Subunit Structure of Cell-specific E Box-binding Proteins Analyzed by Quantitation of Electrophoretic Mobility Shift*

Cheol Won Park and Michael D. Walker‡
From the Biochemistry Department, Weizmann Institute of Science, Rehovot, Israel 76100

Expression of insulin and immunoglobulin genes is dependent on the presence of E boxes (consensus sequence CAXXTG) within the enhancer regions. These sequences are recognized by cell-specific nuclear factors IEF1 (insulin enhancer factor 1) and LEFl (lymphoid enhancer factor 1). Although IEF1 and LEFl are distinct by several parameters, they are both recognized by antisera to the mouse helix-loop-helix (HLH) protein A1 (a homolog of the human protein E47, product of the E2A gene). This suggests that A1/47 or a close relative is a component of both complexes. In order to further characterize the complexes, we have used in vitro translated DNA-binding proteins of known size to verify that electrophoretic mobility shift analysis can be used to estimate the molecular weight of DNA-binding proteins from both the HLH family and the leucine zipper family. Under the conditions used, migration is relatively insensitive to changes in protein charge. This analysis, in combination with mixing experiments between nuclear extracts and in vitro translated HLH proteins, indicates that IEF1 and LEFl are distinct by several parameters, they are both recognized by antibodies to A1, and the second of which is 25 kDa. LEFl on the other hand, appears to be a complex of two proteins of 67 kDa. The size of the 67-kDa subunits is consistent with that reported for the full-length E2A gene products. The 25-kDa subunit of IEF1 forms DNA-binding heterodimers with A1 but not MyoD and is present in a limited range of cell types, features characteristic of class B HLH proteins such as MyoD and achaete-scute. Taken together, the data support the idea that the E2A gene products are involved directly in regulation of insulin and immunoglobulin gene expression; regulation of the insulin gene apparently requires, in addition, the 25-kDa HLH protein (designated IEF1 for insulin enhancer-specific factor 1).

Cell-specific expression of the insulin gene is transcriptionally controlled by cis-elements located in the 5'-flanking region (1–3). Two closely similar 9-bp sequences (IEFl and LEFl) containing an E box consensus sequence (CAXXTG) are located at −108 and −233 in the rat insulin I gene; these sequences (4, 5) and related sequences from rat insulin II (6, 7) and human insulin genes (8) are essential for efficient transcription from the insulin flank. IEF1 and LEF2 are recognized by a nuclear factor (IEFl) present in β cells (9) but absent from all other non-endocrine cell types tested (10). The DNA binding specificity (11) and distribution of IEF1 has been suggested previously that the electrophoretic mobility shift analysis can be used to estimate the molecular weight of DNA-protein complexes. Using a series of in vitro translated DNA-binding proteins of known molecular weight derived from both HLH and leucine zipper families, we were able to confirm that molecular weight can be reliably estimated by measuring electrophoretic mobility on nondenaturing acrylamide gels. We extended the procedure by using an in vitro subunit exchange reaction for characterizing multimeric complexes. Using this approach, we showed that IEF1 contains two proteins of 67 kDa and 25 kDa, whereas LEFl behaves as a dimer of two proteins of molecular mass ≈ 67 kDa.

MATERIALS AND METHODS AND RESULTS

Estimation of Molecular Weight on Nondenaturing Gels—It has been suggested previously that the electrophoretic mobility of protein-DNA complexes in nondenaturing acrylamide gels can be used to estimate the molecular weight of the protein (26, 27). This may result from the strongly negative charge of the DNA, which minimizes the effects of net charge of the protein and effectively makes mobility dependent on protein molecular weight. To investigate the generality of this approach, we measured the mobility of DNA-protein complexes and found that the electrophoretic mobility is directly proportional to the logarithm of the molecular weight of the protein, which is the basis of the method.

1 Portions of this paper (including "Materials and Methods," part of "Results," Fig. 5, and Table 1) are presented in miniprint at the end of this paper. Miniprint is easily read with the aid of a standard magnifying glass. Full size photocopies are included in the microfilm edition of the Journal that is available from Waverly Press.
plexes derived from in vitro translated DNA-binding proteins of known molecular weight. For this purpose, we generated two series of proteins, one based on the HLH protein A1 (10) and the other based on the leucine zipper protein JunB (28) (Fig. 1). Proteins were synthesized from appropriate DNA constructions by in vitro transcription-translation reactions and allowed to interact with cognate radioactive DNA probes of equal length (IEB1 probe for HLH proteins and AP-1 DNA for JunB protein (oligonucleotides A, B, D, and E, respectively; see "Materials and Methods"). Since homodimeric JunB alone shows relatively inefficient DNA binding activity (28), JunB-derived proteins were mixed with separately translated c-Fos in order to generate efficient DNA-binding complexes. For each complex, mobility was plotted against log molecular weight, based on the assumption that the DNA-binding form of both HLH and leucine zipper proteins is a dimer. This assumption has been shown to be valid for the leucine zipper family of proteins (29). For the HLH family, although multimerization is clearly a prerequisite for DNA binding (30), the stoichiometry of the DNA-binding complexes is not established. For the six protein complexes tested, a linear relationship was seen between mobility and log molecular weight (Fig. 2, A and B). This substantiates the validity of the approach for estimation of molecular weights and supports the notion that HLH proteins bind DNA as dimeric complexes.

In order to verify the molecular weights of the in vitro translated proteins used in the study, [35S]methionine-labeled proteins were fractionated by SDS-PAGE and visualized by autoradiography. The proteins migrated linearly according to the molecular weights calculated from inferred protein sequence (Fig. 2, C and D), indicating that the translated proteins correspond to the species predicted from the DNA sequences.

**In Vitro Subunit Exchange Reaction—IEF1 and LEF1/BCF1 are E box binding factors identified in endocrine cells (9, 10) and lymphoid cells (10, 14), respectively. To examine the structure of these complexes, we combined the procedure for molecular weight estimation with an in vitro subunit exchange reaction using in vitro translated proteins derived from two truncated forms of A1, designated here as A1 (281 amino acids) and TA1 (106 amino acids) (Fig. 1A). Consistent with previous observations (10), these truncated proteins generated characteristic bands on EMSA analysis with labeled IEB1 (Fig. 3A, lanes 7 and 9). When A1 and TA1-containing lysate were mixed, a single additional EMSA band was observed, presumably corresponding to an A1/TA1 heterodimer (Fig. 3A, lane 8). This is consistent with the idea that A1 and TA1 bind DNA as homodimers; upon mixing, a single heterodimeric species of intermediate size is generated. When the in vitro translated proteins were mixed with β cell extracts, additional bands D and E were seen (lanes 2 and 3). Likewise with B cell extract, additional bands (B and C) were observed (lanes 5 and 6). In this experiment, incubations were performed at 23 °C to promote subunit exchange; at this temperature the binding efficiency of the LEF1 is lower (lane 4) as compared with incubations performed at 0 °C (e.g. Fig. 4, lane 2). Using a calibration curve (Fig. 3B) based on known molecular weights (A1, TA1, and the intermediate complex of A1 and TA1) versus the electrophoretic migration of these proteins, the intermediate complexes D and E derived from β cells have molecular masses of 56 and 37 kDa, respectively. Since these complexes contain A1-derived proteins of 31 and 12 kDa, respectively, the subunit derived from IEFl is calculated in both cases as 25 kDa.

The molecular mass values of LEF1, IEF1, and complexes B and C were estimated as approximately 134, 92, 98, and 79 kDa, respectively, by extrapolation based on the sizes of the three closest migrating complexes (A1 + A1, D, and A1 + TA1) (Fig. 3C). Since complexes B and C apparently derive from one subunit of LEF1 together with one subunit of in vitro translated protein of known size, the molecular mass of the subunit derived from LEF1 could be calculated (98 = 31 + X for complex B, 79 = 12 + X for complex C, therefore X = 67). Since the apparent molecular mass of LEF1 is 134 kDa, the second subunit of LEF1 is estimated also as 67 kDa. Likewise, since one subunit of IEFl was estimated above to be 25 kDa, the remaining subunit is 67 kDa. The above data indicate that LEF1 is a dimer composed of 67- and 25-kDa proteins, whereas LEF1 is a dimer composed of two 67-kDa proteins. The molecular masses of the 67-kDa subunits of LEF1 and IEF1 correspond closely to that reported for the putative full-length E2A gene products for human (67.6 kDa) (31) and rat (67.4–67.7 kDa) (17).

**Artificial Reconstitution for Analysis of Size of Cell-specific E Box Proteins IEFl and LEF1—Since several of the above estimates were based on extrapolations from faster migrating complexes, we used additional higher molecular weight markers for more reliable molecular weight determinations. For this, an additional plasmid (267) was prepared using a clone encoding a near full-length A1 protein fused in frame to a fragment of the β-lactamase gene of pUC18 encoding a 39-amino acid polypeptide (Fig. 1A). In vitro transcription-translation reactions yield a 67-kDa protein. This protein should form a DNA-binding homodimer of molecular mass = 134 kDa. Likewise upon mixing with the A125 protein (25 kDa), a heterodimeric species of 92 kDa should be formed. Therefore if the values deduced above for IEFl and LEF1 are accurate, it should be possible to use the new construct to reconstitute complexes of identical electrophoretic mobility as IEFl and LEF1 in vitro translated proteins generated from plasmid 267 (67 kDa) and A125 (25 kDa) each generated a distinct EMSA band of characteristic mobility (Fig. 4, lanes 5 and 6). When the two lysates were mixed, an additional band was produced (lane 9) which was absent from control lanes (lanes 5, 7, and 8). The mobility of this additional band is consistent with its identification as a heterodimer of the 67- and 25-kDa species. As predicted, LEF1 was seen to comigrate with the complex derived from 67-kDa protein while IEFl comigrated with the putative 67/25 heterodimer. Therefore, these data support the notion that LEF1 is a dimer containing two 67-

---

**Fig. 1. Schematic representation of proteins used for EMSA analysis.** A, deletion mutation derivatives of the A1 gene. The filled region corresponds to the bHLH domain required for DNA binding and dimerization. The hatched area corresponds to the region derived from the β-lactamase gene (see "Materials and Methods"). B, deletion mutation derivatives of the junB gene. The filled region corresponds to the basic-zipper domain required for DNA binding and dimerization.
kDa proteins and that IEF1 is a dimer of 67- and 25-kDa proteins.

Comparison of Intermediate Species Using TMyoD and TA1—It has been previously shown that HLH proteins have a restricted range of dimerization capacities. For example HLH proteins which are confined to specific cell types (class B) (e.g. MyoD, achaete-scute) homodimerize inefficiently but can heterodimerize with constitutively expressed members of the family (class A) (e.g. E47, daughterless); class A, in addition to heterodimerizing with class B, can also homodimerize (15, 30, 32). Thus by performing subunit exchange reactions with nuclear extracts and in vitro translated class A or class B HLH proteins, one may be able to analyze the structure of nuclear HLH complexes. We first verified that valid molecular weight estimates could be obtained based on the migration of complexes containing MyoD. We used the in vitro synthesized protein from a truncated form (termed TMyoD, 112 amino acids, 12.8 kDa) of a full-length MyoD protein. In parallel, in vitro translated TA1 protein was utilized as a molecular marker. As shown in Fig. 6, TMyoD protein alone did not show significant binding activity (lane 2). However, upon mixing of TMyoD protein and 267 protein, an intermediate complex (lane 5) was generated, whose migration is indistinguishable from that of the complex gener-
The presence of radiolabeled IEBl as described. The figure shows nuclear extract alone protein were translated in the reticulocyte lysate cell-free system and labeled A1 and nuclear complexes IEFl and LEF1. A1 protein and TA1 translation products.

FIG. 4. Artificial reconstitution of complexes of identical mobility to IEFl and LEF1. EMSA was carried out with nuclear extracts from HIT cells (lane 1), WEHI 231 cells (lane 2) and reticulocyte lysates containing 267 and A125 proteins (lane 3), 267 protein alone (lane 4), 267 protein and control lysate (lane 5), A125 protein alone (lane 6), A125 protein and control lysate (lane 7), and control lysate (no added RNA) alone (lane 8).

FIG. 6. Comparison of migration of complexes containing TMyoD (12.8 kDa) and TA1 proteins (12.2 kDa). EMSA was carried out with radiolabeled IEBl and reticulocyte lysate directing 267 protein (lane 1), TMyoD protein (lane 2), TA1 protein (lane 3), and control lysate (no RNA added, lane 4). Lanes 5–7 show the binding pattern of in vitro translated 267 protein preincubated in the presence of reticulocyte lysate containing TMyoD protein (lane 5), TA1 protein (lane 6), and control (no added RNA) lysate (lane 7). Lanes 8 and 9 show DNA binding of control lysate preincubated with in-vitro translated TMyoD protein (lane 8) and TA1 protein (lane 9). Intermediate complexes (267/TMyoD (lane 5) and 267/TA1 (lane 6)) are indicated by an arrow.

Ated from mixing of 267 and TA1 (lane 6). Control reactions did not give these complexes (lanes 4, 7, 8, and 9). Thus, estimation of molecular weights can be applied to MyoD, a class B HLH protein.

We then performed a mixing experiment with HIT nuclear extract in the absence or in the presence of reticulocyte lysate containing TMyoD protein (Fig. 7A) or TA1 (Fig. 7B). In order to detect DNA binding of homodimer TMyoD protein, we used as probe an E box sequence derived from the MCK gene, which is a preferred binding site for MyoD (33). With from inferred protein sequence and based on the assumption that the binding species are dimers. The migration of TA1 was measured using shorter exposures than shown here. The relative migration of complexes D and E was used to interpolate the apparent molecular weight of these complexes (rectangles). Molecular weights (× 10^{-7}) of complexes are indicated in parentheses. C, calibration curve to extrapolate the molecular weights of IEFl, LEF1, and complexes B and C (rectangles) based on the molecular weights of the three closest-migrating complexes (A1, complex D, and A1/TA1 (circles)) versus their migration. Molecular weights of complexes (× 10^{-7}) are indicated in parentheses.
Characterization of Enhancer-binding Proteins

**Fig. 7.** Comparison of intermediate bands formed upon mixing of nuclear extracts from β cells (HIT) (panels A and B) or B lymphoid cells (WEHI 231) (panels C and D) together with either TMyoD (panels A and C) or TA1 (panels B and D). A and B, EMSA was carried out with radiolabeled MCK E box (A) and IEB1 E box (B) and HIT nuclear extracts alone at 0 °C (lane 1) or preincubated in the presence of in vitro translated TMyoD protein (panel A) or TA1 protein (panel B) at the indicated temperatures (lanes 2-5 in panels A and B). Lane 6 in panels A and B represents the binding pattern of reticulocyte lysate containing TMyoD or TA1, respectively, at 0 °C. Intermediate bands are indicated by an arrow. C and D, EMSA analysis as above except that nuclear extract from WEHI 231 cells was used. Probes used were radiolabeled MCK E box (C) and IEB1 E box (D).

Both TA1 and TMyoD, mixing resulted in formation of an intermediate species. However, the mobility of the intermediate complex with TMyoD was substantially less than that produced with TA1. The migration of these species corresponds to dimerization between TMyoD and TA1 and nuclear proteins of 67 and 25 kDa, respectively, according to calibration curve analysis (data not shown), and is consistent with the postulated structure of IEF1 and the known dimerization properties of HLH proteins. The ability of TA1 to generate heterodimers with the 25-kDa species and not the 67-kDa species of IEF1 (Fig. 7B) may be due to lower dimerization ability and/or DNA binding capacity of homodimer A1 as compared with heterodimers. For example, it has been shown that homodimerization of E47 is about 10-fold less efficient than heterodimerization between E47 and MyoD (32). The formation of intermediate complexes is temperature-dependent; when mixed extracts were incubated at 0 °C, little or no intermediate complex was formed. Following incubation at higher temperatures, intermediate complexes were observed; near-optimal levels were seen with TMyoD at 23 °C, whereas highest levels with A1 required incubation at 37 °C (Fig. 7, A and B). Although several bands derived from the nuclear extract disappear with similar kinetics to MyoD, their electrophoretic migration (hence molecular weight) and inability to be recognized by anti-A1 antibody (Fig. 8, A and B) make it unlikely that they participate in the intermediate complex.

A similar mixing experiment was performed with lymphoid cell extract. In this case, the intermediate species observed with TMyoD (Fig. 7C) was identical in mobility to that seen with TA1 (Fig. 7D). The migration of the complex corresponds to interaction with a 67-kDa protein. As before, incubation at 0 °C led to inefficient complex formation; some complex was observed with TMyoD, essentially none with TA1. Following incubation at 23 °C, optimal complex formation was seen in both cases.

**Anti-A1 Antibodies Recognize the 67-kDa Subunits of IEF1 and LEF1**—The above results indicate that IEF1 is composed of 67- and 25-kDa proteins, and LEF1, two 67-kDa proteins. Based on this molecular mass and the previously described ability of anti-A1/E47/E12 antibodies to recognize IEF1 (10, 18, 24) and LEF1 (10, 14), it seems likely that the 67-kDa protein of IEF1 and LEF1 is A1/E47 or a closely related protein. To further address this issue, we used TMyoD protein as a partner for formation of the intermediate complexes with the 67-kDa proteins of IEF1/LEF1. Mixing reactions were carried out to generate the intermediate complexes from reticulocyte lysate containing TMyoD protein with HIT nuclear extract (Fig. 8A, lanes 4–6) and WEHI 231 nuclear extract (Fig. 8B, lanes 4–6). These intermediate complexes (lane 6 in Fig. 8, A and B), as well as IEF1 and LEF1 (lane 3 in Fig. 8, A and B), were reduced in intensity by anti-A1 antibody but not by control antibody (lanes 2 and 5 in Fig. 8, A and B).

The estimated molecular masses of the intermediate complexes are about 80 kDa (most probably one subunit of A1 protein (67 kDa) and one subunit of TMyoD protein (12.8 kDa), based on electrophoretic mobility. These deduced molecular masses argue against the possibility that these complexes are composed of other ubiquitously expressed members of HLH family, such as USF (43 kDa; Ref. 22), TFE3 (59 kDa; Ref. 20), and AP4 (48 kDa; Refs. 21 and 34). Furthermore TFE3, AP4, and Myc are unable to heterodimerize in vitro with E47 or MyoD (21, 30, 35).
Characterization of Enhancer-binding Proteins

DISCUSSION

EMSA has been extensively used as a convenient, sensitive method of detecting the presence and binding specificity of DNA-binding proteins (36). In general, the intensity of binding complexes is considered; quantitation of relative mobility is less frequently used as an informative parameter. Bading (26) and Schreiber et al. (27) have suggested that the molecular weight of proteins in protein-DNA complexes can be estimated by measurement of electrophoretic mobility in non-denaturing gels. Here we have confirmed and extended this application of EMSA and applied the methodology to analysis of two binding activities previously identified in pancreatic β and lymphoid β cells by virtue of specific interaction with E boxes (CA/transfected). Both IEF1 (10, 18, 24) and LEF1/BCF1 (10, 14) are recognized by antibodies directed against the HLH protein A1/E47, indicating that both complexes contain A1 or a closely related protein. However, the structure and subunit molecular weights of these complexes have not been reported. A standard approach to characterizing DNA-binding complexes involves UV cross-linking of protein to radiolabeled DNA. Following irradiation, the complex is fractionated on SDS-polyacrylamide gels (37). However, the method has limitations. First, the complex as fractionated on SDS gels contains DNA as well as protein and therefore migrates more slowly than the corresponding free protein, complicating the determination of molecular weight. Second, the cross-linking is not equally effective for different DNA-protein complexes. Indeed, using this approach to analyze IEF1 and LEF1 interaction with E boxes, we have been unable to observe specific cross-linked proteins (not shown).

As an alternative, we analyzed the structure of IEF1 and LEF1 using EMSA-based estimation of molecular weight. The method was first validated using a series of in vitro translated DNA-binding proteins which revealed a linear dependence of electrophoretic mobility on log molecular weight of the binding protein. This was true both for proteins of the HLH family (A1 and MyoD) and members of the leucine zipper family (JunB and Fos). Although we observed a consistent dependence of mobility on molecular weight in our experiments, clearly this may not apply for all DNA-binding proteins. In particular, highly charged proteins may migrate anomalously. However, based on the broad range of charges of proteins tested in this study (see Table 1 in the Miniprint) and previous reports (26, 27) it seems likely that many, if not most, DNA-binding proteins will migrate in accordance with molecular weight. Another parameter which can influence mobility of protein DNA complexes in non-denaturing gels is DNA bending (38). However, the effects of bending become negligible with short DNA probes such as those used in this study (20–28 bp); standard procedures for visualizing the effects of DNA bending involve the use of much longer DNA

To exclude the possibility that the antibody recognition may be due to cross-reactivity between TMyoD protein and anti-A1 antibodies, rather than recognition of the 67-kDa subunits of IEF1 and LEF1, binding reactions between TMyoD protein with the MCK E box were carried out. Anti-MyoD antibodies recognized TMyoD protein efficiently (Fig. 8C, lane 5). Neither anti-A1 nor anti-TrpE antibodies recognized the protein (lanes 3 and 4). These data strongly suggest that the recognition of IEF1, LEF1, and the intermediate complexes by anti-A1 antibodies is due to the fact that anti-A1 antibodies recognize the 67-kDa subunits of IEF1 and LEF1. We have recently shown that polyclonal antibodies directed against the N-terminal portion of A1, like the anti-C terminal antibodies used here and previously (10), interact with both IEF1 and LEF1.3 This confirms the above contention that both complexes contain E2A gene products rather than immunologically cross-reacting proteins.

3 R. Shiran and M. D. Walker, unpublished observations.
Characterization of Enhancer-binding Proteins

probes (>150 bp) (38-40). Assuming a negligible contribution of protein charge and DNA bending, the major source of error in molecular weight determination derives from measurements of electrophoretic mobility of complexes. Based on the widths of autoradiographic bands in our experiments, the error in measurement of migration is ±1 mm, corresponding to an error of ±5% in molecular weight estimation.

We extended the EMSA-based approach by combination with an in vitro subunit exchange reaction for characterizing multimeric complexes. Mixing of in vitro synthesized HLH proteins of defined size with nuclear extracts led to the generation of novel intermediate species whose molecular weight was determined, thereby allowing analysis of the sizes of subunits. By selecting cloned genes with limited heterodimerization potential, e.g. MyoD, one can obtain information about the dimerization capacity of the unknown proteins. Furthermore, the identity of constituents can be investigated by incubation with antibodies directed against potential constituents of the binding complex. The efficiency of in vitro subunit exchange reactions between in vitro translated proteins can be increased by co-translating the proteins (13). For subunit exchange reactions involving cellular extracts, this option is not available. As an alternative, optimization of incubation temperature may be helpful. In our studies, optimal formation of heterodimeric complexes required a step of incubation at 20-37 °C, perhaps indicating a temperature-dependent dissociation step. Thus the in vitro subunit exchange approach permits characterization of the subunit structure of dimeric DNA-binding complexes. In addition it may be of value in identifying in cell extracts, the dimerization partners of cloned genes.

Our analysis indicates that the lymphoid-specific factor LEFl contains two subunits of molecular mass ≈ 67 kDa. This may represent a homodimer of the A1/E47 protein as previously suggested (14), or a heterodimer between A1 and a distinct, possibly lymphoid-specific, HLH protein. If LEFl is indeed a homodimer, it remains to be explained why this binding activity is observed only in lymphoid cells, given the fact that A1/E47 RNA expression is observed at similar levels in multiple cell types (15, 17, 19, 41).

IEFl appears to be a heterodimer of A1 (or a closely related protein), complexed to a distinct endocrine cell-specific HLH protein of molecular mass ≈ 25 kDa (IESFl). IESFl is lineage-restricted protein able to form DNA-binding heterodimers with A1/E47 but not MyoD and therefore can be considered to be a member of the so-called class B HLH proteins such as MyoD and achaete-scute (30), which are expressed in cell-specific fashion. Pancreatic β cells are thought to derive from terminal differentiation of a pluripotent pancreatic islet cell precursor (42). It has not been conclusively determined whether this precursor derives from neuroectoderm or endoderm (reviewed in Ref. 43). By analogy to the activities of MyoD and related myogenic regulatory genes (44), IESFl may play a role in developmental decisions leading to the generation of the β cell. Alternatively, its role may be more restricted to transcriptional activation of the insulin gene in the fully differentiated β cell. The observation that IEF1 or a very similar binding activity is present in both endocrine α cells (45) and pituitary somatotrophs cells (10) is consistent with this view. An important goal in further understanding insulin gene transcription is the isolation of cDNA clones encoding IESFl.

Acknowledgments—We thank Drs. Y. Citri, A. Aronheim, and A. Rosen for valuable discussions, Drs. C. Kabanna, A. Lassar, and H. Weintraub for generously providing plasmids, Drs. A. Lassar and H. Weintraub for anti-MyoD antisera, and Drs. U. Nudel and Y. Shaul for oligonucleotides.

REFERENCES

SUPPLEMENTAL MATERIAL TO

Subunit Structure of Cell-Specific E-box Binding Proteins
Analyzed by Quantitation of Electrophoretic Mobility Shift
Cheryl Won Park and Michael D. Walker

**RESULTS**

Deletion mutation of TA1 protein in the MLH1 motif. The TA1 protein encompasses the basic HLH (bHLH) domain (55 amino acids) of A1 together with an additional 37 flanking amino acids (Fig. 4A). Therefore the bHLH domain of the TA1 protein is homologous to and is a subset of the bHLH domain of the TA1 protein. However, two cell extracts suggests that the bHLH domain is required for formation of intermediates. To confirm this, we used the basic and bHLH domains of the TA1 protein in the DNA binding analysis, we generated deletion mutants lacking portions of the bHLH motif (AvvTA1) or the helix-II epitope (TAIPst). (Fig. 5A).

Mixing of in vitro translated TA1 protein and either HEY or WEHI 231 nuclear extract, generated two bands (3 and 4), the intermediate species (lanes 2 and 4) 11 such bands. The heterodimeric intermediate species, which are present in the TA1 DNA construct prior to in vitro transcription, the region representing to basic, helix and loop domains are indicated. The N terminal 14 amino acids of TA1 and TAIPst are derived from phage/bacteriophage sequences.

**Figure 5B** - EMSA using IEB1 probe with nuclear extracts (HEI and WEHI 231, lanes 1-2, resp.) and in vitro translated proteins (AvvTA1, TAIPst, and TA1, lanes 3-5, resp). Lane 1 shows the binding activity of in-vitro translated protein TA1. TAIPst and TA1 probed in the presence of NIH NRK extract (lanes 4-5) and WEHI 231 nuclear extract (lanes 9-11). The intermediate bands are indicated by arrows on the right.

To confirm that these deletions do in fact reduce dimerization and/or DNA binding with bona fide bHLH proteins, we used in vitro translated A125 protein as a heterodimerization partner for TA1 and its deletion mutant derivatives (Fig. 5C). Mixing of in vitro translated A125 and TA1 protein generated an intermediate complex (lane 8), absent from control reactions (lanes 1, 2, and 3). This intermediate complex did not appear upon mixing of in vitro translated AvvTA1 and A125 proteins (lane 8), while in vitro translated TAIPst protein did generate the complex, but with much reduced intensity (lane 7), indicating that this protein while incapable of homodimerization retains residual ability to heterodimerize with an intact HLH protein

**Figure 5C - EMSA using IEB1 probe with nuclear extracts (HEI and WEHI 231, lanes 1-5, resp.) and in vitro translated proteins (AvvTA1, TAIPst, and TA1, lanes 3-5, resp). Lane 1 shows the binding activity of in-vitro translated protein TA1. TAIPst and TA1 probed in the presence of NIH NRK extract (lanes 4-5) and WEHI 231 nuclear extract (lanes 9-11). The intermediate bands are indicated by arrows on the right.

To test the translational efficiency of the RNAs, we performed in vitro translations in the presence of 3H-methionine, and fractionated the products by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) (Fig. 5D). Mixing of in vitro translated A125 and TA1 proteins generated an intermediate complex (lane 8), absent from control reactions (lanes 1, 2, and 3). This intermediate complex did not appear upon mixing of in vitro translated AvvTA1 and A125 proteins (lane 8), while in vitro translated TAIPst protein did generate the complex, but with much reduced intensity (lane 7), indicating that this protein while incapable of homodimerization retains residual ability to heterodimerize with an intact HLH protein.