The Polycomb-group gene, extra sex combs, encodes a nuclear member of the WD-40 repeat family

Thomas Gutjahr, Erich Frei, Carl Spicer, Stefan Baumgartner, Robert A.H. White and Markus Noll

Institute for Molecular Biology, University of Zürich, Winterthurerstrasse 190, CH-8057 Zürich, 2Friedrich Miescher Institut, Ciba-Geigy, CH-4002 Basel, Switzerland and 3Department of Anatomy, University of Cambridge, Downing Street, Cambridge CB2 3DY, UK

We have delimited the extra sex combs (esc) gene to <4 kb that include a single transcript and are able to rescue both the maternal and zygotic esc phenotypes. Several mutations have been identified within the esc transcript. In agreement with earlier genetic studies, esc is expressed maternally and its product is most abundant during the early embryonic stages. It encodes a protein of the WD-40 repeat family, which localizes predominantly to the nucleus. During germ band extension, it is expressed in a stereotypic pattern of neuroblasts. We propose a model in which Esc is recruited by gap proteins both to act as a corepressor that competes with the TAF180 coactivator to block transcription and also to mediate the transition to permanent repression by Polycomb-group proteins.

Keywords: extra sex combs gene silencing/homeosis/Polycomb group/WD-40 repeat

Introduction

In Drosophila development, the expression patterns of the homeotic genes are stably maintained by a mechanism that requires two sets of genes. On the one hand, the trithorax-group (trx-G) genes are responsible for sustaining the active state of expression; on the other hand, the Polycomb-group (Pc-G) genes are required for the maintenance of the repressed state. The underlying molecular mechanism is largely unknown, but recent evidence points to a role for the higher order chromatin structure (for a review see Orlando and Paro, 1995). Several Pc-G gene products have been shown to be associated with specific regions of the salivary gland chromosomes (Zink and Paro, 1989; Zink et al., 1991; DeCamillis et al., 1992; Franke et al., 1992; Rastelli et al., 1993) and the protein product of the Polycomb (Pc) gene covers regulatory regions of inactive homeotic genes (Orlando and Paro, 1993). In nuclear extracts, the Pc protein is present in large multiprotein complexes together with other Pc-G products (Franke et al., 1992). In addition, the Pc protein shows sequence similarity to HP1, a protein associated with heterochromatin (James and Elgin, 1986; Paro and Hogness, 1991). HP1 is encoded by the Su(var)205 gene which regulates position-effect variegation (Eissenberg et al., 1990), a phenomenon which is believed to have its basis in the control of heterochromatin assembly (reviewed by Reuter and Spierer, 1992). Furthermore, a regulatory element controlled by Pc and polyhomeotic proteins can induce the variegation of neighbouring transcription units (Fauvarque and Dura, 1993).

A model for the establishment of stable states of homeotic gene expression is emerging (Gaunt and Singh, 1990; Gould et al., 1990; Paro, 1990, 1993). Early in development, the patterns of homeotic gene expression are dependent on activation and repression by the products of the segmentation genes. However, these regulators are expressed only transiently, but at ≈6 h of development the gene expression states that they have set up are frozen by the activity of the Pc-G genes. Genes that are repressed by the segmentation genes are recognized as targets by the Pc-G genes and may be packaged as condensed chromatin. This packaging results in the stable and heritable silencing of gene activity, possibly by excluding transcriptional activators of the trx-G.

A major question in this scheme is how the Pc-G machinery is targeted to repressed homeotic genes. It has been proposed that early repressors, such as gap gene products, could interact directly with particular Pc-G products tethering them to specific regulatory regions (Zhang and Bienz, 1992). This could provide the signal for subsequent packaging by Pc-G proteins of chromatin into a repressed state (Bienz, 1992). One candidate for a Pc-G gene with a special role in the initiation of the maintenance complex is extra sex combs (esc) (Silber, 1942; Struhl, 1981). Although a typical Pc-G gene, in the sense that esc mutations cause the ectopic expression of homeotic genes at ≈6 h of development, esc has some distinguishing characteristics. A temperature-sensitive esc mutation demonstrates that the critical requirement for esc function is early in embryogenesis during the extended germ band stage (Struhl and Brower, 1982). In contrast, both the Pc and Enhancer of zeste [E(z)] genes appear to be continuously required during development (Lewis, 1978; Struhl, 1981; Rastelli et al., 1993). Unlike mutations in several other Pc-G genes (Breen and Duncan, 1986; Dura et al., 1987; Martin and Adler, 1993), the phenotype of esc mutations appears to depend exclusively on mis-regulation of the homeotic genes; in particular, esc is not involved in the regulation of engrafted (en) (Moazed and O’Farrell, 1992) or gap genes (Pelegri and Lehmann, 1994).

We are interested in defining further the role of esc. Here we present the sequence of the esc gene and an analysis of its expression. Its early expression, nuclear localization and sequence homology to the WD-40 repeat family are consistent with a specific early role for esc in forging a link between the regulators encoded by segmentation genes and the Pc-G machinery.
Results

Rescue of the esc− phenotype by a 4 kb genomic fragment

Previously, we have isolated and localized the esc gene to a DNA region of 12 kb that was able to support the normal development of esc− embryos derived from transheterozygous esc− mothers carrying this DNA as a transgene (Frei et al., 1985a,b). Within this 12 kb region shown to include the esc gene, Northern blot analysis detected a 1.8 kb RNA with a developmental profile consistent with the known requirements for the esc product (Frei et al., 1985b). To corroborate that this transcript is derived from esc, the genomic DNA was first reduced to a 6.5 kb SalI fragment (esc−E) and subsequently to a 3.9 kb SalI-XbaI fragment (Figure 1A; esc−F). Both genomic fragments encode the entire 1.8 kb mRNA and rescue esc− offspring from transgenic mothers carrying no functional endogenous esc gene (see Materials and methods). The 3.9 kb genomic fragment includes only 0.26 kb of upstream and 1.4 kb of downstream sequences (Figure 1A and B), yet it completely rescues, as esc−F transgene, both the maternal esc phenotype of embryos (Struhl, 1981) and the zygotic esc phenotype of adults (Slifer, 1942; Tokunaga and Stern, 1965).

Wild-type and mutant esc sequences

Several cDNAs of the 1.8 kb mRNA were isolated from cDNA libraries of mature oocytes and 0–4 h old embryos. Comparison of nearly full-length cDNAs with the corresponding genomic DNA sequence revealed four exons interrupted by three short introns of 72, 62 and 364 bp (Figure 1A and B). Figure 1B shows the genomic DNA sequence of the 5′ portion of the 3.9 kb SalI-XbaI fragment used as the esc transgene and the translation of the longest open reading frame of the four exons into the putative 425 amino acid protein. This protein is composed of several domains, as illustrated schematically in Figure 1C and discussed below. The transcriptional start, which has been mapped by the RACE technique (Frohman, 1990), does not appear to be defined precisely but occurs at several sites scattered over a region of ~40 bp. The longest transcripts are initiated at a site located only 260 bp downstream from the SalI site (Figure 1B), resulting in an mRNA length of 1.7 kb without poly(A) tail, which is consistent with its size derived from Northern blot analysis.

In addition, we demonstrated directly that this transcript is the esc mRNA by DNA sequence analysis of three esc mutant alleles. Sequencing all four exons of esc1 (Slifer, 1942), esc2 and esc6 (Struhl, 1981) revealed two point mutations in esc1 and one point mutation each in esc2 and esc6 (Figure 2A). In addition, a more complex mutational event was apparent from the esc2 sequence that replaced 10 bp of the wild-type sequence by 12 bp in the mutant (Figure 2B). Since the parental stock from which the viable esc1 allele originated spontaneously is no longer available, we do not know which of the two point mutations detected in esc1 generates the mutant phenotype, although it is probably caused by the change from Leu to Arg at position 240 rather than by the conservative amino acid substitution of Tyr by Phe at position 64 (see also below).

Similarly, two mutations were found in the esc2 allele that arose spontaneously on the CyO balancer chromosome (Struhl, 1981). The resulting mutant protein carries an amino acid substitution (Gln to Lys) at position 184, and the 22 C-terminal amino acids of the wild-type protein are replaced by a different stretch of 48 amino acids in the Esc2 protein. This C-terminal alteration might have arisen by two consecutive slippings by 18 and 2 bp of the DNA polymerase during replication in the manner illustrated in Figure 2B. The esc2 mutant phenotype is caused by this second mutational event rather than by the point mutation because the Gln to Lys substitution is also present in CyO, esc+ chromosomes, and hence is a polymorphism (Figure 2A). Finally, the most severe mutation was found as a single point mutation in esc6. In this case, the conversion of the first intron’s splice donor site from GT to GA gives rise to an esc6 mRNA that retains the first intron, as verified by the isolation and sequencing of an esc6 cDNA (Figure 2A). Failure to splice the first intron would result in the premature release of a truncated Esc protein of 24 amino acids caused by the presence of a stop codon after 12 nucleotides of the first intron. It is reasonable to assume that this protein retains no wild-type function and, accordingly, that esc6 is a null allele.

The Esc protein includes tandemly repeated WD-40 domains

A database search for proteins homologous to the amino acid sequence of wild-type Esc protein revealed that the fourth exon of esc encodes six tandemly repeated C-terminal domains (Figures 1B and C and 3) of the so-called β-transducin-like or WD-40 repeat (Fong et al., 1986; Dalrymple et al., 1989; Simon et al., 1991; van der Voorn and Ploegh, 1992; Neer et al., 1994). The WD-40 domain is found as a tandem repeat in a variety of proteins, some of which are indicated below those of Esc in Figure 3. Among these, the β-subunits of G proteins (as, for example, β-transducin) consist entirely of WD-40 domains that are highly conserved between Drosophila and man. This domain is not restricted to membrane-associated proteins, but also occurs as a tandem repeat in the C-terminal portion of nuclear proteins like Groucho (Gro; Hartley et al., 1988), dTAF680 (Dynlacht et al., 1993), CDC4 (Yochem and Byers, 1987; Choi et al., 1990) and Tup1 (Williams and Trumbly, 1990; Zhang et al., 1991b; Komachi et al., 1994). The Tup1 protein has been shown to mediate the repression of several specific gene sets in yeast, including the sets of a-specific, haploid-specific and glucose-repressible genes (Williams et al., 1991; Keleher et al., 1992). As esc is required for the repression of homeotic genes (Struhl, 1981, 1983; Struhl and Akam, 1985; Struhl and White, 1985), the finding of sequence similarity to the yeast repressor Tup1 is particularly interesting.

The WD-40 domain consists of two elements (A and B) that are separated by short stretches of amino acids that vary in length and sequence (van der Voorn and Ploegh, 1992). In many WD-40 repeats, the A element has diverged considerably while the B element is usually well conserved. As evident from Figure 3, this is also the case for the first and the last two WD-40 domains of Esc, although some of the characteristic amino acids of the A element are conserved. The observation that one of the two point mutations found in esc1 alters a Leu to an Arg at a position in the B element of the third WD-40 domain that is well conserved (Figures 2A and 3) further supports
Fig. 1. The esc gene and its protein product. (A) Map of the esc gene and two esc+ transgenes. At the top, an EcoRI restriction map of the chromosomal region that includes the esc gene is shown. The scale (in kb) refers to positions within a particular chromosomal walk of which the cloned DNA segment of recombinant phage C2.1 is indicated (Frei et al., 1985a). Two esc+ transgenes derived from C2.1 are shown below. The 12 kb region of the C2.1.5 (Car4) construct, shown previously to rescue esc mutants (Frei et al., 1985a), has been reduced further to 3.9 kb of the esc+ F transgene in this study. A map of the esc transcript and its three introns is illustrated at the bottom. Abbreviations of restriction sites: B, BamHI; P, PstI; R, EcoRI; S, SalI; X, XbaI. (B) DNA sequence of the esc gene and corresponding amino acid sequence of its protein product. The genomic esc DNA sequence from the SalI site, located 0.26 kb upstream of the transcriptional start to 0.04 kb downstream of the poly(A) addition site, is shown. Sequences present in the mature esc mRNA are shown in capitals, while upstream, intron and downstream sequences are indicated in lower-case letters. The number of nucleotides (shown in both margins) refers to the transcriptional start site as position +1. The amino acid sequence of the putative Esc protein is abbreviated in a single-letter code above the DNA sequence and is numbered only in the right margin. The 5' and 3' ends of the sequenced esc+ cDNAs, c322.2, c323.2 and c323.4, the 5' end of an esc+ cDNA and the poly(A) addition site are indicated by vertical arrows. The canonical poly(A) addition site is underlined. (C) Schematic representation of Esc protein domains. The Esc protein consists of 425 amino acids and includes a stretch of acidic amino acids (---) at its N-terminus (1–22), followed by a Ser/Thr-rich sequence (23–42), a nuclear localization signal (NLS) and basic (+) region (43–63) and six tandem repeats of WD-40 domains at its C-terminus (123–425).
in embryos derived from esc<sup>6</sup>/Df(2L)esc<sup>10</sup> mothers and homozygous esc<sup>6</sup> fathers. In wild-type embryos, shown in Figure 4, the Esc protein is localized to the nucleus throughout embryogenesis. The Esc protein, which is abundant in unfertilized eggs (Figure 4A), is increasingly taken up by the accumulating nuclei during cleavage and syncytial blastoderm stages (Figure 4B–D), and is found in all nuclei of cellular blastoderm embryos (Figure 4E). The concentration of uniformly distributed Esc protein decreases during gastrulation (results not shown). During germ band elongation, Esc appears to become restricted to certain nuclei (Figure 4F), which is particularly obvious from a stereotypic pattern of labelled neuroblasts (Figures 4G). A segmentally repeated pattern of Esc-positive nuclei persists in the central nervous system during the extended germ band stage (Figures 4H and 5A and B) and germ band retraction (Figure 4I). Subsequently, this pattern disappears and only a few nuclei in the brain retain Esc protein (Figure 4K).

**Specific zygotic expression of esc in neuroblasts and in the brain**

It is surprising that Esc is detected only in specific neuroblasts after germ band elongation because the wild-type cuticular phenotype depends on the presence of a functional Esc protein between cellular blastoderm and the extended germ band stage (Struhl and Brower, 1982), implying its presence in the epidermis. Apparently, Esc occurs in epidermal cells in amounts that are much smaller than those in esc-expressing neuroblasts, and hence fails to be detected by the monoclonal antibody (Figure 5C and D). Since Esc seems to accumulate de novo in neuroblasts, it is probably the result of zygotic esc activity. At very late stages, esc is again expressed zygotically in specific regions of the brain (Figures 4K and 5E and F).

**Perdurance of maternal Esc protein and paternal rescue**

Although Esc protein is found in much lower amounts in neuroblasts of esc<sup>-</sup> embryos derived from heterozygous mothers carrying a mutant allele and one or two wild-type copies of esc (results not shown; see Materials and methods), such embryos develop into viable adults with an esc leg phenotype (Slifer, 1942; Struhl, 1981). This result confirms earlier reports that there is only an early absolute requirement for Esc (Struhl and Brower, 1982), which is met by the perduring maternal Esc protein (Struhl, 1981). Further, it supports the view that the accumulation of Esc in neuroblasts of wild-type embryos probably results from zygotic esc activity (Figure 5A and B). Consistent with this conclusion, embryos derived from esc<sup>-</sup> mothers but carrying two paternal copies of the esc<sup>-</sup> gene show the zygotic expression of Esc protein in neuroblasts in a pattern similar to that observed in embryos derived from wild-type females (results not shown). Although such embryos show a complete paternal rescue of their cuticular phenotype (Struhl, 1981), we again failed to detect esc expression in the epidermis, confirming that relatively low amounts of Esc are sufficient for cuticular rescue.
Discussion

We have definitively mapped the esc locus to a single 1.8 kb transcript. We have shown that a 3.9 kb genomic fragment, encompassing the transcription unit of the 1.8 kb transcript, is sufficient to rescue esc mutants. Furthermore, the esc mutants esc¹, esc² and esc⁶ are all associated with sequence disruptions that affect the open reading frame encoded by the 1.8 kb transcript. This esc transcript encodes a predicted protein of 425 amino acids, of which
Esc is a nuclear protein with WD-40 domains

the C-terminal 310 amino acids contain six repeats of the WD-40 or β-transducin repeat motif. This motif is found in a number of functional classes of proteins with a variety of subcellular localizations (Fong et al., 1986; Duronio et al., 1992; van der Voorn and Ploegh, 1992; Neer et al., 1994). The WD-40 protein motif is thought to function as a protein–protein or protein–nucleic acid interaction interface, and secondary structure predictions suggest a series of β-strands separated by turns. To study the subcellular localization of the esc product, we raised a specific monoclonal antibody which was used to demonstrate that the Esc protein is localized in nuclei. The characteristics of nuclear localization and WD-40 repeats are shared with a small number of previously identified
proteins. These include Drosophila TAF180 (Dynlacht et al., 1993), Drosophila Groucho (Hartley et al., 1988), yeast Tup1 (Williams and Trumbly, 1990; Komachi et al., 1994) and Arabidopsis COP1 (Deng et al., 1992). Of these, the Tup1 protein has been the most studied. It is particularly intriguing as a relative of Esc because it acts as a corepressor of several gene sets in yeast and so suggests some interesting functional parallels with Esc. Tup1 and Esc share a common structure of C-terminal WD-40 repeats, but outside this region we do not observe significant sequence similarity other than a region of high Ser/Thr content in Esc compared with three in Tup1.

A model of Esc acting as a corepressor and mediating transition to permanent repression by Pc-G proteins

Elegant genetic studies by Struhl (1981, 1983) have shown that the wild-type product of esc is required to ensure the proper expression of the homeotic genes of the bithorax (BX-C) and Antennapedia complexes (ANT-C). In the absence of Esc, these genes are indiscriminately expressed in segments in which they are normally repressed. Subsequent experiments confirmed that, in esc⁻ embryos, the BX-C gene Ultrabithorax (Ubx) is indeed expressed in segments both anterior and posterior to its wild-type expression domain, although its initial activation is not affected and ectopic expression does not occur until the extended germ band stage (Struhl and Akam, 1985). These results are consistent with a transient requirement for the Esc protein after germ band extension (Struhl and Brower, 1982) to repress Ubx as well as other genes of the BX-C and ANT-C (Struhl, 1981, 1983). The idea that esc might serve a specific early role is also supported by our analysis of protein expression showing the early uniform nuclear accumulation of Esc, which declines following gastrulation.

We have shown here that the Esc protein is located in the nucleus (Figures 4 and 5) and contains six tandemly repeated C-terminal WD-40 domains (Figures 1C and 3). This suggests a possible analogy to the yeast protein Tup1, which also includes tandemly repeated WD-40 domains in its C-terminal portion (Williams and Trumbly, 1990; Zhang et al., 1991b). The analogy between Esc and Tup1 may be extended even further because Tup1 protein has been shown to act as a corepressor when recruited by DNA binding proteins, such as the homeodomain protein

Fig. 5. Expression of esc in neuroblasts and in the brain. Anterior (A and C) or posterior portions (B and D) of stage 11 embryos, unfolded along the amnioserosa and stained with anti-Esc monoclonal antibodies, are shown in focus for the neuroblast (A and B) or the epidermal layer (C and D). Note that the label visible in (C) and (D) derives from the neuroblast layer directly below. Stage 13 embryos stained for esc RNA by a digoxigenin-labelled probe are shown as a lateral view with dorsal side up (E) or as a dorsal view (F). All embryos are orientated with their anterior to the left.
Esc is a nuclear protein with WD-40 domains

Fig. 6. Model of Esc protein acting as a corepressor by interfering with basal transcription machinery during the assembly of repressed chromatin by Pc-G proteins. It is proposed that Esc blocks transcription by competing through its WD-40 domains with those of TAF180 for interaction with other components of the TFIIID complex or the basal transcription apparatus. The model shows repression of the homeotic gene Ubx to illustrate a general mechanism. Esc is recruited to act as a competitive corepressor by the binding of gap gene products, e.g. Hb protein, to specific DNA sites. Esc may then catalyse the nucleation of a repressed chromatin structure that depends on the presence of Pc-G proteins but not of Esc. The model does not wish to imply that all six WD-40 domains of Esc interact with TFIIID.TAF180, the TFIIID complex lacking TAF180. For a detailed discussion of the model, see the text.

α2, to repress specific gene sets like those of the haploid-or a-specific genes (Johnson and Herskowitz, 1985; Goutte and Johnson, 1988; Kelche et al., 1988, 1992; Komachi et al., 1994). Although it is unclear how Tup1 represses the transcription of these genes, it has been shown that Tup1 interferes with the general transcription machinery (Herschbach et al., 1994; Komachi et al., 1994; Tzamarias and Struhl, 1994) and appears to induce a repressed state of chromatin structure (Cooper et al., 1994; reviewed in Roth, 1995). The interference of Tup1 with the assembly of an active transcription complex depends on the WD-40 domains (Komachi et al., 1994).

Similarly, we propose that Esc is recruited by specific enhancer binding proteins to repress the homeotic genes of the BX-C and ANT-C by interference with the assembly of the basal transcription machinery (Figure 6). During early embryogenesis, candidates for such enhancer binding proteins are the products of the gap genes (White and Lehmann, 1986; Harding and Levine, 1988; Irish et al., 1989; Reinitz and Levine, 1990; Bustinia and Bienz, 1993; Shimell et al., 1994). For example, in hunchback (hb) embryos, Ubx is derepressed anterior to its wild-type expression pattern (White and Lehmann, 1986) in a way similar to that observed in esc embryos (Struhl and Akam, 1985). However, Ubx is derepressed earlier in hb than in esc embryos, indicating that the initial repression of Ubx depends on Hb but not on Esc (Irish et al., 1989). Indeed, it has been shown that this early repression of Ubx is achieved by Hb binding to Ubx control elements, displacing pair-rule activator proteins from overlapping binding sites (Zhang et al., 1991a; Müller and Bienz, 1992; Qian et al., 1993). This type of early repression, mediated by Hb (or other gap proteins) through competitive binding, is short range and independent of esc and Pc, whereas the subsequent maintenance of the repressed state, termed silencing, acts over large distances and depends on both Hb and Pc-G proteins (Muller and Bienz, 1992; Zhang and Bienz, 1992; Simon et al., 1993; Chan et al., 1994).

Accordingly, repression depends on two different cis-regulatory elements, one that binds Hb which is required for early repression, and another, the Pc-G response element (PRE; Simon et al., 1993), through which Pc and Pc-G proteins are anchored to DNA, dictating heritable repression (Zink et al., 1991; DeCamillis et al., 1992; Chan et al., 1994; Müller, 1995). Although the initial high concentrations of Hb protein are able to repress Ubx in the absence of Esc, later, when its concentration declines, Hb may become dependent on Esc for efficient repression.

The direct target of Tup1, as well as of Esc, in the basal transcription machinery remains a mystery (Komachi et al., 1994). However, a possible clue is provided by the finding in Drosoptila that TAF180, one of the factors associated with the TATA binding protein (TBP) in the TFIIID complex, includes seven C-terminal WD-40 repeats (Figure 3; Dynlacht et al., 1993). We therefore propose that the corepressor activity of Esc is mediated through its WD-40 domains that compete with those of TAF180 for binding to other TAFs, TBP or other parts of the basal transcription apparatus, and thus displaces TAF180 from the TFIIID complex (Figure 6). This interaction of Esc with the basal transcription machinery might be stabilized by, or only be possible after, the binding of Esc to specific enhancer binding proteins such as Hb, ensuring that only specific gene sets are repressed by Esc. In such a model, Esc is a corepressor and is recruited by gap gene products bound to specific DNA control regions, the Esc-recruiting control elements (ERCEs), and acts antagonistically to TAFs that have been shown to be coactivators (Chen et al., 1994; Tjian and Maniatis, 1994).

After the extended germ band stage, Esc is no longer required for the repression of homeotic genes (Struhl and Brower, 1982). However, continued repression of these genes in the same regions is accomplished by the Pc-G genes (Jurgens, 1985) that are thought to alter chromatin structure to a repressed state (Gaunt and Singh, 1990; Gould et al., 1990; Paro, 1990, 1993). In analogy to the yeast Tup1 protein, which acts not only as a corepressor but also alters chromatin structure to a state favouring repression (Cooper et al., 1994), we propose that Esc mediates the transition to this repressed state. The function of Esc during this transitory stage is to inhibit transcription by competitive corepression and to interact with Pc-G proteins to stabilize their nucleation on the PRE (Figure 6). If we assume that only transcriptionally inactive chromatin can be assembled into a Pc-G protein-dependent inactive state of chromatin structure, it follows that the corepressor function of Esc is a prerequisite for its subsequent function in silencing. Thus, the role we propose for Esc is twofold. First, it inhibits transcription by bridging and locking a complex between a DNA-bound repressor and the basal transcription apparatus. This inhibition of transcription becomes crucial when the repressor concentration declines and is no longer able to repress activation by competitive binding (Muller and Bienz, 1992; Qian et al., 1993). Second, Esc protein may direct the assembly of the Pc-G proteins and stabilizes the interaction of one or several of these proteins with the PRE. The resultant tethering of the PRE to the ERCE (Figure 6) might provide the signal for the Pc-G protein-mediated assembly of silenced chromatin between the tethered elements. Once repression through Pc-G proteins has been established (Orlando and
Paro, 1993), this repressed state (and its inheritance) no longer depends on Esc but only on the continued presence of Pc-G proteins and is likely to involve a higher order chromatin structure (Paro, 1990, 1993).

Recent experiments have shown that silencing by a GAL4–Pc fusion protein of a reporter gene, under the control of GAL4 binding sites, requires the additional presence of a PRE and depends on Pc-G proteins as well as Esc, but not on Hb (Müller, 1995). At first sight, these results, which fail to distinguish between the roles of Esc and other Pc-G proteins, appear to imply that Esc, like the other members of the Pc-G proteins, is required for the elaboration of the Pc-G protein complex. However, if repression of transcription is a prerequisite for the assembly of chromatin with Pc-G proteins into an inactive state, as we propose, our model offers an alternative interpretation of Müller’s (1995) results. In this artificial situation, GAL4–Pc protein bound to GAL4 sites substitutes for the function of Hb bound to ERCE and recruits Esc directly or indirectly through the interaction with other Pc-G proteins that bind Esc. The assumption that GAL4–Pc recruits Esc is consistent with our model, which proposes that Esc permits the assembly of, and interacts with, Pc-G proteins. Subsequent silencing of chromatin depends on the additional presence of a PRE and endogenous Pc-G proteins (Müller, 1995) that permit the tethering of two regions by the interaction of GAL4–Pc bound to the GAL4 sites and Pc-G proteins bound to PRE.

Our model is also consistent with the observation that the initial ectopic activation of AbdB occurs in a pair-rule pattern in esc− or most severe Pc-G mutant embryos (Simon et al., 1992), indicating that ectopic activation results from a failure to maintain repression by competition with the normal pair-rule activation. Interestingly, this ectopic activation depends not only on the absence of Esc or Pc-G proteins but also on the presence of trx-G proteins, because the esc− cuticular phenotype is almost completely suppressed by the absence of the trx gene in esc− trx− embryos (Ingham, 1983). This finding is consistent with a requirement for the Trx protein and other trx-G proteins to maintain a transcriptionally active state of chromatin (Orlando and Paro, 1995) after the pair-rule activator proteins decrease below a threshold concentration.

Esc is not the only Drosophila protein that contains WD-40 repeats and may act as a co-repressor. A similar mechanism of repression has been observed to be mediated by the Gro protein, which is recruited by bHLH enhancer binding proteins to repress specific gene sets such as that of the achaete-scute complex (AS–C; Paroush et al., 1994). The interaction of Gro with Hairy-related bHLH proteins does not depend on its WD-40 domains (Paroush et al., 1994), which thus could bind to the basal transcription machinery, in agreement with our model (Figure 6).

In cells destined for the epidermal as opposed to the neural fate, the repression of the AS-C by Gro may be followed by stable silencing as part of the cellular commitment to epidermal development. Therefore, it would be interesting to know whether the specific gene sets repressed by Gro are repressed later by Pc-G genes as well, and it might be significant that Pc has been reported to bind in section 1B of polytene chromosomes, the locus of the AS-C (Zink and Paro, 1989). In such a case, the proposed mechanism of repression might be quite general, with different co-repressor proteins mediating the interaction between enhancer binding proteins and the general transcription machinery that leads to a permanent state of chromatin repression effected by Pc-G proteins.

**Materials and methods**

**General procedures**

Genomic and cDNA libraries were prepared from esc mutant and wild-type stocks in EMBL4 and Agt10, respectively. Genomic and cDNA clones were isolated and sequenced according to standard procedures, as described previously (Frei et al., 1985a; Frigerio et al., 1986). The S′ RACE technique (Frohman, 1990) was applied to poly(A)+ RNA from 0.4 h old embryos using the S′ Amplifier RACE kit and following the instructions of Clontech.

**Preparation of fly stocks carrying esc− transgenes**

The DNA region, which is sufficient for the rescue of esc− embryos derived from esc− mothers and hence includes the entire esc gene, was reduced in two steps from an earlier rescue construct C2.1.5 (Card; Frei et al., 1985a; Figure 1A). First, a 6.5 kb Xbal fragment, derived from the insert of phage C2.1 of a genomic walk (Frei et al., 1985a; Figure 1A), was subcloned into the Xbal site of the P-element vector cp20.1 carrying the ry+ gene (Ogg et al., 1985) to generate the escE construct. From this construct, most of the esc upstream sequences were eliminated by digestion with SalI and religation to obtain the escF construct, which includes only ~260 bp of esc upstream sequence (Figure 1A). Both constructs were used in germline transformations (Spradling and Rubin, 1982; Karess and Rubin, 1984) to generate several independent transgenic lines of P[esc+ E ry+] or P[esc− F ry+] flies, carrying the esc− E or esc− F transgene. Since both transgenes were able to rescue both the maternal and zygotic esc phenotypes, only homozygous P[esc− F ry+] lines were maintained. Subsequently, the P-element on the third chromosome of one of the P[esc− F ry+] stocks was remobilized by a cross with flies carrying the Δ2.3 P-element transposase on a w+ transposon with hobo ends (Tn7.1) on the second chromosome (kindly provided by Brian Calvi and Bill Gelbart), to produce a stock that carries two esc− genes on the CO fragment balancer chromosome. By crossing such flies with esc− mutants, an esc−/CyO, P[esc− F ry+] stock was derived.

**Paternal rescue and perdurance of maternal Esc protein**

Paternally rescued embryos were obtained from homozygous esc− mothers and esc−/CyO, P[esc− F ry+] fathers. The homozygous esc− embryos produced by this cross served as a negative control for staining with the Esc monoclonal antibodies when the zygotic Esc protein appears in the paternally rescued embryos.

The perdurance of maternal Esc protein was examined in embryos obtained from two crosses in which the mothers carried one or two esc− genes. Either Δ[2L]Prl/CyO flies were crossed inter se or esc−/CyO, P[esc− F ry+] virgins were crossed with homozygous esc− males. A quarter of the embryos from the first cross were homozygous for Δ[2L]Prl and thus lacked esc and pdr that were uncovered by the deficiency. They could be distinguished from prd− embryos on the basis of their segmentation phenotype. Half of the embryos from the second cross were homozygous for esc− and showed disappearance of maternal Esc protein.

**Preparation of monoclonal antibodies against Esc antigen and immunocytochemical staining of embryos**

The Esc antigen used for immunization was expressed in Escherichia coli from the 0.37 kb BsrBI–BamHI fragment of c323.2 subcloned into the BamHI site of the pAR3040 vector (Studier and Moffat, 1986) by BsrBI and subsequent blunt-end ligation of the filled ends. This esc cDNA fragment spans the region from the BsrBI site immediately preceding the ATG initiation codon of Esc (Figure 1A) to the BamHI site generated by the splicing of exon 3 to exon 4 (Figure 1A and B) and thus encodes the N-terminal portion of Esc truncated immediately before the C-terminal WD-40 domains (Figure 1C). Expression and purification of this Esc antigen was as described previously (Bopp et al., 1989). Balb/c mice were immunized intraperitoneally with 50 µg Esc antigen in complete Freund’s adjuvant and boosted twice intraperitoneally with 50 µg in incomplete Freund’s adjuvant. At 5 days prior to fusion, a further intraperitoneal boost was given with 50 µg Esc antigen in PBS. Fusion supernatants were screened on protein blots of lysates of bacteria carrying the esc BsrBI–BamHI fragment expressed as a fusion protein.
in pGEX-2T (Smith and Johnson, 1988), i.e. in a different fusion protein background than was used for immunization. Only one monoclonal antibody, E53.1, showed binding to both the pAR and pGEX Esc fusion proteins. Using this monoclonal antibody, we have been unable to detect Esc protein on Western blots of Drosophila embryonic proteins.

Embryos were collected, fixed and stained for Esc protein essentially as described (Gutjahr et al., 1993). The anti-Esc monoclonal antibody was used without preabsorption at a 1:10–1:20 dilution (culture supernatant) or at a 1:300 dilution (ascites fluid). As secondary antibody a 1:300 diluted and preabsorbed horse anti-mouse antibody conjugated to biotin (Vector) was used.

In situ hybridization to whole-mount embryos with a digoxigenin-labelled esc cdNA probe [c3234 (Figure 1B) without untranslated leader] was performed as described previously (Tautz and Pfeifle, 1989).

For photography by Nomarski optics, the embryos were mounted in glycerol as described previously (Gutjahr et al., 1993).

Fly stocks
The main fly stocks used were: (i) P[C2.1.5(Car4) esc+]/+; In(2)Li, esc1 c splpd/esc6 b cn bw; (ii) P[C2.1.5(Car4) esc+]/+; CyO, esc-1 dpw/esc6 b cn bw; (iii) esc/esc; P[escF r y+] ry+P(escF r y+)/P[escF r y+]; (iv) esc/esc; P[escF r y+] ry+P[escF r y+]; (v) Df(2)L.2P[Ubx];CyO; (vi) w; Tn70.1 (H[w+] Y2.3)/CyO; and (vii) y; w; Ly ry+P/TM3, Sb ry+.

GenBank accession number
The gene accession number for the esc gene is L41867.

Acknowledgements
We are grateful to Maya Burri for considerable help in the early part of this project and to Fritz Ochsbein for his expert photographic and art work. We thank Patrick Spielberg for help in antigen preparation, Peter Lawrence for the original esc- and esc6 stocks (Struhl, 1981), Rolf Näthiger for the esc- stock (Slier, 1942), and Brian Calvi and Bill Gelbart for the Tn7.1 stock. This work was supported by the Swiss National Science Foundation grants 3.348-0.86 and 31-26652.89, the Kantons Basel and Zürich, and the Wellcome Trust.

References
Duronio,R.J., Gordon,J.I. and Boguski,M.S. (1992) Comparative analysis of the β transducin family with identification of several new members including WP1, a nonessential gene of Saccharomyces cerevisiae that is divergently transcribed from NMT1. Protein, 13, 41–56.

Esc is a nuclear protein with WD-40 domains.

Received on May 3, 1995; revised on June 21, 1995.