The *Drosophila melanogaster* suppressor of Hairy-wing protein binds to specific sequences of the gypsy retrotransposon

Carl Spana, Douglas A. Harrison, and Victor G. Corces

Department of Biology, The Johns Hopkins University, Baltimore, Maryland 21218 USA

Mutations at the suppressor of Hairy-wing [su(Hw), 3-54.8] locus reverse the phenotype of second-site mutations induced by the gypsy transposable element in *Drosophila melanogaster*. This gene encodes a protein with a predicted molecular weight of 109,000 that contains an acidic domain and 12 copies of the DNA-binding 'Zn finger' motif. The su(Hw) protein was overexpressed in *Escherichia coli* and *Drosophila* cells, and partially purified. It was shown to interact specifically in vitro with a 367-bp DNA fragment that contains 12 copies of the DNA-binding repeat (LTR) and the first ATG initiation codon. This sequence shows striking homology to some mammalian transcriptional enhancer elements, supporting a role for the su(Hw) protein in the control of gypsy transcription. In addition, the su(Hw) protein is present at ~100–200 sites on *Drosophila* polytene chromosomes, suggesting that it also interacts in vivo with DNA and might be involved functionally in the regulation of normal cellular genes.

[Key Words: Transposable element; suppression; DNA–protein interaction; regulatory factor; transcriptional control]

Received June 17, 1988; revised version accepted September 18, 1988.

Transposable element insertions are the major source of spontaneous mutations in *Drosophila* and yeast [Roeder and Fink 1983; Rubin et al. 1983; Levis et al. 1984]. The mechanisms by which these elements elicit a mutant phenotype are varied. For example, transposons can cause abnormal expression of genes by disrupting the protein-coding region, they might affect the normal processing of the RNA by inserting into sequences required for splicing, or they may interrupt sequences necessary for transcription initiation, thus affecting the rate of this process [for review, see Finnegan and Fawcett 1986]. Other transposons, such as the Ty element of yeast [Winston et al. 1984a,b] and the copia [Levis et al. 1984; Zachar et al. 1985; Mount et al. 1988], 412 [Searles and Voelker 1986], and gypsy [Modolell et al. 1983; Parkhurst and Corces 1986a] elements of *Drosophila*, also utilize more sophisticated mechanisms to create mutant phenotypes that can be modified by second-site mutations at unlinked loci [Rutledge et al. 1988].

The gypsy retrotransposon causes mutant alleles in at least 11 different genes whose phenotype is reversed by mutations at the suppressor of Hairy-wing [su(Hw), 3-54.8] locus [Modolell et al. 1983; Parkhurst and Corces 1986a; Rutledge et al. 1988]. These suppressible alleles result from the insertion of gypsy in the 5' region [Biessman 1985; Chia et al. 1986; Geyer et al. 1986; Parkhurst and Corces 1986b], in an intron [Parkhurst and Corces 1985; Peifer and Bender 1986], or in an exon [Campuzano et al. 1980] of the mutated gene. The molecular basis of the mutant phenotype has been studied in most detail in the case of the y^2^ allele that originates from the insertion of gypsy at ~700 bp from the transcription start site of the yellow [y, 1-0.0] gene. This insertion leads to a temporally and spatially restricted phenotype consisting of mutant wing blades and body cuticle, but wild-type bristles in the adult fly. The expression of the yellow gene in the body and wings of adult *Drosophila* is governed by two independent tissue-specific transcriptional enhancers located upstream of the insertion site of gypsy in the yellow locus [Geyer and Corces 1987]. We have suggested that this phenotype emerges from the interaction of the su(Hw)-encoded protein with specific sequences in the central region of the gypsy element [Geyer et al. 1988a,b]. This interaction, which is concomitant with and probably necessary for the transcriptional activation of gypsy during pupal development, disrupts the effect of tissue-specific transcriptional enhancers on the yellow promoter and ultimately leads to the y^2^ phenotype. The basis for the mutagenic effect of the gypsy element in *forked* [f, 1-56.7], *bithorax* [bx, 3-58.8], and *bithoraxoid* [bxd, 3-58.8] alleles that result from the insertion of this element in the 5' region or an intron of the respective genes has been attributed to a similar mechanism [Parkhurst and Corces 1985; Peifer and Bender 1986, 1988]. The insertion of gypsy into a *forked* intron does not result in accumulation of aberrant size transcripts, suggesting that the low levels of normal size *forked* RNAs observed in...
this mutant are the consequence of the interference of the gypsy element with the transcription or stability of the forked message [Parkhurst and Corces 1985]. An apparently different mechanism is responsible for the dominant gain-of-function phenotype of the hairy-wing (Hw, 1-0.0) allele Hw1, which results from the insertion of the gypsy element into an exon of one of the transcription units of the achaete-scute locus. This insertion gives rise to a truncated RNA that terminates in the 5’ LTR of gypsy and accumulates at higher levels than in wild-type flies, suggesting that the gypsy element is affecting the proper termination of transcription of the gene [Campuzano et al. 1986]. Whatever the mechanism of gypsy mutagenesis, these results suggest that the su(Hw)-encoded protein interacts with gypsy DNA to regulate transcription initiation, elongation, or termination of this transposon, and, as a consequence, it affects the expression of adjacent genes. Recent developments emerging from the isolation and characterization of the su(Hw) locus support this hypothesis. This gene encodes a protein that contains an acidic domain and 12 repeats of the Zn finger motif, a modular structure typical of transcriptional activators such as GAL4 [Ma and Ptashne 1987], suggesting that the su(Hw)-encoded protein might interact with DNA [Parkhurst et al. 1988]. Here we report the partial purification of this protein and analyze its in vitro interaction with specific sequences in the gypsy element involved in the mutagenic effect of this retrotransposon.

Results

Expression of su(Hw) protein in Escherichia coli and Drosophila cells

To obtain the su(Hw) protein in sufficient amounts to carry out biochemical studies, we constructed a plasmid containing the XbaI–HindIII fragment from a su(Hw) cDNA clone [cDNA-1, see Parkhurst et al. 1988] fused to the E. coli trpE promoter [Dieckmann and Tzagoloff 1985]. This cDNA fragment encodes a truncated su(Hw) protein that lacks the amino-terminal region containing the acidic domain and one of the Zn finger motifs [Fig. 1A]. After transformation into E. coli, the resulting plasmid, denominated pTE-suHw, can be induced with 3-β-indolacrylic acid to synthesize large amounts of a trpE-su(Hw) hybrid protein of the predicted 115,000 m.w. [Fig. 1B], originating from the fusion of the 37,000-m.w. trpE fragment and the 78,000-m.w. truncated su(Hw) protein. The large amount of fusion protein synthesized by the bacteria is usually partially degraded, and it complexes in large aggregates called inclusion bodies. The band containing this protein was isolated from an acrylamide gel and injected into rabbits to obtain polyclonal antibodies. These antibodies react with a series of bands on Western blots containing extracts from E. coli transformed with the fusion protein-encoding expression vector (Fig. 1C). Some of these bands are probably degradation products of the full-length 115,000-m.w. fusion protein, whereas others are likely

Figure 1. Overexpression of su(Hw) protein in E. coli and Drosophila cells. (A) Structure of the su(Hw) locus. Indicated are restriction sites used to construct different expression vectors, restriction fragments used for this purpose were obtained from cDNA clones and therefore lacked introns. The acidic domain is indicated by a solid circle, and the Zn finger motifs are represented by vertical lines. (B) SDS-polyacrylamide gel electrophoresis of protein extracts obtained from E. coli after induction with 3-β-indolacrylic acid. [Lane 1] Proteins obtained from cells transformed with the pATH2 vector, which encodes a 37,000-m.w. truncated trpE protein, [lane 2] the protein profile of cells producing the trpE-su(Hw) 115,000-m.w. fusion protein. [C] Western analysis of protein extracts obtained from E. coli producing the trpE-su(Hw) fusion protein [lane 1] and Drosophila adults [lane 2]. [D] Western analysis of nuclear protein extracts prepared from a Drosophila cell line transformed with a plasmid containing a metallothionein promoter–su(Hw) hybrid gene, before [lane 1] and after [lane 2] induction with CuSO4.
due to antibodies raised against contaminating peptides present in the protein preparation injected into the rabbits.

To avoid problems derived from the use of nonspecific antibodies that could cross-react with proteins other than su(Hw) in Drosophila extracts, antisera against the trpE–su(Hw) protein was preabsorbed with extracts made from E. coli transformed with the expression vector encoding the trpE polypeptide alone, to eliminate antibodies against bacterial-encoded proteins. The resulting antiserum was further affinity-purified against gel-isolated su(Hw) fusion protein. The affinity-purified antibodies recognize a single polypeptide of 110,000 m.w., similar to that predicted from the DNA sequence of the su(Hw) gene on Western blots containing protein extracts from adult Drosophila melanogaster flies [Fig. 1C]. This result indicates that the affinity-purified antibody is specific for the su(Hw) protein and does not cross-react with other Zn finger-containing polypeptides present in Drosophila.

The bacteria-produced protein might not contain the proper post-translational modifications that occur in Drosophila during the synthesis of the su(Hw) protein; therefore, its biological properties might be different from those of the Drosophila-made protein. To circumvent possible artifactual results in the DNA-binding experiments described below, su(Hw) protein was also expressed in Drosophila tissue culture cells. A full-length su(Hw) cDNA (cDNA-1; Parkhurst et al. 1988) was inserted into the expression vector pRMHa-3, which contains the metallothionein promoter [Bunch et al. 1988]. The resulting plasmid, designated pTC-suHw, was then cotransformed into Drosophila cells with a clone containing the DHFR gene, which confers resistance to methotrexate. The transformed cells can be induced with Cu²⁺, giving rise to >10-fold overexpression of the intact 110,000-m.w. su(Hw) protein [Fig. 1D].

An E. coli-produced truncated su(Hw) protein can bind DNA

Partially purified trpE–su(Hw) fusion protein was obtained from extracts of E. coli transformed with the pTE-su(Hw) expression vector described above. The protein contained in the inclusion bodies was resuspended in 8 M urea and then renatured by dialysis against decreasing urea concentrations. The partially purified polypeptide was tested for its ability to interact with gypsy DNA using a filter-binding assay. A plasmid containing the gypsy transposon was digested with Avall (Fig. 2A), and the resulting fragments were end labeled with [³²P]dNTP and the Klenow fragment of E. coli DNA polymerase I, and incubated with E. coli-produced trpE–su(Hw) hybrid protein. Only one restriction fragment, 746 bp long, was retained on the filter [Fig. 2B], suggesting that this hybrid protein binds to specific sequences of the gypsy element. This fragment is located immediately adjacent to the 5' long terminal repeat (LTR), between nucleotides 459 and 1205 [Marlor et al. 1986], and its position on the gypsy restriction map is indicated in Figure 2A, with a solid black box representing the sequences that interact with the trpE–su(Hw) fusion protein (see below). Extracts obtained from bacteria carrying the trpE vector alone and treated in the same fashion were unable to bind gypsy DNA, suggesting that the observed interaction was due to the trpE–su(Hw) fusion protein. In addition, DNA obtained from the pUC18 vector alone was not retained in the nitrocellulose filter after incubation with su(Hw) protein [data not shown].

An intact su(Hw) protein produced in a Drosophila cell line binds specific gypsy sequences

The bacterial fusion protein used in the studies described above lacked the acidic domain and the first Zn finger motif. Attempts to express an intact su(Hw) product using the same bacterial vector failed, probably due to a deleterious effect of the complete protein on E. coli. To overcome this problem, full-length su(Hw) protein was obtained from Drosophila cells transformed with the metallothionein promoter–su(Hw) expression vector pTC-suHw. Nuclear extracts prepared from transformed cells by standard procedures [Heberlein et al. 1985] were chromatographed on a DEAE–Sepharose column. Proteins retained in the column were eluted with a salt step gradient, using 0.2, 0.4, 0.6, 0.8, and 1.0 M KCl successively. Protein present in the different fractions was electrophoresed on a 5% acrylamide gel and subjected to Western analysis using anti-su(Hw) antibodies. Fractions eluted with 0.8 M KCl contained the su(Hw) protein [data not shown], and they were pooled, stored at −70°C, and used in all experiments described below.

This preparation was first tested for its ability to interact with gypsy DNA using the filter binding assay. Figure 2C shows that the Drosophila-produced protein binds to the same 746-bp Avall fragment of the gypsy element. This fragment was further digested with XmnI and Bsp1286, giving rise to 211-bp Avall–XmnI, 367-bp XmnI–Bsp1286, and 118-bp Bsp1286–Avall fragments. These restriction products were labeled with [³²P]dNTP and T4 DNA polymerase at the 3' overhangs and the Klenow fragment of E. coli DNA polymerase I at the 5' overhangs, and subjected to the same type of analysis [Fig. 2D]. The 211-bp Avall–XmnI fragment appears to label with a higher specific activity than the other two, explaining the difference in intensities of the various bands in lane 1 of Figure 2D. Of the three DNA fragments present in the reaction mix, only the 367-bp XmnI–Bsp1286 fragment was retained on the filter after incubation with the su(Hw) protein preparation, suggesting that this protein binds to sequences contained within this fragment [Fig. 2D]. A residual amount of the 211-bp fragment was also retained on the filter; competition experiments suggest that this interaction was very weak, and it was not studied further.

To analyze in more detail the specificity of the interaction of the su(Hw) protein with gypsy DNA, we carried out competition experiments. The labeled 746-bp
Figure 2. Specificity of the su(Hw) protein–gypsy DNA interaction. (A) Restriction map of the AvaII digest of the gypsy element; the specific fragment containing the sequences that interact with the su(Hw) protein is indicated by a solid box. (B) Filter-binding assay of the interaction between su(Hw) protein and gypsy DNA. (Lane 1) AvaII-digested gypsy DNA, end labeled with \[^{32}P\]dNTP; (lane 2) the DNA retained in the filter after incubation of the AvaII digest with trpE-su(Hw) fusion protein. (C) Filter-binding assay of partially purified full-length su(Hw) protein produced in Drosophila cells and gypsy DNA. (Lane 1) AvaII-digested, \[^{32}P\]labeled gypsy DNA; (lane 2) DNA retained in the filter after incubation with the Drosophila-made protein. (D) The 746-bp AvaII fragment bound by su(Hw) protein was digested with XmnI and Bsp1286 and end labeled with \[^{32}P\]dNTP and T4 DNA polymerase and Klenow fragment (lane 1); DNA retained in the filter after incubation with partially purified su(Hw) protein is shown in lane 2. (E) Competition experiment. A fixed amount (1 ng) of \[^{32}P\]labeled 746-bp AvaII gypsy fragment was incubated with 5 μg of partially purified Drosophila produced su(Hw) protein in the presence of increasing concentrations (lanes 2–5) of cold 367-bp XmnI–Bsp1286 fragment; the total amount of poly(dI-dC) carrier plus cold competitor was kept constant at 3 μg in the different lanes. (Lane 2) 0 ng of cold competitor; (lane 3) 75 ng; (lane 4) 300 ng.

AvaII fragment was incubated with partially purified Drosophila su(Hw) protein and increasing amounts of unlabeled 367-bp XmnI–Bsp1286 fragment. Figure 2E shows that the cold 367-bp fragment is able to compete with the 746-bp labeled one for the su(Hw) protein, suggesting that the binding is specific for sequences contained within the smaller fragment. It should be noted that this competition is complete, and no residual binding of labeled 746-bp fragment could be detected on the filter. This suggests that the affinity of the su(Hw) protein for the 367-bp XmnI–Bsp1286 fragment is much higher than its affinity for the 211-bp AvaII–XmnI fragment, as this smaller fragment is contained within the 746-bp piece (see above). The specificity of this interaction was studied further by gel mobility shift experiments. The 367-bp XmnI–Bsp1286 fragment was incubated with Drosophila extract and electrophoresed on a 5% acrylamide gel. The presence of the protein extract causes a retardation in the mobility of the DNA fragment because of a change in the effective molecular weight resulting from the binding of the su(Hw) protein (Fig. 3, lane 2). Three bands of different molecular weights can be observed after the binding of this protein. These various bands are probably the consequence of the interaction of increasing amounts of protein with a single DNA molecule. In fact, the slowest moving band disappears, and only one or two faster moving DNA–protein complex bands are observed using lower amounts of protein (data not shown); this suggests that at least three molecules of su(Hw) might bind consecutively to gypsy DNA. This effect is specifically due to the su(Hw) protein present in the protein extract, as antibodies against this polypeptide eliminate the anomalously moving bands and cause the su(Hw)–DNA complex to be retained in the well of the acrylamide gel (Fig. 3, lanes 3 and 4). The antibodies do not interfere with the DNA-binding capacity of the su(Hw) protein but specifically precipitate the complex of this protein with gypsy DNA. This effect is specific rather than due to random trapping of the DNA by the antiserum, because the mobility of the protein-free DNA is not affected, and only the protein–DNA complexes are retained in the well. In addition, goat anti-rabbit IgG has no effect on the mobility of the su(Hw)–DNA complex, even when a
Figure 3. Gel mobility shift assay of the interaction of su(Hw) protein with gypsy DNA. Labeled 367-bp XmnI–Bsp1286 gypsy fragment was incubated with 5 μg of partially purified su(Hw) protein and electrophoresed on a 5% acrylamide gel. (Lane 1) 1 ng of 32P-labeled DNA with no protein added; (lane 2) the same DNA incubated with 5 μg of su(Hw) protein; (lanes 3 and 4) the same as lane 2, with two different concentrations of affinity-purified su(Hw) antibody; (lane 5 and 6) 2 ng of labeled DNA and 5 μg of protein extract with two different concentrations [1:1 dilution in lane 1, 1:2 dilution in lane 2] of goat anti-rabbit IgG; (lanes 7 and 8) the same as lane 2, but they also contain two different concentrations of the buffer in which the su(Hw) antibody was stored.

twofold excess of protein extract is used (Fig. 3, lanes 5 and 6).

The su(Hw) protein binds to sequences homologous to mammalian transcriptional enhancers

The nucleotide sequence of the 367-bp DNA fragment to which the su(Hw) protein binds, located between nucleotides 670 and 1037 in the gypsy element, is given in Figure 4A. This fragment contains 12 copies of the consensus sequence PyPuTTGCATACCPy separated by stretches of highly AT-rich sequences [Marlow et al. 1986]. To determine accurately the gypsy sequences that interact with the su(Hw) protein, we carried out footprinting experiments. The 367-bp fragment was blunt ended and cloned into the Smal site of pUC18. The orientation of the insert was such that the XmnI site was closer to the EcoRI site of pUC18. The resulting plasmid was digested with EcoRI, end labeled with Klenow fragment and [32P]dNTP, and incubated with protein extract, and the complex was digested with DNase I and electrophoresed on an 8% acrylamide–urea gel (Fig. 5). The same end-labeled DNA was sequenced following the procedure of Maxam and Gilbert [1980], and the sequencing reactions were electrophoresed in the same gel to obtain an accurate ladder of molecular weight markers [data not shown]. Figure 5 shows the result obtained by digestion of DNase I of the labeled DNA alone [lane 1] and after incubation with a protein extract that included the fractions of the DEAE-Sepharose column containing su(Hw) protein [lane 2] or fractions lacking this activity [lane 3]. These results indicate that the su(Hw) protein interacts specifically with the repeated sequences PyPuTTGCATACCPy contained in the XmnI–Bsp1286 367-bp fragment and that the affinity of the protein is different for the various copies of the repeat. The protection from digestion by DNase I can be eliminated by adding increasing amounts of unlabeled competitor 367-bp XmnI–Bsp1286 fragment [data not shown].

We have divided the protected region into several units, or binding domains, designated with roman numerals I–VI, based on the extent of the protection to digestion by DNase I by the su(Hw) protein [Fig. 5]. Each of these units comprises an average of 55 bp centered around one of the PyPuTTGCATACCPy repeats and including half of one repeat upstream and half of one repeat downstream plus the AT-rich spacer sequences. The affinity of the su(Hw) protein for each of these units alternates between weak [boxes I, III, and V, designated with open rectangles in Figs. 4A and 5] and strong [boxes II, IV, and VI, designated with solid rectangles in Figs. 4A and 5]. Part of this repeated sequence shows striking homology with the octamer motif present in mammalian transcriptional enhancer elements such as the SV40, immunoglobulin heavy-chain, U2 snRNA gene, and histone H2B gene enhancers [Bohmann et al. 1987; Rosales et al. 1987]. Figure 4B shows a comparison of the octamer motif from some of these enhancers and promoters with the gypsy consensus repeat. The combination of structural data suggesting a similarity of the su(Hw) protein to transcriptional activators and the fact that this protein binds to an enhancer-like sequence support previously proposed roles for su(Hw) as a positive regulator of gypsy transcription [Parkhurst and Corces 1985, 1986a].

The su(Hw) protein interacts in vivo with DNA

To establish the significance of the observed interaction of the su(Hw) protein with gypsy DNA in vitro and a putative role for this protein as a transcription factor for cellular genes, we tested the possibility that DNA binding also occurs in vivo. It is likely that the su(Hw) gene product has evolved to play a role in cell function other than the regulation of gypsy transcription, as this transposon is present in some Drosophila strains but absent in others. In addition, mutations at the su(Hw) locus result in female sterility, suggesting that the protein product of this gene plays a fundamental role in the normal physiology of the cell and should therefore interact with other DNA sequences in the cellular genome. We used antibodies specific to the su(Hw) protein to detect its presence on polytene chromosomes of third-instar larvae from a stock carrying the gypsy-induced mutation y^sc^ct9f. The existence of these alleles in the stock assu...
**Figure 4.** Sequence structure of the gypsy region that interacts with the *su(Hw)* protein. (A) Location of the different enhancer-like repeats on the gypsy fragment bound by *su(Hw)* protein. The different domains that interact with the *su(Hw)* protein are indicated by boxes above the DNA sequence; open boxes denote regions of weak binding, whereas solid boxes indicate domains of higher affinity for the *su(Hw)* protein. These various domains are numbered I–VI. (B) Comparison of the consensus sequence of the gypsy repeats and the octamer motif of the SV40, immunoglobulin heavy chain (IgH), and U2 snRNA enhancers and the histone H2B and immunoglobulin heavy chain (IgVH) promoters (data obtained from Rosales et al. 1987).

| Region | Description |
|--------|-------------|
| I      | IgH enhancer |
| II     | SV40 enhancer |
| III    | H2B promoter |
| IV     | IgVH promoter |
| V      | U2 RNA enhancer |

where those genes map; this particular strain contains seven euchromatic copies of gypsy: two at subdivision 1B and one each at 7B, 13A, 14C, 15F, and 16E (Parkhurst and Corces 1986b); therefore, antibody staining should be detectable at least at these chromosomal locations. A caveat to the use of polytene chromosomes for this type of study is that they originate from a very specific tissue, the salivary gland, with a distinct pattern of gene expression. Given the structure of the *su(Hw)* protein and its putative role as a regulator of gypsy transcription, we should expect this protein to play a similar role for cellular genes and, therefore, interact with genes that are transcribed actively. Unfortunately, at present, we do not have information on the expression of gypsy in larval salivary glands, but other genes might be expressed in this tissue whose transcription is influenced by the *su(Hw)* protein.

Figure 6 shows a photograph resulting from the immunofluorescence analysis carried out with *su(Hw)* antibodies and indicates that the *su(Hw)* protein is present in ~100–200 sites on polytene chromosomes. Preincubation of the antibody with gel-purified trpE-*su(Hw)* fusion protein eliminates labeling of the bands (data not shown), suggesting that the observed fluorescence is specifically due to the *su(Hw)* protein. Although the resolution of the analysis was not sufficient to distinguish whether the chromosomal locations containing copies of the gypsy element were labeled, it allows us to state that the localization of the antigen is specific for certain chromosomal sites and that these include more than just the seven copies of the gypsy element that are present in the genome of this stock. This result suggests that the *su(Hw)* protein interacts with other DNA sequences in the *Drosophila* genome and that this protein may be a general transcription factor that plays an important role in the regulation of normal cellular genes.

**Discussion**

The gypsy retrotransposon causes mutant phenotypes that can be reversed by nonallelic mutations at the *su(Hw)* locus. We have noted previously a correlation between the developmental time of expression of gypsy and of the genes it mutates in a suppressible manner. This correlation led to the suggestion that the transcription of the gypsy element is involved directly in creating the mutant phenotype (Parkhurst and Corces 1985). More recent evidence from the analysis of the gypsy-in-
duced y² allele indicates that transcription of the gypsy element might not be necessary per se to create a mutant phenotype but that the mere interaction of the su(Hw) protein with sequences located in the central region of gypsy may be sufficient to alter the normal interplay between tissue-specific transcriptional enhancers and the yellow promoter (Geyer and Corces 1987; Geyer et al. 1988a, b). In support of this hypothesis, the cloning and characterization of the su(Hw) locus has shown that the su(Hw)-encoded protein is structurally homologous to transcriptional activators such as GAL4 (Ma and Ptashne 1987), with an acidic domain and Zn finger motifs (Klug and Rhodes 1987), and therefore has the potential to bind DNA and interact with other proteins involved in the formation of transcription complexes (Parkhurst et al. 1988).

The results presented here support this presumption, because they indicate that the su(Hw) protein interacts with specific sequences of the gypsy element. Partially purified su(Hw) protein obtained from both E. coli and

**Figure 5.** Footprint analysis of the su(Hw)-gypsy DNA interaction. The XmnI–Bsp1286 367-bp fragment located in the 5’-untranslated region of gypsy, between nucleotides 670 and 1037, was cloned into the SmaI site of pUC18. The fragment was then labeled at the EcoRI site present in the vector with [³²P]dNTP and Klenow fragment, and incubated with 1 µg/ml of DNase I in the absence (lane 1) or presence of 50 µg of protein extract containing the su(Hw) protein (lane 2) or fractions in which this protein was absent (lane 3). [C, B, and A] represent successive loadings of the same samples.
Drosophila cells transformed with different expression vectors is able to bind specifically a 367-bp fragment of gypsy DNA. This fragment is located in the 5' untranslated region of gypsy, downstream of the 5' LTR, and contains 12 copies of the sequence PyPuTTGCA-TACCPy, which is homologous to the octamer motif of mammalian enhancers [Bohmann et al. 1987, Rosales et al. 1987]. An additional copy of this repeat overlaps a portion of the most 3' member of the family, but its sequence partly deviates from the consensus [Fig. 4A]. Footprinting experiments indicate that the binding domain of the su(Hw) protein is 55 bp long, close in size to the 53-bp region protected by the Xenopus transcription factor TFIIA, which contains nine Zn finger motifs [Vrana et al. 1988]. Each of the su(Hw)-binding domains is centered around one of the enhancer-like repeats, and the protected unit extends upstream and downstream to the center of the neighboring repeats [Fig. 4A]. The region of gypsy that interacts with the su(Hw) protein contains six of these domains, and their affinity for the su(Hw) protein is not uniform but, rather, alternates between weak and strong. This differential affinity is probably responsible for the various bands observed in the gel mobility shift experiments [Fig. 3].

Further support for the involvement of this sequence in the interaction of gypsy with the su(Hw) protein comes from recent experiments demonstrating that the structural integrity of these repeated sequences is involved directly in the intensity of the gypsy-induced mutant phenotype. Two partial revertants of the gypsy-induced bithoraxoid allele bxdI arose as a consequence of the deletion of four of these repeats [Peifer and Bender 1988], and a partial revertant of the cut (ct, 1-20.0] ct6 allele originated as a consequence of the insertion of the jockey transposable element into these sequences [Mizrokhi et al. 1988]. Also, two partial revertants of y and the Hw1 revertant Hw85 were due to the insertion of diverse transposable elements into the region containing the repeated enhancer core-like sequences [Geyer et al. 1988b]. The combination of revertant analysis and in vitro DNA-binding studies strongly suggests that the mutant phenotype in gypsy-induced alleles is a consequence of the interaction of the su(Hw) protein with specific repeated sequences in the gypsy element. In fact, one could say that the mutant phenotype in these alleles is not caused by the gypsy element only but that this transposon is merely a mediator of the mutagenic effect of the su(Hw) protein.

A comparison of the results obtained with the bacterial- and Drosophila-produced proteins gives some insights into the nature of the regions of the su(Hw) protein involved in DNA binding. The amino-terminal region of this protein, containing the acidic domain and one of the Zn fingers, has been replaced in the E. coli-produced polypeptide by part of the trpE-coding region. This altered protein is able to bind with the same specificity as the intact Drosophila-made protein, although the strength of the interaction was not compared between protein preparations obtained from both sources. This result suggests that not all fingers are necessary for binding and that the acidic domain is not involved in the interaction with DNA. This domain, as has been proposed for GCN4, might be involved in the interaction with RNA polymerase II [Hope et al. 1988].

The structural similarity of the su(Hw) protein to transcriptional activators and its interaction with enhancer-like sequences are in accordance with previously proposed roles for this protein in the transcription of the gypsy element. The presence of this protein in ∼100–200 loci in larval polytene chromosomes suggests that it might actually be a transcription factor involved in the expression of normal cellular genes and that the gypsy element has borrowed this factor for its own transcription. The only known phenotype of su(Hw) mutations is female sterility, suggesting that this locus might be involved in the regulation of genes that control oogenesis. Alternatively, because the gypsy element contains 12 copies of the su(Hw) protein-binding domain, it might be possible that this transposon interacts more strongly with the su(Hw) product than cellular genes containing only 1 copy of the binding motif. This would explain the fact that small alterations in the levels or structure of the su(Hw) protein in various mutant alleles only give rise to female sterility while still being able to reverse the phenotype of gypsy-induced mutations. The ability to isolate large amounts of normal and in vitro altered protein will allow us to approach these questions and to study the mechanisms involved in the expression of this retrotransposon.

Materials and methods

Expression of su(Hw) protein in E. coli and Drosophila cells

A 2.1-kb XbaI–HindIII DNA fragment containing part of the coding region of the su(Hw) protein was obtained by digestion of su(Hw) CDNA-1 [Parkhurst et al. 1988], filled in at the XbaI site, and cloned into the blunt-ended BamHI–HindIII sites of the pATH2 vector [Dieckmann and Tzagoloff 1985]. Protein synthesis was induced with 3-β-indolacrylic acid, and cells were harvested and lysed with lysozyme and 8 M urea. The pro-
tein extract was incubated at room temperature for 1 hr, and the insoluble material was removed by centrifugation. The supernatant was dialyzed against 100 mM KCl, 50 mM ZnCl₂, 5 mM MgCl₂, 1 mM dithiothreitol (DTT), 1 mM phenylmethylsulfonyl fluoride (PMSF), 10% glycerol, and 10 mM Tris-HCl (pH 7.6), containing decreasing concentrations [4, 2, and 0 mM] of urea. Insoluble material was eliminated by centrifugation, and the supernatant was used for binding studies.

A full-length cDNA clone, including the entire coding region of the su(Hw) gene (cDNA-1; Parkhurst et al. 1988), was cloned into the EcoRI–SalI sites of the pRMHa-3 vector containing the metallothionein promoter (Bunch et al. 1988). The resulting plasmid, denominated pTC-suHw, was cotransformed into Drosophila-made cells with plasmid pHGC0, which contains a DHFR gene (Bourouis and Jarry 1983), using the calcium phosphate procedure. Transformed cells were selected with 75 μg of protein in Freund’s incomplete adjuvant.

White female rabbits were immunized intradermally using the fusion protein. This preparation was used in all the experiments described in which Drosophila-made protein is mentioned.

Antibody production

E. coli cells overexpressing the trpE-su(Hw) fusion protein were lysed in a French press, diluted with 2 × Laemmli buffer, and boiled for 10 min, and the insoluble material was induced by centrifugation, and nuclear protein extracts were made following the procedure of Heberlein et al. (1985), with minor modifications; EDTA and EGTA were omitted, and 50 μM ZnCl₂ and 5 mM DTT were added to all solutions. The protein extracts were further purified by chromatography on DEAE–Sephadex CL-6B. Protein retained in the column was eluted with a step gradient using 0.2, 0.4, 0.6, 0.8, and 1.0 M KCl successively. Fractions [0.5 ml] were assayed for the presence of su(Hw) protein by measuring binding activity using gel mobility shift assays and on Western blots using affinity-purified su(Hw) antibodies. Both types of assays gave consistent results. Fractions containing binding activity were pooled and stored at ~70°C. This procedure resulted in 10-fold purification of the su(Hw) protein, and this preparation was used for the sequence analysis of the su(Hw)-binding site, Dr. L. Goldstein for a gift of plasmid pRMHa-3, Dr. R. Lebovitz for help with the immunofluorescence experiments, and Dr. P. Smith for suggestions on the manuscript. This work was supported by U.S. Public Health Service award GM-35463 and American Cancer Society grant NP-546.

Acknowledgments

We would like to thank Dr. J. Bishop for help with the sequence analysis of the su(Hw)-binding site, Dr. L. Goldstein for a gift of plasmid pRMHa-3, Dr. R. Lebovitz for help with the immunofluorescence experiments, and Dr. P. Smith for suggestions on the manuscript. This work was supported by U.S. Public Health Service award GM-35463 and American Cancer Society grant NP-546.

References

Biessman, H. 1985. Molecular analysis of the yellow gene [y] region of Drosophila melanogaster. Proc. Natl. Acad. Sci. 82: 7369–7373.

Bolhmann, D., W. Keller, T. Dale, H.R. Scholer, G. Tebb, and J.W. Mattaj. 1987. A transcription factor which binds to the enhancers of SV40, immunoglobin heavy chain and U2 snRNA genes. Nature 325: 268–272.

Bourouis, M. and B. Jarry. 1983. Vectors containing a prokaryotic dihydrofolate reductase gene transform Drosophila cells to methotrexate-resistance. EMBO J. 2: 1099–1104.

Bunch, T.A., Y. Grinblat, and L.S.B. Goldstein. 1988. Characterization and use of the Drosophila metallothionein promoter in cultured Drosophila melanogaster cells. Nucleic Acids Res. 16: 1043–1061.

Campuzano, S., L. Balcells, R. Villares, L. Carramolino, L. Garcia-Alonso, and J. Modeell. 1986. Excess function hypothesis. The effect of mutations caused by gypsy and capsa insertions within structural genes of the achaete-scute locus of Drosophila. Cell 44: 303–312.

Chia, W., G. Howes, M. Martin, Y.B. Meng, K. Moses, and S. Tsubota. 1986. Molecular analysis of the yellow locus of Drosophila. EMBO J. 5: 3597–3605.
Cohen, L.H. and B.V. Gotchel. 1971. Histones of polytene and nonpolytene nuclei of Drosophila melanogaster. J. Biol. Chem. 246: 1841–1848.

Dieckmann, C.L. and A. Tzagoloff. 1985. Assembly of the mitochondrial membrane system. J. Biol. Chem. 260: 1513–1520.

Fingerman, D.J. and D.H. Fawcett. 1986. Transposable elements in Drosophila melanogaster. In Oxford surveys in eukaryotic genes [ed. N. Maclean], vol. 3, pp. 1–62. Oxford University Press, Oxford.

Fried, M. and D.M. Crothers. 1981. A gel electrophoresis method for quantifying the binding of proteins to specific DNA regions: Application to components of the Escherichia coli lactose operon regulatory system. Nucleic Acids Res. 9: 3047–3060.

Garner, M.M. and A. Revzin. 1981. A gel electrophoresis method for quantifying the binding of proteins to specific DNA regions: Application to components of the Escherichia coli lactose operon regulatory system. Nucleic Acids Res. 9: 3047–3060.

Heberlein, U., B. Englund, R. Tjian. 1985. Characterization of Drosophila transcription factors that activate the tandem promoters of the alcohol dehydrogenase gene. Cell 41: 965–977.

Heller, R.A., E.R. Shelton, V. Dietrich, S.C.R. Elgin, and D.L. Brutlag. 1986. Multiple forms and cellular localization of Drosophila DNA topoisomerase II. J. Biol. Chem. 261: 8063–8069.

Hope, I.A., S. Mahadevan, and K. Struhl. 1988. Structural and functional characterization of the short acidic transcriptional activation region of yeast GCN4 protein. Nature 335: 635–640.

Klug, A. and D. Rhodes. 1987. ‘Zinc fingers’: A novel protein motif for nucleic acid recognition. Trends Biochem. Sci. 12: 464–469.

Levis, R., K. O’Hare, and G.M. Rubin. 1984. Effects of transposable element insertions on RNA encoded by the white gene of Drosophila. Cell 38: 471–481.

Ma, J. and M. Ptashne. 1987. Deletion analysis of GAL4 defines two transcriptional activating segments. Cell 48: 847–853.

Maniatis, T., E.F. Fritsch, and J. Sambrook. 1982. Molecular cloning. A laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

Marlor, R.L., S.M. Parkhurst, and V.G. Corces. 1986. The Drosophila melanogaster gypsy transposable element encodes putative gene products homologous to retroviral proteins. Mol. Cell. Biol. 6: 1129–1134.

Maxam, A.M. and W. Gilbert. 1980. Sequencing end-labeled DNA with base-specific chemical cleavages. Methods Enzymol. 65: 499–560.

McDonough, A.A. 1984. Use of antibodies in the study of Na+, K+ -ATPase biosynthesis and structure. Curr. Top. Membr.

Mizrokhi, L.J., S.G. Georgieva, and Y.V. Ilyin. 1988. Drosophila mobile element jockey is similar to LINEs and is transcribed by RNA polymerase II from the internal promoter. Cell 54: 685–691.

Parkhurst, S.M. and V.G. Corces. 1985. forked, gypsy, and suppressors in Drosophila. Cell 41: 429–437.

Parkhurst, S.M. and V.G. Corces. 1985. forked, gypsy, and suppressors in Drosophila. BioEssays 5: 52–57.

Parkhurst, S.M., D.A. Harrison, M.P. Remington, C. Spana, R.L. Kelley, R.S. Coyne, and V.C. Corces. 1988. The Drosophila su(Hw) gene, which controls the phenotypic effect of the gypsy transposable element, encodes a putative DNA binding protein. Genes Dev. 2: 1205–1215.

Peifer, M. And W. Bender. 1986. The anterothorax mutations of the bithorax complex. EMBO J. 5: 2293–2303.

Roeder, G.S. and G.R. Fink. 1983. Transposable elements in yeast. In Mobile genetic elements [ed. J.A. Shapiro], Academic Press, New York. pp. 300–328.

Rutledge, B.J., M.A. Mortin, E. Schwarz, D. Thierry-Mieg, and M. Meselson. 1988. Genetic interactions of modifier genes and modifiable alleles in Drosophila melanogaster. Genetics 119: 391–397.

Searles, L. and R.A. Voelker. 1986. Molecular characterization of the Drosophila vermilion locus and its suppressible alleles. Proc. Natl. Acad. Sci. 83: 404–408.

Vrana, K.E., M.E.A. Churchill, T.D. Tullius, and D.B. Brown. 1988. Mapping functional regions of the transcription factor TFIIA. Mol. Cell. Biol. 8: 1684–1696.

Winston, F., K.F. Durbin, and G.R. Fink. 1984a. The SPT7 gene is required for normal transcription of Ty elements in S. cerevisiae. Cell 39: 675–682.

Winston, F., D.T. Challeff, B. Valent, and G.R. Fink. 1984b. Mutations affecting Ty-mediated expression of the HIS4 gene in Saccharomyces cerevisiae. Genetics 107: 179–197.

Zachar, Z., D. Davison, D. Garza, and P.M. Bingham. 1985. A detailed developmental and structural study of the transcriptional effects of insertion of the copia transposon into the white locus Drosophila melanogaster. Genetics 111: 495–515.
The Drosophila melanogaster suppressor of Hairy-wing protein binds to specific sequences of the gypsy retrotransposon.

C Spana, D A Harrison and V G Corces

Genes Dev. 1988, 2:
Access the most recent version at doi:10.1101/gad.2.11.1414