The EMBO Journal vol.13 no.2 pp.400–406, 1994

Compatibility between enhancers and promoters determines the transcriptional specificity of *gooseberry* and *gooseberry neuro* in the *Drosophila* embryo

Xuelin Li and Markus Noll

Institute for Molecular Biology II, University of Zürich, Winterthurerstrasse 190, CH-8057 Zürich, Switzerland

Communicated by M.Noll

The two *Drosophila* genes *gooseberry* (*gsb*) and *gooseberry neuro* (*gsbn*) are closely apposed and divergently transcribed. While *gsb* is a segment-polarity gene and mainly expressed in the epidermis, *gsbn* is expressed in the central nervous system. An intriguing question is how their transcriptional specificity arises. Here we show that different non-overlapping enhancer or upstream control elements drive the specific expression of *gsb* and *gsbn*. Specificity of these enhancers for their genes is achieved by their inability to activate transcription in combination with the heterologous promoter of the other gene. These results therefore suggest that compatibility between the enhancer and its cognate promoter is a mechanism ensuring transcriptional specificity.

Key words: enhancer/*gooseberry*/*gooseberry neuro*/promoter/transcriptional regulation

Introduction

Development of a multicellular organism is controlled by an elaborate genetic program. Its successful execution largely depends on the coordinate regulation of gene expression by mechanisms that control transcription precisely in time and space. This regulation operates through *cis*-regulatory enhancer and promoter sequences which, in cooperation with protein factors, initiate transcription. Promoters consist of a so-called basal promoter, which includes the transcriptional start site and an optional TATA box sequence, and of at least another *cis*-regulatory element in close vicinity, usually upstream, of the basal promoter. Promoters by themselves are unable to support efficient transcription, which is achieved only in combination with enhancers [Seipel *et al.* (1993), for reviews see Roeder (1991), and Gill and Tjian (1992)]. Enhancers are defined as *cis*-regulatory sequences that are able to activate transcription only in combination with a promoter and may be located at close as well as large distances from the promoter. Accordingly, sequences located in close proximity of the basal promoter that are unable to support efficient transcription in combination with the basal promoter should not be considered as enhancer elements even though they may contain promoter elements that are required for proper tissue-specific transcription.

It is generally assumed that the TATA box of each promoter assembles the same transcriptional machinery. Therefore, enhancers should activate not only their cognate promoter but also heterologous promoters such as the *hsp70* promoter, if they are combined for *in vivo* analysis. In other words, enhancers rather than promoters mediate the specificity of transcriptional regulation. By the same argument, one expects genes that are sufficiently close to one or several enhancers to be expressed in the same pattern. While this rule indeed appears to be followed by several pairs of adjacent genes in *Drosophila*, such as *zerknüllt* (*zen*) *z1* and *z2* (Rushlow *et al.*, 1987), *engrailed* (*en*) and *invected* (*inv*) (Coleman *et al.*, 1987), *yp1* and *yp2* (Logan *et al.*, 1989), *sloppy paired* (*slp*)*1* and *slp2* (Grossniklaus *et al.*, 1992), in other examples, such as *gooseberry* (*gsb*) and *gooseberry neuro* (*gsbn*) (Bopp *et al.*, 1986), neighbouring genes exhibit quite different expression patterns, raising the question by what mechanism these patterns are restricted to one of the genes.

The *Drosophila* genes *gsb* and *gsbn* are divergently transcribed and separated by ~10 kb of a common upstream region (Baumgartner *et al.*, 1987; Li *et al.*, 1993). Both genes encode transcription factors containing a paired-domain and a homeodomain and are important for proper segmentation and neurogenesis in the embryo (Noll, 1993). During embryogenesis, *gsb* acts as a segment-polarity gene and begins to be transcribed at late blastoderm in a segmentally repeated epidermal pattern (Baumgartner *et al.*, 1987). In contrast, *gsbn* is activated later and participates in the control of the development of the central nervous system (CNS) in which it is mainly expressed (Baumgartner *et al.*, 1987; Gutjahr *et al.*, 1993; Patel *et al.*, in preparation). The regulation of *gsb* and *gsbn*, therefore, provides an excellent paradigm for two neighbouring genes expressed in different patterns that makes it possible to study the nature of their differential transcriptional control.

Here we show that non-overlapping enhancers located in the common upstream region of *gsb* and *gsbn* regulate their specific expression. Moreover, we demonstrate that the action of the enhancers is restricted to their cognate promoter and gene because the heterologous components are incompatible.

Results

Control of gooseberry and gooseberry neuro expression by non-overlapping cis-regulatory regions

In previous experiments, the enhancer elements regulating the expression of *gsb* in the embryo (gsbE) have been shown to be confined to a 3.8 kb *Eco*RI fragment, located 1.9–5.7 kb upstream of the *gsb* transcriptional start site (Figure 1I; Li *et al.*, 1993). The enhancer region that controls the embryonic expression of *gsbn* was determined in a similar fashion by the fusion of genomic *gsbn* sequences with a *lacZ* reporter gene. Thus, 5′ transcribed sequences with progressively reduced adjacent upstream regions were fused in-frame to *lacZ*. The 5′ transcribed portion included the sequence up to the *Eco*RI site at the end of the paired-box in the third exon. These *gsbn–lacZ* constructs were used for P element-mediated transformation, and their expression was examined in embryos of the resulting transgenic lines.

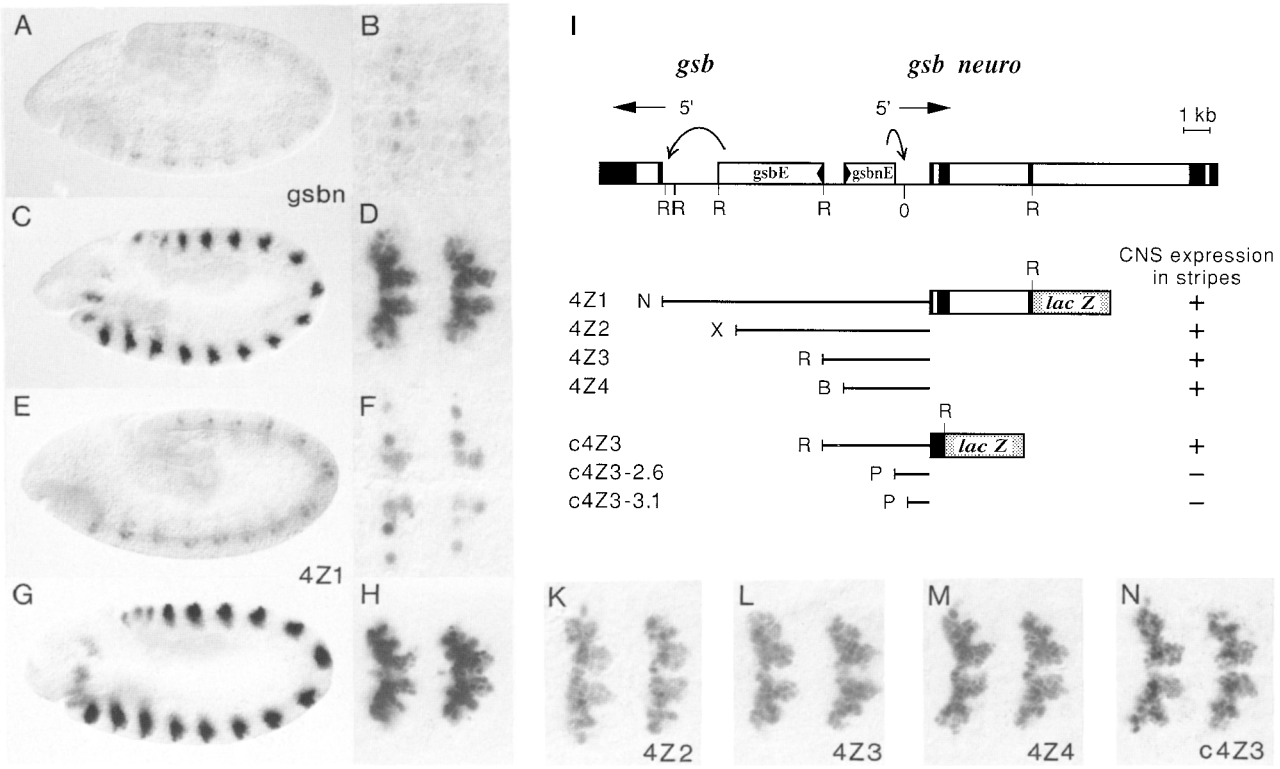

Fig. 1. Identification of enhancer regulating *gsbn* expression in the CNS. (**A**–**H** and **K**–**N**) Expression of *gsbn*–*lacZ* in transgenic 4Z1 embryos at stage 10 (E and F) and late stage 11 (G and H) is restricted to the CNS and identical to that of *gsbn* in wild-type embryos at comparable stages (A–D). The *gsbn*–*lacZ* expression patterns of transgenic 4Z2 (K), 4Z3 (L), 4Z4 (M) or c4Z3 (N) embryos are the same as that of transgenic 4Z1 embryos (H) or as that of *gsbn* in wild-type embryos (D). Embryos in A, C, E and G are shown in lateral views with their anterior to the left. B, D, F, H and K–N show magnified ventral views focussed on the CNS of two segments (labial and prothoracic segments in B and F; metathoracic and first abdominal segments in D, H and K–N). Expression of *gsbn* or *gsbn*–*lacZ* was visualized by immunostaining with anti-gsbn or anti-lacZ antibodies. (**I**) *gsbn*–*lacZ* constructs used for identification of *gsbn cis*-regulatory region, gsbnE, activating transcription in the CNS. The genomic organization of *gsb* and *gsbn* is illustrated at the top. Introns and the coding regions of exons are shown as open and filled boxes, respectively, while the directions of transcription are indicated by arrows. The *cis*-regulatory regions of *gsb*, gsbE (Li *et al.*, 1993) and of *gsbn*, gsbnE, act on their own promoters as symbolized by the curved arrows. Two start sites of *gsbn* transcription, one at 0 (corresponding to nucleotide 290 of the *gsbn* sequence; Baumgartner *et al.*, 1987) and one, used at lower frequency, 179 bp downstream of it, were determined by extension of a primer (Kingston, 1989) complementary to nucleotides 611–640 of the *gsbn* sequence (Baumgartner *et al.*, 1987; data not shown). The *Eco*RI map is complete only in the common upstream region of *gsb* and *gsbn*. Below the genomic map, the *gsbn*–*lacZ* constructs used for the identification of gsbnE are shown. Genomic *gsbn* sequences corresponding to those shown in the map above were fused in-frame to *lacZ* at the *Eco*RI site of the third exon (Baumgartner *et al.*, 1987). In the constructs c4Z3, c4Z3-2.6, and c4Z3-3.1, the two *gsbn* introns were removed by replacing the genomic DNA with the corresponding cDNA. The ability of the *gsbn*–*lacZ* constructs to be expressed in the CNS like *gsbn* is indicated in the column on their right. Abbreviations of restriction sites: R, *Eco*RI; N, *Nru*I; X, *Xho*I; B, *Bam*HI; P, *Pst*I.

Expression patterns identical to that of *gsbn* were observed in the CNS with four *gsbn*–*lacZ* fusion lines containing 8.9 kb (4Z1), 6.3 kb (4Z2), 3.1 kb (4Z3) or 2.3 kb (4Z4) of the *gsbn* upstream region (Figure 1A–M).

Expression of both *gsbn* and *gsbn*–*lacZ* is first detected during stage 10 (late germ band extension; Campos-Ortega and Hartenstein, 1985) in several neuroblasts in each segment (compare Figure 1A and B with E and F). During the extended germ band stage (stage 11), expression in the CNS increases and shows, by the end of stage 11, a characteristic L-shape in each hemi-segment (compare Figure 1C and D with G–M; Gutjahr *et al.*, 1993). Therefore, the shortest fusion construct 4Z4 contains all *gsbn* sequences required for the *gsbn*-specific expression in the CNS.

To test whether the expression of the analysed *gsbn*–*lacZ* constructs depends on the first two introns of *gsbn*, the transcribed genomic region of *gsbn* was replaced by the corresponding cDNA fragment in 4Z3 (c4Z3 in Figure 1I). As is evident from Figure 1N, this construct still produces the same *gsbn*–*lacZ* expression pattern as *gsbn* in wild-type embryos, suggesting that its expression depends solely on *cis*-regulatory elements of the *gsbn* upstream region. Since the removal of 2.6 kb from the distal end of c4Z3 (c4Z3-2.6 in Figure 1I) completely abolishes its expression (not shown), *gsbn* enhancer elements are contained within the 1.9 kb *Bam*HI–*Pst*I fragment of 4Z4 (gsbnE; Figure 1I), located between 2.6 and 0.7 kb upstream of the *gsbn* transcriptional start site.

This conclusion is further supported by our findings that the *gsbn*–*lacZ* fusions show the same regulation as *gsbn*. The activation of *gsbn* depends on *gsb*, both of which are probably expressed in the same neuroblasts (Gutjahr *et al.*, 1993). For the same reason, *gsbn*–*lacZ* expression is almost completely abolished in *Df(2R)gsb*^{IIX62} embryos (Figure 2A) which are deficient for both *gsb* and *gsbn* (Bopp *et al.*, 1986; Côté *et al.*, 1987). In contrast, *gsbn*–*lacZ* expression is largely restored in *Df(2R)gsb*^{IIX62} embryos carrying an exogenous *gsb* gene (Figure 2B). Some remaining irregularity of the *gsbn*–*lacZ* expression pattern (Figure 2B) might result from the deletion of *gsbn* in these embryos.

It follows from all these results that the upstream control

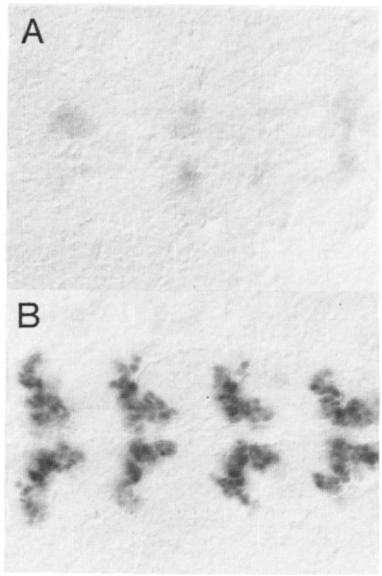

Fig. 2. Dependence of *gsbn−lacZ* expression on *gsb*. (**A**) Expression of 4Z3 in a homozygous *Df(2R)gsb^{IIX62}* embryo is strongly reduced as compared with its expression in a wild-type embryo (Figure 1L). (**B**) The expression of 4Z3 is largely rescued in a homozygous *Df(2R)gsb^{IIX62}* embryo containing an exogenous *P[gsb]* gene (Gutjahr *et al.*, 1993) but still displays some irregularities (compare with Figure 1L). Ventral views focussed on the CNS of the thoracic and first abdominal segments are shown. Anterior is to the left.

region of *gsbn*, gsbnE, does not overlap with gsbE (Figure 1), and thus separable enhancers act on *gsb* and *gsbn*.

Restriction of gsbE and gsbnE enhancer activities to their cognate promoters

Since *gsb* and *gsbn* share the same upstream sequence which includes both gsbE and gsbnE (Figure 1I), the question arises as to why these enhancers activate only their own and not also the other gene. As we have shown, gsbE and gsbnE still preserve their distinct regulatory specificities in the corresponding *gsb−lacZ* (9Z1 in Figure 3A; Li *et al.*, 1993) and *gsbn−lacZ* (4Z1 in Figures 1I and 3B) fusion constructs that contain both gsbE and gsbnE. The functions of gsbE and gsbnE are easily distinguished. While gsbnE activates *gsbn* only in the CNS after stage 10 (Figure 1), gsbE activates *gsb* mainly in the epidermis already much earlier by the successive action of its two elements, GEE and GLE (Li *et al.*, 1993). GEE, the *gsb* early element, begins to act on *gsb* during syncytial blastoderm whereas GLE, the *gsb* late element, takes over *gsb* activation after stage 10. As shown in Figure 4A and C, the 9Z1 *lacZ* fusion construct is expressed in the epidermis before and after stage 10 in a pattern resembling that of *gsb* (Figure 4B and D). The additional weak expression of 9Z1 in the CNS (Figure 4C) does not result from activation by gsbnE since smaller *gsb−lacZ* constructs that lack gsbnE still display this weak neural expression (Li *et al.*, 1993). Hence, gsbnE is inactive in 9Z1. Similarly, gsbE has no effect in 4Z1 as its expression of *gsbn−lacZ* remains largely restricted to the CNS and is not detected before stage 10 (Figure 1E−H). Therefore, the information restricting the activity of gsbE and gsbnE to their cognate genes is included in the common upstream sequence of *gsb* and *gsbn*.

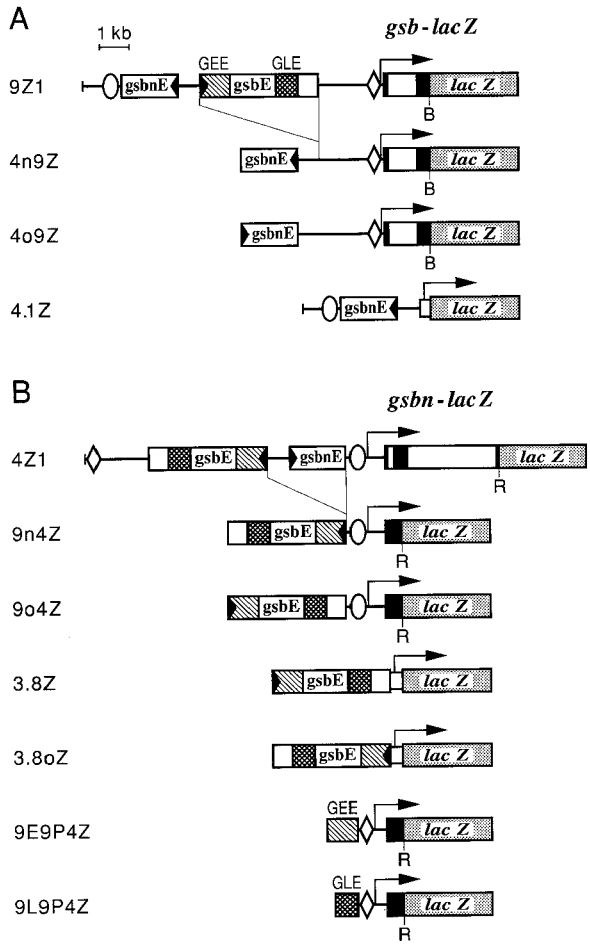

Fig. 3. Constructs used to demonstrate the incompatibility between gsbE or gsbnE and their heterologous *gsbn* or *gsb* promoter. *gsb−lacZ* (**A**) and *gsbn−lacZ* (**B**) fusions are shown with reference to those of 9Z1 (Li *et al.*, 1993) or 4Z1 (Figure 1I) and were constructed as described in Materials and methods. The oval (in 9Z1, 4.1Z, 4Z1, 9n4Z and 9o4Z) and diamond symbols (in 9Z1, 4n9Z, 4o9Z, 4Z1, 9E9P4Z and 9L9P4Z) represent the *gsbn* and the *gsb* promoter while the *hsp70* promoter is indicated as an open box preceding *lacZ* (in 4.1Z, 3.8Z and 3.8oZ). Within the gsbE enhancer the locations of the GEE and GLE elements (Li *et al.*, 1993) are shown (labeled in 9Z1). The orientation of gsbE and gsbnE is indicated by a filled arrowhead at one end. The genomic *gsb* DNA is fused to *lacZ* at the *Bam*HI site (B) of the second exon, the *gsbn* DNA at the *Eco*RI site (R) of the third exon.

This restriction could be explained in several ways. First, gsbE and gsbnE might block each other's action and thus be unable to act on their distal promoters. Second, the activities of gsbE and gsbnE might be dependent on their orientation. Third, a sequence between gsbE and gsbnE may function as a boundary element restricting their action to the gene located on the same side of the boundary (for a review, see Eissenberg and Elgin, 1991). Finally, the gsbE or gsbnE enhancer might be unable to interact with and thus activate the promoter of the other gene.

To distinguish between these mechanisms, the embryonic expression patterns of the *gsb−lacZ* and *gsbn−lacZ* constructs illustrated in Figure 3 were analysed. First, we tested whether gsbnE can activate the *gsb* promoter in the absence of the *gsb* enhancer, by examining the expression of 4n9Z in which gsbE has been removed from 9Z1. As

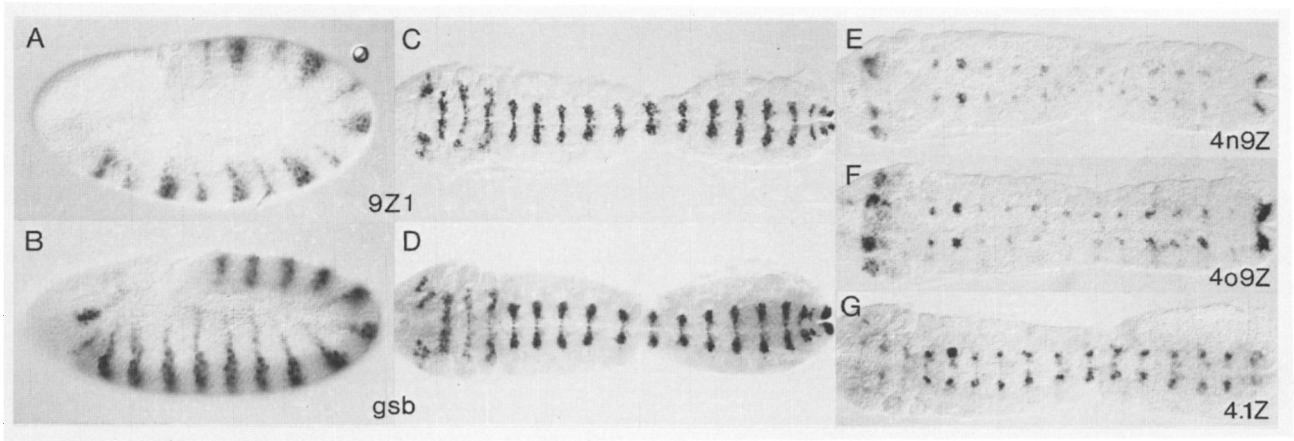

Fig. 4. Incompatibility between gsbnE enhancer and *gsb* or *hsp70* promoter. (**A**–**D**) Expression of *gsb*–*lacZ* in 9Z1 embryos (A and C) is similar to *gsb* expression in wild-type embryos (B and D), but unlike *gsbn* expression (Figure 1). The epidermal expression at stage 9 (A and B) is regulated by GEE (Li *et al.*, 1993). At this stage, not all gsb–lacZ stripes have reached the same intensity yet (A). Expression at stage 11 (C and D) is regulated by GLE (Li *et al.*, 1993). Since the plane of focus is in the epidermis of the unfolded embryos in C and D, some underlying cells of the CNS that also express *gsb*–*lacZ* (C) or *gsb* (D) are out of focus. (**E** and **F**) Expression of *gsb*–*lacZ* in 4n9Z (E) and 4o9Z (F) embryos at stage 11. (**G**) Expression of *lacZ* in 4.1Z embryos at stage 11. Note that the weak expression pattern is different from that of *gsbn* expression (Figure 1). Embryos are shown as lateral views (A and B) or unfolded along the amnioserosa (C–G) with their anterior to the left.

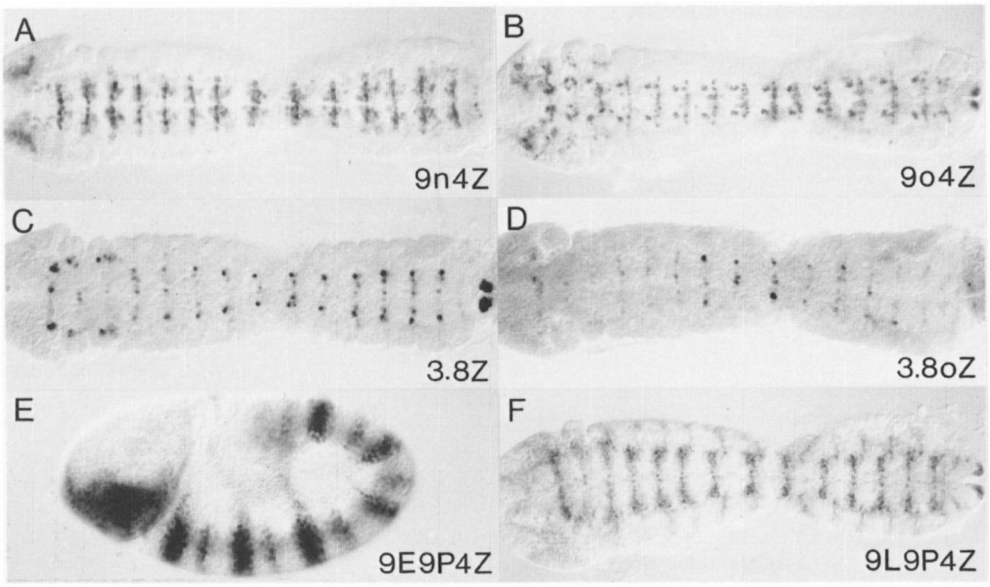

Fig. 5. Incompatibility between gsbE enhancer and *gsbn* or *hsp70* promoter. (**A** and **B**) Expression of *gsbn*–*lacZ* in 9n4Z (A) and 9o4Z (B) embryos at stage 11. Weak expression is visible in a row of several epidermal cells in each segment. Some ectopic expression in the CNS is also seen. (**C** and **D**) Expression of *lacZ* in 3.8Z (C) and 3.8oZ (D) embryos at stage 11. Expression is not detected in the epidermis but only in several neuroblasts of each segment. (**E** and **F**) Expression of *gsbn*–*lacZ* in 9E9P4Z embryos at stage 8 (E) and in 9L9P4Z embryos at stage 11 (F). Note that the functions of GEE (E) and GLE (F) are restored in 9E9P4Z and 9L9P4Z embryos. Embryos, oriented with their anterior to the left, are shown in a lateral view (E) or unfolded along the amnioserosa (A–D and F).

shown in Figure 4E, 4n9Z is only weakly expressed in a small set of internal cells of either neural or mesodermal origin in each segment after stage 11. This expression pattern differs dramatically from that of *gsbn* (Figure 1C and D) and is not affected by the orientation of gsbnE (4o9Z in Figures 3A and 4F). Hence, these results suggest that gsbnE cannot act properly on the *gsb* promoter to activate the *gsbn*-specific CNS expression although they do not strictly eliminate the possibility of a boundary element located between gsbnE and gsbE. Similarly, gsbnE cannot function properly in combination with the *hsp70* instead of the *gsb* promoter (4.1Z in Figures 3A and 4G).

In analogous experiments, we tested whether gsbE can activate the *gsbn* promoter. As shown in Figure 5A, when the region between gsbE and the *gsbn* promoter as well as the first two introns of *gsbn* were removed from 4Z1, the resulting 9n4Z construct (Figure 3B) is only weakly expressed in a row of epidermal and underlying neural cells in each segment after stage 11. In addition, the weak epidermal expression of 9n4Z differs from the characteristic barbell-shaped expression pattern of *gsb* or 9Z1 (Figure 4C and D). Few cells of the CNS that normally do not express *gsb* also express 9n4Z (Figure 5A). These results suggest that the elimination of sequences that might block the interaction of the *gsb* enhancer with the *gsbn* promoter does not restore its activity in 9n4Z. Neither is the expression

pattern of 9n4Z affected by the inversion of gsbE (9o4Z in Figures 3B and 5B), indicating that orientation is not the cause of its inactivity. Similarly, if gsbE is placed in either orientation upstream of the *hsp70* promoter (3.8Z and 3.8oZ in Figure 3B), *lacZ* expression is never detected in the epidermis at any stage. However, its expression in several neuroblasts or ganglion mother cells in each segment (Figure 5C and D) may correspond to part of the normal *gsb* activity in the CNS (Gutjahr *et al.*, 1993). We conclude that gsbE can activate neither the *gsbn* nor the *hsp70* promoter properly.

Taken together, these results clearly show that activation by the gsbE and gsbnE enhancers requires interaction with their cognate promoters.

Restoration of gsbE enhancer function with its cognate promoter

If the elements GEE and GLE of gsbE indeed fail to activate transcription of *gsbn−lacZ* in 9n4Z or 9o4Z because of their improper interaction with the *gsbn* promoter, we expect to be able to restore their function by exchanging the *gsbn* with the *gsb* promoter. Hence, 9E9P4Z and 9L9P4Z (Figure 3B) were constructed by combining GEE or GLE with the *gsbn−lacZ* portion of 9n4Z, in which the *gsbn* promoter and most of the *gsbn* leader (the *gsbn* sequence between −0.5 kb and +0.7 kb) were replaced by a 260 bp *gsb* region (from −155 bp to +104 bp), including the *gsb* promoter (Li *et al.*, 1993). As expected, the activity of GEE and GLE is restored in these constructs. In 9E9P4Z embryos, GEE activates *gsbn−lacZ* expression during the blastoderm up to the extended germ band stage in a pattern of segmentally repeated epidermal stripes (Figure 5E) very similar to that of 9Z1 (Figure 4A) which is driven by the original *gsb* upstream region. Similarly, the segmentally repeated stripes of *gsbn−lacZ* expression activated by GLE in 9L9P4Z embryos (Figure 5F) resemble those of 9Z1 (Figure 4C) after late germ band extension. These results further corroborate our conclusion that the *gsbn* promoter cannot replace the *gsb* promoter and thus mediate transcriptional activation by the *gsb* enhancer elements GEE and GLE. Moreover, the 260 bp *gsb* fragment is sufficient to act as *gsb* promoter in response to the GEE or GLE enhancers.

Discussion

A basic problem in the understanding of enhancers is how they recognize the genes that they control. An important aspect of this problem is the cooperation between enhancer and promoter. Given the well known specificity of enhancers, the question remains as to what, if any, is the contribution of the promoter. It is generally assumed that enhancers are indiscriminate in their choice of a promoter. This idea arose from studies in which one particular promoter was combined *in vivo* with many different enhancers, a strategy generally known as enhancer trap method (Weber *et al.*, 1984; O'Kane and Gehring, 1987; Bellen *et al.*, 1989; Bier *et al.*, 1989; Wilson *et al.*, 1989). If this has general validity, enhancers in the vicinity of more than one gene would be expected to act on each of them (Coleman *et al.*, 1987; Rushlow *et al.*, 1987; Logan *et al.*, 1989; Grossniklaus *et al.*, 1992) unless certain barriers restrict their action (Kellum and Schedl, 1991; Galloni *et al.*, 1993; for a review, see Eissenberg and Elgin, 1991). An alternative mechanism would be a specific interaction between an enhancer and its cognate promoter. The present study of the control of *gsb* and *gsbn* transcription is the first example that such a mechanism operates *in vivo*.

What then is the nature of the specificity in the interaction between the enhancer and its cognate promoter uncovered by our results? In the case of the *gsb* promoter, we have reduced its specificity to 260 bp, which include 155 bp of the *gsb* upstream region and thus may contain in addition to the basal promoter other promoter elements. Both the basal promoter as well as those putative additional promoter elements might contribute to the specificity of the promoter by restricting its activity to the combination with only certain enhancers. Our results cannot decide which of these, basal promoter, additional promoter elements or their combination, determine specificity. If specificity is conferred to the promoter by elements different from the basal promoter, one might argue that these elements constitute a closely located enhancer. However, this possibility is ruled out because these elements in combination with the *gsb* basal promoter cannot support transcription (Li *et al.*, 1993) as would be expected if they constituted an enhancer.

Assuming that both enhancers, gsbE and gsbnE, and promoters, gsbP and gsbnP, are active only after their assembly with certain sets of transcription factors, we may envisage two possible explanations. The enhancers are unable to activate the gene of the heterologous promoter because the promoter lacks its full complement of required factors to interact with the active enhancer. Alternatively, there is an intrinsic inability of the active heterologous promoter to interact with the proteins of an active enhancer. In this second case, the heterologous promoter and enhancer are incompatible *per se*.

If the lack of a factor is the reason in the case of the *gsb* locus, we would expect that each enhancer is able to activate transcription of the heterologous gene at least in those cells

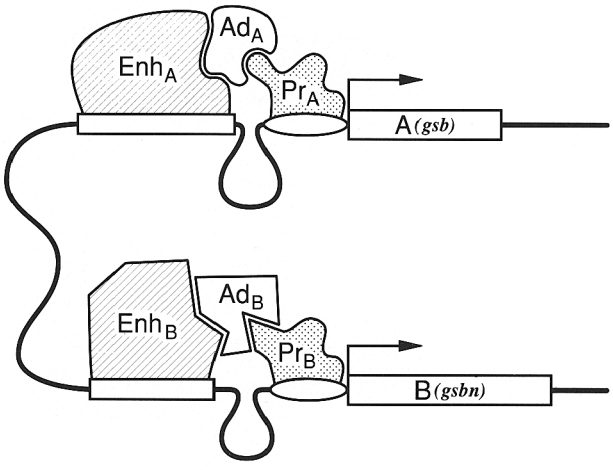

Fig. 6. Model illustrating restricted compatibility between enhancers and promoters. The interaction between transcription factors bound to the enhancer, Enh$_A$, and promoter, Pr$_A$, of gene A (e.g. *gsb*) and between those bound to the enhancer, Enh$_B$, and promoter, Pr$_B$, of gene B (e.g. *gsbn*) activates transcription of these genes. The interaction between enhancer and promoter binding factors may be direct or, as illustrated here, be mediated by adaptor proteins, Ad$_A$ and Ad$_B$, that do not interact with DNA. In this way, each enhancer is restricted to interact with its cognate promoter and thus is unable to support gene activation from the heterologous promoter.

that simultaneously express *gsb* and *gsbn*, i.e. in a specific set of neuroblasts and perhaps ganglion mother cells (Gutjahr *et al.*, 1993). While this is clearly not the case for the *gsbn* enhancer in combination with the *gsb* promoter (Figure 4E and F), the *gsb* enhancer in combination with the *gsbn* promoter seems partially active in such cells at late stage 11 (Figure 5A and B). Thus, the heterologous promoters and enhancers are largely incompatible *per se*, as illustrated in Figure 6, although there is a weak activation of the late element of the *gsb* enhancer, GLE, in combination with the *gsbn* promoter.

The reason for this incompatibility may reside in the extensive differences between the *gsb* and *gsbn* promoter sequences (Baumgartner *et al.*, 1987; Li *et al.*, 1993). Specificity of an enhancer for its cognate promoter and its failure to interact with a heterologous promoter has been previously reported in cell transfection assays and attributed to the highly divergent TATA box sequences of the two promoters (Wefald *et al.*, 1990). Clearly, this specificity must be mediated by factors that bind specifically to the different promoters as has been shown in another example (Parvin *et al.*, 1992).

Some enhancers, e.g. the *ftz* zebra element (Hiromi and Gehring, 1987), function with both their natural and the *hsp70* promoter, presumably because of their sequence similarity. Conversely, the gsbE and gsbnE enhancers are incompatible with the *hsp70* promoter probably because of its dissimilarity with the *gsb* and *gsbn* promoters. In mammalian cells, random combination of several enhancers and promoters did not reveal a preferential activity of any particular combination (Kermekchiev *et al.*, 1991). However, these experiments tested simple enhancer and promoter constructions in transient transfections of cells in culture. Therefore, these results, obtained with only one or a few DNA binding factors, might not be representative of *in vivo* situations and thus might not reveal specific interactions between promoters and enhancers.

The emergence of new combinations of compatible enhancers and promoters might be an important mechanism to drive evolution by their generation of new expression patterns (Jacob, 1977). This mechanism is a variation of the analogous strategy that combines coding regions with new *cis*-regulatory elements (Li and Noll, 1993).

Materials and methods

Plasmid constructions and generation of transgenic flies

To obtain the *gsbn*–*lacZ* constructs shown in Figure 1I, *gsbn* DNA was first subcloned into the Bluescript-derived vector pKSpL2 (Gutjahr *et al.*, 1993), in which the cloning site is flanked by an *Xba*I and a *Not*I site. Subsequently, the *gsbn* DNA was excised as a *Xba*I–*Not*I fragment and ligated in-frame to *lacZ* at the *Not*I site of the P element/*lacZ* expression vector CZ.1 (Li *et al.*, 1993). Thus, to obtain 4Z3, the 8.2 kb *Eco*RI fragment of *gsbn* was subcloned into the *Eco*RI site of pKSpL2 to prepare 4Z3′, from which the *gsbn* DNA was transferred as a *Xba*I–*Not*I fragment into CZ.1. The unique *Bss*HII site in the *gsbn* leader sequence (Baumgartner *et al.*, 1987) was then used to replace the distal *Eco*RI–*Bss*HII *gsbn* fragment of 4Z3′ by the *Nru*I–*Bss*HII or *Bam*HI–*Bss*HII *gsbn* fragment to generate 4Z1′ and 4Z4′, from which 4Z1 and 4Z4 were obtained by transfer of the *gsbn* DNA into CZ.1. 4Z2 was prepared directly by ligation of the *Xho*I–*Not*I *gsbn* fragment of 4Z1′ into the *Xba*I/*Not*I sites of CZ.1. To prepare c4Z3, in which the genomic *gsbn* DNA of 4Z3 is replaced by the corresponding cDNA, a 0.8 kb *Eco*RI fragment of the *gsbn* cDNA BSH4c4 (Baumgartner *et al.*, 1987) was subcloned into pKSpL2. Subsequently, the *Xba*I–*Bss*HII fragment of this construct was replaced by the corresponding fragment of 4Z3′, generating c4Z3′, from which c4Z3 was obtained by transfer of the *gsbn* DNA into CZ.1. The c4Z3-2.6 and c4Z3-3.1 constructs were derived from c4Z3′ by partial *Pst*I digestion to prepare c4Z3′-2.6 and c4Z3′-3.1, from which the *gsbn* sequences were transferred into CZ.1 as outlined above.

The *lacZ* constructs illustrated in Figure 3 were prepared as follows. The 3.8Z and 3.8oZ constructs were obtained by subcloning the 3.8 kb *Eco*RI fragment corresponding to gsbE (Li *et al.*, 1993) in both orientations into the *Xba*I/*Not*I sites of the HZ50pL vector (Hiromi and Gehring, 1987), in which the *hsp70* basal promoter is fused to the *lacZ* gene. Similarly, 4.1Z was prepared by subcloning the 4.1 kb *Eco*RI–*Nco*I *gsbn* fragment, containing gsbnE and the entire *gsbn* leader, into HZ50pL. To obtain 9n4Z and 9o4Z, the 3.8 kb *Eco*RI fragment (gsbE) was inserted in either orientation into the unique *Xba*I site of c4Z3-2.6. Similarly, 4n9Z and 4o9Z were prepared by inserting the 1.9 kb *Bam*HI–*Pst*I fragment (corresponding to gsbnE) into the unique *Xba*I site of the *gsb*–*lacZ* fusion construct 9Z2-3.8 (Li *et al.*, 1993). Finally, the *Xba*I–*Bss*HII fragment of c4Z3-2.6′, which contains the *gsbn* promoter and most of its leader sequence, was replaced by the *Xba*I–*Nru*I fragment of 9ZE″ or 9ZL″ (consisting of the *gsb* promoter and GEE and GLE, respectively; Li *et al.*, 1993) to generate constructs from which 9E9P4Z and 9L9P4Z were obtained by transfer of the *Xba*I–*Not*I fragments into CZ.1. The construction of 9Z1 has been described previously (Li *et al.*, 1993).

All constructs were injected into homozygous ry^{506} embryos (Rubin and Spradling, 1982). For each construct, at least two transgenic lines were obtained. Different lines of the same construct show essentially the same *lacZ* expression pattern except in a few instances where positional effects have been observed. In these cases, the pattern expressed by the majority of the lines is shown.

Immunostaining of embryos

Immunostaining of embryos was carried out as described (Li *et al.*, 1993), using the rabbit anti-gsb, anti-gsbn (Gutjahr *et al.*, 1993) or anti-lacZ (Cappel) antisera.

Acknowledgements

We thank Thomas Gutjahr for his invaluable help in the initial stage of this work and Fritz Ochsenbein for his expert artwork. We thank Walter Schaffner and Hans Noll for discussions and comments on the manuscript. This work was supported by the Swiss National Science Foundation grant 31-26652.89 and by the Kanton Zürich.

References

Baumgartner,S., Bopp,D., Burri,M. and Noll,M. (1987) *Genes Dev.*, **1**, 1247–1267.

Bellen,H.J., O'Kane,C.J., Wilson,C., Grossniklaus,U., Pearson,R.K. and Gehring,W.J. (1989) *Genes Dev.*, **3**, 1288–1300.

Bier,E. *et al.* (1989) *Genes Dev.*, **3**, 1273–1287.

Bopp,D., Burri,M., Baumgartner,S., Frigerio,G. and Noll,M. (1986) *Cell*, **47**, 1033–1040.

Campos-Ortega,J.A. and Hartenstein,V. (1985) *The Embryonic Development of Drosophila melanogaster*. Springer-Verlag, Berlin.

Coleman,K.G., Poole,S.J., Weir,M.P., Soeller,W.C. and Kornberg,T. (1987) *Genes Dev.*, **1**, 19–28.

Côté,S., Preiss,A., Haller,J., Schuh,R., Kienlin,A., Seifert,E. and Jäckle,H. (1987) *EMBO J.*, **6**, 2793–2801.

Eissenberg,J.C. and Elgin,S.C.R. (1991) *Trends Genet.*, **7**, 335–340.

Galloni,M., Gyurkovics,H., Schedl,P. and Karch,F. (1993) *EMBO J.*, **12**, 1087–1097.

Gill,G. and Tjian,R. (1992) *Curr. Opin. Gen. Dev.*, **2**, 236–242.

Grossniklaus,U., Pearson,R.K. and Gehring,W.J. (1992) *Genes Dev.*, **6**, 1030–1051.

Gutjahr,T., Patel,N.H., Li,X., Goodman,C.S. and Noll,M. (1993) *Development*, **118**, 21–31.

Hiromi,Y. and Gehring,W.J. (1987) *Cell*, **50**, 963–974.

Jacob,F. (1977) *Science*, **196**, 1161–1166.

Kellum,R. and Schedl,P. (1991) *Cell*, **64**, 941–950.

Kermekchiev,M., Pettersson,M., Matthias,P. and Schaffner,W. (1991) *Gene Expr.*, **1**, 71–81.

Kingston,R.E. (1989) In Ausubel,F.M. *et al.* (eds), *Current Protocols in Molecular Biology*. Wiley & Sons, New York, Vol. 1, pp. 4.8.1–4.8.3.

Li,X. and Noll,M. (1993) *Nature*, **366**, in press.

Li,X., Gutjahr,T. and Noll,M. (1993) *EMBO J.*, **12**, 1427–1436.

Logan,S.K., Garabedian,M.J. and Wensink,P.C. (1989) *Genes Dev.*, **3**, 1453–1461.

Noll,M. (1993) *Curr. Opin. Gen. Dev.*, **3**, 595–605.

O'Kane,C.J. and Gehring,W.J. (1987) *Proc. Natl Acad. Sci. USA*, **84**, 9123–9127.
Parvin,J.D., Timmers,H.Th.M. and Sharp,P.A. (1992) *Cell*, **68**, 1135–1144.
Roeder,R.G. (1991) *Trends Biochem. Sci.*, **16**, 402–408.
Rubin,G.M. and Spradling,A.C. (1982) *Science*, **218**, 348–353.
Rushlow,C., Doyle,H., Hoey,T. and Levine,M. (1987) *Genes Dev.*, **1**, 1268–1279.
Seipel,K., Georgiev,O. and Schaffner,W. (1992) *EMBO J.*, **11**, 4961–4968.
Weber,F., de Villiers,J. and Schaffner,W. (1984) *Cell*, **36**, 983–992.
Wefald,F.C., Devlin,B.H. and Williams,R.S. (1990) *Nature*, **344**, 260–262.
Wilson,C., Pearson,R.K., Bellen,H.J., O'Kane,C.J., Grossniklaus,U. and Gehring,W.J. (1989) *Genes Dev.*, **3**, 1301–1313.

Received on September 6, 1993; revised on October 18, 1993