MOLECULAR AND CELLULAR BIOLOGY, Aug. 1995, p. 4518–4524 Vol. 15, No. 8
0270-7306/95/$04.00+0

B-Cell-Specific DNA Binding by an E47 Homodimer

CHUN-PYN SHEN AND TOM KADESCH*

Howard Hughes Medical Institute and Department of Genetics, University of Pennsylvania School of Medicine, Philadelphia, Pennsylvania 19104

Received 14 February 1995/Returned for modification 18 April 1995/Accepted 2 May 1995

B cells express a unique E-box-binding activity that contains basic helix-loop-helix (bHLH) proteins encoded by the E2A gene. E2A proteins play a central role in immunoglobulin gene transcription and are also required for the generation of the B-lymphocyte lineage. In muscle, E2A proteins bind DNA as heterodimers with muscle-specific bHLH partners, such as MyoD and myogenin, and these heterodimers are thought to be both necessary and sufficient for muscle determination in cultured cells. Our results indicate that in B cells, the bHLH partners for E2A proteins are not B-cell-restricted proteins, but are the E2A proteins themselves. UV cross-linking, gel purification, and the analysis of "forced heterodimers" indicate that BCF1 is primarily a homodimer of the E2A protein E47. Since E47 is widely expressed, our results argue for a difference in the inherent DNA-binding properties of the E47 protein in B cells and may help explain the restricted B-lineage defect observed in E2A-deficient mice.

The basic helix-loop-helix (bHLH) family of DNA-binding proteins is represented by both cell type-restricted and widely expressed members. The myogenic proteins MyoD, myogenin, MRF4, and myf-5 are examples of cell type-specific bHLH proteins whose expression is limited to skeletal muscle. They are also examples of master regulatory proteins that, individually, have the capacity to initiate the expression of a cascade of muscle-specific genes that imposes muscle phenotypes on a variety of nonmuscle cell lines (for reviews, see references 20 and 27). The E proteins E2A, E2-2, and Heb are examples of widely expressed bHLH proteins (11, 12, 16, 19). There are three E2A proteins that arise from alternative splicing of the E2A gene's primary transcript. E47 and E2-5 differ only in their N-terminal amino acids; a functional difference between them has not been observed. The E12 protein differs from E47 and E2-5 in its bHLH region, which is encoded by alternative E2A gene exons (14). E12 has been reported to bind DNA poorly as a homodimer as a consequence of a negative domain immediately N-terminal of the bHLH motif (25). Indeed, early studies with E12 led to the observation that MyoD-E12 heterodimers bound DNA more avidly than either MyoD-MyoD homodimers or E12-E12 homodimers (17). This, along with the finding that MyoD and myogenin in muscle cells exist as heterodimers with E proteins (15), led to the hypothesis that the widely expressed bHLH proteins may typically bind DNA as heterodimers with cell type-restricted bHLH proteins. Genetic data with *Drosophila* bHLH proteins also support a general involvement of heteromeric interactions among the members of this protein family.

The E2A proteins were initially isolated as immunoglobulin enhancer-binding proteins (11, 16) and insulin promoter-binding proteins (19). Indeed, transfection studies indicated that they could activate the immunoglobulin heavy-chain (IgH) enhancer through a mechanism involving both derepression and direct activation (7, 23). Interestingly, the E2-5 used in the latter studies was able to activate transcription in fibroblasts and in yeast cells, for which cell type-restricted partners probably do not exist. Whether the potent activity of these presumed homodimers was a consequence of gross overexpression—imparting characteristics not normally exhibited by endogenous proteins—or due to some innate property of the E2-5 protein is not known. Recent experiments showing that MyoD cannot activate the IgH enhancer in transfection experiments despite its ability to bind the enhancer in vitro suggest that the functional properties of an E2A homodimer are different from those of a MyoD-E2A heterodimer (28). The involvement of E2A proteins in immunoglobulin gene expression is further supported by the existence of a B-cell-specific DNA-binding activity, denoted BCF1, that contains E2A proteins (3, 13, 18). Because of its high sensitivity to an anti-E2A antiserum and its comigration with E47 homodimers in native gels, BCF1 was postulated to be an E2A homodimer (18). However, the wide cellular distribution of E2A RNA, combined with the B-cell-specific expression of BCF1 and the finding of MyoD-E2A heterodimers in muscle cells, makes it equally reasonable to speculate that BCF1 is composed of heterodimers of E2A and B-cell-restricted bHLH proteins.

Recent work indicates that E2A proteins are absolutely required for the B lineage in mice (4, 29). However, these findings do not address the possibility of a B-cell-restricted bHLH partner. The work presented here addresses this issue directly. We find that BCF1 does not contain a B-cell-restricted bHLH partner. Rather, BCF1 appears to be a homodimer of the E2A protein E47, suggesting a unique difference in the properties of E47 found in B cells.

MATERIALS AND METHODS

Analyses of nuclear DNA-binding activities. HeLa and RD rhabdomyosarcoma cells were grown in Dulbecco's modified Eagle's medium (DMEM) containing 10% fetal bovine serum, 100 U of penicillin per ml, and 100 μg of streptomycin per ml. Namalwa, Raji, and Jurkat cell lines were grown in RPMI containing 10% fetal bovine serum, 1× nonessential amino acids, 100 U of penicillin per ml, and 100 μg of streptomycin per ml. Aza-myoblast P-2 cells (a gift of Harold Weintraub, Hutchinson Cancer Research Center) were maintained in DMEM containing 20% fetal bovine serum, 100 U of penicillin per ml, and 100 μg of streptomycin per ml. P-2 cells were differentiated at 80% confluence by changing the medium to DMEM containing 10 μg of insulin and 10 μg of transferrin per ml for 48 h. Nuclear extracts were prepared by the procedure described before (24), modified to include the following protease inhibitors in the nuclear resuspension buffer: leupeptin, 10 μg/ml; pepstatin A, 1 μg/ml; aprotinin, 1 μg/ml; and phenylmethylsulfonyl fluoride, 1 μM.

The probes used for electrophoretic mobility shift assays (EMSAs) are shown in Table 1. Probes were labeled with Klenow fragment (Promega), [α-^{32}P]dCTP, [α-^{32}P]dATP, dGTP, and either dTTP (for octamer, Pu.1, μE4, and μE5) or 5-bromo-dUTP (Sigma) (for μE5-UV).

* Corresponding author. Phone: (215) 898-1047. Fax: (215) 898-9750. Electronic mail address: tk@hmivax.humgen.upenn.edu.

TABLE 1. Probes used for EMSAs

Probe	Sequence
μE4	5′-GATCAAACACCACCTGGGTAAT-3′
	3′-TTTGTGGTGGACCCATTACTAG-5′
Mutant μE4	5′-GATCAAACTCTAGATGGGTAAT-3′
	3′-TTTGAGATGTACCCATTACTAG-5′
μE5	5′-GATCCCAGAACATTTGCAGCAG-3′
	3′-GGTCTTGTGGACGTCGTCCTAG-5′
Mutant μE5	5′-GATCCCAGAACATTTGCAGCAG-3′
	3′-GGTCTTGTAAACGTCGTCCTAG-5′
Octamer	5′-GATCCCTGGGTAATTTGCATTTCTAA-3′
	3′-GGACCCATTAAACGTAAAGATTCTAG-5′
Mutant octamer	5′-GATCCCTGGGTAATGTTCAGTTCTAA-3′
	3′-GGACCCATTACAAGTCAAGATTCTAG-5′
PU.1	5′-GATCCCTCTGAAAGAGGAACTTGG-3′
	3′-GGAGACTTCGATAGCGAACCCTAG-5′
Mutant PU.1	5′-GATCCCTCTGAAGCTATCGCTTGG-3′
	3′-GGAGACTTCGATAGCGAACCCTAG-5′
μE5-UV	5′-CCAGAACACC-3′
	3′-GGTCTTGTGGACGTCGTCCTAG-5′

Binding reaction mixes (15 to 20 μl) contained 1 to 3 μl of nuclear extract or bacterially produced ΔE47 (where specified), 1 ng of labeled probe, 25 ng of unlabeled competitor (where specified), 1 μl of antibody (where specified), and 2 μg of poly(dI:dC) in 12 mM HEPES (*N*-2-hydroxyethylpiperazine-*N*′-2-ethanesulfonic acid)-NaOH (pH 7.9)–4 mM Tris-HCl (pH 7.9)–60 mM KCl–1 mM EDTA–1 mM dithiothreitol (DTT)–12% (vol/vol) glycerol. Samples were resolved on 5% polyacrylamide gels (29:1, acrylamide-bisacrylamide) at room temperature with 0.5× TBE (45 mM Tris-borate, 1 mM EDTA [pH 8]). For the UV cross-linking analysis of complexes formed on the μE5-UV probe, the gel was exposed to UV light (300 nm; Fotodyne transilluminator) for 15 min, and protein-DNA complexes within individual gel slices were then resolved by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and visualized by autoradiography.

Preparation of SDS-PAGE-fractionated extracts. Wet cell pellets were dissolved in 2× SDS loading buffer, boiled for 5 min, and resolved by SDS-PAGE. Molecular mass standards were used to estimate the molecular masses of proteins in each part of the gel, and slices were isolated accordingly (see figure legends). The proteins in each slice were electroeluted with an Electro-Eluter (Bio-Rad model 422), dialyzed against 50 mM Tris-HCl (pH 8)–150 mM NaCl, and precipitated with acetone solution (80% acetone, 10 mM Tris-HCl [pH 8], 30 mM NaCl) overnight at −80°C. Proteins were collected at 4°C by centrifugation (15 min, 10,000 rpm; Beckmann JA-20) and suspended in 6 M guanidinium-HCl (pH 8). Proteins were renatured at 4°C by dialyzing against TM buffer (50 mM KCl, 50 mM Tris-HCl [pH 7.9], 1 mM EDTA, 1 mM DTT, 20% [vol/vol] glycerol). For fraction mixing experiments, eluted proteins were mixed prior to dialysis against TM buffer.

Expression of transfected proteins. The E2-5 expression plasmid pCMVE2-5 was made by inserting the *Hin*dIII-*Bam*HI fragment from pSVAE2-5 (11) into pCMV5 (1). The expression vector for the N-terminally truncated E2-5 (ΔE47), pCHNE2-5C, was constructed by inserting the *Not*I-*Bam*HI fragment from the E2-5 cDNA into pCHN. pCHN was made by placing oligonucleotides coding for a hemagglutinin epitope (YPYDVPDYA) and simian virus 40 T antigen nuclear localization signal (PKKKRKV) downstream of the cytomegalovirus promoter in pCMV5. pCHNE2-5C expresses a truncated E2-5 with a hemagglutinin epitope and nuclear localization signal at its N terminus. Transfections of NIH 3T3 cells were carried out essentially as described before (8) with 5 μg of expression vector and 10 μg of an unrelated plasmid to bring the total plasmid amount to 15 μg per 100-mm plate. ΔE47 was expressed in bacteria as an oligohistidine fusion protein by using the bacterial expression plasmid pDSE2-5C (map available upon request). Purification of ΔE47 from bacterial extracts was achieved by nickel chelate chromatography (Novagen) per the manufacturer's instructions.

Analyses of endogenous E2A proteins and RNAs. For Western (immunoblot) analysis (26), nuclear extracts of Namalwa, Raji, Jurkat, and HeLa cells were mixed with an equal volume of 2× SDS loading buffer and boiled for 5 min. After resolution by SDS-PAGE, proteins were transferred to nitrocellulose membranes (BA85; Schleicher & Schuell). Membranes were then exposed to the anti-E2A monoclonal antibody YAE (13) (generous gift of C. Nelson, University of California at Riverside), followed by horseradish peroxidase-conjugated goat anti-mouse immunoglobulin antibody (ECL; Amersham) per the manufacturer's instructions.

For phosphorylation analysis, cells were pretreated with phosphate-free medium for 2 h and incubated with ^{32}P (1 μCi/ml) for 3 h. Nuclear extracts were prepared, and E2A proteins were precipitated with the monoclonal antibody MAb G98-271.1.3 (3) (generous gift of C. Murre, University of California at San Diego). Pellets were washed three times with RIPA buffer (150 mM NaCl, 1% Nonidet P-40, 0.5% deoxycholate, 0.1% SDS, 50 mM Tris-HCl [pH 8]), resolved by SDS-PAGE, and visualized by autoradiography.

For analysis of E47 (E2-5)- and E12-specific mRNAs, polyadenylated RNA was isolated from Namalwa and HeLa cells (FastTrack; Invitrogen) per the manufacturer's instructions. Reverse transcription reactions were carried out with 100 and 200 ng of each RNA, and PCRs were carried out with E2A primer 1 (5′-GGAGAAGAAGGAGCTGAAGG-3′) and primer 2 (5′-AGGCTGCTT TGGGATTCAGG-3′) end labeled with ^{32}P. PCR products were cut with *Pst*I and *Xma*I, resolved by PAGE, and visualized by autoradiography.

RESULTS

BCF1 DNA-binding activity does not correlate with levels of individual E2A proteins. It has been shown previously that E2A proteins from B cells bind DNA as part of a B-cell-restricted activity denoted BCF1. To examine E2A in cells in which DNA binding cannot be detected, we used monoclonal antibodies raised against human E2A. Because these antibodies recognize the human protein only, we resorted to studying E2A in human cell lines. In agreement with the results obtained with mouse cell lines (3, 13), human cells also identified BCF1 as a B-cell-specific E-box-binding activity (Fig. 1A). BCF1 was detected in two mature B-cell lines (Namalwa and Raji) and not in muscle-derived rhabdomyosarcoma (RD) cells, T cells (Jurkat), or cervical carcinoma cells (HeLa). The BCF1 in Namalwa cells bound to probes corresponding to both the μE5 and μE4 E boxes within the IgH enhancer (Fig. 1B) and reacted completely with antibodies specific for the E2A protein and not with antibodies directed against the related bHLH protein E2-2 (Fig. 1B) or Heb (data not shown).

Western analysis detected E2A protein in the nuclei of all five of the human lines (top band, Fig. 1C). Although the levels of E2A protein were observed to vary somewhat from cell line to cell line, we found no correlation between the levels of E2A protein and DNA binding by BCF1. Raji and HeLa cells, for example, expressed similar levels of the E2A protein, but BCF1 activity was only detected in Raji cells. Likewise, Namalwa, RD, and Jurkat cells expressed comparably high levels of E2A protein, but only Namalwa cells contained BCF1.

Human E2A proteins include three alternatively spliced forms: E12, E47, and E2-5. E47 and E2-5 have the same bHLH domain and differ in their N-terminal amino acids. Since a functional difference between E47 and E2-5 has not been observed, we will henceforth use the name E47 for both of these proteins as a matter of convenience. E12 differs from E47 in that it contains a distinct bHLH domain. It has been reported that E12 does not bind DNA as a homodimer because of the presence of a negative domain immediately N-terminal to the bHLH (25). This region is different in E47, which binds DNA well as a homodimer in vitro. It is formally possible that cell type-specific differences in DNA binding by E2A could be a consequence of cell type-specific differences in the expression of E12 versus E47. For example, B cells may express E47, while other cell types express E12. We therefore examined the relative levels of E12 and E47 transcripts in HeLa cells and Namalwa cells by a PCR-based protocol that distinguishes the two transcripts (Fig. 2). E12 and E47 RNAs were reverse transcribed and amplified with a single set of primers, and the PCR products were subsequently distinguished by restriction analysis. Two concentrations of cDNAs were used to verify the linearity of the assay. We found that both HeLa and Namalwa cells expressed predominantly E47 and that E12 constituted a relatively minor species in both cell types. We conclude that cell type-specific differences in DNA binding are not a consequence of cell type-specific expression of differently spliced variants of E2A.

We also looked for gross differences in phosphorylation between E2A proteins expressed in the various cell lines. E2A

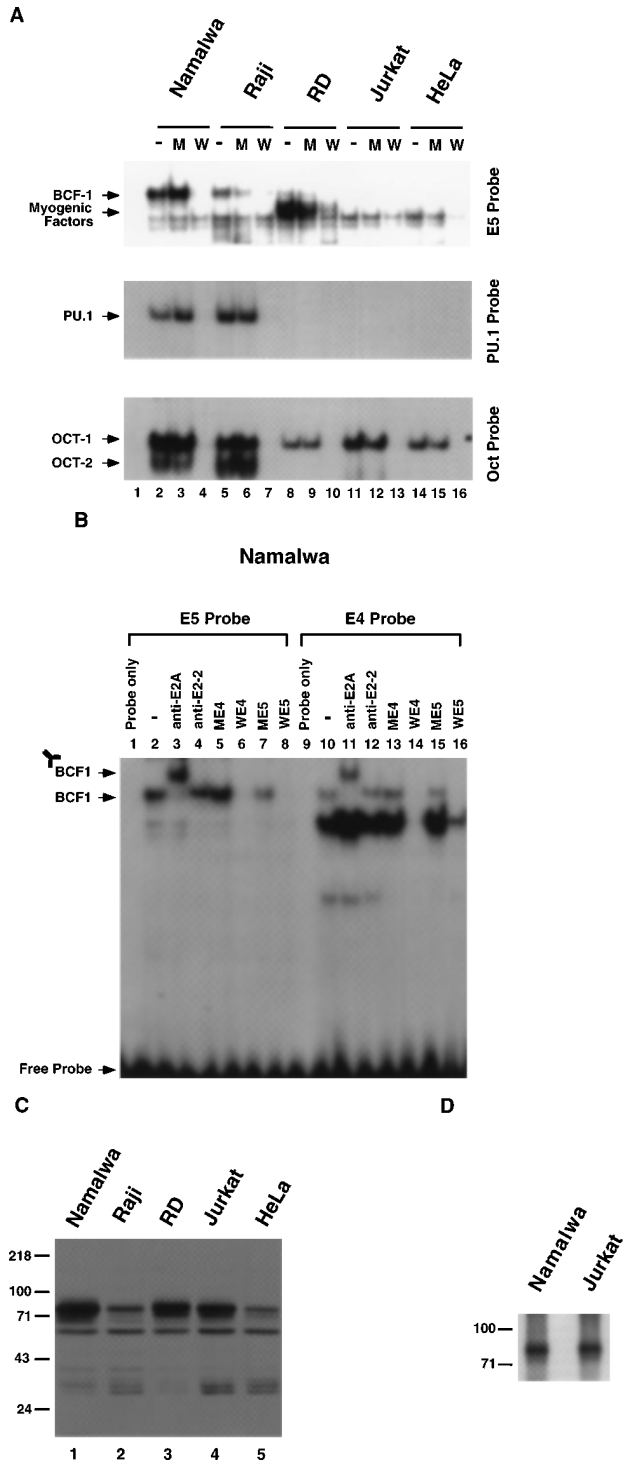

FIG. 1. BCF1 in human cell lines contains E2A protein but does not correlate with E2A expression levels. (A) BCF1 is detected in human B-cell lines (Namalwa and Raji), but not in muscle-derived rhabdomyosarcoma cells (RD), T cells (Jurkat), or cervical carcinoma cells (HeLa). DNA-binding reactions were carried out as described in Materials and Methods. Each panel depicts DNA-binding activities with different radiolabeled probes in an electrophoretic mobility shift assay. Lanes: —, no competitor DNA; M, mutant competitor DNA; W, wild-type competitor DNA added to the binding reaction mixes. (B) Namalwa BCF1 binds two E boxes in the IgH enhancer and contains E2A but not E2-2 proteins. Namalwa cell nuclear extracts were incubated with probes corresponding to the μE5 sites and μE4 sites of the IgH enhancer alone or in the presence of anti-E2A monoclonal antibody, anti-E2-2 monoclonal antibody, or competitor DNA corresponding to a mutant μE4 site (ME4), a wild-type μE4 site (WE4), a mutant μE5 site (ME5), or a wild-type μE5 site (WE5). (C) Western analysis of E2A in nuclear extracts from Namalwa, Raji, RD, Jurkat, and HeLa cells. The expected molecular mass of E2A is consistent with migration of the upper band. Other bands likely represent other proteins that are also recognized by the YAE monoclonal antibody. Sizes are shown in kilodaltons. (D) Autoradiogram of E2A protein immunoprecipitated from ^{32}P-labeled Namalwa and Jurkat cells with monoclonal antibody MAb G98-271.1.3. Similar results were obtained with the YAE monoclonal antibody.

was seen to be phosphorylated in both Jurkat cells and Namalwa cells (Fig. 1D) and also in the Raji, HeLa, and RD cell lines (data not shown). We conclude that phosphorylation at this level of analysis does not regulate E2A's DNA-binding activity, although we cannot discount the possibility that phosphorylation of a subset of amino acids may be important.

An E47 fraction from B cells is sufficient to generate E-box-binding activity. It has been proposed that certain bHLH proteins bind DNA as obligate heterodimers. Indeed, heterodimers of MyoD and E12 produced in vitro bind DNA better than homodimers of either protein alone (17). Moreover, the major E-box-binding species in muscle extracts are heterodimers of E2A proteins and MyoD and/or myogenin (15). Thus, DNA binding by E2A proteins can be activated by cell type-specific bHLH proteins. We therefore asked whether B cells also harbor a B-cell-specific bHLH protein that stimulates DNA binding by E47.

We first examined the composition of BCF1 by UV cross-linking analysis. Figure 3 shows the proteins cross-linked to a labeled μE5 probe when a gel slice containing various E-box-binding complexes (lanes 1 to 3) was exposed to UV light. E47 produced in *Escherichia coli* bound DNA as a homodimer (complex B, lane 1, Fig. 3A; note that this E47 was produced from a cDNA that is just under full length). When a gel slice containing this complex was exposed to UV light and labeled proteins were resolved by SDS-PAGE, a single protein was detected (Fig. 3B, lane B). Similarly, a deleted version of E47 (ΔE47) produced in *E. coli* also bound DNA as a homodimer (complex A, lane 3, Fig. 3A), and UV cross-linking identified a single protein species, as expected (Fig. 3B, lane A). When the BCF1 complex from B-cell extracts (complex D, lane 2, Fig. 3A) was similarly analyzed, three labeled proteins were observed (Fig. 3B, lane D). One of these had the expected mobility of native E47. The second and third species had apparent molecular masses of approximately 30 and 190 kDa.

To assess whether the 30-kDa and 190-kDa proteins were tightly associated with the BCF1 complex or, instead, were cross-linked as a consequence of coincidental comigration, we altered the mobility of the BCF1 components in the gel. We did this by adding the bacterially produced ΔE47 to the B-cell extract. This resulted in the appearance of a heteromeric complex between the components of BCF1 and ΔE47 (Fig. 3A, complex C, lane 3). This complex is not formed when HeLa or NIH 3T3 cell extracts are mixed with ΔE47, indicating that it retains the B-cell-specific components of BCF1 (data not shown). When a gel slice containing the altered-mobility complex was exposed to UV light, three proteins were detected (Fig. 3B, lane C). One had the mobility expected of native E2A, one had the mobility of ΔE47, and a third appeared to be the ~30-kDa protein. We conclude that the mobility of the ~30-kDa protein shifted when ΔE47 was added to the B-cell extract, either as a direct consequence of heterodimer formation (with ΔE47) or through stable association with E47 or other components of BCF1. We are unable to draw any conclusions concerning the ~190-kDa protein, which was not retained in the altered-mobility complex.

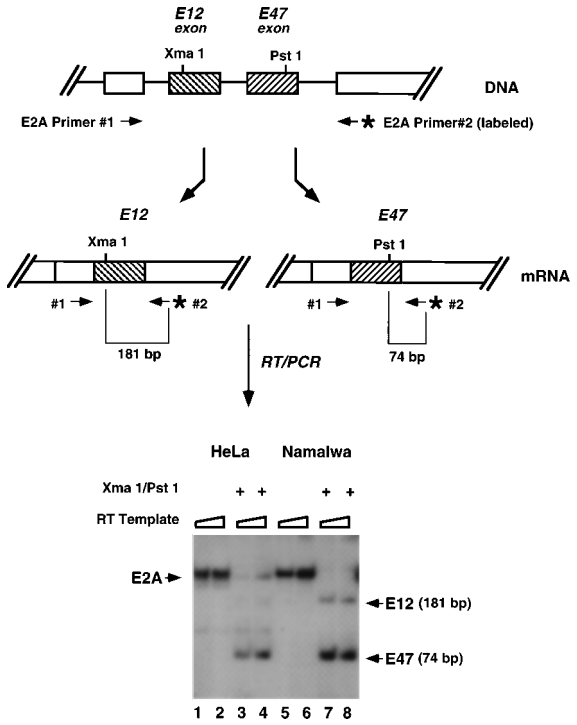

FIG. 2. E47 (E2-5) is the major bHLH splice variant in HeLa and Namalwa cells. Polyadenylated RNA was isolated from HeLa cells and Namalwa cells and subjected to reverse transcription (RT)-PCR analysis as diagrammed in the upper portion of the figure. The structure of the E2A gene, with E12- and E47-encoding exons, is shown, along with the structures of the corresponding alternatively spliced mRNAs. Cutting the mixture of end-labeled PCR products with *Xma*I and *Pst*I gives rise to products diagnostic of E12 and E47 mRNAs, respectively. The lower portion of the figure indicates the results obtained for each cell line with two different amounts of total polyadenylated RNA.

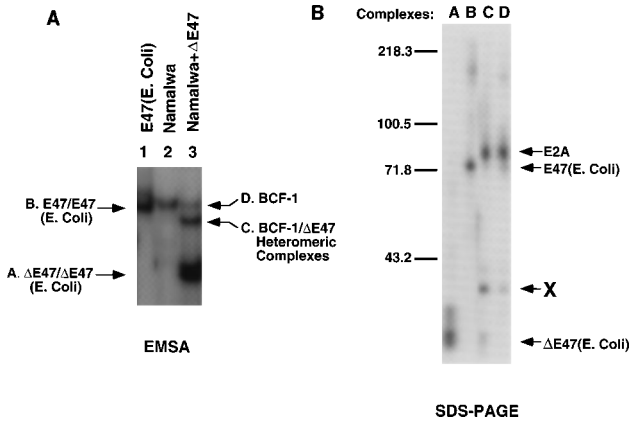

FIG. 3. UV cross-linking identifies E47 and possibly additional proteins in the BCF1-DNA complex. (A) EMSA of various proteins binding to the μE5-UV probe. Lane 1, *E. coli*-produced E47; lane 2, Namalwa BCF1; lane 3, Namalwa BCF1 mixed with *E. coli*-produced ΔE47. (B) Each complex (A to D) was cross-linked with UV light and resolved by SDS-PAGE. The expected mobilities of DNA-bound cellular E47, *E. coli*-produced E47 (note that the cDNA is not full length), and ΔE47 are indicated. X, unknown protein found in both BCF1 and the BCF1-ΔE47 heteromeric complex. Sizes are shown in kilodaltons.

We next tested the idea that the 30-kDa protein functions like MyoD to stimulate the DNA-binding activity of native E2A. To do this, we resolved cell extracts by SDS-PAGE and then tested the DNA-binding activity of proteins eluted and renatured from individual gel slices representing different-sized proteins. Since we had a reasonable expectation for the results of this experiment with muscle cells (i.e., E-box-binding activity should be enhanced upon heterodimer formation), we used differentiated P-2 muscle cells as a positive control (Fig. 4A). When unfractionated P-2 extracts were tested for E-box-binding activity, a major cluster of complexes was observed (Fig. 4A, lane 2). These complexes were shown previously to contain MyoD-E2A and myogenin-E2A heterodimers (15). When the P-2 extracts were first resolved by SDS-PAGE and the DNA-binding activity of renatured proteins eluted from individual gel slices was examined, very little activity was recovered (Fig. 4A, lanes 3 through 8). The gel slice that contained E2A (slice A) gave rise to a small amount of binding activity that comigrated with a minor, slow-mobility complex observed in the unfractionated extract (Fig. 4A, lane 3). The gel slices that contained the myogenic bHLH proteins (slices C and D) gave rise to no DNA-binding activity on their own (Fig. 4A, lane 6). However, when the proteins of slice A and slice D were mixed prior to renaturation, a strong stimulation of DNA-binding activity was seen, and this activity comigrated with the major complexes observed with the unfractionated extract (Fig. 4A, lane 11). These results confirm that muscle cell-derived E2A proteins bind DNA better as heterodimers with myogenic bHLH proteins.

We performed the same analysis with B-cell extracts (Fig. 4B). In the particular experiment shown, the B-cell extract gave rise to BCF1 along with some additional E-box-binding complexes whose presence varies from extract to extract. Because of this variability, these complexes have not been characterized further. Unlike the situation seen with the muscle extracts, the E2A-containing gel slice from B cells gave rise to an E-box-binding activity that comigrated with BCF1 (Fig. 4B, lane 16). We estimated that this represented very good recovery of DNA-binding activity, approximately 10% of the input BCF1. Proteins from other gel slices, including the slice predicted to contain the 30-kDa protein (slice E), had no effect on the DNA-binding activity generated from the E2A-containing slice (Fig. 4B, lanes 21 through 24). This result, taken with the result from the UV cross-linking, suggests that the E2A from B cells binds DNA well either as a homodimer or, alternatively, as a heterodimer with a similar-sized protein. Although the 30-kDa protein may be a part of the BCF1 complex and important for E2A function, it does not appear to regulate E2A's DNA-binding activity. These results are consistent with those of Park and Walker (21), who likewise concluded that BCF1 (LEF1) is a dimer of similar-sized proteins (21).

Lack of DNA binding by E2A in cells other than muscle cells and B lymphocytes (e.g., Jurkat and HeLa cells) could be due to association of inhibitory proteins such as those represented by the Id family. We tested this by fractionating HeLa cell extracts by SDS-PAGE and examining DNA binding by E2A. Similar to the results obtained with the P-2 muscle extracts, the DNA binding by E2A from HeLa cells was poor, with an estimated specific activity 1/20 that observed in Namalwa cells (data not shown). This indicates that the lack of binding by E2A may be a property intrinsic to the protein in certain cell types.

BCF1 is an E47 homodimer. We exploited ΔE47's ability to generate B-cell-specific complexes of intermediate mobility to further assess the composition of BCF1. Our reasoning, illustrated in Fig. 5A, was as follows. If BCF1 contains two distinct bHLH proteins, then the forced heterodimers generated upon addition of ΔE47 should be of two types: one would be the E47-ΔE47 dimer, and the other would be a dimer of ΔE47 and the "other" bHLH partner within BCF1. This reasoning as-

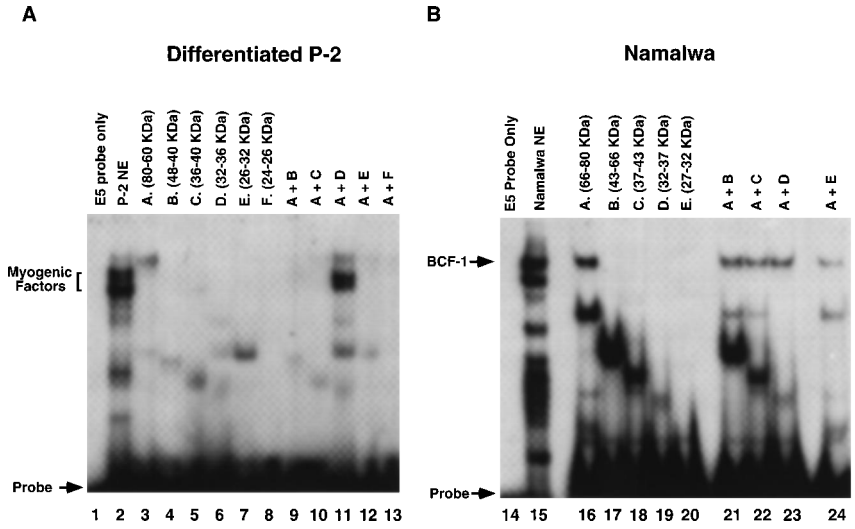

FIG. 4. BCF1 can be generated by SDS-PAGE-purified E47. Nuclear extracts from either (A) P-2 myotubes or (B) Namalwa cells were purified by SDS-PAGE, and the E-box-binding activities of proteins eluted from individual gel slices were analyzed by EMSA. Lanes 2 and 15 show binding activity of unfractionated nuclear extracts. Lanes 3 to 8 and 16 to 20 show binding activities from gel slices representing proteins in the indicated molecular mass ranges. Lanes 9 to 13 and 21 to 24 indicate the binding activities of mixed samples.

sumes that both the endogenous E47 and the "other" bHLH protein dimerize equally well with bacterial ΔE47, an assumption supported by the observation that the intermediate complex is also cell type restricted. If BCF1 contains an E47 homodimer, then the intermediate complexes would contain only E47-ΔE47 dimers. We used a monoclonal antibody (YAE) that recognizes an epitope within the N-terminal region of E47 and thus is missing in the ΔE47 protein. By Western analysis, this antibody is specific and does not identify any B-cell-specific proteins (Fig. 1C) (12). If the intermediate complex contains only E47-ΔE47 dimers, the intermediate complex should be completely supershifted when the antibody is added to the DNA-binding reaction mix. If the intermediate complex represents a mixture of dimers, then only half of them should be shifted by the antibody. Those that are dimers of ΔE47 and the other bHLH protein would be resistant to the antibody, since neither protein would contain the epitope.

As a positive control for the experiment, we transiently transfected E47 and ΔE47 expression vectors into NIH 3T3 cells and analyzed E-box-binding activities in nuclear extracts. As shown in Fig. 5B, these extracts contained E47 homodimers, ΔE47 homodimers, and E47-ΔE47 heterodimers. When the extracts were treated with increasing amounts of the antibody prior to the DNA-binding assay, both the E47 homodimers and E47-ΔE47 heterodimers were completely supershifted, resulting in two corresponding complexes of slower mobility. As expected, the ΔE47 homodimers were completely resistant to the antibody.

When the same analysis was carried out on B-cell extracts that contained BCF1 and added bacterial ΔE47, exactly the

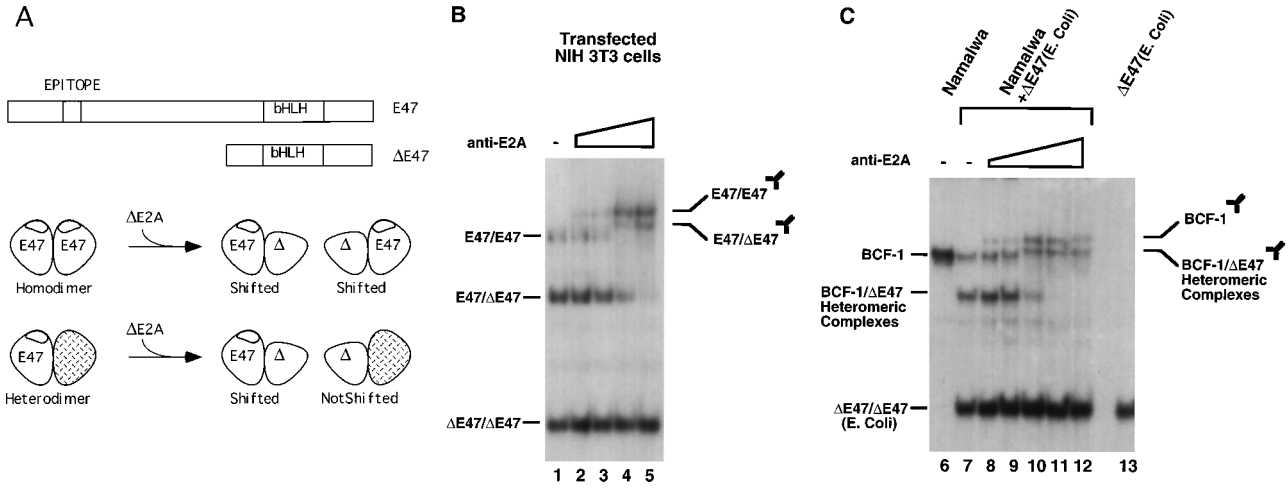

FIG. 5. BCF1 is an apparent E47 homodimer. (A) Rationale. The YAE monoclonal antibody recognizes an N-terminal epitope in E47 that is not present in ΔE47. The reactivity of BCF1-ΔE47 heteromeric complexes to the antibody will therefore indicate if any complexes lack the epitope because of dimerization of ΔE47 with a novel bHLH protein. (B) Results obtained with nuclear extracts of NIH 3T3 cells transfected with both an E47 cDNA (E2-5) and ΔE47 in the absence (lane 1) or presence (lanes 2 to 5) of increasing amounts of the YAE antibody. (C) Results of the same experiment with Namalwa nuclear extracts mixed with *E. coli*-produced ΔE47 (lanes 7 to 12). Lanes 6 and 13 show DNA binding from Namalwa nuclear extract alone and *E. coli*-produced ΔE47 alone, respectively.

same result was seen (Fig. 5C). Both the BCF1 complex and the intermediate complex were completely shifted by the antibody, while the ΔE47 homodimers were resistant. This result is exactly that expected if BCF1 is an E47 homodimer.

DISCUSSION

A central question in developmental biology addresses the issue of how cell lineages are determined. Although it is well established that transcription factors play a major role, little is known about the nuclear events that trigger lineage commitment in mammalian systems. Studies of mice carrying gene knockouts have shown that particular transcription factors may be necessary for a given lineage, but for many cases, the question remains whether the transcription factor is involved in the decision making per se. The myogenic bHLH proteins are clear exceptions. Myogenin, for example, is absolutely required for myoblast differentiation, but is also sufficient for myogenic determination in many established cell lines (10). Similarly, the combined efforts of MyoD and myf-5 are also both necessary and sufficient for myogenic determination (22, 24). The classification of MyoD and myogenin as bHLH proteins has focused attention on other cell type-specific bHLH proteins and their potential roles in lineage commitment. One such example is Mash-1, a bHLH protein that displays a restricted tissue distribution and is necessary for the formation of olfactory and autonomic neurons in the mouse (9). However, forced expression of Mash-1 in embryonal carcinoma cells does not drive neuronal differentiation. Hence, Mash-1 falls short of the myogenic bHLH proteins in dominantly reprogramming gene expression.

We have focused our attention on the role of BCF1 in B-cell-specific transcription and in B-cell determination. BCF1 has been shown previously to contain E2A proteins (3, 13, 18) and has been implicated in both activating and derepressing the IgH enhancer (7, 23). Recent results with E2A knockout mice indicate that the E2A protein is required for the B-cell lineage (4, 29). It is likely that the lack of B cells in E2A-deficient mice is due to the loss of BCF1. Nevertheless, the B-cell phenotype is puzzling given the wide tissue distribution of the E2A protein. Something besides E2A expression per se must also be important for B-cell determination. That BCF1 activity is restricted to B cells indicates that DNA binding by E2A is activated in a novel way in B cells. The mechanism of this activation will be key to understanding why E2A is required only for the B-cell lineage.

A central issue addressed in this study concerns the possible existence of a B-cell homolog of MyoD. Our finding that DNA binding by E2A does not correlate with levels of the various E2A proteins suggests a posttranslational mechanism of regulation. The paradigm established in muscle cells is that DNA binding by E2A is stimulated (posttranslationally) by the myogenic bHLH proteins. Our results with SDS-PAGE-fractionated muscle extracts confirm this. Moreover, our UV cross-linking experiments initially suggested an involvement of additional proteins in the BCF1 complex. However, our results do not support a role for these proteins in stimulating DNA binding by E2A proteins in a manner analogous to MyoD. Our analysis of SDS-PAGE-fractionated B-cell extracts indicates that E47 alone may be sufficient to generate BCF1. Our analysis of forced heterodimers with ΔE47 supplies direct evidence that E47 is the only bHLH protein component of BCF1.

Our results support the idea that E2A proteins—E47 in particular—vary from cell type to cell type. The E2A proteins in B cells may be unique in their ability to bind DNA as homodimers; in muscle cells and pancreatic cells, they clearly prefer to bind DNA as heterodimers (2, 15). The E2A proteins in other cell types may not bind DNA under any circumstances. Indeed, our experiments with HeLa cell-derived E47 and NIH 3T3 cell-derived E47 have failed to reveal DNA binding even in the presence of ΔE47 or MyoD (C.-P.S., unpublished). A non-DNA-binding form of E2A has also been observed in B cells, but the physiological significance is a mystery (13). These properties of endogenous E2A proteins are not reflected in transfection studies. Both E12 and E47 expression plasmids have been shown to activate transcription in both B cells and non-B cells. It is likely that this reflects the artificial nature of these assays due to overexpression of these proteins.

Benezra has recently shown that transcriptional activation by E47 alone requires the formation of a disulfide-linked homodimer (5). Indeed, in support of the conclusions of this work, Benezra detected disulfide-linked homodimers in B-cell extracts. However, these constituted only a minor fraction of the overall BCF1 activity and thus cannot account for its B-cell-specific expression. Moreover, B cells are not unique in their ability to oxidize E47, since the protein is active when overexpressed in a variety of cell types, including yeast cells. Also, we observed covalently linked dimers directly after transfection of NIH 3T3 cells (24a).

Although our results indicate that gross differences in phosphorylation are not involved in regulating E2A, it is possible that specific phosphorylation events are. When the bHLH-zip protein Max is phosphorylated by casein kinase II at a single serine residue, it loses its ability to bind DNA as a homodimer but not as a heterodimer with Myc (5). Perhaps a similar situation exists for E2A. The facts that E47 isolated from *E. coli* binds DNA well in the absence of additional factors and that E2A overexpressed in transient assays of NIH 3T3 cells or in stable transformants of yeast cells argue that a specific E2A modification machinery may be involved in turning expression off in inappropriate cell types. We propose either that such machinery does not exist in B cells or that its activity is neutralized.

ACKNOWLEDGMENTS

We thank members of the Kadesch laboratory for criticisms and helpful comments and especially Lily Wan for a very critical reading of the manuscript. We also thank Kes Murre and Chris Nelson for antibodies.

This work was supported by funds from the Howard Hughes Medical Institute (T.K.).

REFERENCES

1. **Andersson, S., D. Davis, H. Dahlback, H. Jornvall, and D. W. Russell.** 1989. Cloning, structure, and expression of the mitochondrial cytochrome P-450 sterol 26-hydroxylase, a bile acid biosynthetic enzyme. J. Biol. Chem. **264:** 8222–8229.
2. **Aronheim, A., H. Ohlsson, C. W. Park, T. Edlund, and M. D. Walker.** 1991. Distribution and characterization of helix-loop-helix enhancer-binding proteins from pancreatic b cells and lymphocytes. Nucleic Acids Res. **19:**3893–3899.
3. **Bain, G., S. Gruenwald, and C. Murre.** 1993. E2A and E2-2 are subunits of B-cell-specific E2-box DNA-binding proteins. Mol. Cell. Biol. **13:**3522–3529.
4. **Bain, G., E. C. Robanus Maandag, D. J. Izon, D. Amsen, A. M. Kruisdeek, B. C. Weintraub, I. Krop, M. S. Schlissel, A. J. Feeney, M. van Roon, M. van der Valk, H. P. J. te Reile, A. Berns, and C. Murre.** 1994. E2A proteins are required for proper B cell development and initiation of immunoglobulin gene rearrangements. Cell **79:**885–892.
5. **Benezra, R.** 1994. An intermolecular disulfide bond stabilizes E2A homodimers and is required for DNA binding at physiological temperature. Cell **79:**1057–1067.
6. **Berberich, S. J., and M. D. Cole.** 1992. Casein kinase II inhibits the DNA-binding activity of Max homodimers but not Myc/Max heterodimers. Genes Dev. **6:**166–176.
7. **Genetta, T., D. Ruezinsky, and T. Kadesch.** 1994. Displacement of an E-box-binding repressor by helix-loop-helix proteins: implications for B-cell speci-

ficity of the immunoglobulin heavy-chain enhancer. Mol. Cell. Biol. **14:**6153–6163.
8. **Graham, F. L., and A. J. van der Eb.** 1973. A new technique for the assay of infectivity of human adenovirus 5 DNA. Virology **52:**456–467.
9. **Guillemot, F., L. C. Lo, J. E. Johnson, A. Auerbach, D. J. Anderson, and A. L. Joyner.** 1993. Mammalian achete-scute homolog 1 is required for the early development of olfactory and autonomic neurons. Cell **75:**463–476.
10. **Hasty, P., A. Bradley, J. H. Morres, D. G. Edmondson, J. M. Venuti, E. N. Olson, and W. H. Klein.** 1993. Muscle deficiency and neonatal death in mice with a targeted mutation in the myogenin gene. Nature (London) **346:**501–506.
11. **Henthorn, P., M. Kiledjian, and T. Kadesch.** 1990. Two distinct transcription factors that bind the immunoglobulin enhancer mE5/kE2 motif. Science **247:**467–470.
12. **Hu, J.-S., E. N. Olson, and R. E. Kingston.** 1992. HEB, a helix-loop-helix protein related to E2A and ITF2 that can modulate the DNA-binding ability of myogenic regulatory factors. Mol. Cell. Biol. **12:**1031–1042.
13. **Jacobs, Y., C. Vierra, and C. Nelson.** 1993. E2A expression, nuclear localization, and in vivo formation of DNA- and non-DNA-binding species during B-cell development. Mol. Cell. Biol. **13:**7321–7333.
14. **Kamps, M. P., A. T. Look, and D. Baltimore.** 1991. The human t(1;19) translocation in pre-B ALL produces multiple nuclear E2A-Pbx1 fusion proteins with differing transformation potentials. Genes Dev. **5:**358–368.
15. **Lassar, A. B., R. L. Davis, W. E. Wright, T. Kadesch, C. Murre, A. Voronora, D. Baltimore, and H. Weintraub.** 1991. Functional activity of myogenic HLH proteins requires hetero-oligomerization with E12/E47-like proteins in vivo. Cell **66:**305–315.
16. **Murre, C., P. S. McCaw, and D. Baltimore.** 1989. A new DNA binding and dimerization motif in immunoglobulin enhancer binding, daughterless, MyoD, and Myc proteins. Cell **56:**777–783.
17. **Murre, C., P. S. McCaw, H. Vaessin, M. Caudy, L. Y. Jan, Y. N. Jan, C. V. Carbrera, J. N. Buskin, S. D. Hauschka, A. B. Lassar, H. Weintraub, and D. Baltimore.** 1989. Interactions between heterologous helix-loop-helix proteins generate complexes that bind specifically to a common DNA sequence. Cell **58:**537–544.
18. **Murre, C., A. Voronova, and D. Baltimore.** 1991. B-cell- and myocyte-specific E2-box-binding factors contain E12/E47-like subunits. Mol. Cell. Biol. **11:**1159–1160.
19. **Nelson, C., L.-P. Shen, A. Meister, and W. Rutter.** 1990. Pan: a transcriptional regulator that binds chymotrypsin, insulin, and AP-4 enhancer motifs. Genes Dev. **4:**1035–1043.
20. **Olson, E. N.** 1994. bHLH factors in muscle development: dead lines and commitments, what to leave in and what to leave out. Genes Dev. **8:**1–8.
21. **Park, C. W., and M. D. Walker.** 1992. Subunit structure of cell-specific E box-binding proteins analyzed by quantitation of electrophoretic mobility shift. J. Biol. Chem. **267:**15642–15649.
22. **Rudniki, M. A., P. N. J. Schnegelsberg, R. H. Stead, T. Braun, H. H. Arnold, and R. Jaenisch.** 1993. MyoD or Myf-5 is required for the formation of skeletal-muscle. Cell **75:**1351–1359.
23. **Ruezinsky, D., H. Beckmann, and T. Kadesch.** 1991. Modulation of IgH enhancer's cell type specificity through a genetic switch. Genes Dev. **5:**29–37.
24. **Schreiber, E., P. Matthias, M. M. Muller, and W. Schaffner.** 1989. Rapid detection of octamer binding proteins with "mini-extracts," preparation from a small number of cells. Nucleic Acids Res. **17:**6419.

24a. **Sloan, S., and T. Kadesch.** Unpublished data.

25. **Sun, X.-H., and D. Baltimore.** 1991. An inhibitory domain of E12 transcription factor prevents DNA binding in E12 homodimer but not in E12 heterodimers. Cell **64:**459–470.
26. **Vierra, C., Y. Jacob, L. Ly, and C. Nelson.** 1994. Patterns of Pan expression and role of Pan proteins in endocrine cell-type-specific complex formation. Mol. Endocrinol. **8:**197–209.
27. **Weintraub, H.** 1993. The MyoD family and myogenesis: redundancy, networks, and thresholds. Cell **75:**1241–1244.
28. **Weintraub, H., T. Genetta, and T. Kadesch.** 1994. Tissue-specific gene activation by MyoD: determination of specificity by cis-acting repression elements. Genes Dev. **8:**2203–2211.
29. **Zhuang, Y., P. Soriano, and H. Weintraub.** 1994. The helix-loop-helix gene E2A is required for B cell formation. Cell **79:**875–884.