carel, H. Prydz, and A. Kolsto.** 1994. Molecular cloning of a putative novel human bZIP transcription factor on chromosome 17q22. Genomics **22:**553–562.
21. **Mignotte, V., J. F. Eleouet, N. Raich, and P.-H. Romeo.** 1989. 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.
22. **Mignotte, V., L. Wall, E. deBohr, F. Grosveld, and P.-H. Romeo.** 1989. Two tissue-specific factors bind the erythroid promoter of human porphobilinogen deaminase gene. Nucleic Acids Res. **17:**37–54.
23. **Mizushima, S., and S. Nagata.** 1990. pEF-BOS, a powerful mammalian expression vector. Nucleic Acids Res. **18:**5332.
24. **Mohler, J., K. Vani, S. Leung, and A. Epstein.** 1991. Segmentally restricted, cephalic expression of a leucine zipper gene during Drosophila embryogenesis. Mech. Dev. **34:**3–10.
25. **Moi, P., K. Chan, I. Asunis, A. Cao, and Y. W. Kan.** 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.

26. **Ney, P. A., N. C. Andrews, S. M. Jane, B. Safer, M. E. Purucker, S. Weremowicz, C. C. Morton, S. A. Goef, S. H. Orkin, and A. W. Nienhuis.** 1993. Purification of the human NF-E2 complex: cDNA cloning of the hematopoietic cell-specific subunit and evidence for an associated partner. Mol. Cell. Biol. **13:**5604–5612.
27. **Ney, P. A., B. P. Sorrentino, K. T. McDonagh, and A. W. Nienhuis.** 1990. Tandem AP-1-binding sites within the human β-globin dominant control region function as an inducible enhancer in erythroid cells. Genes Dev. **4:**993–1006.
28. **Ney, P. A., P. Sorrentino, C. H. Lowrey, and A. W. Nienhuis.** 1990. Inducibility of the HS II enhancer depends on binding of an erythroid specific nuclear protein. Nucleic Acids Res. **18:**6011–6017.
28a. **Nishizawa, M.** Unpublished data.
29. **Peters, L. I., N. C. Andrews, E. M. Eicher, M. B. Davidson, S. H. Orkin, and S. E. Lux.** 1993. Mouse microcytic anaemia caused by a defect in the gene encoding the globin enhancer-binding protein NF-E2. Nature (London) **362:**768–770.
30. **Reitman, M., and G. Felsenfeld.** 1988. Mutational analysis of the chicken β-globin enhancer reveals two positive-acting domains. Proc. Natl. Acad. Sci. USA **86:**6267–6271.
31. **Reitman, M., E. Lee, H. Westphal, and G. Felsenfeld.** 1990. Site-independent expression of the chicken β^A-globin gene in transgenic mice. Nature (London) **348:**749–752.
32. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
33. **Strauss, E. C., N. C. Andrews, D. R. Higgs, and S. H. Orkin.** 1992. In vivo footprinting of the human alpha-globin locus upstream regulatory element by guanine and adenine ligation-mediated polymerase chain reaction. Mol. Cell. Biol. **12:**2135–2142.
34. **Trainor, C. D., T. Evans, G. Felsenfeld, and M. S. Boguski.** 1990. Structure and evolution of human erythroid transcription factors. Nature (London) **343:**92–96.
35. **Vinson, C. R., T. Hai, and S. M. Boyd.** 1993. Dimerization specificity of the leucine zipper-containing bZip motifs on DNA binding: prediction and rational design. Genes Dev. **7:**1047–1058.
36. **Yamamoto, M., L. J. Ko, M. W. Leonard, H. Beug, S. H. Orkin, and J. D. Engel.** 1990. Activity and tissue-specific expression of the transcription factor NF-E1 multigene family. Genes Dev. **4:**1650–1662.
37. **Yamamoto, M., N. S. Yew, M. Federspeiel, J. B. Dodgson, N. Hayashi, and J. D. Engel.** 1985. Isolation of recombinant cDNAs encoding chicken erythroid δ-aminolevulinate synthase. Proc. Natl. Acad. Sci. USA **82:**3702–3706.