Az ENHANCER OF ZESTE fehérje szerepe az epigenetikus represszió létrehozásában *Drosophila melanogasterben*

Ph.D thesis

Bajusz Izabella
Supervisor: Dr. Gyurkovics Henrik

Hungarian Academy of Sciences
Biological Research Center
Institute of Genetics

2003.

INTRODUCTION AND OBJECTIVES

In multicellular organisms, the division of a single totipotent cell, the fertilized zygote, ultimately results in a myriad of cells with different structures and functions. The basis of this diversity is differential gene expression, and different cell types are characterized by distinct gene expression patterns. Once established, the cell type-specific gene expression patterns have to be maintained throughout life. Wide range of morphogenic factors and signaling mechanisms are known which are responsible for establishing specific gene expression patterns and recent findings indicate the existence of a complex epigenetic maintenance system conserved in eukaryots which is responsible for the maintenance of this pattern through numerous mitotic cycles. These epigenetic mechanisms are thought to establish transcriptionally active or silent state of genes by dynamic changes in chromatin structure.

The factors mediating epigenetic activation or repression of gene expression were first identified in *Drosophila* as regulators of homeotic gene complexes, and have been divided in two antagonistic groups, the *Polycomb*- and the *trithorax*-group. POLYCOMB-group proteins are thought to establish and maintain inactive transcriptional states. In contrast, TRITHORAX-group (TRX-G) establishes and maintains active transcriptional states. Although PC-G and TRX-G homologues are highly conserved, and a substantial number of papers describe their mode of action, it remains unknown how any of these proteins recognizes and distinguishes chromatin domains containing genes to be activated or repressed. Homeotic genes furnish an excellent model system to tackle this problem.

The segment-specific expression of Drosophila homeotic genes is governed by complex DNA-regions, called *cis*-regulators. *Cis*-regulators are activated in a sequential order along the anterior-posterior axis in the early embryo by the transiently expressed segmentation genes. Each *cis*-regulator is activated early in development in one specific parasegment and control the level of expression of the appropriate homeotic gene. After the cessation of segmentation gene expression, the activity pattern of *cis*-regulators is epigenetically maintained through the remainder of development. The appropriate level of homeotic genes control morphogenetic differentiation of body segments, and their misexpression produces homeotic transformations.

We have identified a dominant mutation with a strong trx-like phenotype, which we termed *Trithorax-mimic* (*Trm*). The mutation was localized to the third chromosome and was not allelic to any of the known *trx*-G mutations, but strongly enhanced their phenotype. Recombination experiments located it to the 34,25 region not harboring any known *trx*-G allele, but overlapping the location of a *Pc*-G gene *Enhancer of zeste* [*E(z)*]. However the homeotic phenotype of our allele was the opposite as the phenotype of the characterized *E(z)* alleles.

The X-ray reversion of the *Trm* mutation proved that its trx-like phenotype could be reverted by inducing a new loss of function *E(z)* mutation on the *Trm* chromosome, indicating that *Trm* is really an *E(z)* allele and should be termed $E(z)^{Trithorax\text{-}mimic}$ [$(E(z)^{Trm}$)]. The *E(z)* is an important factor responsible for the repression of homeotic selector genes during *Drosophila* embryogenesis, and if its activity is lost these genes are ectopically activated after the disappearance of transient morphogenic factors. In contrast, the presence of the isolated $E(z)^{Trm}$ allele results in the ectopic repression of normally active *cis*-regulators.

The mutant allele is fully capable of performing the normal inactivation function of *E(z)*, hence there is no sign of weakening of the repressor function even in $E(z)^{Trm}$ hemizygous or homozygous conditions. The existence of intact repressor function in $E(z)^{Trm}$ is further supported by the result that the allele is capable of inducing the repression of the *white* gene in $zeste^{1}$ (z^{1}) background in hemizygous condition, in contrast to another gain of function allele, $E(z)^{1}$.

In spite of its gain of function nature the isolated allele is not conventional neomorphic, hence the homeotic transformations caused by the *Trm* mutants depends on the dosage of wild type *E(z)*. Introducing chromosomal duplications of the *E(z)* region or transgenic *E(z)* rescue constructs strongly alleviates the mutant phenotype. This finding indicates that the mutant protein competes with the wild type E(Z) and suggest that it targets much the same genes as the wild type E(Z). Supporting this hipothesis we also found that immunohistochemical detection of the E(Z) protein on the salivary gland giant chromosomes does not reveal any major difference of the binding pattern of the wild and mutant version of *E(z)*. Based on these findings we propose that the $E(Z)^{TRM}$ fully retains the epigenetic repressor activity but lacks target recognition specificity. This is the reason why $E(Z)^{TRM}$ also represses active domains. Based on our results we propose that the wild type E(Z) protein is involved in the early decision of which domains should be inactivated.

E(Z) protein contains a number of characterized protein-interacting domains in its NH_2-terminal half. At least two PC-G proteins (EXTRA SEX COMBS, ESC, and POLYCOMB LIKE, PCL) are known to interact with this part of E(Z).

Additionally, E(Z) contains a SET domain at its COOH-terminal end. SET domain was identified as a conserved motif found in three chromatin-associated proteins of Drosophila, SUPPRESSOR OF VARIEGATION 3-9, ENHANCER OF ZESTE and TRITHORAX. Nowadays the number of the known SET domain proteins is over 600. The name giving three chromatin associated proteins [SU(VAR)3-9, E(Z), TRX] perform different regulatory functions. The SU(VAR)3-9 is important for the silencing of heterochomatin, E(Z) is involved in epigenetic silencing of homeotic genes, while TRX is an epigenetic activator. The mammalian homologue of SU(VAR)3-9 was shown to specifically methylate histone H3 on lysine 9, whereas TRITHORAX, although has no proven enzymatic activity, is able to interact specifically with histone tails.

Sequencing the $E(z)^{Trm}$ allele revealed a single amino acid exchange point mutation in the SET domain that converts arginine (R) at position 741 into lysine (K) (R741K). It is noteworthy that all E(Z) homologues contain R at position 741, while TRX homologues contain K.

Crossing the $E(z)^{Trm}$ with different other *E(z)* alleles leads to surprising results. First, four of the homozygous lethal antimorphic alleles are not only viable with the *Trm* mutant but suppress its trx-like phenotype. The suppression is moderate in case of the transposon induced $E(z)^{60}$ mutation which codes for a truncated E(Z) protein lacking the C-terminal SAC- (cystein rich, pre SET) and SET-domains. Stronger suppression is detected using the three $E(z)^{son}$ alleles which carry point mutations in the SAC-SET-region. The suppressor effect is proportional to the distance of the mutated positions of the antimorphic alleles from the 741 arginin. The suppressor effect of the $E(z)^{son1}$ and $E(z)^{son3}$ alleles, in which the mutated positions only 60 and 106 amino acid away from the $E(z)^{Trm}$ transition, are so strong that most of the $E(z)^{Trm}$ $E(z)^{son1}$ and $E(z)^{son3}$ trans-heterozygotes look wild type. This effect does not depend on the Pc-enhancer phenotype of the examined *son* alleles. The strongest Pc-enhancer $E(z)^{son2}$ allele in fact the weakest suppressor. These results indicate a specific interaction between the mutated SET-domain regions and suggest that E(Z) proteins inactivate as homodimers or homomultimers, in which the SET-regions are closely contacted.

This hypothesis is further supported by the independent experiment in which the *nos*-suppressor phenotype of the above mentioned $E(z)^{son}$ alleles was examined in $E(z)^{Trm}$ background. Similarly to the homeotic interactions, in these experiment the suppressor effect of $E(z)^{Trm}$ is also proportional to the distance of the affected mutations from the 741. position, and does not depend on how strong *nos*-suppressors the $E(z)^{son}$ alleles are on their own. Surprisingly the strongest *nos*-suppressor $E(z)^{son3}$ allele is most suppressible by $E(z)^{Trm}$. Taken together these results lead to the conclusion, that the SAC-SET regions of at least two E(Z) proteins form an interactive surface functioning in the epigenetic repression of both homeotic and *gap*-genes.

Our results implied that the mutated part of the SET domain of E(Z) plays an important role in distinguishing silent and active *cis*-regulators. Based on this we concluded that $E(Z)^{TRM}$ may represent a valuable tool to identify factors and mechanisms that are involved in the recognition process.

We have employed genetic analyses to decipher how this single point mutation in the SET domain leads to repression of normally active *cis*-regulators in Drosophila. We were able to prove that the ectopic repression initiated by the mutant protein is dependent on the same factors involved in the normal repression function of E(Z). Mutations in many of the known *Pc*-G genes suppress the phenotype of $E(z)^{Trm}$. Reduction of the dosage of *Polycomblike* (*Pcl*), *extra sex combs* (*esc*), *Suppressor of zeste 12* [*Su(z)12*], *polyhomeotic* (*ph*), *pleiohomeotic* (*pho*), and *Suppressor of zeste 5* [*Su(z)5*] almost completely abolished the mutant phenotype. Of these, PCL, ESC and SU(Z)12 are known to interact directly with the wild type E(Z) both *in vitro* and *in vivo*, whereas PHO is the only conventional PC-G protein with proven DNA binding properties. These results suggest that the $E(Z)^{TRM}$-mediated ectopic repression requires the same epigenetic partners which are involved in the normal silencing function of the wild type E(Z), and indicate a high degree of similarity between normal and ectopic inactivation. Therefore, factors identified by using the sensitive $E(z)^{Trm}$ phenotype as a screening system are likely to play a similar role in normal PC-G dependent silencing.

By genetic means, we have identified four factors, which suppress ectopic repression by $E(Z)^{TRM}$ and have no previously demonstrated effect on homeotic gene regulation. Protein phosphorylation, histone gene dosage, mutations in two histone modifying enzymes, and deacetylase- inhibitor (Na-butirate) feeding affect ectopic repression by $E(Z)^{TRM}$. These were the first *in vivo* results implying that specific histone modifications play an important role in the recognition of different transcriptional states and the consequent formation of epigenetically activated and silenced chromatin domains.

We were able to show the followings:

-Reducing the dosage of histone genes by half, significantly suppresses the $E(z)^{Trm}$ phenotype.

-Ectopic repression of homeotic genes by $E(Z)^{TRM}$ is suppressed by mutant alleles of the *Su(var)2-*

1 gene. This gene has been described as a modifier of position effect variegation (PEV), and the specific alleles used were shown to cause significant hyperacetylation of the bulk of histone H3 and H4 proteins in homozygous conditions. These results indicate that histone acetylation and deacetylation plays an important role in E(Z)^{TRM}-mediated repression of transcription. This hypothesis is further supported by our independent result that increasing the overall level of acetylated histones by feeding the mutant flies with histone deacetylas inhibitor Na-butyrate, also suppressed the E(z)^{Trm} phenotype.

-We have shown that an another PEV modifier with no known homeotic function, *Su(var)3-6* also modifies the E(z)^{Trm} phenotype. The *Su(var)3-6* (*Pp1^{87D}*) gene product is responsible for about 80% of the total protein-phosphatase-1 activity in the fly. Mutations in the *Pp1^{87D}* gene suppress the E(z)^{Trm} phenotype, while extra wild type copy of the same gene enhances it. On the other hand, loss-of-function mutations of the genes encoding the *aurora* kinase and the *jil* kinase enhanced the E(z)^{Trm} phenotype. These results raise the possibility that phosphorylation/dephosphorylation of some protein substrates is a part of the mechanism which marks active and inactive domains in the homeotic gene complexes, and that the E(Z) protein is able to recognize this mark. In contrast to the phenotype of heterozygous E(z)^{Trm}, the phenotype of the homozygous mutant is not sensitive to the dosage of *Pp1^{87D}*. This indicates that the mutant E(Z)^{TRM} protein is unable to respond to the presence of the same phosphorylated compound. However, it is not due to a general insensitivity of the homozygous E(z)^{Trm} phenotype, because it is strongly suppressed by the overall increase of the level of histone acetylation caused by *Su(var)2-1* mutations. Taken together, our results indicate that the E(Z)^{TRM} is wild type in every aspect of the E(z) function, except for its inability to respond to the presence of a phosphorylated compound.

In order to reveal the phosphorylated compound marking active chromatin domains we were looking for nuclear kinases modifying the *Trm* specific repression in the opposite direction to the PP1^{87D} phosphatase. We have been able to find two kinases with significant enhancer effect on the homeotic phenotype of E(z)^{Trm} *aurora* (*aur*) and *jil*.

Recently *Pp1*-type phosphatases *aur*-type kinases and the Drosophila *jil* kinase itself are shown to influence the level of H3 phosphorylation. Based on these findings we speculate that the phosphorylation of histone tails in the active chromatin domains may be responsible for inhibiting the inactivating function of wild type E(Z).

To reach our final aim and elucidate the molecular mechanisms underlying the recognition of appropriate chromatin domains by E(Z), besides *in vivo* genetical interaction experiments we analyzed and compared the biochemical properties of wild type E(Z) and the E(Z)^{TRM} mutant. These experiments provided valuable clues towards the understanding of the molecular mechanisms underlying not only the recognition, but also the establishment of E(Z)-dependent epigenetic repression.

The genetic interactions of *E(z)^{Trm}* with the *E(z)^{son}* alleles indicated *in vivo* interaction between the E(Z) SAC-SET regions. In case of two other SET-domain proteins ASH1 and TRX the existence of specific homo-, and heteromer interactions were proven both *in vivo* and *in vitro*. In order to look for similar interactions in the case of E(Z) we performed specific binding experiments using Sepharose linked E(Z) C-terminal SAC-SET fragment and radioactively labeled *in vitro* translated E(Z), TRX and ASH1, protein. We could detect significant binding in all examined combination. In order to determine the interaction domain we could show that this interaction does not require the full length E(Z) and ASH1. The radioactively labeled E(Z) and ASH1 pre SET-SET fragment was also capable of interaction. We could not detect significant differences in the specificity or strength of the binding between the wild type protein and the appropriate E(Z)^{TRM} fragments in any of the examined cases.

Among the strong suppressors of the homeotic phenotype of *E(z)^{Trm}* we have found the S*u(z)5* gene which codes for an enzyme responsible for the biosynthesis of S-adenosil methionin, the important cofactor for all methyl transferase reactions. The SET-domains of two other chromatin associated proteins, SU(VAR)39H and ASH1, possess histone-methyl-transferase activity. Therefore, we speculated that E(Z)^{TRM} could possess a novel, or, compared to E(Z), altered histone methyl-transferase (HMT) activity, which may contribute to the repression of active transcription domains. We therefore examined whether E(Z) is a HMT. We generated a recombinant baculovirus which expresses epitope-tagged (FLAG) versions of wild type and mutant E(Z) in *Sf9* cells. Recombinant E(Z) derivatives were purified from *Sf9* cells using affinity-chromatography techniques.

To test whether wild type and mutant E(Z) proteins are able to methylate histones, we used conventional HMT-activity assays. Recombinant E(Z) proteins were incubated with the methyl-group donor [^{3}H]-adenosyl-methionine (SAM), and individual recombinant core histones (H2A, H2B, H3, H4), purified histone core-octamer, reconstituted nucleosomes, and polynucleosomes. After the reaction, proteins were separated by SDS-PAGE, and radiolabeled proteins were detected by fluorography. These assays revealed that both the wild type E(Z) and the E(Z)^{TRM} mutant has H+ specific methyl-transferase activity, as expected based on the genetic experiments.

A recent study implies that the SET-domain of TRX, which possesses no HMT-activity, binds nucleosomes, and specific mutations in TRX SET domain abolish the interaction of TRX with nucleosomes. Preliminary results implied that histone acetylation and phosphorylation play an important role in marking active domains for E(Z) in Drosophila. This raised the possibility that $E(Z)^{TRM}$ interacts differently with covalently modified histones.

To test this hypothesis, we performed protein-protein interaction assays, using N-terminal histone H3 tail peptides, and in vitro translated wild type and Trm mutant E(Z) proteins. Sepharose-beads loaded with differently modified H3 peptides were incubated with radio-labeled E(Z) and $E(Z)^{TRM}$. Interactions were detected by fluorography. We found that E(Z) and $E(Z)^{TRM}$ binds to the used H3 peptides and in both cases the K9-dimethilated peptide showed the strongest binding. As opposed to this the binding of the S10 phosphorilated version was always weak and $E(Z)^{TRM}$ binds less effectively than the wild type version to this phosphorilated tail peptide. These experiments elucidated that $E(Z)^{TRM}$ has similar overall bindig specifity to different covalently modified H3 peptides but interacts with modified, serin 10 phosphorylated, NH_2-terminal histone tail different to wild type E(Z).

In conclusion, our results imply that the interaction and modification of histones plays an important role in the repression of gene expression by E(Z). The C-terminal SAC-SET-domains are capable of specific histone H3 interaction, and posess intrinsic histone-methyl-transferase activity. The C-terminal half of the SET-domain plays an important role in recognizing the specific chromatin domains to be methylated, and finally repressed.

LIST OF PUBLICATIONS

PUBLICATIONS:

Bristow DR, Banford PC, Bajusz I, Vedat A, Young JM. : Desensitization of histamine H1 receptor-mediated inositol phosphate accumulation in guinea pig cerebral cortex slices.Br J Pharmacol. 1993 Sep; 110 (1): 269-74.

Bajusz I, Sipos L, Gyorgypal Z, Carrington EA, Jones RS, Gausz J, Gyurkovics H. : The *Trithorax-mimic* allele of *Enhancer of zeste* renders active domains of target genes accessible to *Polycomb*-group-dependent silencing in *Drosophila melanogaster*. Genetics. 2001 Nov; 159 (3): 1135-50

IN PREPARATION:

Bajusz I., Sauer F., Gyurkovics H. : The ENHANCER OF ZESTE protein is capable of stabilizing the inactive state of chromatin domains by histone methylation

Bajusz I., Sipos L., Pintér L., Gyurkovics H.: The role of the POLYCOMB-group proteins in the active chromatin domains

Honti V., Blastyák A., Pintér L., Bajusz I., Gausz J., Gyurkovics H.: Mapping of the *iab-7* TRE in *Drosophila melanogaster*