Proc. Natl. Acad. Sci. USA
Vol. 90, pp. 8063–8067, September 1993
Biochemistry

The *E2A* gene product contains two separable and functionally distinct transcription activation domains

(pancreatic beta cells/helix–loop–helix/insulin)

AMI ARONHEIM, RINAT SHIRAN, ADA ROSEN, AND MICHAEL D. WALKER*

Biochemistry Department, Weizmann Institute of Science, Rehovot, Israel 76100

Communicated by William J. Rutter, May 10, 1993

ABSTRACT **The *E2A* gene encodes transcription factors of the helix–loop–helix (HLH) family which are implicated in cell-specific transcriptional control in several cell lineages, including pancreatic beta cells. In the present work, we show by deletion mapping of both the E2A protein itself and the Gal4-E2A fusion protein that the protein contains at least two distinct activation domains. One domain (located between amino acids 1 and 153) functions efficiently in a variety of mammalian cell lines. The second domain (located between amino acids 369 and 485) functions preferentially in pancreatic beta cell lines. The latter domain shows a pattern of heptad repeats of leucine residues characteristic of "leucine zipper" transcription factors; site-directed mutagenesis of leucines within this repeat led to substantial reductions in activity. The selective properties of this activation domain may contribute to cell-specific transcription directed by the *E2A* gene.**

Cell-specific transcription of many genes is controlled in part by the cis elements known as E boxes (consensus sequence CANNTG) (1–4). The E boxes are recognized by lineage-restricted nuclear protein complexes such as insulin enhancer binding factor 1 (IEF1) in pancreatic beta cells (5, 6), muscle enhancer binding factor 1 (MEF1) in muscle cells (7, 8), and B-cell factor 1/lymphocyte enhancer binding factor 1 (BCF1/LEF1) in lymphoid cells (6, 8). These complexes most likely are positively acting cell-specific transcription factors. It has recently been shown that the E boxes can be specifically bound *in vitro* by proteins of the helix–loop–helix (HLH) family; for example, the *E2A* gene encodes, by means of alternative splicing, two proteins, E12 and E47, that can recognize the κE2/μE5 sequence elements located in the enhancers of immunoglobulin genes (9–13). *In vitro* studies have shown that HLH proteins are capable of homo- and heterodimerization that is essential for DNA binding (14). The muscle-specific HLH protein MyoD homodimerizes poorly yet heterodimerizes efficiently with E2A protein (14, 15). These and other observations led to the notion that cell-specific transcription factors such as MEF1 are composed of heterodimeric HLH complexes containing a cell-specific [or class B (14)] HLH partner and a ubiquitous [class A (14)] HLH partner, the E2A protein (7). Such a pattern is also observed in pancreatic beta cells (6, 13, 16, 17): the factor IEF1 appears to be a heterodimer of E2A protein with IESF1, a 25-kDa endocrine-specific protein (18). By contrast, the lymphoid-specific E box binding factor BCF1/LEF1 appears to be a homodimeric complex of E2A proteins (6, 8, 19). The precise mechanism by which these complexes control cell-specific transcription is not completely understood: homo- and heterodimeric HLH complexes containing E2A protein can bind efficiently *in vitro* to E boxes of genes expressed in multiple cell types, including muscle, lymphoid, and pancreatic cells. How is specificity achieved?

It has previously been shown that the E2A proteins are able to activate transcription through multiple E box sequences *in vivo* (11, 12, 20); the activity was observed in both yeast and mammalian cells and was attributed to the N-terminal ⅔ of the protein, distinct from the HLH domain (11). More recently, a transcriptional activation domain was localized to a 108-amino acid fragment of the human E2A protein with a putative loop–α-helical (LH) structure (21). We now extend these observations. We have identified two separable activation domains, designated AD I and AD II, that show distinct characteristics: AD I is highly active in all cell types tested, whereas AD II, which overlaps the previously defined LH fragment (21), shows strong preferential activity in pancreatic beta cells.

MATERIALS AND METHODS

Plasmids. The *E2A* gene encodes two major gene products, the result of alternative splicing events (9, 15). The proteins are designated E47 and E12 (human) (9, 15, 22), A1 and A7 (mouse) (10), Pan1 and Pan2 (rat) (12), and shPan1 and shPan2 (hamster) (13), respectively. An additional human cDNA (ITF-1) has been described which is identical to E47, except for a splice variation leading to a substitution of the N-terminal 100 amino acids of the encoded protein (11, 23). Our mapping analysis was performed with the A1 cDNA (10). Using standard procedures (24), we introduced portions of the A1 cDNA into the mammalian expression vector pRSV.HB, which contains the Rous sarcoma virus (RSV) promoter and simian virus 40 (SV40) virus splice and polyadenylylation signals (25, 26). Expression of protein from these plasmids utilized the endogenous translation stop codons, except for construct TA1, for which a stop codon was present in the primer used for construction of the expression plasmid. The reporter plasmid used [IEBx6-TK-CAT (20)] contained six copies of the IEB2 sequence upstream of the herpes simplex virus thymidine kinase (TK) promoter controlling the chloramphenicol acetyltransferase (CAT) (25) gene. IEB ("insulin enhancer box") is a key E box element found twice in the rat insulin I gene regulatory region (3, 27). All Gal4 hybrid constructs were generated by polymerase chain reaction (PCR) using appropriate primers. All fragments were designed to contain an in-frame *Bgl* II site and were introduced into an expression vector encoding Gal4 DNA-binding domain (residues 1–147) (28) under the control of the RSV promoter (26). In addition, all bottom-strand primers include a stop codon to ensure translation termination at the correct location. SV40 splice sites and polyadenylylation signals were present at the 3′ end

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: HLH, helix–loop–helix; LH, loop–α-helix; AD, activation domain; RSV, Rous sarcoma virus; TK, thymidine kinase; CAT, chloramphenicol acetyltransferase; EMSA, electrophoretic mobility shift analysis.

*To whom reprint requests should be addressed.

of all constructs (25). Appropriate regions of the plasmids were verified by DNA sequencing. The reporter plasmid used (5Gal4.E1b.CAT) (29) contained five copies of a Gal4 DNA-binding site upstream of the adenovirus E1b promoter controlling the CAT gene.

Site-Directed Mutagenesis. An oligonucleotide spanning the region to be mutated was synthesized and annealed with single-strand DNA corresponding to the *E2A* gene. Elongation and ligation were performed with T4 DNA polymerase and T4 DNA ligase, respectively (30). The reaction products were used to transform *Escherichia coli* strain HB101 cells, and mutated plasmids were isolated and verified by sequencing. PCR was used to amplify the desired segment, which was then subcloned in the Gal4 expression vector.

Cell Culture. The following established cell lines were used in this study: HIT M2.2.2 and βTC1 (hamster and mouse insulin-producing beta cells, respectively), CHO and BHK (hamster fibroblast cells), Ltk$^-$ (mouse fibroblast cells), SK-Hep1 (human undifferentiated liver cells), HepG2 (human hepatoma cells), and HeLa (human epithelial cells).

Transfection. Cells were cotransfected with the indicated amount of CAT (25, 31) reporter plasmid, expression plasmid, and internal control plasmid, either pRSV-βgal or pRSV-luc (32) (for HeLa cells). CAT activities were normalized by using the activity of β-galactosidase or luciferase obtained for each extract.

Electrophoretic Mobility Shift Analysis (EMSA). To compare the extent of accumulation of proteins encoded by expression plasmids, cells were transfected with 15 μg of the corresponding plasmid and nuclear extract was prepared 36–48 hr later (6, 33). EMSA was performed (6) with either IEB1 (GATCCGCCATCTGCCA) or Gal4 (GATCCCGGAGTACTGTCCTCCGA) double-stranded DNA probes.

RESULTS

Trans-Activation by E2A Protein. Cultured cells were cotransfected with expression plasmids and a reporter plasmid containing six repeats of an E box-binding sequence fused to the CAT reporter controlled by the TK promoter (20). The near-full-length E2A protein (amino acids 55–652) was able to efficiently activate expression of the reporter in both fibroblasts (CHO cells) (11, 12, 20) and pancreatic beta (HIT) (20) cells as previously reported (Fig. 1*A*, construct 1). N-terminal truncations showed a loss of trans-activation potential. However, these proteins differed in their capacity to trans-activate in distinct cells; for example, construct 2 (amino acids 369–652) showed reduced but still substantial trans-activation capacity in HIT cells but virtually no activity in CHO cells (Fig. 1*A*). Further deletions led to complete loss of activity in both cell types (Fig. 1*A*, constructs 3 and 4). Since variations in the observed trans-activation capacity could result from alterations in stability, nuclear localization, or DNA-binding activity, we compared the accumulation of the truncated proteins by using EMSA with the E box probe IEB1 (Fig. 1 *B* and *C*). We observed comparable signals for the different constructions, indicating the differences in trans-activation capacity are most likely due to removal of amino acids involved in trans-activation *per se*. Apparently the N-terminal 368 amino acids contain one or more domains which function in both fibroblasts and beta cells. A domain important for activity in beta cells but not fibroblasts may be located C-terminal of Arg-369.

Transactivation by Gal4-E2A Hybrid Proteins. To characterize the precise regions involved in trans-activation, we generated constructs which encode fusion proteins between domains of E2A and the Gal4 DNA-binding domain (28). We cotransfected the Gal4 expression plasmids with a CAT reporter plasmid bearing five copies of a Gal4 binding site linked to the E1b promoter. The N-terminal region of the

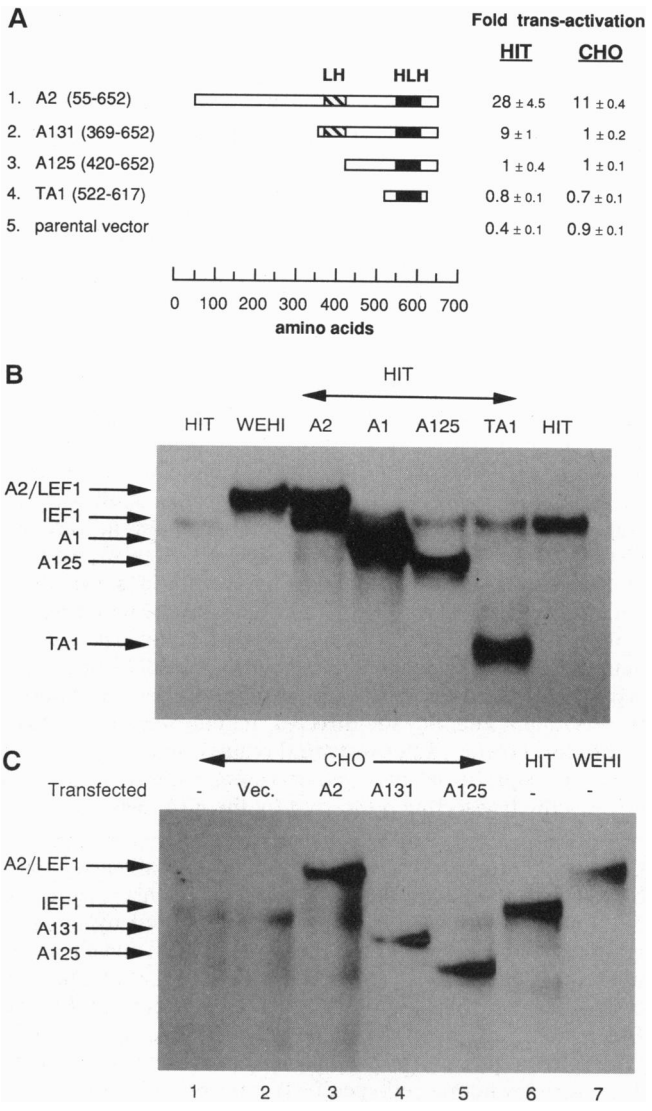
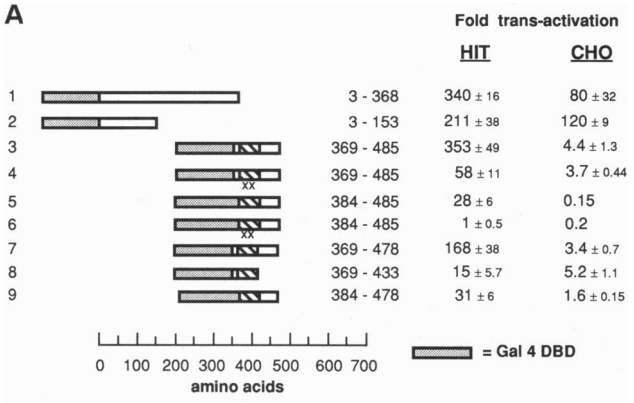


FIG. 1. (*A*) Trans-activation by A1 deletion mutants. The portions of A1 cDNA included in expression plasmids are indicated. For the trans-activation assay hamster insulin-producing cells (HIT) and fibroblast cells (CHO) were transiently cotransfected with 4 μg of pRSV.HB expression vector encoding the different A1 deletion mutants, 1 μg of CAT reporter plasmid under the control of either TK promoter (TK CAT) or TK promoter and enhancer consisting of six copies of IEB elements (IEBx6-TK-CAT), and 5 μg of pRSV-β-gal. The results shown are normalized CAT activities obtained in cotransfection assays with the indicated plasmid. Each number represents the ratio between the trans-activation of a plasmid containing six copies of the IEB2 element linked to the TK promoter and the activity obtained with the TK promoter alone. For each value SEM is shown from three to five independent experiments (except construct 3 in CHO cells, two experiments). The filled region corresponds to the basic-HLH domain required for DNA binding and dimerization. The hatched region corresponds to the leucine heptad repeat. (*B* and *C*) EMSA of expression of A1 deletion mutants. Cultured cells were transfected with 15 μg of expression plasmids encoding one of the series of A1 deletion mutants. Nuclear extracts were prepared and analyzed by EMSA, using the IEB1 probe. Extracts from nontransfected HIT M2.2.2 and WEHI231 (lymphoid) cells were used to provide migration markers for the complexes IEF1 and LEF1, respectively. The migration of homodimeric complexes A2, A131, A125, and TA1 is indicated. (*B*) Transfection in HIT cells. (*C*) Transfection in CHO cells. Vec., parental vector.

protein (amino acids 3–368) functioned very efficiently in HIT cells and CHO cells (340-fold and 80-fold, respectively); this activity could be further localized to the N-terminal 153



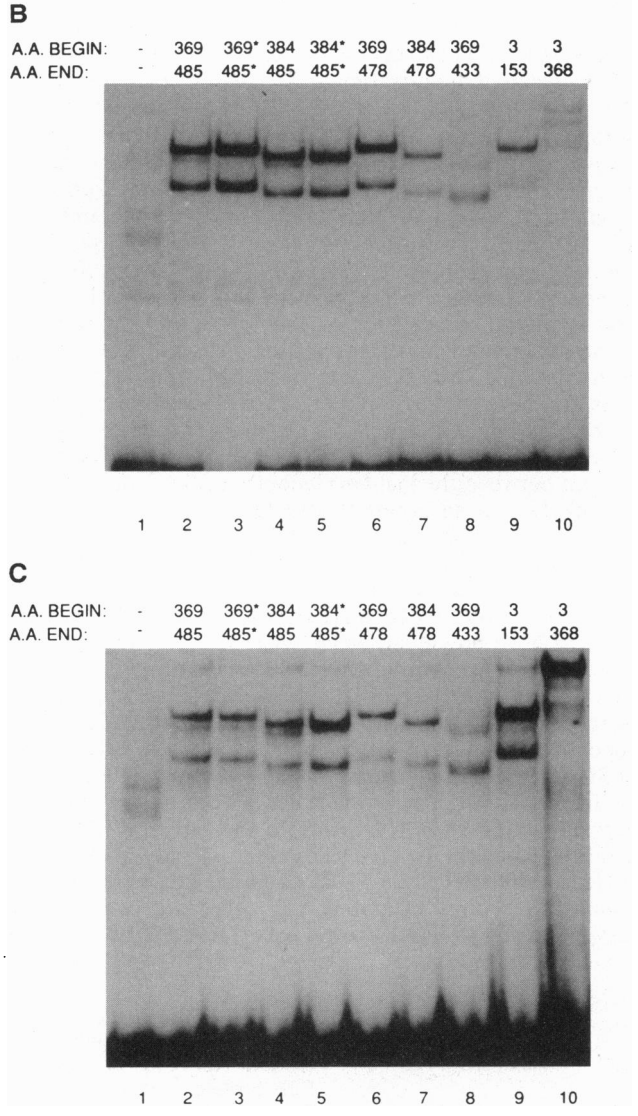


FIG. 2. (*A*) Trans-activation by Gal4 chimeric proteins. Hybrid plasmids (8 μg) encoding Gal4 DNA-binding domain (1–147) fused to the indicated protein sequences derived from the A1 protein were cotransfected into HIT and CHO cells together with 2 μg of CAT reporter plasmid under the control of five copies of Gal4 DNA-binding sites (5Gal4.E1b.CAT) and 5 μg of the pRSV-β-gal control plasmid. The symbol XX for constructs 4 and 6 indicates the mutation of two leucine residues of the putative zipper (Leu-395 and Leu-402) to phenylalanine. The results shown are normalized CAT activities obtained in cotransfection assays with the indicated plasmid. The resulting trans-activation is calculated as relative activity compared with the activity of the Gal4 DNA-binding domain (1–147) devoid of any trans-activation domain. For each value SEM is shown for three to five independent experiments (except constructions 5 and 6 in CHO cells—two experiments). The hatched region corresponds to the leucine heptad repeat. (*B* and *C*) EMSA of expression of Gal4-E2A deletion mutants. Cells were transfected with 15 μg of expression plasmid encoding Gal4-E2A deletion mutants. Nuclear extracts were prepared and analyzed by EMSA using the Gal4 DNA probe. (*B*) Transfection in HIT cells. (*C*) Transfection in CHO cells. The asterisks indicate transfection of plasmids bearing double Leu-to-Phe mutations at amino acids 395 and 402.

amino acids (3–153) (Fig. 2*A*, constructs 1 and 2). On the other hand, the region 369–485, while showing similar activity to the N-terminal domain in beta cells, showed little or no activity in CHO cells (Fig. 2*A*, construct 3). Further removal of amino acids from N- or C-terminal portions led to substantial loss of activity (constructs 5, 7, 8, and 9). The results are consistent with those obtained with the native E2A protein and indicate the presence of two separable and functionally distinct transcription activation domains that can function independently of the basic-HLH domain. We designate the N-terminal domain (amino acids 3–153) as AD I and the more C-terminal domain (amino acids 369–485) as AD II. To compare the relative strengths of the activation domains, we tested a plasmid expressing Gal4-VP16 fusion protein (VP16 being a potent viral trans-activator protein), controlled by the RSV promoter; we observed activation of 401-fold and 86-fold in HIT and CHO cells, respectively (data not shown).

Within AD II is located a region of heptad repeats of leucine residues (34, 35) with the sequence $LX_6LX_6TX_6LX_6L$. We performed site-directed mutagenesis to replace the first two leucine residues of the potential zipper (Leu-395 and Leu-402) with phenylalanine, a mutation that should substantially reduce the potential for protein–protein interaction (36). Indeed these mutations in the context of either the fragment 369–485 or the fragment 384–485 had substantial effects on the activity of AD II in HIT cells (Fig. 2*A*, constructs 4 and 6).

Extracts of transfected cells were used to determine the binding activity of the expressed fusion protein, using a Gal4 DNA-binding site probe (Fig. 2 *B* and *C*). Extracts from transfected cells consistently showed doublet bands of binding activity, presumably reflecting alternative multimerization of the Gal4 fusion proteins by means of the Gal4-(1–147) region. Comparison of the intensity of the bands observed among the various constructs revealed only minor differences that cannot explain the differential activation potential of the constructs and thus indicates that the deletion analysis defines bona fide activation domains.

The striking functional differences between domains AD I and AD II in HIT compared with CHO cells (Fig. 2*A*) prompted us to test their action in additional pancreatic beta cells as compared with other cell types. Using the mouse beta cell line βTC1 we observed, as with HIT cells, a strong action of AD II (4.7-fold higher than AD I) (Fig. 3). On the other hand, five additional non-beta cells—BHK, Ltk$^-$, SK-Hep1, HepG2, and HeLa—showed a much lower activity of AD II, 0.01–0.22 of that observed with AD I (Fig. 3). Thus AD II appears to show preferential function in beta cells as compared with non-beta cells.

DISCUSSION

Eukaryotic transcription factors typically show a modular structure, containing separable DNA-binding and transcription-activation domains (37); domain-swap experiments have shown that functional molecules can be generated by interchanging domains (38, 39). However, the ability to bind DNA, though necessary for function, is not a sufficient condition for transcription activation *in vivo*. Thus, transcription factors with similar DNA-binding specificities may pro-

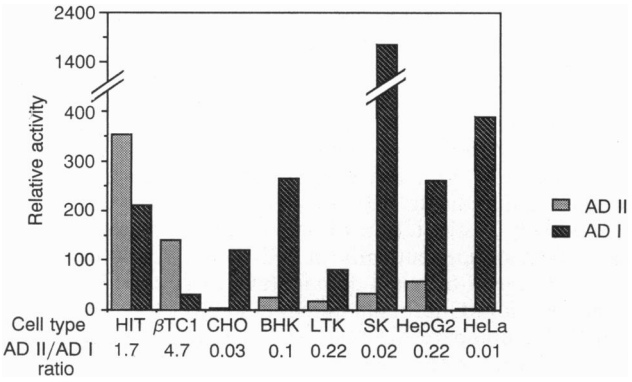


FIG. 3. Trans-activation by Gal4 chimeric proteins. Gal4-E2A expression plasmids encoding AD I or AD II were transfected into the following cell types: HIT, βTC1, CHO, BHK, Ltk⁻, SK-Hep1, HepG2, and HeLa. The trans-activation capacity of each plasmid was determined (see legend to Fig. 2*A*). For each cell type, the ratio of activity AD II/AD I is shown. The values represent the mean of two to five independent experiments. The relative activities of AD II and AD I respectively for each cell type are as follows: HIT, 353/211; βTC1, 141/30.0; CHO, 4.4/120; BHK, 25.5/265; Ltk⁻, 17.6/80.5; SK-Hep1, 35.0/1743; HepG2, 57/262; HeLa, 4.5/390.

duce widely diverse effects on transcription; mutations of homeobox genes, for example, lead to distinct and characteristic phenotypes in *Drosophila*, yet the encoded proteins show relatively low DNA-binding specificity *in vitro* (reviewed in ref. 40). Therefore, additional specificity must be generated through functional domains distinct from the DNA-binding capacity of the protein (41–43).

The HLH proteins contribute toward diverse developmental programs in *Drosophila* (44) and in higher eukaryotes (45); distinct mammalian cell lineages appear to contain characteristic heterodimeric complexes consisting of a lineage-restricted HLH protein together with a widely distributed HLH protein, the E2A protein. The precise role of the E2A protein is not established; conceivably its major function could be to permit efficient DNA binding of the cell-specific HLH protein—typically such proteins bind weakly to E boxes as homodimers; alternatively, the E2A protein could participate actively by providing activation domains with either ubiquitous or lineage-restricted potential. For example, it has been shown that an E2A protein containing a portion of the basic-HLH region of MyoD is able to activate the myogenic program, albeit at reduced efficiency (46). To further examine these issues we have analyzed the transcription activation domains of the E2A protein.

We have identified and mapped two distinct transcription activation domains in the E2A protein, designated AD I and AD II. Neither domain shows close similarity to previously reported activation domains of other genes. AD I has a high concentration of serine (34/153) and glycine (19/153) residues (Fig. 4*A*). Interestingly, one of the human E2A proteins, ITF-1, differs extensively from E12/E47 and mouse E2A in the first 100 amino acids at the N terminus (Fig. 4) (11, 23). This region of ITF-1 shows little or no trans-activation capacity (21). AD II has a high concentration of serine (17/117) and leucine (18/117) residues (Fig. 4*B*). In addition, AD II contains a heptad repeat of leucine residues (Fig. 4*B*) characteristic of "leucine zipper" proteins such as C/EBP, Jun, and Fos; the repeat is involved in generating the coiled-coil dimer structure required for function of this class of proteins (47). A heptad repeat of leucines has also been identified in a class of HLH proteins designated HLH-zip and including the proteins myc (48), USF (49), and TFE3 (50). In the HLH-zip subfamily, the zipper region is located adjacent to the HLH and has been shown to play a role in stabilizing HLH-mediated dimerization and to restrict the dimerization specificity of these proteins (51, 52). By contrast, in the E2A proteins, the heptad repeat is separated from the HLH region by over 100 amino acids and has not yet been assigned a clear functional role in the molecule; it is clearly not essential for E box DNA binding (9). Both in the HLH-zip subfamily and in other zipper transcription factors, the transcription activation domain maps to regions of the protein distinct from the zipper (50). In our experiments, site-directed mutation of the first two leucine residues in the repeat of AD II led to substantial reduction in trans-activation potential, suggesting an important role in directing transcription activation, perhaps involving interaction with other zipper proteins. However, additional complexity is probably involved, since the α-helix-destabilizing amino acids proline and glycine are present between the last two leucines of the repeat, making it unlikely that an α-helical structure could form at the C terminus of the leucine repeat. Furthermore, in our experiments, the leucine repeat region alone is not an efficient trans-activator: downstream sequences are required in addition. Interactions between HLH proteins and leucine zipper proteins have previously been suggested—for example, the functional antagonism between the c-Jun protein and myogenesis may involve an interaction between the zipper region of the c-Jun protein (53); however the biochemical target appears to be not the E2A protein but rather the HLH region of MyoD (53) and myogenin (54).

The activation domain identified by our studies as AD II overlaps extensively (Fig. 4) with the putative LH region identified in the human E2A protein as an activation domain in HeLa cells and yeast (21). By comparing the activity of this region in a variety of mammalian cell types, we have observed striking selectivity in the action of AD II; although AD

A AD I

```
mouse E2A    1   MMNQSQRMAPVGSDKELSDLLDFSMMFPLPVANGKSRPASLGGTQFAGSGLEDRPSSGSWGSSDQNSSSFDPSRTYS  77
                                                                 | |       |
human E2A    1                                                GGGECL-------AW------CGPSAVHRCAD  19

mouse E2A   78   EGAHFSDSHSSLPPSTFLGAGLGGKGSERNAYATFGRDTSVGTLSQAGFLPGELSLSSPGPLSPSGIKSSSQYYPS  153
                                                 ||  || ||| |||| || | ||||| ||| | |||||||||| |  ||||||
human E2A   20   VGLGMVSARTA----------PGKSGERGAYASFGRDAGVGGLTQAGFLSGELALNSPGPLSPSGMKGTSQYYPS  83
```

B AD II

```
mouse E2A  323  asssgdalgkalasiyspdhssnnfspspstpvgspqglpgtsqwpRAGAPSALSPNYDAGLHGLQSKMEDRLDEAIHVLRS  404
                | |||||||||||||||||||||||| |||||||||| |||||||||| |||| || |||||||| || ||||||||||
human E2A  259  AGSSGDALGKALASIYSPDHSSNNFSSSPSTPVGSPQGLAGTSQWPRAGAPGALSPSYDGGLHGLQSKIEDHLDEAIHVLRS  340
                 .    .    . /.....................LOOP.....................| .HELIX.|

mouse E2A  405  HAVGTASDLHGLLPGHGALTTSFTGPMSLGGRHAGLVGGSHPEEGLTSGASLLHNHASLPSQPSSLPDLSQRPP  478
                |||||| | | ||||||||   ||||||||||||||||||||||| ||    |||||||| |||||  |||||| |||
human E2A  341  HAVGTAGDMHTLLPGHGALASGFTGPmslggrhaglvggshpedglagstslmhnhaalpsqpgtlpdls-rpp  413
```

FIG. 4. Sequence of the transcription activation domains AD I (*A*) and AD II (*B*) of the mouse E2A protein. The sequences are aligned with the human E2A protein [ITF-1 (23)] used in ref. 21. Hyphens indicate gaps introduced to optimize the sequence alignment. Vertical lines indicate amino acid identity. The putative LH region (21) in the human E2A protein is indicated. In *B*, domains defined in functional assays are shown in uppercase letters; flanking sequences are shown in lowercase letters. The heptad leucine repeat is indicated by underscored letters.

II functions detectably in non-beta cells, the activity is consistently lower than observed in beta cells; in non-beta cells the ratio of activity AD II/AD I ranges from 0.01 to 0.22, whereas in beta cells the ratio is 1.7–4.7 (Fig. 3). We tested a plasmid encoding amino acids 330–485 of mouse E2A, which include the entire putative loop (21), and we observed no significant differences in activity in HIT, HeLa, or CHO cells as compared with AD II (369–485) (data not shown).

Typically activation domains, even those derived from cell-specific transcription factors, are functional in nonselective fashion. For example, the transcription factor Oct2, though expressed in lineage-restricted fashion (mainly lymphoid cells), contains an activation domain that functions efficiently in nonlymphoid cells (55). The effect we observe is also clearly distinct from the specificity inherent in the action of the muscle activators MyoD and myogenin: in these cases, specificity involves the presence of the DNA-binding regions of these proteins and is believed to involve action of an accessory transcription factor (recognition factor) which interacts with the DNA-binding domain (56). The selectivity of AD II can be most easily explained by postulating either cell-specific post-translational modification of the protein or the presence of a transcriptional coactivator/corepressor which is distributed in cell-specific fashion (57). Such a factor may interact with AD II, perhaps via the leucine heptad, to increase the efficiency of transcription through protein–protein contacts with the basal transcription machinery. It remains to be established how the function of AD II is integrated in the context of the heterodimeric beta cell HLH complex. This will require analysis of trans-activation by E2A protein with a mutated AD II both in isolation and in combination with the beta cell HLH partner. Such experiments using native E box binding sites will indicate the relative contribution of AD I and AD II to cell-specific expression in the context of the intact insulin regulatory region. The present data provide the basis for these approaches by showing that inherent specificity of an activation domain of the E2A protein may contribute toward determining the expression pattern of target genes.

We thank Drs. Y. Shaul, Y. Citri, I. Haviv, C. W. Park, Y. Arava, and F. Argenton for valuable discussions; E. Or and D. Vaisel for assistance with DNA sequencing and construction of plasmids; and Drs. Y. Shaul, V. Daniel, and I. Haviv for cell lines and plasmids. The work was supported by grants from the Juvenile Diabetes Foundation International, the Israel Academy of Sciences and Humanities, the Henry and Francoise Glasberg Foundation, and the Forchheimer Center for Molecular Genetics. M.D.W. is the incumbent of the Maurice M. Bookstein Career Development Chair.

1. Ephrussi, A., Church, G. M., Tonegawa, S. & Gilbert, W. (1985) *Science* **227**, 134–138.
2. Lenardo, M., Pierce, J. & Baltimore, D. (1987) *Science* **236**, 1573–1577.
3. Karlsson, O., Edlund, T., Moss, J. B., Rutter, W. J. & Walker, M. D. (1987) *Proc. Natl. Acad. Sci. USA* **84**, 8819–8823.
4. Meister, A., Weinrich, S. L., Nelson, C. & Rutter, W. J. (1989) *J. Biol. Chem.* **264**, 20744–20751.
5. Ohlsson, H., Karlsson, O. & Edlund, T. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 4228–4231.
6. Aronheim, A., Ohlsson, H., Park, C. W., Edlund, T. & Walker, M. D. (1991) *Nucleic Acids Res.* **19**, 3893–3899.
7. Lassar, A. B., Davis, R. L., Wright, W. E., Kadesch, T., Murre, C., Voronova, A., Baltimore, D. & Weintraub, H. (1991) *Cell* **66**, 305–315.
8. Murre, C., Voronova, A. & Baltimore, D. (1991) *Mol. Cell. Biol.* **11**, 1156–1160.
9. Murre, C., McCaw, P. S. & Baltimore, D. (1989) *Cell* **56**, 777–783.
10. Walker, M. D., Park, C. W., Rosen, A. & Aronheim, A. (1990) *Nucleic Acids Res.* **18**, 1159–1166.
11. Henthorn, P., Kiledjian, M. & Kadesch, T. (1990) *Science* **247**, 467–470.
12. Nelson, C., Shen, L. P., Meister, A., Fodor, E. & Rutter, W. J. (1990) *Genes Dev.* **4**, 1035–1043.
13. German, M. S., Blanar, M. A., Nelson, C., Moss, L. G. & Rutter, W. J. (1991) *Mol. Endocrinol.* **5**, 292–299.
14. Murre, C., McCaw, P. S., Vaessin, H., Caudy, M., Jan, L. Y., Jan, Y. N., Cabrera, C. V., Buskin, J. N., Hauschka, S. D., Lassar, A. B., Weintraub, H. & Baltimore, D. (1989) *Cell* **58**, 537–544.
15. Sun, X.-H. & Baltimore, D. (1991) *Cell* **64**, 459–470.
16. Cordle, S. R., Henderson, E., Masuoka, H., Weil, P. A. & Stein, R. (1991) *Mol. Cell. Biol.* **11**, 1734–1738.
17. Shieh, S.-Y. & Tsai, M.-J. (1991) *J. Biol. Chem.* **266**, 16708–16714.
18. Park, C. W. & Walker, M. D. (1992) *J. Biol. Chem.* **267**, 15642–15649.
19. Aronheim, A., Shiran, R., Rosen, A. & Walker, M. D. (1993) *Nucleic Acids Res.* **21**, 1601–1606.
20. Leshkowitz, D., Aronheim, A. & Walker, M. D. (1992) *DNA Cell Biol.* **11**, 549–558.
21. Quong, M. W., Massari, M. E., Zwart, R. & Murre, C. (1993) *Mol. Cell. Biol.* **13**, 792–800.
22. Kamps, M. P., Murre, C., Sun, X. & Baltimore, D. (1990) *Cell* **60**, 547–555.
23. Henthorn, P., McCarrick-Walmsley, R. & Kadesch, T. (1990) *Nucleic Acids Res.* **18**, 677.
24. Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab. Press, Plainview, NY), 2nd Ed.
25. Gorman, C. M., Moffat, L. F. & Howard, B. H. (1982) *Mol. Cell. Biol.* **2**, 1044–1051.
26. Gorman, C. M., Merlino, G. T., Willingham, M. C., Pastan, T. R. & Howard, B. H. (1982) *Proc. Natl. Acad. Sci. USA* **79**, 6777–6781.
27. Karlsson, O., Walker, M. D., Rutter, W. J. & Edlund, T. (1989) *Mol. Cell. Biol.* **9**, 823–827.
28. Ma, J. & Ptashne, M. (1987) *Cell* **48**, 847–853.
29. Lillie, J. W. & Green, M. R. (1989) *Nature (London)* **338**, 39–44.
30. Kunkel, T. A. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 488–492.
31. Edlund, T., Walker, M. D., Barr, P. J. & Rutter, W. J. (1985) *Science* **230**, 912–916.
32. De Wet, J. R., Wood, K. V., DeLuca, M., Helinski, D. R. & Subramani, S. (1987) *Mol. Cell. Biol.* **7**, 725–737.
33. Schreiber, E., Matthias, P., Muller, M. M. & Schaffner, W. (1989) *Nucleic Acids Res.* **17**, 6419.
34. Landschulz, W. H., Johnson, P. F. & McKnight, S. L. (1988) *Science* **240**, 1759–1764.
35. O'Shea, E. K., Rutkowski, R. & Kim, P. S. (1989) *Science* **243**, 538–542.
36. Smeal, T., Angel, P., Meek, J. & Karin, M. (1989) *Genes Dev.* **3**, 2091–2100.
37. Keegan, L., Gill, G. & Ptashne, M. (1986) *Science* **231**, 699–704.
38. Brent, R. & Ptashne, M. (1985) *Cell* **43**, 729–736.
39. Ptashne, M. (1988) *Nature (London)* **335**, 683–689.
40. Hayashi, S. & Scott, M. P. (1990) *Cell* **63**, 883–894.
41. Johnson, P. F. & McKnight, S. L. (1989) *Annu. Rev. Biochem.* **58**, 799–839.
42. Thompson, C. C. & McKnight, S. L. (1992) *Trends Genet.* **8**, 232–236.
43. German, M. S., Wang, J. H., Chadwick, R. B. & Rutter, W. J. (1992) *Genes Dev.* **6**, 2165–2176.
44. Caudy, M., Vassin, H., Brand, M., Tuma, R., Jan, L. Y. & Jan, Y. N. (1988) *Cell* **55**, 1061–1067.
45. Kadesch, T. (1993) *Cell Growth Differ.* **4**, 49–55.
46. Davis, R. L. & Weintraub, H. (1992) *Science* **256**, 1027–1030.
47. Lamb, P. & McKnight, S. L. (1991) *Trends Biochem. Sci.* **16**, 417–422.
48. Blackwood, E. M. & Eisenman, R. N. (1991) *Science* **251**, 1211–1217.
49. Gregor, P. D., Sawadogo, M. & Roeder, R. (1990) *Genes Dev.* **4**, 1730–1740.
50. Beckmann, H., Su, L.-K. & Kadesch, T. (1990) *Genes Dev.* **4**, 167–179.
51. Hu, Y.-H., Luscher, A., Admon, A., Mermod, N. & Tjian, R. T. (1990) *Genes Dev.* **4**, 1741–1752.
52. Beckmann, H. & Kadesch, T. (1991) *Genes Dev.* **5**, 1057–1066.
53. Bengal, E., Ransone, L., Scharfmann, R., Dwarki, V. J., Tapscott, S. J., Weintraub, H. & Verma, I. M. (1992) *Cell* **68**, 507–519.
54. Li, L., Chambard, J. C., Karin, M. & Olson, E. N. (1992) *Genes Dev.* **6**, 676–689.
55. Gerster, T., Balmaceda, C.-G. & Roeder, R. G. (1990) *EMBO J.* **9**, 1635–1643.
56. Davis, R. L., Cheng, P.-F., Lassar, A. B. & Weintraub, H. (1990) *Cell* **60**, 733–746.
57. Baichwal, V. R. & Tjian, R. T. (1990) *Cell* **63**, 815–825.