Proc. Natl. Acad. Sci. USA
Vol. 89, pp. 6152–6156, July 1992
Developmental Biology

Enhancer of split^D, a dominant mutation of *Drosophila*, and its use in the study of functional domains of a helix–loop–helix protein

(DNA binding)

KYRIA TIETZE*, NADJA OELLERS, AND ELISABETH KNUST†

Institut für Entwicklungsbiologie, Universität zu Köln, Gyrhofstrasse 17, 5000 Köln 41, Federal Republic of Germany

Communicated by C. Nüsslein-Volhard, March 27, 1992

ABSTRACT **Helix–loop–helix proteins play important roles in developmental processes, such as myogenesis, neurogenesis, and sex determination. The gene Enhancer of split [*E(spl)*] of *Drosophila*, a member of a gene complex that is involved in early neurogenesis, encodes a protein with a basic domain and a helix–loop–helix motif. We took advantage of a dominant mutation of this gene, *E(spl)*^{D}, to define *in vivo* structural features of this protein for proper function. The mutation renders the otherwise recessive eye phenotype of *spl* dominant. By germ-line transformation of different *in vitro* mutagenized versions of the *E(spl)* gene, we could demonstrate that the basic domain of this helix–loop–helix protein is functional and necessary for expression of the dominant phenotype. These results are supported by *in vitro* DNA-binding assays, which showed that the basic domain is in fact necessary for DNA binding, despite the presence of a proline residue. Furthermore, we could show that the dominant enhancement of *spl* is caused by truncation of the E(SPL)^D protein in combination with deletion of a putative regulatory element.**

The gene Enhancer of split [*E(spl)*] of *Drosophila* is part of a gene complex [*E(SPL)–C*], the function of which is required for control of the binary decision between neural and epidermal cell fates in uncommitted cells of the neurogenic ectoderm (1). Results of genetic and cell transplantation experiments suggest that *E(SPL)–C* is a key element in establishment of epidermal commitment of neuroectodermal cells (2, 3). Seven members of this complex, including the *E(spl)* gene, encode proteins characterized by basic and helix–loop–helix domains (the bHLH motif) (ref. 4; E.K., H. Schrons, F. Grawe, and J. A. Campos-Ortega, unpublished data), suggesting a function in DNA binding and transcriptional regulation. bHLH proteins have been found in many organisms, ranging from yeast (5, 6) and plants (7) to animals, and some of them are of crucial importance for control of developmental pathways, such as myogenesis or neurogenesis (see refs. 1, 8, and 9 for reviews). The common characteristic of these proteins is their suggested DNA-binding capacity, mediated by the bHLH domains (10), and in several cases bHLH proteins have been shown to act as transcriptional activators (e.g., see refs. 11–13). The HLH domain is required for dimerization, which is a prerequisite for DNA binding, whereas target specificity is mediated by the basic domain, which lies adjacent to the HLH motif (14, 15). Unlike most other bHLH proteins, the E(SPL) protein and the other six bHLH proteins of the E(SPL)–C complex each contains a proline residue in its basic domain (ref. 4; E.K., H. Schrons, F. Grawe, and J. A. Campos-Ortega, unpublished data), which, in the case of MyoD (15), has been reported to be incompatible with DNA binding. Thus, the DNA-binding capabilities of the E(SPL)–C products are open to question.

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.

In many instances, conclusions drawn from *in vitro* DNA-binding assays suffer from the drawback that no *in vivo* test is available to prove the relevance of the observed effects for the organism or the cell itself. In the case of the *E(spl)*^{D} mutation, a dominant mutation of the gene *E(spl)*, we have an ideal tool in hand with which to study both *in vivo* and *in vitro* the effects of modifications of the gene. This mutation renders expression of the otherwise recessive phenotype of split (*spl*), an allelomorph of the neurogenic gene Notch (*N*), dominant (16–19). *N* encodes a transmembrane protein with 36 epidermal growth factor (EGF)-like repeats in the extracellular domain and is involved in receiving the lateral inhibition signal that suppresses the neural and allows the epidermal fate in cells of the neurogenic ectoderm. Furthermore, *N* is involved during several other processes that require cell fate choices—e.g., development of the sensilla or the compound eye. The recessive *spl* mutation is due to a single amino acid exchange in one of the EGF-like domains and leads to development of fewer and split bristles (hence the name) and failure of the photoreceptor cells to form proper ommatidial clusters (for recent review, see ref. 20 and references therein). In combination with *E(spl)*^{D}, the latter phenotype is particularly enhanced and is accompanied by severe cell death (19).

The *E(spl)*^{D} chromosome carries lesions in two genes, two deletions, one within and one upstream of the gene *E(spl)* (which corresponds to transcription unit *m8*), and an insertion in the adjacent gene groucho (*gro*; transcription unit *m9/m10*) (refs. 4, 21, and 22; Fig. 1*A*). We have previously shown by germ-line transformation that the ability to enhance the split phenotype is associated with the mutant *E(spl)*^{D} gene (ref. 4; Fig. 2*C*). Here we used the mutant gene as well as several *in vitro* mutagenized versions to demonstrate the importance of the basic domain for the enhancing effect of the *spl* phenotype, and we confirmed these results by *in vitro* DNA-binding experiments. Furthermore, we established a causal relationship between molecular lesions in the *E(spl)*^{D} gene and the dominant phenotype.

MATERIALS AND METHODS

***Drosophila* Stocks.** Flies were maintained according to standard laboratory procedures. All crosses were performed at room temperature or at 25°C. Oregon^R was used as wild type; for description of balancer chromosomes and markers see refs. 23 and 24. The following alleles of E(SPL)–C were used: *gro*^{B48} (mutant for *gro*; ref. 25), *gro*$^{ry78-r171.1}$ (deficient for *gro* and other complementation groups distal to it; H. Schrons, E.K., and J. A. Campos-Ortega, unpublished data), *E(spl)*^{D} (16), P[(*w*$^+$) *m8*^{D}; *T18.2*], and P[(*w*$^+$) *m8*^{D}; *T18.3*] (4).

Abbreviation: bHLH, basic domain and helix–loop–helix motif.

*Present address: Howard Hughes Medical Institute, Research Laboratories, Division of Biology, 156-29, California Institute of Technology, Pasadena, CA 91125.
†To whom reprint requests should be addressed.

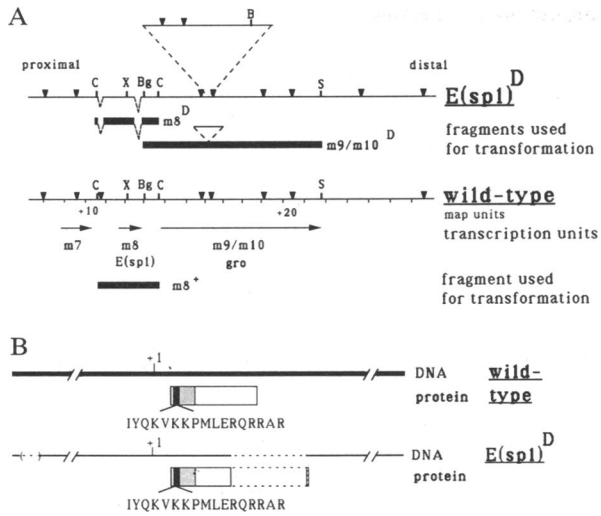

FIG. 1. The genes *E(spl)* and *gro* in wild type and $E(spl)^D$. (*A*) Part of the genomic region of the E(SPL) complex with the defects mapped in the $E(spl)^D$ chromosome (4, 21, 22). Dotted lines indicate deletions or insertions relative to wild-type DNA. All *Eco*RI sites (arrowheads) but only some of the other restriction sites (B, *Bam*HI; Bg, *Bgl* II; C, *Cla* I; S, *Sal* I; X, *Xho* I) are shown. Map units in kilobases (kb) are according to ref. 21. Arrows, transcription units; solid bars, fragments used for transformation. In the $E(spl)^D$ chromosome, the gene *gro* carries a 5-kb insertion of a middle repetitive element in an intron. Furthermore, two deletions have been mapped, one 5′ to the $E(spl)^D$ gene and the other at the 3′ end of its transcription unit. (*B*) Organization of the predicted E(SPL) and $E(SPL)^D$ protein products. Solid lines, genomic DNA; thick lines, wild-type DNA; thin lines, DNA from $E(spl)^D$; broken lines, deletions. Bars, putative translation products with basic domain (solid) and HLH domain (stippled). The predicted $E(SPL)^+$ and $E(SPL)^D$ proteins are identical up to the breakpoint of the deletion, but in the $E(SPL)^D$ protein the 56 C-terminal amino acids, including a conserved tetrapeptide (WRPW), are deleted and replaced by a different nine-amino acid sequence (hatched) (4).

Construction of Plasmids. *m8* and $m8^D$ constructs were subcloned as *Cla* I fragments from genomic clones (map units 10.7–13.8, according to ref. 21; see also Fig. 1) into the Bluescript vector (Stratagene) and cloned as *Pst* I/*Kpn* I fragments (using the restriction sites of the polylinker) into the transformation vector pW8 (26). To produce the recombinant inserts $m8^{+/D}$ and $m8^{D/+}$, we used the *Xho* I site at nucleotide 3523 (4). For the $m9,m10^D$ construct, phage clones were isolated from a genomic library made from $E(spl)^D$ DNA in the EMBL-4 phage vector (27). Two fragments from two different phages, a *Bgl* II (map unit 13.1)/*Bam*HI (site within the inserted DNA) fragment and a *Bam*HI/*Sal* I (map unit 21.9) fragment (map units according to ref. 21) were fused to build a complete $m9,m10^D$ fragment (Fig. 1*A*).

***In Vitro* Mutagenesis.** *In vitro* mutagenesis essentially followed the method described in ref. 28, with modifications as described by the supplier of the system (Amersham). Primers used are summarized below, numbering (in parentheses) is according to ref. 4, << indicates primers in antisense orientation. Nucleotide sequences that do not correspond to wild type are underlined, nucleotides in boldface represent introduced *Kpn* I restriction sites used for cloning or introduced stop codons. *In vitro* mutagenesis was confirmed by sequencing, using the chain-termination method (29), or by testing for the newly introduced restriction site, followed by sequencing. For construction of $m8^{+,del}$, two *Kpn* I sites were introduced into the $m8^+$ construct at the borders of the desired deletion by *in vitro* mutagenesis, using the primers CTACAAGAACTTGGTACCATTCCACGAAGC (nucleotides 3717–3746) and GGAAATCTATTTTGGTACCGACCGAGTTG (nucleotides 4200–4228). After removing the fragment between the two *Kpn* I sites, the remaining *Kpn* I site was eliminated by further *in vitro* mutagenesis, using primer GCGTCTACAAGAACTTGGCAATGACCGAGTTGA (nucleotides 3713–3729 and 4214–4229). The nucleotides at positions 4225 and 4226 of this construct are the same as in the $m8^D$ gene, thus leading to the same nine-amino acid sequence as in the dominant mutation. For construction of $m8^{+,stop}$, different termination codons were used: AAGAACTTGTAGCAATTCCA (nucleotides 3721–3740), CAAGAACTTGTAACAATTCCAC (nucleotides 3720–3741), and CTACAAGAACTTGTAATAATAACACGAAGCACAG (nucleotides 3717–3750). For the remaining constructs, the following primers were used: for $m8^{D,stop}$, CTCGGTCATTTACAAGTTCTTG (nucleotides <<4229–4214 and <<3729–3719); for $m8^{D,bd-}$, CACTTGTTCATGTCGGCAAGTAGCTGGTCCACCAG-CATTGGC (nucleotides <<3446–3405); for $m8^{D,P\text{-}T}$, CCAGCATTGTCTTCTTCACC (nucleotides <<3415–3396); for $m8^{D,P\text{-}N}$, CGCTCCAGCATGTTCTTCTTCACCTTCTGG (nucleotides <<3419–3390). The same primers were used to create the corresponding mutations in the wild-type construct.

Germ-Line Transformation. Germ-line transformation experiments were carried out essentially as described (30), using the transformation vector pW8 (26). Transposase was supplied either by coinjection of the Δ2-3 helper plasmid (31) or by using *w/w; Δ2-3/Δ2-3* embryos as hosts (32). Trans-

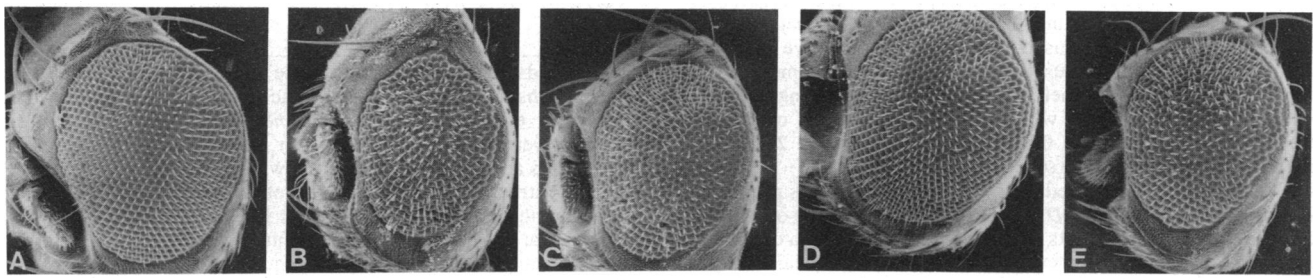

FIG. 2. Effects of different constructs on the eye phenotype of *spl*/+ females. (*A*) *y w spl*/+. Note the regular, wild-type patterning of the ommatidia. (*B*) *y w spl*/+; $E(spl)^D$/+. Although the *spl* mutation also affects bristle development, we analyzed only the effect of $E(spl)^D$ and the different transgenes on the eye phenotype. As the *spl* phenotype is variable and easily affected by modifying influences, we assayed the effects of our constructs in heterozygous females, which themselves display a wild-type, regularly spaced ommatidial pattern (*A*). The *spl* phenotype, however, becomes dominant in an $E(spl)^D$/+ background (*B*); this effect is referred to in the text as *spl* enhancement. (*C*) *y w spl*/+; $P[(w^+)m8^D$ *E6.1*$]$/+. The fragment encoding the $E(SPL)^D$ protein (see Fig. 3*b*) enhances the *spl* phenotype to nearly the same extent as the original mutation. (*D*) *y w spl*/+; $P[(w^+)m8^{+,stop}St25.1]$/+. This construct encodes a truncated $E(SPL)^+$ protein (see Fig. 3*f*), which does not lead to an enhancement of the *spl* phenotype when present in one copy. (*E*) *y w spl*/+; $P[(w^+)m8^{+,stop}St25.1]/P[(w^+)m8^{+,stop}St25.1]$. The same construct as in *D* results in a distinct enhancement when present in two copies, although the degree of enhancement is not as strong as with the $m8^D$ construct (see *C*).

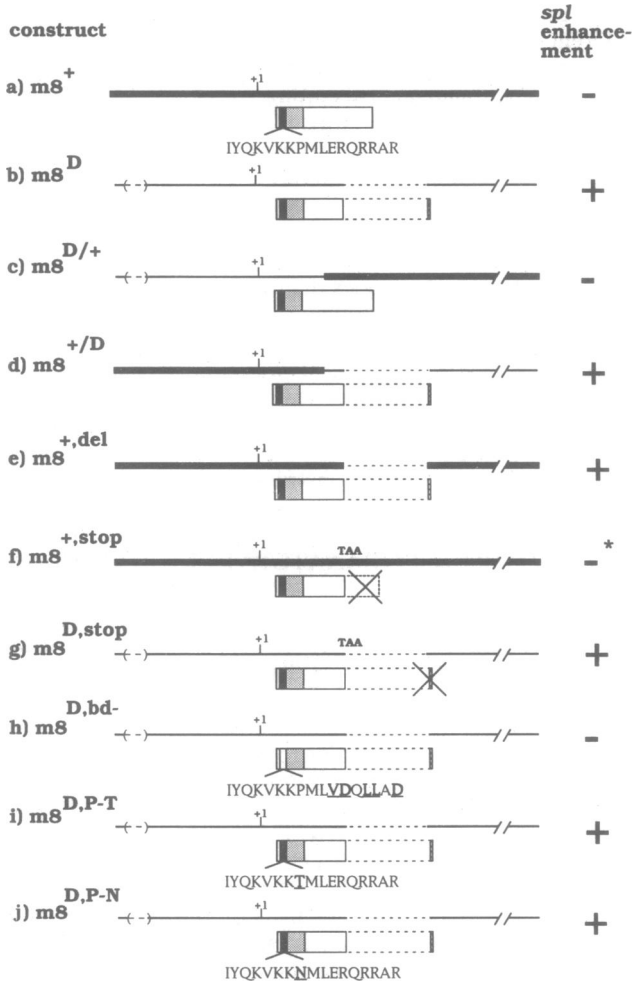

FIG. 3. Different $m8^{+}$ and $m8^{D}$ constructs and their effects on the *spl* phenotype. Shown are various *Cla* I fragments (see Fig. 1; not drawn to scale). Thick lines, wild-type DNA; thin lines, DNA from $E(spl)^{D}$; broken lines, deletions. +1, Putative transcription start site. Bars, translation products with the basic region (solid), the HLH domain (stippled), and the nine-amino acid sequence added at the C terminus of the $E(SPL)^{D}$ protein (hatched). Parts of the mutated proteins that were eliminated by the introduction of a termination codon are indicated (×). The amino acid sequence of the basic domains is as shown in *a* for the wild type, except in *h–j*, where it has been mutated (altered amino acids are in boldface and underlined). Assessment of the degree of *spl* enhancement. *E(spl)* is characterized by (*i*) smaller eyes and (*ii*) rough eyes. As in wild-type and *spl* eyes the number of ommatidia varies considerably (with a deviation of ≈15%), which makes it an unsuitable criterion for estimation of the degree of enhancement, we used the roughening of the eye as a more reliable measure. This is caused by an irregular arrangement of the ommatidia due to a loss of their hexagonal array. +, Strong enhancement of the *spl* phenotype, with no continuous lines of ommatidia left. For example, although the size of the eyes in Fig. 2 *B* and *C* differs, the roughening is the same. −, No enhancement; −*, distinct enhancement can be detected only when the construct is present in two copies (compare Fig. 2 *D* and *E*). In this case, some ommatidia are still aligned, which results in a weaker roughening. To determine the degree of *spl* enhancement, all transgenic lines obtained from each construct were evaluated (for numbers of lines tested, see *Materials and Methods*). In principle, the same result was obtained in each case. (*a* and *b*) *Cla* I fragments as in Fig. 1*A*, encoding a protein of 179 amino acids (*a*) and 132 amino acids (*b*), 1–123 from the wild type plus nine new amino acids (hatched box). (*c* and *d*) Recombinant fragments $m8^{D/+}$ and $m8^{+/D}$, encoding a wild-type (*c*) and a truncated protein with the nine additional amino acids (*d*) as shown in *b*. (*e*) For the construction of $m8^{+,del}$, two *Kpn* I sites were introduced into the $m8^{+}$ construct at the borders of the desired deletion, and, after removing the fragment between the two *Kpn* I sites, the remaining *Kpn* I site was eliminated. The C-terminal nine amino acids correspond to those of the $m8^{D}$ construct. In *g*, however, we demonstrate that these amino acids did not contribute to the dominant enhancement. (*f*) For $m8^{+,stop}$, different termination codons were used—TAG, TAA, or TAATAATAA. The protein represents amino acids 1–123 of the wild-type protein. (*g*) $m8^{D,stop}$. This construct encodes the same protein as in *f*. (*h*) $m8^{D,bd-}$. The protein encoded by this construct is the same as that shown in *b*, except that it contains a partially neutralized basic domain. Amino acid alterations were based on the sequences of either Id (33) or emc (34, 35), HLH proteins that do not contain a functional basic domain. (*i* and *j*) The proteins encoded by these constructs correspond to the $E(SPL)^{D}$ protein (*b*), except that the proline in the basic domain has been exchanged for either a threonine ($m8^{D,P-T}$; *i*) an amino acid found at this position in all members of the myogenic bHLH family (see ref. 8 for review), or for an asparagine ($m8^{D,P-N}$; *j*), an amino acid found at this position in proteins of the proneural genes of *Drosophila* (36–39), or the proteins of the *myc* family (see ref. 40 for review). For primers used to create the constructs shown in *e–j*, see *Materials and Methods*. The same primers as in *h–j* were used to create the corresponding mutations in the wild-type construct.

genic stocks were established over the appropriate balancers or were kept as homozygotes.

Numbers of lines obtained were as follows: $m9,10^{D}$, 2; $m8^{+}$, 4; $m8^{D}$, 1; [in addition to two lines previously described (4)]; $m8^{D/+}$, 3; $m8^{+/D}$, 5; $m8^{+,stop}$ (TAG, 6; TAA, 6; TAATAATAA, 10), $m8^{D,stop}$, 6; $m8^{+,del}$, 9; $m8^{+,bd-}$, 5; $m8^{+,P-T}$, 7; $m8^{+,P-N}$, 3; $m8^{D,bd-}$, 11; $m8^{D,P-T}$, 8; $m8^{D,P-N}$, 6.

Electrophoretic Mobility-Shift Assays. For construction of expression vectors, the coding regions of the different constructs shown in Fig. 3 were amplified by PCR, using primers at positions 3361–3379 and 3889–3907 for the wild-type protein and its derivatives and primers at positions 3361–3379 and 4221–4241 for the $E(SPL)^{D}$ protein (numbered according to ref. 4). The fragments were cloned into the *Bam*HI site of the pET-3a expression vector (41). Transformation, induction, and purification of proteins were performed according to ref. 42. The protein was purified by sequential detergent washes of the insoluble inclusion bodies. The final pellet was solubilized in a buffer containing 5 M urea, and, after dialysis, aliquots were analyzed by SDS/polyacrylamide gel electrophoresis. A double-stranded oligonucleotide (top strand, 5′-GATCACGCCACGAGCCACAAGGATTG-3′; bottom strand, 5′-GATCCAATCCTTGTGGCTCGTGGCGT-3′) was labeled at both ends by filling in with Klenow fragment in the presence of [α-^{32}P]dATP to a specific activity of 5×10^{6} cpm/μg. Protein–DNA complexes were formed by incubation of 0.5 μg of protein with 1.5 ng (7500 cpm) of labeled oligonucleotide in 10 μl of buffer [25 mM Hepes·KOH, pH 7.5/100 mM KCl/20% (vol/vol) glycerol/0.1% (vol/vol) Nonidet P-40/10 μM $ZnSO_4$/1 mM dithiothreitol] at 4°C for 10 min and separated by electrophoresis on a 5% polyacrylamide gel in 0.25× TBE buffer (1× TBE buffer = 90 mM Tris/64.6 mM boric acid/2.5 mM EDTA, pH 8.3).

RESULTS

We have previously shown that the ability to enhance the *spl* phenotype is associated with the mutant $E(spl)^{D}$ gene: transgenic flies carrying one copy of the mutant gene ($m8^{D}$) in addition to two wild-type copies of the gene show the same degree of enhancement of the *spl* phenotype as do animals bearing the original $E(spl)^{D}$ chromosome (ref. 4; see also Fig. 2 *B* and *C*). In contrast, transgenic flies containing one or two copies of either the corresponding wild-type construct ($m8^{+}$) or the mutant *gro* gene ($m9/m10^{D}$, see Fig. 1*A*) do not express the *spl* phenotype in a heterozygous *spl* background (data not shown). In addition, the $m9/m10^{D}$ transgene fails to rescue the lethality associated with a loss-of-function mutation of *gro* (gro^{B48}; ref. 25). In this respect, it differs from the corresponding wild-type $m9/m10^{+}$ fragment, which is capa-

ble of rescuing the embryonic lethality of this or other *gro* mutations or the lethality of hemizygous $E(spl)^D$ embryos (ref. 22; H. Schrons, E.K., and J. A. Campos-Ortega, unpublished data). Although we cannot exclude the presence of minor defects at the amino acid level not detectable by conventional Southern blot analysis, we suggest that it is the insertion present in the *gro* gene on the $E(spl)^D$ chromosome that leads to a partial loss of function of *gro*. Therefore, we conclude that the $E(spl)^D$ chromosome carries two mutations—one enhances the phenotype of *spl* and the other reduces the activity of *gro*.

As the $m8^D$ fragment used earlier carries several lesions (Fig. 1), we have used different hybrid constructs and mutations for germ-line transformation to localize the region responsible for the enhancement of *spl* to the distal part of the mutant *E(spl)* gene (Fig. 3 *a–d*). This region carries a deletion of 483 base pairs, which removes 3′ noncoding DNA as well as a sequence encoding the last 56 amino acids of the wild-type product, which are replaced by a sequence of 9 amino acids (4) (Fig. 1*B*). Hence, the mutant phenotype may be due to (*i*) truncation of the protein *per se*, (*ii*) the additional amino acid sequence, (*iii*) deletion of the 3′ region of the transcription unit, (*iv*) other changes not associated with the deletion, or (*v*) any combination of these effects. To exclude the possibility that other minor sequence differences detected in the $m8^D$ construct (4) might be responsible for this effect, we reconstructed the same deletion in an $m8^+$ fragment ($m8^{+,del}$), which produced the same protein as $m8^D$, and we obtained the same result as with $m8^D$ (Fig. 3*e*). To distinguish between the other possibilities, we carried out a set of manipulations on the $m8^+$ and the $m8^D$ constructs and analyzed their effects on the *spl* phenotype in transgenic females (Fig. 3 *e–g*). Strikingly, although the introduction of a termination codon at the corresponding position in the wild-type and mutant genes ($m8^{+,stop}$ and $m8^{D,stop}$; Fig. 3 *f* and *g*) should result in production of the same truncated protein, these two constructs differ significantly in the degree to which they enhance *spl*. Whereas one copy of the $m8^{D,stop}$ construct is sufficient to render *spl* dominant and thus behaves like the $m8^D$ construct (data not shown), the $m8^{+,stop}$ construct gives rise to dominant expression of *spl* only when present in two copies (Fig. 2 *D* and *E*). Thus the truncation of the E(SPL) protein is not sufficient to provoke the same strong enhancement as observed with $m8^D$. As the only difference between the $m8^{+,stop}$ and $m8^{D,stop}$ constructs resides at the nucleic acid level (see Fig. 3 *f* and *g*), we conclude that the DNA, which is deleted in the mutant, carries the responsible element and we assume that the strong dominant enhancement of *spl* is caused by the loss of a regulatory sequence in the mutant DNA in combination with a truncation of the protein. This regulatory element could act on either the transcriptional or the posttranscriptional level—e.g., by changing the stability of the RNA—but its precise definition as well as analysis of the molecular basis of its function await further experiments.

The high degree of amino acid sequence similarity between the E(SPL) protein and other members of the bHLH class, which bind via the basic domain to DNA and can activate transcription (14), led us to assume that the wild-type and mutant versions of this protein are also involved in DNA binding. However, the E(SPL) protein as well as all other members of the E(SPL)–C complex contain a proline residue in the basic domain (4) at a position where most other bHLH proteins, except the *Drosophila* protein hairy (H) (43), contain either a threonine or an asparagine residue. Replacement of an alanine adjacent to this threonine in the MyoD protein by a proline has been shown to abolish DNA binding to its target sequence (15). To address the question of the role of the basic domain of the $E(SPL)^D$ protein, we introduced various amino acid exchanges (see *Material and Methods*; Fig. 3 *h–j*) and tested their effects on the *spl* phenotype after germ-line transformation. In addition, we directly assayed bacterially expressed wild-type and $E(SPL)^D$ protein and some modified versions of the wild-type protein for their ability to bind specifically to DNA *in vitro*. By band-shift assays, we could demonstrate that both the wild-type and the $E(SPL)^D$ proteins specifically bind to a sequence in the *E(spl)* promoter *in vitro*. Furthermore, DNA binding is totally abolished after neutralization of part of the basic domain (Fig. 4). *In vivo*, neutralization of part of the basic domain of the $E(SPL)^D$ protein ($m8^{D,bd-}$; Fig. 3*h*) leads to complete loss of *spl*-enhancing capacity, suggesting loss of an essential function of the protein. Replacement of the proline by a threonine ($m8^{D,P\text{-}T}$; Fig. 3*i*) or asparagine residue ($m8^{D,P\text{-}N}$; Fig. 3*j*), the two residues that are found at the equivalent position in the basic domains of most of the other bHLH proteins, results in only minor modification of the strong *spl* enhancement: transgenic flies carrying the $m8^{D,P\text{-}N}$ construct exhibit a slightly stronger enhancement, judged by more severe irregularities of the ommatidia, those with the $m8^{D,P\text{-}T}$ construct have a slightly reduced enhancement compared to those with the $m8^D$ construct itself (for assessment of the *spl* phenotype; see legend to Fig. 3). Proteins carrying these exchanges [$E(SPL)^{+,P\text{-}N}$ and $E(SPL)^{+,P\text{-}T}$] are still able to bind to the specific target sequence *in vitro* (Fig. 4). These results clearly demonstrate that the basic domains of the $E(SPL)^+$ and $E(SPL)^D$ proteins are functionally important and bind to DNA, despite the presence of a proline residue, and that neutralization of part of this domain not only abolishes DNA binding *in vitro* but also leads to loss of *in vivo* function of the $E(SPL)^D$ protein, as shown by its inability to enhance the *spl* phenotype.

DISCUSSION

The bHLH proteins of the E(SPL)–C differ from most of the other members of this family by the presence of a proline residue in the basic domain. This has led to speculation as to whether this domain can function in DNA binding at all, since a proline residue can confer a kink in α-helical structures (44). In fact, two HLH proteins without a basic domain have been

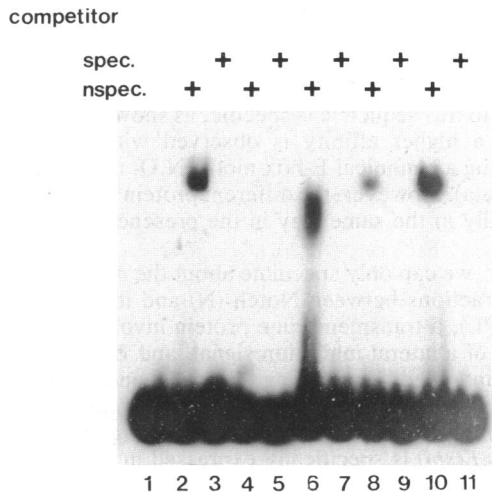

FIG. 4. Mobility-shift assay with different versions of the E(SPL) protein. The ^{32}P-labeled oligonucleotide, derived from the *E(spl)* promoter (4), was incubated without protein (lane 1) or with proteins, expressed in *Escherichia coli* and solubilized, renatured, and purified from inclusion bodies. Lanes: 2 and 3, $E(SPL)^+$; 4 and 5, $E(SPL)^{+,bd-}$; 6 and 7, $E(SPL)^D$; 8 and 9, $E(SPL)^{+,P\text{-}N}$; 10 and 11, $E(SPL)^{+,P\text{-}T}$ (designations according to those in Fig. 3). Reactions were carried out with the same amount of protein in the presence of either 2 μg of unspecific competitor [poly(dI-dC)] (lanes 2, 4, 6, 8, and 10) or 0.5 μg of specific competitor (lanes 3, 5, 7, 9, and 11).

described—Id (33) and the protein encoded by the *Drosophila* gene *emc* (34, 35), which are not capable of DNA binding but act as negative regulators by forming heterodimers with other bHLH proteins, thus preventing them from binding DNA (33, 45). In the case of the bHLH proteins of the E(SPL)–C, analysis of structure–function relationships is hampered by the fact that, due to the redundancy of the functions encoded in this gene complex, no point mutations, apart from the dominant $E(spl)^D$ allele, have ever been found whose effects would have been indicative of functional domains within the proteins. To overcome this problem, we designed mutations in the *E(spl)* gene and analyzed their effects on the dominant phenotype. Our results clearly demonstrate the importance of the basic domain of the $E(SPL)^D$ protein for proper function, since $m8^D$-induced enhancement of the *spl* phenotype and DNA binding are abolished after neutralization of the basic domain. Recent results have shown that this mutant version antagonizes the *in vitro* DNA-binding activity of wild-type E(SPL) and M5, another bHLH protein of the E(SPL)–C (N.O. and E.K., unpublished data). This indicates that it acts as negative regulator in the same way as the *Drosophila* EMC or the mouse Id protein (33, 45)—i.e., by formation of heterodimers that are no longer able to bind to DNA, probably because two intact basic domains are required for this process.

Although the *in vivo* assay (enhancement of the *spl* phenotype) could only be performed with a truncated version of the E(SPL) protein, we assume that similar conclusions about the functional importance of the basic domain *in vivo* can be applied to the wild-type protein as well. This is particularly supported by the observation that the wild-type and truncated proteins behave similarly *in vitro*: both bind to the same target sequence (cf. Fig. 4). In none of the transgenic stocks carrying a mutated version of either $m8^D$ or $m8^+$ did we observe any obvious mutant phenotype apart from the enhancement of *spl*. However, further functional assays are required to determine whether any of the mutant E(SPL) proteins acts as a dominant negative regulator *in vivo*, as has been described for a mutant MyoD protein, in which the basic domain has been deleted (15).

Strikingly, the binding site of the different forms of the E(SPL) protein tested here contains not the consensus CANNTG (E-box) present in the target sites of all other bHLH proteins described so far but a tandem array of two slightly divergent motifs (CACGAGCCACAAG). Although binding to this sequence is specific, as shown by footprinting assays, a higher affinity is observed with a binding site containing a canonical E-box motif (N.O. and E.K., unpublished data). However, the different protein versions behaved essentially in the same way in the presence of either target site.

So far, we can only speculate about the molecular basis of the interactions between Notch (N) and its mutant version split (SPL), a transmembrane protein involved in the transduction of a lateral inhibition signal, and E(SPL), a nuclear protein involved in mediating this signal and reprogramming the cell's fate, probably by transcriptional regulation. During eye development, *N* is expressed throughout all stages, whereas *E(spl)* is specifically expressed in the morphogenic furrow (19), a region of the eye imaginal disc in which the founder cells of the ommatidial clusters are determined. Actually, nothing is known about changes induced in the N protein in response to the signal, which could include structural changes of the cytoplasmic domain or the initiation of a signal transduction cascade, which, in turn, could modulate the activity of the E(SPL) protein in one way or the other. For example, one can imagine that phosphorylation or structural modification of the E(SPL) protein may induce changes in the choice of the dimerization partner or binding specificity and thereby alter the transcription program.

We are indebted to J. A. Campos-Ortega for many helpful discussions throughout the work and thank him and P. Hardy for constructive criticisms on the manuscript, F. Grawe for expert technical assistance, and F. W. Studier for the gift of pET expression vectors. This work was supported by a grant from the Deutsche Forschungsgemeinschaft (SFB 243).

1. Campos-Ortega, J. A. (1991) *Comments Dev. Neurobiol.* **1**, 287–310.
2. Vässin, H., Vielmetter, J. & Campos-Ortega, J. A. (1985) *J. Neurosci.* **2**, 291–308.
3. de la Concha, A., Dietrich, U., Weigel, D. & Campos-Ortega, J. A. (1988) *Genetics* **118**, 499–508.
4. Klämbt, C., Knust, E., Tietze, K. & Campos-Ortega, J. A. (1989) *EMBO J.* **8**, 203–210.
5. Mellor, J., Jiang, W., Funk, M., Rathjen, J., Barnes, C. A., Hinz, T., Hegemann, J. H. & Philippsen, P. (1990) *EMBO J.* **9**, 4017–4026.
6. Cai, M. & Davis, R. W. (1990) *Cell* **61**, 437–446.
7. Ludwig, S. R. & Wessler, S. R. (1990) *Cell* **62**, 849–851.
8. Olson, E. N. (1990) *Genes Dev.* **4**, 1454–1461.
9. Garrell, J. & Campuzano, S. (1991) *BioEssays* **13**, 493–498.
10. Murre, C., McCaw, P. S. & Baltimore, D. (1989) *Cell* **56**, 777–783.
11. Weintraub, H., Tapscott, S. J., Davis, R. L., Thayer, M. J., Adam, M. A., Lassar, A. B. & Miller, A. D. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 5434–5438.
12. Gregor, P. D., Sawadogo, M. & Roeder, R. G. (1990) *Genes Dev.* **4**, 1730–1740.
13. Hu, Y. F., Lascher, B., Admon, A., Mermod, N. & Tjian, R. (1990) *Genes Dev.* **4**, 1741–1752.
14. Murre, C., McCaw, P. S., Vässin, 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. Davis, R. L., Cheng, P. F., Lassar, A. B. & Weintraub, H. (1990) *Cell* **60**, 733–746.
16. Welshons, W. J. (1965) *Science* **150**, 1122–1129.
17. Knust, E., Bremer, K. A., Vässin, H., Ziemer, A., Tepass, U. & Campos-Ortega, J. A. (1987) *Dev. Biol.* **122**, 262–273.
18. Shepard, S. B., Broverman, S. A. & Muskavitch, M. A. T. (1989) *Genetics* **122**, 429–438.
19. Campos-Ortega, J. A. & Knust, E. (1990) *Roux's Arch. Dev. Biol.* **198**, 286–294.
20. Artavanis-Tsakonas, S. & Simpson, P. (1991) *Trends Genet.* **7**, 403–408.
21. Knust, E., Tietze, K. & Campos-Ortega, J. A. (1987) *EMBO J.* **6**, 4113–4123.
22. Preiss, A., Hartley, D. A. & Artavanis-Tsakonas, S. (1988) *EMBO J.* **7**, 3917–3928.
23. Lindsley, D. L. & Zimm, G. (1985) *Drosoph. Inf. Serv.* **62**.
24. Lindsley, D. L. & Zimm, G. (1990) *Drosoph. Inf. Serv.* **68**.
25. Ziemer, A., Tietze, K., Knust, E. & Campos-Ortega, J. A. (1988) *Genetics* **119**, 63–74.
26. Klemenz, R., Weber, U. & Gehring, W. J. (1987) *Nucleic Acids Res.* **15**, 3947–3959.
27. Tietze, K. (1987) Thesis (University of Cologne).
28. Sayers, J. R., Schmidt, W. & Eckstein, F. (1988) *Nucleic Acids Res.* **16**, 791–802.
29. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 5463–5467.
30. Spradling, A. C. (1986) in *Drosophila: A Practical Approach*, ed. Roberts, D. B. (IRL, Oxford), pp. 175–197.
31. Laski, F. A., Rio, D. C. & Rubin, G. M. (1986) *Cell* **44**, 7–19.
32. Robertson, H. M., Preston, C. R., Phillis, R. W., Johnson-Schlitz, D., Benz, W. K. & Engels, W. R. (1988) *Genetics* **118**, 461–470.
33. Benezra, R., Davis, R. L., Lockshon, D., Turner, D. L. & Weintraub, H. (1990) *Cell* **61**, 49–59.
34. Ellis, H. M., Spann, D. R. & Posakony, J. W. (1990) *Cell* **61**, 27–38.
35. Garrell, J. & Modolell, J. (1990) *Cell* **61**, 39–48.
36. Villares, R. & Cabrera, C. V. (1987) *Cell* **50**, 415–424.
37. Alonso, M. C. & Cabrera, C. V. (1988) *EMBO J.* **7**, 2585–2591.
38. Caudy, M., Vässin, H., Brand, M., Tuma, R., Jan, L. Y. & Jan, Y. N. (1988) *Cell* **55**, 1061–1067.
39. González, F., Romani, S., Cubas, P., Modolell, J. & Campuzano, S. (1989) *EMBO J.* **8**, 3553–3562.
40. Lüscher, B. & Eisenmann, R. N. (1990) *Genes Dev.* **4**, 2025–2035.
41. Studier, F. W. & Moffatt, B. A. (1986) *J. Mol. Biol.* **189**, 113–130.
42. Kerkhoff, E. & Bister, K. (1991) *Oncogene* **6**, 93–102.
43. Rushlow, C. A., Hogan, A., Pinchin, S. M., Howe, K. M., Lardelli, M. & Ish-Horowicz, D. (1989) *EMBO J.* **8**, 3095–3103.
44. Yun, R. H., Anderson, A. & Hermans, J. (1991) *Proteins* **10**, 219–228.
45. van Doren, M., Ellis, H. M. & Posakony, J. W. (1991) *Development* **113**, 245–255.