A Rat Homolog of the Drosophila Enhancer of Split (groucho) Locus Lacking WD-40 Repeats*

(Received for publication, June 8, 1993)

Carl J. Schmidt and Theresa E. Sladek
From the Department of Medicine, Cardiology Division, Brigham and Women's Hospital, Boston, Massachusetts 02115

In Drosophila, neurogenic loci function in defining cellular fate by interpreting the identity of other cells in the immediate environment. To begin studies of mammalian homologs of these genes, we have isolated two rat homologs of the neurogenic locus Enhancer of split. The protein encoded by the Drosophila Enhancer of split locus is complex and contains five distinct regions based on amino acid composition. One region contains six WD-40 repeats, which were first described in the β subunit of the heterotrimeric guanine nucleotide-binding protein. One of the rat cDNAs we isolated, R-espl, encodes a novel form that lacks the WD-40 repeating units. Data are presented demonstrating that the R-espl cDNA is a full-length clone encoding an expressed 24-kDa protein. Antibodies raised against this protein stain the nucleus of both PC-12 and GH3 cells. The second clone, R-esp2, encodes a full-length homolog containing WD-40 repeats. The hydrodynamic properties of in vitro translated R-espl and R-esp2 proteins indicate that they do not stably self-associate or form heterodimers. A model is presented for the possible role of the R-esp1 protein in the negative regulation of Enhancer of split proteins containing WD-40 repeats.

The fate of cells during development can be affected by the identity of surrounding cells. In the fly, neurogenic gene products play important roles in defining cellular fate by interpreting the identity of other cells in the immediate environment (for review, see Artavanis-Tsakonas et al., 1991). Mutations in each of these loci cause most ventral ectodermal cells to become neuroblasts with little or no formation of epidermis. Hence, the phenotype of mutants is one of neuronal hypertrophy and epidermal hypotrophy. The neuronal hypertrophy does not arise from unregulated growth of the neuroblasts. Instead, cells normally destined to become epidermis are diverted to become neuroblasts. The neurogenic genes not only affect fate decisions in ectodermal cells but also affect mesodermal and endodermal development (Corbin et al., 1991; Hartenstein et al., 1992). At least six different neurogenic genes have been defined by mutations that affect Drosophila embryonic development. These include: Enhancer of split, Notch, Delta, mastermind, big brain and neuralized. Classical genetic approaches in Drosophila were used to identify this group of genes but such studies are difficult to conduct in mammals. The cloning of mammalian homologs of Drosophila genes involved in determination of cell fate may allow identification of mammalian genes that play a similar role.

We have used this approach to study vertebrate homologs of the neurogenic class of Drosophila genes. This paper focuses on identification of mammalian homologs of the Drosophila neurogenic gene Enhancer of split (E(sp1)). In Drosophila the E(sp1) locus is complicated, giving rise to two distinct classes of proteins which are the product of multiple genes (Hartley et al., 1988; Klambt et al., 1989). The E(sp1)-HLH class encodes proteins containing the helix-loop-helix motif characteristic of gene regulatory proteins. The E(sp1)-WD class contains m3/10 groucho, encodes a protein that contains a series of six WD-40 repeating units. The WD-40 repeat typically ends in the sequence tryptophan (W)-aspartate (D), and the consensus repeat is approximately 40 amino acids long (for review, see Duronio et al., 1992). This repeating unit was first recognized in the guanine nucleotide-binding protein β subunit, but it has been found in at least 20 other regulatory proteins, including the transcription regulatory factor TUP1 (Williams and Trumbly, 1990), the splicing factor PRP4 (Bjorn et al., 1989), and cleavage stimulation factor which modulates polyadenylation (Takagaki and Manley, 1992). Although the function of the WD-40 repeat is still uncertain, the diversity of proteins that contain this motif suggest that its presence does not necessarily confer participation in signal transduction pathways mediated by G-proteins. In addition to the conserved WD-40 repeating units, full-length E(sp1)-WD proteins contain two other conserved regions. The first spans the amino-terminal 120 amino acids and is rich in glutamine. The second spans the central region of the protein, contains a nuclear localization signal, and is rich in proline, serine, and glycine.

We have isolated two different cDNAs encoding rat E(sp1)-WD homologs which we term R-espl and R-esp2. Both clones show great similarity to both the Drosophila E(sp1)-WD gene product (Hartley et al., 1988) and the four human E(sp1)-WD products (Stifani et al., 1992). However, R-espl encodes a novel member of this gene family lacking the WD-40 repeating units. No equivalent to R-espl has been described in Drosophila. R-esp2 encodes a full-length E(sp1)-WD homolog containing all three domains.

MATERIALS AND METHODS

Cloning Rat E(sp1)-WD Homologs—A rat hippocampal cDNA library in Agt-10 (Diyao Zhao and William Chin, Harvard Medical School) was screened with an Ssp1-Xba1 fragment isolated from the Drosophila m8/10 cDNA (C. Delidakis and S. Artavanis-Tsakonas, Yale University) using standard procedures (Maniatis et al., 1982). The fragment used as probe contains 250 nucleotides of S'-noncoding and the first 180 codons of the Drosophila E(sp1)-WD cDNA. Filters were hybridized and washed under reduced stringency conditions as described in Schmidt et al. (1989).

DNA Sequencing and Analysis—DNA sequence was determined using the dideoxy chain termination technique (Sanger et al., 1977) and specific DNA primers. Computer analysis was done on a Vax using the
Vogelstein (1983). Hybridization was carried out at 45 °C in a solution gation at 14,000 rpm in a microcentrifuge, and the pellet was then lysed in EBC buffer (50
washed once in IB  followed  by  two washes in 25
followed  by three washes with blocking  buffer. Fluoroscein-conjugated protein A-Sepharose was added to the  supernatant followed  by incuba-
phy.

Stifani et al. (1992) reported the sequence of four human homologs of the E(sp)-WD locus which they term transducin-
like enhancer of split (TLE1–4). Comparison of the human and Drosophila sequences defined five different regions based on amino acid composition: an amino-terminal region rich in glutamine (Q); a glycine/proline-rich region (GP); a Cn region which contains a nuclear localization signal along with potential casein kinase II and cdc2 phosphorylation sites; a proline-
serine (PS)–rich region; and the six WD-40 repeats which make up the carboxyl-terminal region of the E(sp)-WD homologs (Fig. 3). R-espl contains only the amino-terminal glutamine and glycine/proline-rich regions. R-espl contains five of these regions.

The five regions defined above were based upon amino acid composition. Another way to compare and classify proteins is to

Immunohistochemical Staining—Cells were grown on laminin-coated glass coverslips. Cells were fixed in 4% paraformaldehyde precipitated at 4 °C for 15 min and then the lysis was scraped into a 50-ml Falcon tube. Debris was removed by centrifugation at 2000 x g for 20 min, and the supernatant was concentrated to 2 ml by centrifugation through a Centricon 30 microcentrifuge. For immunoprecipitation, 450 μl of IB (25 μm Tris, pH 7.8, 150 μm NaCl, 4 μm EDTA, 0.1% SDS, 0.5% deoxycholate, 1% Nonidet P-40 was added to 50 μl of the concentrated supernatant, and the sample was incubated at 95 °C for 2 min. After cooling, antibody and a 50% slurry of Sepharose beads suspended in IB was added, and the sample was mixed for 2 h at room temperature. The Sepharose was removed by centrifugation, and protein A-Sepharose was added to the supernatant followed by incubation overnight at 4 °C. Precipitated material was pelleted by centrifuga-

Stokes radii of the autoradiogram, and Stokes radii were determined by comparison with the elution positions of the added markers of

The abbreviations used are: PBS, phosphate-buffered saline; NGF, nerve growth factor.

RESULTS

Rat E(sp)-WD Homologs—We have used the cDNA encoding the Drosophila E(sp)-WD locus to identify rat homologs from an adult hippocampal cDNA library. Two of these cDNAs have been sequenced revealing that one clone, R-espl, encodes a novel member of the E(sp)-WD family, whereas the other, R-
esp2, encodes a protein that is highly similar to the Drosophila E(sp)-WD gene product. A sequence comparison of R-espl and R-esp2 with the Drosophila E(sp)-WD protein is shown in Fig. 1.

The Drosophila E(sp)-WD protein is 720 amino acids long and the six WD repeats are found in the carboxyl-terminal region of the protein. The first clone, Resp-1, encodes a protein of only 197 amino acids which is similar to the amino terminus of the Drosophila E(sp)-WD protein. Data presented below indicates that R-espl is a full-length cDNA clone. The second clone, R-esp2, encodes a 741-amino acid protein, which, like Drosophila E(sp)-WD, contains six WD-40 repeats at the carboxyl terminus (Fig. 2).

Linear sucrose density gradients were prepared in TED 14,000 rpm in a microcentrifuge, and the pellet was then

Immunoblot Analysis—Total RNA was isolated from tissue by gua-

Physical Properties of R-esp Proteins—Stokes radii and sedimentation coefficients were determined as described in Schmidt and Neer (1991). Stokes radii were determined by gel filtration over either Ultragel AAc 34 or Aca 22 (IBF Inc.) equilibrated with 50 μl Tris-HCl, pH 7.5, 1 mm EDTA, 1 mm diethiothreitol (TED) and 0.9% Lubrol PX. Markers included β-galactosidase (Stokes radius = 69 Å), fumarase (Stokes radius = 52 Å), bovine serum albumin (Stokes radius = 37 Å), carbonic anhydrase (Stokes radius = 24 Å), and cytochrome c (Stokes radius = 17 Å). Fractions from the column that contained in vitro translated R-esp proteins were identified by SDS-gel electrophoresis followed by autoradiography. Elution profiles were analyzed by scanning densitometry of the autoradiogram, and Stokes radii were determined by comparison with the elution positions of the added markers of known Stokes radius. Linear sucrose density gradients were prepared in TED + 0.3% Lubrol PX. Gradients were centrifuged at 95,000 rpm for 6 h in a Beckman SW 60 rotor at 4 °C. The elution of R-esp was determined as described above for Stokes radius determination. The markers used were β-galactosidase (ε280 = 15.99), bovine serum albumin (ε280 = 4.2), carbonic anhydrase (ε280 = 2.9), and cytochrome c (ε280 = 1.8).

The abbreviations used are: PBS, phosphate-buffered saline; NGF, nerve growth factor.
identify regions where the amino acid sequences have been conserved. This is particularly important when comparing sequences that contain regions rich in particular amino acids, such as glutamine, proline, glycine, or serine. For example, it is possible for two proteins to be similar merely because they both contain a high glutamine content. However, if the two proteins are homologs, residues other than glutamine would also be conserved. The PILEUP and PLOTSIMILARITY programs of the Genetics Computer Group sequence analysis package (Devereux et al., 1987) allows a graphical presentation of the sequence identity between R-esp1, R-esp2, D-esp, and the human E(spl)-WD homologs, TLE 1–4 (Fig. 3). This analysis indicates that these proteins contain three discrete domains based upon the extent of amino acid identity. Domain I extends from the amino terminus to approximately amino acid 120 and corresponds to the Q-rich region. The level of conservation drops in domain II which corresponds to amino acids 120–450 and spans the GP, CCN, and PS regions. Inspection reveals that the high proline, glycine, and serine content is best conserved in domain II. Many other amino acids have been substituted within this domain. Domain III corresponds to the region containing the six WD-40 repeating units. The highest degree of amino acid identity between these proteins is within domains I and III, suggesting that these two may play an important role in E(spl)-WD structure or function.

FIG. 2. WD-40 repeating units. The WD-40 repeating units of R-esp2 and Drosophila E(spl)-WD (D-esp) are aligned to maximize amino acid identity between the different repeats. The consensus indicates the predominant amino acid present at that position.

FIG. 3. Conservation of E(spl)-WD homologs. The top shows a stick diagram of the R-esp clones indicating the conserved amino acid regions defined by Stifani et al. (1992). Q = glutamine-rich; GP = glycine/proline-rich region; CCN = nuclear localization signal, casein kinase I alpha kinase site; PS = proline/serine-rich region; and WD-40, which contains a series of six WD = 40 repeating units. The plot shows the identity score assigned for a comparison of the available rat, human, and Drosophila E(spl)-WD proteins using a sliding window of 10 amino acids. Three domains, I, II, and III, can be recognized based upon the extent of conservation. Both domains I and III contain regions that are conserved perfectly (score = 1) between the rat, human, and Drosophila sequences. Domain II is less well conserved.

Inspection of domain I reveals that the linear sequence of this region, including those amino acids that are not glutamine, are well conserved (Fig. 1, underlined region). This is important to the classification of R-esp1 as a member of the E(spl)-WD family, even though it lacks WD-40 repeats. This classification would not be justified if R-esp1 was related to other E(spl)-WD proteins solely on the basis of high glutamine content.

Only one E(spl)-WD gene has been identified in Drosophila, whereas mammalian species appear to contain multiple homologs. Comparison of the amino acid and nucleotide sequences (Tables I and II) of the rat, Drosophila, and human sequences reveals that the mammalian sequences are more similar to one another than to the Drosophila E(spl)-WD protein. These observations indicate an amplification of the E(spl)-WD loci occurred after the divergence of the arthropods and chordates. At the nucleotide level (Table II) R-esp1 is most similar to an expressed sequence tag (EST00256) which was identified by randomly selecting and sequencing clones from a human cDNA library (Adams et al., 1991). In addition, R-esp1 shows greater similarity to the amino-terminal region of the rat and human E(spl)-WD homologs than to the Drosophila E(spl)-WD gene product. Apparently, the R-esp1 gene also arose after the divergence of arthropods from chordates, suggesting that an E(spl) gene exactly homologous to R-esp1 may not exist in Drosophila. Given the similarity of R-esp1 with domain I of the E(spl)-WD gene family, it is likely that R-esp1 arose in chordates by mutation of a full-length E(spl)-WD gene. R-esp2 shows 96% amino acid identity with the human E(spl)-WD homolog termed TLE4 (Table I). This human cDNA is only a partial clone, so it is unclear if this great similarity extends throughout the entire protein.

Northern Blot Analysis of R-esp1—To determine if R-esp1 is expressed in adult tissues, total RNA was extracted from heart, brain, lung, and kidney, electrophoresed through a formaldehyde-agarose gel, and transferred to nitrocellulose filters. The filters were hybridized with [32P]dCTP-labeled R-esp1 and then washed at high stringency. R-esp1 anneals to a 1.3-kilobase mRNA that is expressed in heart, brain, lung, and kidney (Fig. 4). The length of the mRNA is similar to that of the cDNA (1350 base pairs), suggesting that the cDNA clone is probably full length. This supports the idea that R-esp1 is a unique clone rather than a truncation artifact of a larger cDNA. The level of mRNAs varied among tissues. Typically, the R-esp1 mRNA
program of the Wisconsin Genetics Computer Group sequence analysis software package (Devereux et al., 1984).

Amino acid sequences were aligned using the Pileup program and then the amount of amino acid identity was determined using the Distances program of the Wisconsin Genetics Computer Group sequence analysis software package. References are given in Table I.

| R-esp1 | R-esp2 | EST^a | TLE1^b | TLE2^b | TLE3^b | TLE4^b | D-esp^d |
|--------|--------|-------|--------|--------|--------|--------|--------|
| R-esp1 | 1.00   | 0.87  | 0.83   | 0.85   | 0.85   | 0.85   | 0.72   |
| R-esp2 | 1.00   | 0.80  | 0.87   | 0.74   | 0.85   | 0.98   | 0.67   |
| EST    | 1.00   | 0.74  | 0.74   | 0.74   | 0.74   | 0.66   |
| TLE1   | 1.00   | 0.75  | 0.76   | 0.86   | 0.84   | 0.71   |
| TLE2   | 1.00   | 0.76  | 0.82   | 0.74   | 0.66   |
| TLE3   | 1.00   | 1.00  | 0.90   | 0.71   |
| TLE4   | 1.00   | 0.76  |
| D-esp  |       |       |

^a EST = EST00256 expressed sequence tag (Adams et al. 1991).
^b TLE1-4 = human E(spl)-WD homologs (Stifani et al. 1992).
^c TLE4 is a partial cDNA clone which spans the coding region from the prolindserine-rich region through the six WD-40 repeating units.
^d D-esp = Drosophila E(spl)-WD (Harley et al. 1988).

TABLE II
Comparison of the nucleotide sequence of E(spl)-WD homologs
Sequences were aligned and then the extent of nucleotide identity was determined using the Distances program of the Wisconsin genetics computer group sequence analysis software package. References are given in Table I.

| R-esp1 | R-esp2 | EST^a | TLE1 | TLE2 | TLE3 | D-esp |
|--------|--------|-------|------|------|------|-------|
| R-esp1 | 1.00   | 0.76  | 0.80 | 0.77 | 0.73 | 0.74  | 0.69  |
| R-esp2 | 1.00   | 0.73  | 0.77 | 0.73 | 0.71 | 0.72  | 0.66  |
| EST*   | 1.00   | 0.71  | 0.71 | 0.71 | 0.66 |
| TLE1   | 1.00   | 0.75  | 0.77 | 0.70 |
| TLE2   | 1.00   | 0.76  | 0.75 |
| TLE3   | 1.00   | 0.76  | 0.75 |
| D-esp  |       |       |

^a EST = EST00256 expressed sequence tag.

Fig. 4. Northern blot analysis of R-esp1 in adult tissue. Total RNA was separated by formaldehyde gel electrophoresis, transferred to nitrocellulose, and probed with [32P]dCTP labeled R-esp1. K = kidney, L = lung, B = brain, H = heart.

The level was 3-fold higher in kidney than lung and 2-fold higher in kidney than heart or brain.

Immunoprecipitation—Antipeptide antibodies have been raised against R-esp1 and have been used to show that the R-esp1 cDNA encodes an expressed protein of the size predicted from the nucleotide sequence. The R-esp1 peptide, spanning amino acids 12–26, was chosen from a region that has only 7 out of 14 residues identical to R-esp2. The R-esp1 antisera have been tested for their ability to immunoprecipitate protein from either in vitro translated R-esp1 or from proteins extracted from PC-12 cells that were metabolically labeled with [35S]methionine (Fig. 5). In vitro translation of R-esp1 yields a 24-kDa protein that is immunoprecipitated by the anti-R-esp1 antisera but not by preimmune sera. In addition, this antisera does not immunoprecipitate in vitro translated R-esp2 (not shown). One protein is specifically precipitated from the [35S]methionine-labeled PC-12 cell extract and comigrates with the in vitro translated R-esp1. Precipitation of this 24-kDa protein was blocked by the peptide used for immunization and it was not precipitated by preimmune serum. The data show that R-esp1 encodes an expressed protein of the size predicted from the cDNA.

Immunohistochemical Localization—The full-length Drosophila and the human E(spl)-WD homologs contain nuclear localization signals and are found in the nucleus (Hartley et al., 1989; Stifani et al., 1992). However, R-esp1 lacks a recognizable nuclear localization signal, making it difficult to predict the subcellular location of this protein. To determine where R-esp1 protein is found in the cell, the anti-R-esp1 antibodies have been used for immunohistochemical localization in two different rat cell lines: PC-12, a pheochromocytoma line, and GH3, a line established from an anterior pituitary tumor. These cell lines were chosen because they are both derived from ectoderm, a germ layer which is clearly affected by mutations in the Drosophila E(spl)-WD gene. Furthermore, PC-12 cells can be induced to differentiate into neuronal-type cells with nerve growth factor (NGF). Immunohistochemical staining was visualized using a fluorescein-conjugated goat anti-rabbit IgG as a secondary antibody. As seen in Fig. 6, anti-R-esp1 stains the nucleus of these cells. This staining is absent when antisera is first incubated with peptide used to raise the antibody (Fig. 6, control), when primary antibody was omitted or when preimmune sera was used (not shown). The staining pattern within the nuclei of the PC-12 cells had a reticulated or sponge-like appearance. Optical sectioning by confocal microscopy indicates that this pattern extends throughout the nucleus. NGF-treated PC-12 cells that had extended neurite processes have a more diffuse staining appearance, although regions lacking
Fig. 5. Immunoprecipitation of R-esp1. Samples were precipitated with protein A-Sepharose and then applied to an SDS-polyacrylamide gel and the results determined by autoradiography. Lane 1, in vitro translated [35S]methionine-labeled R-esp1. R-esp1 is indicated with an arrowhead. This sample was not immunoprecipitated. Lanes 2-4, protein samples were extracted from [35S]methionine-labeled PC-12 cells and then precipitated as follows: lane 2, preimmune serum; lane 3, anti-R-esp1 immune serum; lane 4, anti-R-esp1 immune serum that was preincubated with peptide that was used as antigen. Lanes 5-7, in vitro translated [35S]methionine-labeled R-esp1. Lane 5, preimmune serum. Lane 6, anti-R-esp1 immune serum; lane 7, anti-R-esp1 immune serum that was pre-incubated with peptide that was used as antigen.

In Vitro Translated R-esp Proteins

The ability of proteins to interact plays a fundamental role in determining the function of the resulting protein complex. To determine whether R-esp1 or R-esp2 are capable of forming homomultimers or whether they may directly interact to form a heterocomplex, we characterized the physical state of the R-esp proteins. R-esp1 and R-esp2 proteins were translated in vitro using the rabbit reticulocyte lysate system, and their apparent molecular weight was determined by a combination of gel filtration and sucrose density gradient centrifugation (see Schmidt and Neer, 1991). The results are summarized in Table III. R-esp1 behaves as a uniform population of molecules and the physical parameters indicate that in vitro translated R-esp1 is a monomer that forms extended rods. This is based on comparison of the observed sedimentation coefficients and Stokes radii with those of globular proteins. For example, R-esp1 has a $s_{20,w}$ similar to that of the cytochrome c which has a molecular weight of 13,370. In contrast, the Stokes radius of R-esp1 is similar to that of glyceraldehyde-3-phosphate dehydrogenase, which has a molecular weight of 145,000. These differences indicate that the in vitro translated proteins do not fold to a globular structure and suggest an extended rod like shape.

In contrast to R-esp1, the in vitro translated R-esp2 separated into two distinct populations on both gel filtration and sucrose density gradient centrifugation (not shown). One population behaved as rod shaped monomers yielding the hydrodynamic data presented in Table III for R-esp2. The second population eluted in the excluded volume of the ACA 22 column (exclusion limit = $1.2 \times 10^6$ daltons), and this material rapidly sedimented as a broad peak during sucrose density gradient centrifugation, indicating that it is not a stable complex (not shown). This high molecular weight population represents ei-
their large aggregates of R-esp2 or aggregation of R-esp2 with other proteins present in the rabbit polyclonal lysate.

To determine if R-esp1 and R-esp2 can form heterodimers the proteins were translated separately, mixed, and incubated at either 4 or 37 °C for 1 h prior to loading on the gel filtration column or sucrose density gradient. No changes in the physical properties of either R-esp1 or R-esp2 were detected indicating that the R-esp proteins do not form heterocomplexes under these conditions.

**DISCUSSION**

The R-esp1 cDNA encodes a novel homolog of the E(spl)-WD locus of *Drosophila*, also known as m9/10 or groucho, which appears to be a truncated form of the previously characterized products of the E(spl)-WD genes. The data presented here indicate that R-esp1 is a full-length cDNA clone, because it hybridized with a 1.3-kilobase mRNA in adult rat tissue. R-esp1 mRNA was detected in kidney, brain, heart, and lung tissue of adult rats. Differences in mRNA level among these tissues suggest that R-esp1 mRNA is subject to tissue-specific regulation in the adult. Immunoprecipitation with antibodies against a peptide from R-esp1 detected R-esp1 protein in PC-12 cells. In vitro translation of R-esp1 yielded a product that comigrated with a 24-kDa protein that was specifically immunoprecipitated from the PC-12 cells. Finally, immunohistochemical staining revealed this protein localized to the nucleus of both PC-12 and GH3 cells. Taken together, these results argue that R-esp1 cDNA is a full-length cDNA encoding a cellular product of a gene also encoding R-esp2. However, R-esp1 could be an alternative splice product of some other full-length E(sp1)-WD homolog. Resolution of this question will require isolation of genomic clones encoding these proteins.

When the available nucleotide sequences of the E(spl)-WD gene family are compared, R-esp1 shows greatest identity with human cDNA EST00256 (Table II). EST00256 was originally isolated as an expressed sequence tag and identified as a membrane-bound WD homolog. Resolution of this question will require isolation of genomic clones encoding these proteins.

The R-esp1 cDNA is a full-length cDNA encoding a cellular protein. R-esp1 may arise from a unique gene or by alternative splicing of a gene that also encodes a protein containing the WD-40 repeating units. Sequence comparison of R-esp1 and R-esp2 suggest that R-esp1 cDNA could not be an alternative splice product of a gene also encoding R-esp2. However, R-esp1 could arise by alternative splicing of some other full-length E(spl)-WD homolog. Resolution of this question will require isolation of genomic clones encoding these proteins.

When the available nucleotide sequences of the E(spl)-WD gene family are compared, R-esp1 shows greatest identity with human cDNA EST00256 (Table II). EST00256 was originally isolated as an expressed sequence tag and identified as a member of the E(spl) gene family by partial sequencing of the cDNA (Adams et al., 1991). In their screening of a human brain cDNA library, Stifani et al. (1992) failed to isolate a clone corresponding to EST00256. The probe used to screen their cDNA library corresponded to a region containing WD-40 repeats, a region absent from the R-esp1 clone. Possibly, EST00256 corresponds to the human homolog of R-esp1.

Stifani et al., (1992) recognized that full-length E(spl)-WD homologs contain five discrete regions based upon amino acid composition and sequence. If these regions are important, they are likely to be conserved. Comparison of the amino acid conservation between the two new rat sequences together with the human and *Drosophila* E(spl)-WD proteins suggests that three discrete domains (I, II, and III in Fig. 3) are present in the full-length proteins. However, R-esp1 only contains amino-terminus domain I and the proline-glycine-rich portion of domain II while lacking the remaining portion of domain II and all of domain III. R-esp2 encodes a full-length E(spl)-WD homolog containing all three domains.

The E(spl)-WD homologs containing WD-40 repeats are nuclear proteins. If R-esp1 affects the function of full-length E(spl)-WD homologs then R-esp1 protein may also be found in the nucleus. Although R-esp1 protein lacks a recognizable nuclear localization signal, immunohistochemical staining clearly shows it in the nucleus of both PC-12 and GH3 cells. The hydrodynamic data suggest that R-esp1 is a rod-shaped monomer that does not form homodimers. R-esp1 monomers, with a predicted molecular mass of 24 kDa, would be small enough to pass into or out of the nucleus by diffusion (Paine and Horowitz, 1980). Alternatively, R-esp1 may associate with another protein which contains a nuclear localization signal. Preliminary results with anti-R-esp2 antibodies indicate that this protein is also found in the nucleus (not shown). Like the human and *Drosophila* E(spl)-WD homologs, R-esp2 contains a nuclear localization signal.

The pattern of nuclear staining was distinguishable between untreated and NGF-treated PC-12 cells. Staining in untreated cells appeared reticulated with the stained portions forming a network containing many punctate regions lacking stain. It is unclear what confers this reticulated appearance, although one possibility is that R-esp1 associates with the nuclear matrix in untreated cells. In PC-12 cells that extended processes due to NGF treatment, the fibrillar appearance was less dramatic with staining broadly distributed throughout the nucleus. However, unstained regions are still clearly visible. Differentiation may change the nuclear targets that bind R-esp1, resulting in altered immunohistochemical staining.

R-esp1 contains all of domain I but lacks most of domain II and all of domain III. Possibly, R-esp1 acts as a negative regulator of full-length E(spl)-WD proteins by competing for factors that interact with domain I. This model predicts that R-esp1 and R-esp2 should share a subset of interacting factors. Current work is focused upon evaluating this possibility.

**Acknowledgment**—We thank Dr. Eva Neer for encouragement during the course of this work and for critically reading the manuscript.

**REFERENCES**

Adams, M. D. Kelley, J. M., Gocayne, J. D., Dubnick, M., Polymeropoulos, M. H., Xiao, H., Merrill, C. R., Wu, A., Olde, B., Moreno, R. F., Kerlavage, A. R., McBroom, W. R., and Venter, J. C. (1991) Science 252, 1651-1656

Artavanis-Takonas, S., Delidakis, C., and Fehon, G. R. (1991) Annu. Rev. Cell Biol. 7, 437-454

Bjorn, S. P., Soltysk, A., Beegs, J. D., and Friesen, J. D. (1988) Mol. Cell. Biol. 8, 3698-3707

Cao, R., and Saccani, N. (1987) Anal. Biochem. 163, 156-159

Cubin, V., Michelson, A. M., Ahmay, S. N., Neel, V., Alcamo, E., Maniatis, T., and Young, M. W. (1991) Cell 67, 311-323

Derveaux, J., Haseber, F., and Smithies, O. (1984) Nucleic Acids Res. 12, 387-395

Durotino, R.J., Gordon, J. L., and Boguski, M. S. (1992) Proteins 13, 41-56

Feinberg, A. P., and Vogelstein, B. (1985) Anal. Biochem. 152, 3-12

Harnett, A. Y., Bugendoff, A., Sgouros, L., and Hartenstein, V. (1992) Development (Camb.) 116, 1203-1220

Hartley, D.A., Preiss, A., and Artavanis-Takonas, S. (1988) Cell 55, 785-795

Kamb, C., Krust, E., Tietz, K., and Campo-Ortega, J. A. (1989) EMBO J. 8, 203-210

Laemmli, U. K. (1970) Nature 227, 680-685

Maniatis, T., Fritsch, E. F., and Sambrook, J. (1982) Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY

Paine, P. L., and Horowitz, S. B. (1980) Cell Biology, A Comprehensive Treatise (Prescott, D. M., and Goldstein, eds) Vol. 4, pp. 299-338, Academic Press, New York

Sanger, F., Nicklen, S., and Coulson, A. R. (1977) Proc. Natl. Acad. Sci. U. S. A. 74, 5463-5467

Schmidt, C. J., and Neer, E. J. (1991) J. Biol. Chem. 266, 4538-4544

Schmidt, C. J., Garen-Fazio, S., Chow, Y.-K., and Neer, E. J. (1989) Cell Regul. 1, 125-134

Stifani, S., Blaumueller, C. M., Redhead, N. J., Hill, R. E., and Artavanis-Takonas, S. (1992) Nature genetics 2, 119-127

Takagaki, Y., and Manley, J. L. (1992) J. Biol. Chem. 267, 23471-23474

Williams, P. R., and Trumbly, R. J. (1990) Mol. Cell. Biol. 10, 6560-6561