Zeste maintains repression of *Ubx* transgenes: support for a new model of Polycomb repression

Man-Wook Hur¹, Jeffrey D. Laney², Sang-Hack Jeon³, Janann Ali⁴, and Mark D. Biggin⁵*

¹Department of Biochemistry and Molecular Biology, Yonsei University School of Medicine, 134 ShinChon-Dong, SeoDaemoon-Ku, Seoul, Korea 120-752.

²Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, CT 06520.

³Department of Biology, Kon-Kuk University, Seoul, Korea 143-701.

⁴Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, 240 Longwood Ave, Boston, MA 02115.

⁵Genome Sciences Department, Lawrence Berkeley National Laboratory, 1 Cyclotron Road, Mail stop 84-171, Berkeley, CA 94720.

* Corresponding author

Running Title: Zeste maintains repression of *Ubx*
SUMMARY

During late embryogenesis, the expression domains of homeotic genes are maintained by two groups of ubiquitously expressed regulators: the Polycomb repressors and the Trithorax activators. It is not known how the activities of the two maintenance systems are initially targeted to the correct genes. Zeste and GAGA are sequence specific DNA binding proteins previously shown to be Trithorax group activators of the homeotic gene Ultrabithorax (Ubx).

Here we demonstrate that Zeste and GAGA DNA binding sites at the proximal promoter are also required to maintain, but not to initiate, repression of Ubx. Further, the repression mediated by Zeste DNA binding site is abolished in zeste null embryos. These data imply that Zeste and probably GAGA mediate Polycomb repression. We present a model in which the dual transcriptional activities of Zeste and GAGA are an essential component of the mechanism that chooses which maintenance system is to be targeted to a given promoter.

**Keywords:** Polycomb Group / Trithorax Group / Homeotic / Zeste / GAGA.
INTRODUCTION

The mRNA expression patterns of the Hox or homeotic genes in Drosophila are initiated by spatially restricted activators and repressors that are transiently expressed in pregastrula embryos. As the expression of these early regulators decays, the transcription patterns of the homeotic genes are maintained by two complementary sets of regulators that are expressed in all cells: the Polycomb Group (PcG) repressors and the Trithorax Group (trxG) activators (Kennison, 1995; Pirrotta, 1998; Tillib et al., 1999).

The PcG are a coherent set of genes whose sole function is to maintain repression of homeotic and other developmental control genes. Repression is maintained only in those cells in which transcription has not been activated in the pregastrual embryo. PcG repression has been likened to form of a molecular memory because the PcG proteins must continuously mark those genes that are initially repressed in the early embryo (Bienz and Muller, 1995; Cavalli and Paro, 1999; Pirrotta, 1998). The mark on the gene must be continuous as there are no spatially restricted regulators in older embryos that can reinitiate the correct pattern of homeotic expression. PcG proteins are physically associated with their target genes, suggesting that they form a stable structure that is propagated through multiple rounds of cell division (Sinclair et al., 1998; Strutt and Paro, 1997). The Polycomb system is conserved in most animals, including mammals, and is important for maintaining the determined state of cells (Hashimoto et al., 1998; Strouboulis et al., 1999).

The trxG are ubiquitously expressed activators of one or more of the homeotic genes and are not as homogenous as the Polycomb group (Kennison, 1995; Tillib et al., 1999). By definition, the trxG are distinct from the spatially restricted activators that initiate the early patterns of homeotic transcription. However, unlike the PcG, which are dedicated to a shared set
of targets, the trxG act on a wide range of different genes; there is much less overlap in the genes that are regulated by each of the trxG members.

It has often been assumed that the PcG and trxG are mutually exclusive. However, several genetic experiments have hinted at the possibility that some trxG proteins are also involved in Polycomb repression (Gildea et al., 2000; Hagstrom et al., 1997; LaJeunesse and Shearn, 1996; Wu et al., 1989). For example, larvae mutant for the PcG gene Enhancer of Zeste show reduced expression of homeotic genes in some modified backgrounds (LaJeunesse and Shearn, 1996). But because the trxG members have broad pleiotropic effects, it could not be ruled out that the regulation seen in these experiments was indirect and that the PcG and trxG are therefore distinct.

In this paper we show that the trxG proteins Zeste and GAGA play a direct role in maintaining repression of the homeotic gene Ultrabithorax (Ubx). Based on these data and on previous results demonstrating that Zeste binds to a promoter regardless of its activation state, we propose a new model for the establishment of Polycomb repression in the early embryo. We suggest that the selective recruitment of the PcG to their targets requires proteins that function in both transcriptional activation and repression in the following manner. On inactive promoters in the early embryo, these duel activity proteins are bound to the DNA but are not sequestered in a regulatory complex with other activators. In cells where the promoters are transcribed before gastrulation, however, we propose that these factors make protein/protein interactions with an activation complex. Thus, surfaces on the dual activity factors would be differently exposed depending on the transcription state, providing a unique tag that the PcG factors could read.
MATERIALS AND METHODS

Plasmid constructs

22UZ Zeste, 22UZ GAGA, 22UZ NTF-1, and 22UZ Deletion were prepared by exchanging the wild type proximal promoter region of 22UZ (Irvine et al., 1991) with that of mutant proximal promoters from constructs Uβ ZESTE, Uβ GAGA, and Uβ NTF-1 and UβΔ–200/-31 respectively (Laney and Biggin, 1992). The 22UZ ZESTE/NTF-1 construct was prepared by inserting five Zeste binding sites at the Asp 718 site’ of the NTF-1 sites in 22UZ NTF-1. Two variants of 22UZ GAGA were prepared similarly that have six GAGA binding sites inserted in the Asp 718 either 3’ or 5’ of the NTF-1 sites in 22UZ NTF-1. The multimerized binding sites used in the preceding three constructs were Asp 718 restriction fragments from the proximal promoters of either 22UZ GAGA or 22UZ ZESTE. A third variant of 22UZ GAGA contains alternating GAGA and NTF-1 binding sites. These sites were introduced as five tandem copies of the following oligonucleotides cloned into the Asp 718 site of 22UZ Deletion: GATCCTGGCTCTCTGTTTCGATCTTGAACCGGTCCTGCGGGTAC, GATCGTACCCGCAGGACCGGTTCAAGATCGAAACAGAGAGCCAG. All three variants of 22UZ GAGA give essentially the same pattern of Lac Z expression, indicating that the precise position and orientation of binding sites is not important.

Drosophila strains and P-element mediated transformation

Germ line transformation, analysis of expression patterns, and crosses into z^{v77h} and z^{ae(bx)} mutant embryos were performed as described previously (Laney and Biggin, 1992; Spradling, 1986; Patel, 1994), except that the host micro injection stock was w^{1118}, and chromosomal linkage of inserts was determined by crosses with the balancer fly, w/Y;
CyO;MKRS/apX. 4 to 13 homozygous, independent transgenic fly lines for each Ubx-LacZ fusion construct were obtained: 6 for 22UZ GAGA, 11 for 22UZ Deletion, 13 for 22UZ ZESTE, 4 for 22UZ NTF-1, 4 for 22UZ ZESTE/NTF–1 and 6 for 22UZ GAGA/NTF-1. All lines for a given construct give essentially the same pattern of expression.

To analyze 22UZ GAGA transgene expression in Pc3 embryos, flies homozygous for a 22UZ GAGA transgene located in the second chromosome were crossed to w; st in ri Pc3 pp/TM3, Sb Ser to produce w/w; 22UZ GAGA; Pc3 flies. These were then self crossed to generate w/w; 22UZ GAGA; Pc3/Pc3 embryos. To analyze expression of 22UZ Zeste transgenes, a 22UZ Zeste transgene inserted on the third chromosome was recombined with Pc3. The resulting w/w; 22UZ Zeste Pc3/TM3 flies were crossed to produce w/w; 22UZ Zeste Pc3/22UZ Zeste Pc3 embryos.
RESULTS AND DISCUSSION.

Previous transgenic experiments indicate that Zeste, GAGA, and a third transcription factor, NTF-1, activate promoter constructs of the \textit{Ubx} gene in embryos via an intermingled cluster of sites between nucleotides \(\text{–}200\) to \(\text{–}31\) (Biggin and Tjian, 1988; Laney and Biggin, 1992). The constructs used in these experiments, however, contain only a small subset of the \textit{Ubx} cis regulatory region, and while they reproduce many features of \textit{Ubx} expression, they do not respond to Polycomb repression when inserted at many chromosomal locations. Consequently, they have not permitted a rigorous analysis of the role of the proximal promoter factors in maintaining repression. To address this question, here we have used larger constructs that contain the 22 kb of DNA upstream of the \textit{Ubx} mRNA start site. These constructs, first employed by Irvine et al (Irvine et al., 1991), do not suffer from significant position effect variation; more closely approximate the expression pattern of the endogenous \textit{Ubx} gene than the shorter constructs; maintain efficient repression in late embryos as shown by the lack of \(\beta\)-galactosidase reporter gene expression in more anterior and posterior regions (Fig 1a, 22UZ Native); and as demonstrated later, are genetically under the control of PcG genes.

Deletion of nucleotides \(-200\) to \(-31\) essentially abolishes transcription from the large \textit{Ubx} promoter constructs (Fig. 1b, 22UZ Deletion), indicating a critical role for factors binding to the proximal promoter. To determine the role of each factor separately, three constructs were prepared, each containing binding sites for either Zeste, or GAGA, or NTF-1 inserted between the deletion end points of the above construct (Fig 1 c, d, and e). Importantly, biochemical, in vivo UV crosslinking, and genetic experiments strongly suggest that the DNA binding sites used in these constructs are only recognized by their cognate factor, and not by any other sequence specific DNA binding activities (Biggin and Tjian, 1988; Laney and Biggin, 1992; Laney and
Biggin, 1996). Binding sites for each factor separately activate transcription of the large constructs during late embryogenesis (Fig. 1, compare b with c, d, and e). Strikingly, constructs containing only GAGA or Zeste binding sites at the proximal promoter are not expressed in the anterior or posterior of the embryo, whereas constructs bearing only NTF-1 sites are strongly transcribed in these terminal regions.

Since ectopic expression of \textit{Ubx} in anterior and posterior regions is generally caused by a failure of the initiating repressors or the Polycomb maintenance system (Chan et al., 1994; Simon et al., 1993; Zhang and Bienz, 1992), one interpretation of the above result is that Zeste and GAGA are required for at least one form of repression, while NTF-1 is not. It is also possible, however, that Zeste and GAGA are not repressors. Instead, it may be that they are unable to activate expression in anterior or posterior regions, even though they are expressed at similar levels throughout the embryo (Bhat et al., 1996; Pirrotta et al., 1988). To distinguish between these two possibilities, we first examined constructs that contained either Zeste and NTF-1 sites or GAGA and NTF-1 sites. These constructs are expressed in the central portion of the embryo; but, importantly, they are not significantly expressed in anterior or posterior regions (Fig. 1, 2, unpublished data). Since NTF-1 can activate \textit{Ubx} transcription in these terminal regions (Fig. 1 e), the absence of terminal expression is consistent with GAGA and Zeste directly repressing transcription in addition to their activation function.

To decisively establish if Zeste and GAGA are repressors, we wished to use a genetic test. Unfortunately, GAGA is lethal gene and a broadly acting regulator required for expression of transcription factors that regulate \textit{Ubx} in early embryos (Bhat et al., 1996; Farkas et al., 1994). Thus, it has not been possible to genetically determine whether GAGA is a direct repressor of
By contrast, \textit{zeste} is a largely redundant gene. \textit{zeste} null embryos and flies are essentially wild type, and the endogenous \textit{Ubx} gene is expressed normally in these animals; but because the 22UZ transgenes lack the cis regulatory elements through which factors that redundantly share \textit{zeste}'s function act, these transgenes should be regulated by \textit{zeste} (Goldberg et al., 1989; Laney and Biggin, 1996).

Consistent with this idea, transgenes containing only Zeste sites at the proximal promoter fail to express in \textit{zeste} mutant embryos, whereas constructs containing only GAGA or NTF-1 binding sites are expressed in this same genetic background (Fig 2 a, b, e, and f). Thus, this genetic experiment confirms that Zeste bound at the proximal promoter is required to activate transcription of the 22UZ constructs in the normal domain of \textit{Ubx} expression. To test the role of Zeste in repression, constructs containing binding sites for both Zeste and NTF-1 at the proximal promoter were compared in wild type and \textit{zeste} mutant embryos. In the normal domain of \textit{Ubx} expression, these constructs are expressed at similar levels in mutant and wild type embryos. Importantly, these constructs are derepressed in anterior and posterior regions of embryos lacking \textit{zeste} (Fig 2 g - 1). Thus, Zeste actively represses transcription in terminal regions of the embryo via binding sites at the proximal promoter.

The embryos shown in Fig 1 and Fig 2 a-i are at late stages of development, well after the Polycomb maintenance system has become active. To distinguish if Zeste is required for the initiation or the maintenance of repression, we examined expression of the 22UZ ZESTE/NTF-1 construct at an earlier stage. As Fig 2 k shows, in embryos lacking \textit{zeste} the 22UZ ZESTE/NTF-1 transgene is almost fully repressed in anterior and posterior regions at this earlier stage. Only weak derepression is observed in a few isolated cells. Thus, the transiently
expressed factors that initiate repression in the early embryo must be active, and the extensive
derepression observed later must be due to a failure in the maintenance system.

The PcG genes are an essential part of system that maintains repression of the
endogenous *Ubx* gene. To confirm that these genes also act on our transgenes, the 22UZ Zeste
and 22UZ GAGA constructs were crossed into PcG mutant embryos. Fig 3 shows that both
transgenes are derepressed in late stage embryos lacking the *Polycomb* gene. Similar results were
obtained in embryos lacking another PcG gene, *extra sex combs* (unpublished data). Thus,
Zeste—and probably also GAGA—act together with the Polycomb system to maintain
repression of *Ubx*.

We suspect that GAGA and Zeste have redundant, overlapping functions in maintaining
repression because the 22UZ Native construct, which contains Zeste, GAGA, and NTF-1 sites, is
not derepressed in *zeste* mutant embryos (Laney and Biggin, 1996), which contrasts with the
behavior of the 22UZ ZESTE/NTF-1 construct. Such redundancy in repression would parallel
the known redundancy between these two transcription factors in activating *Ubx* in the central
portions of the animal (Laney and Biggin, 1996), and helps explain the previous lack of evidence
that Zeste and GAGA are repressors.

The data presented in this paper is consistent with the earlier genetic data that suggested
that some trxG and PcG proteins may have dual activities (Gildea et al., 2000; Hagstrom et al.,
1997; LaJeunesse and Shearn, 1996; Wu et al., 1989). Further support for this idea comes from
recent biochemical experiments that have shown that GAGA is complexed with two PcG
proteins in Drosophila nuclear extracts (Horad et al., 2000) and Zeste is part of a multisubunit
complex that contains Polycomb (Saurin et al., 2001). In addition, PcG proteins are frequently associated in vivo with promoter regions that include Zeste or GAGA DNA recognition sites, including the *Ubx* proximal promoter examined in this paper (Orlando et al., 1998). Most PcG proteins do not recognize specific DNA sequences, thus the interaction with Zeste and GAGA may serve to recruit PcG proteins to promoters.

But is it essential that some proteins, such as Zeste and GAGA, participate in both repression and activation, or is it mere happenstance? We suggest it may be essential. At the transition between the initiating repressors and the Polycomb system, one possibility is it that Polycomb proteins are recruited to or activated on only those genes that are bound by initiating repressors; the initiating repressors may physically bind to PcG proteins to recruit them. However, Poux et al showed that Polycomb repression can be established on *Ubx* promoter constructs that lack initiating repressors elements, provided that initiating enhancer elements are also absent (Poux et al., 1996). In other words, at the transition between the establishment and maintenance of the *Ubx* expression pattern the Polycomb systems reads the absence of activation, rather than the presence of repression or repressors.

Endogenous Zeste protein binds to *Ubx* promoter constructs in vivo whether they are transcribed or not (Laney and Biggin, 1997). We suggest that in the early embryo in the cells in which *Ubx* is activated, Zeste is complexed, directly or indirectly, with initiating activators on the *Ubx* promoter. These complexes masks surface on Zeste that would otherwise be bound by components of the Polycomb system. In contrast, in those cells where *Ubx* is not activated, Zeste is still bound to the promoter but would not be part of an activating complex. Surfaces on Zeste protein would then be exposed and could serve as the signal that the Polycomb system
reads to initiate the maintenance phase of repression. The dual activities of Zeste and GAGA could be a key to understanding this fascinating regulatory mechanism.

ACKNOWLEDGMENTS

We are grateful to Ken Irvine for providing the plasmid 22UZ. This research was supported by a Korea Research Foundation Genetic Engineering Research grant (96-98) and a Life Phenomena & Function Research grant (2000) both awarded to M.-W.H. and by a N.I.H. research grant awarded to M.D.B. Work at the Lawrence Berkeley National Laboratory is performed under Department of Energy Contract DE-AC03-76SF00098, University of California.
REFERENCES


**FIGURE LEGENDS**

Figure 1. Transgenic analysis of *Ubx* proximal promoter mutations. Cartoons of promoter constructs tested are on the left and indicate the mRNA start site at +1 (arrow), the number of factor binding sites present at the proximal promoter (P) of each construct, and the position of the β-galactosidase reporter gene (LacZ). Representative pictures of stage 13-14 embryos showing the pattern of β-galactosidase expression from each transgene are to the right. The normal anterior and posterior boundaries of *Ubx* expression are marked with arrow heads.

a, 22UZ Native. b, 22UZ Deletion. c, 22UZ GAGA. d, 22UZ ZESTE. e, 22UZ NTF-1. f, 22UZ GAGA/NTF-1. g, 22UZ ZESTE/NTF-1.

Figure 2. Comparison of transgene expression in wild type and *zeste* mutant embryos. a and b, 22UZ GAGA. c and d, 22UZ ZESTE. e and f, 22UZ NTF-1. g, h, i, j, and k, 22UZ ZESTE/NTF-1. Expression in wild type embryos is shown to the left (Wt). Expression in *zeste* mutant embryos is shown on the right (Zmt). Essentially identical results were observed in *z*ae(bx) and *z*v77h embryos for all constructs (h and i and unpublished data). All embryos are stage 13-14 except j and k which are stage 11.

Figure 3. Derepression of 22UZ Zeste and 22UZ GAGA transgenes in *Pc*3 mutant embryos. a and b 22UZ GAGA. c and d 22 UZ Zeste. Expression in wild type embryos is shown on the left (Wt). Expression in *Pc*3 embryos in shown on the right (Pc3). The arrow heads mark
the anterior and posterior boundaries of *Ubx* expression in wild type embryos. The increased expression to the left of the anterior arrow head in *Pc*\(^3\) homozygous embryos shows the derepression of the transgenes.