Identification and characterization of human Polycomb-group proteins.

Satijn, D.P.E.

Citation for published version (APA):

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: https://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.
CHAPTER 1

General Introduction
1. General introduction

1. Stable cell differentiation and chromatin structure.

How the identity of a differentiated cell is stably maintained is one of the important problems in biology. During embryogenesis cells become committed to different developmental pathways. Once committed, cells rarely change their fates, even after many cell divisions.

Our understanding of the molecular mechanisms which underlie cellular differentiation, the regulation of individual genes, has increased dramatically in the last decades. It is believed that gene expression is regulated at the level of promoters, enhancers and transcription factors. This level of gene regulation results in the generation of many different cell types, which are characterized by cell type specific expression patterns of distinct sets of active and inactice genes. Subsequently, the specific on/off state of a gene must be faithfully maintained over many cell generations to preserve the cell's identity. It has become increasingly clear that the stable transmission of cell type specific expression patterns is not dependent on the context of the promoter, but is mediated by changes in the structure of DNA and associated proteins, termed chromatin. Gene regulation at the chromosomal level involves modifications of DNA (e.g. methylation), histones (e.g. (de)-acetylation), and long-range interactions between distant chromosomal elements (49).

The chromatin template is a highly condensed complex of DNA, histones and non-histone-proteins, which is able to package the entire genome into the nucleus and simultaneously allow the appropriate expression of specific genes. The eukaryotic chromosome is not a uniform template for the activation of gene transcription. Different types of chromatin and chromatin regions can be distinguished, which differently affect gene expression. The telomeric and centromeric regions of the chromosome are associated with a more condensed and "closed" chromatin structure (based on morphological criteria), which is referred to as heterochromatin. Most genes located in heterochromatin regions are transcriptionally silenced. Euchromatin is associated with a more diffuse and "open" structure in which most of the transcriptionally active genes are found. The euchromatin region can be subject to structural changes, resulting in a higher or less condensed structure, referred to as facultative hetero- and euchromatin. The formation of facultative eu- or heterochromatin is believed to represent the underlying mechanism of chromatin-mediated gene regulation, keeping genes either in an active or a repressed state, in a cell type specific manner (48).

In all eukaryotes several chromatin-associated protein complexes have been identified that are involved in the maintenance of cell type specificity, one of which is the Polycomb group (PcG) complex (61, 65). The PcG complex is involved in the stable repression of genes, in which changes in chromatin structure are believed to play an important role. Similarly, a second class of proteins, named the trithorax group (TrG), has been identified to counteract the action of the PcG proteins (44). TrG proteins are involved in the maintenance of gene expression. Based on their respective modes of action, PcG and TrG proteins therefore represent a cellular memory system that is important for the heritable transmission of gene expression patterns.

2. PcG proteins and their functions.

PcG proteins have initially been identified in Drosophila as a class of approximately 20 genes (table 1) which are involved in the maintenance of the correct expression patterns of homeotic genes. Homeotic genes are regulatory genes that are expressed in complicated patterns in particular embryonic body segments. As such, homeotic genes play a crucial role in the establishment
I. General introduction

and maintenance of a proper body plan, such as the formation of head, leg, wing and tail structures, in both vertebrates and invertebrates. In PcG mutants the expression patterns of homeotic genes are initiated normally, early in embryonic development (43, 61). However, at later developmental stages the homeotic genes become expressed in body segments where they normally remain repressed. This results in dramatic changes in the body plan, leading to the formation of body structures, e.g. legs and wings at an aberrant place of the body, which are called homeotic transformations (6, 50, 81). In PcG mutants homeotic genes are expressed outside the body segments where they normally would remain repressed, and this indicates that PcG proteins are involved in stable repression of homeotic genes.

In Drosophila, Polycomb (Pc) was the first PcG gene, which has been genetically identified. The heterozygous Pc mutation caused a homeotic transformation which results in extra sex combs (51). Several mammalian genes are related to Pc by the presence of two specific protein motifs, the chromodomain and the C-terminal domain (63). The mouse M33 protein is capable of partially rescuing Drosophila Pc mutants and is considered a true functional homolog (56). M33 knock-out mice display homeotic transformations as is demonstrated from skeletal malformations and misexpression of homeotic genes. In addition also hematopoietic defects are observed; a reduced cell proliferation and survival of lymphoid and myeloid cells (17). Other Pc homologues are the mouse MPc2 (2) and the human HPC2 (this thesis) proteins.

The murine Bmi1-gene has been identified through it’s involvement in lymphomagenesis (32, 87, 88). Bmi-1 and the highly related mel-18 protein are homologues to the Drosophila PcG proteins Posterior sex combs (Psc) and Suppressor two of zeste (Su(z)2) (11, 84, 88). These proteins share a RING finger motif which is involved in protein-protein interactions. Bmi-1 homologues have also been found in human (3) and Xenopus, Xbmi1 (69). Murine bmi-1 and mel-18 are involved in the proper development of axial patterns and regulation of cell proliferation during hematopoiesis (37, 40).

The murine eed (embryonic ectoderm development) protein is very similar to the Drosophila PcG protein extra sex combs (esc). In Drosophila esc is unique among PcG proteins in that it is only expressed transiently, early in development. Unlike other PcG proteins eed deficiency in mice is visible before Hox gene expression during early embryonic development. Eed null mutant mice display disrupted anterior-posterior patterning early in development, during gastrulation. Mutant mice lack a notochord, somites and neural structures (73). Embryos carrying hypomorphic eed alleles display the known homeotic transformations. It has been argued that eed may be required both to establish axial pattern formation early in development and maintain correct hox gene expression later in development (23, 73).

Several other PcG proteins have been identified (see table 1). Recently, Drosophila pleiohomeotic (Pho), has been characterized as the first DNA binding PcG protein. Homology searches revealed that Pho is a homolog of the vertebrate protein Yin-Yang (YY1), a ubiquitously expressed zinc finger containing DNA binding protein that is able to act as a transcriptional repressor or activator (10). The murine Rae28/mPh1 (57) and mPh2 (33) and its human counterparts HPH1 and HPH2 (this thesis) share two regions of close similarity to the Drosophila gene polyhomeotic, (Ph). Rae-28 mutant mice display, beside posterior skeletal transformations, defects in neural crest development and parathyroid and thymic hypoplasia. Vertebrate enx/EZH PcG proteins are similar to Drosophila PcG protein Enhancer of Zeste (E(z)) (34, 47). These proteins share a SET protein motif, which is also found in proteins of the TrG complex. In Drosophila E(z) displays features which are characteristic
### Table 1. Polycomb group homologs and function

<table>
<thead>
<tr>
<th>PcG protein homologs</th>
<th>General functions of class of protein</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>-Polycomb</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pc</td>
<td>drosophila</td>
<td></td>
</tr>
<tr>
<td>XPc,</td>
<td>xenopus</td>
<td></td>
</tr>
<tr>
<td>CHCB3,</td>
<td>chicken</td>
<td></td>
</tr>
<tr>
<td>M33,</td>
<td>mouse</td>
<td></td>
</tr>
<tr>
<td>MPC2</td>
<td>mouse</td>
<td></td>
</tr>
<tr>
<td>Cbx2/HPC1,</td>
<td>human</td>
<td></td>
</tr>
<tr>
<td>HPC2,</td>
<td>human</td>
<td></td>
</tr>
<tr>
<td>-Posterior sex combs</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Psc,</td>
<td>drosophila</td>
<td></td>
</tr>
<tr>
<td>Xbmil,</td>
<td>xenopus</td>
<td></td>
</tr>
<tr>
<td>Bmil,</td>
<td>mouse</td>
<td></td>
</tr>
<tr>
<td>mel-18,</td>
<td>mouse</td>
<td></td>
</tr>
<tr>
<td>BMI1,</td>
<td>human</td>
<td></td>
</tr>
<tr>
<td>-Polyhomeotic</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ph</td>
<td>drosophila</td>
<td></td>
</tr>
<tr>
<td>Rae-28/MPh1</td>
<td>mouse</td>
<td></td>
</tr>
<tr>
<td>MPh2</td>
<td>mouse</td>
<td></td>
</tr>
<tr>
<td>HPH1</td>
<td>human</td>
<td></td>
</tr>
<tr>
<td>HPH2</td>
<td>human</td>
<td></td>
</tr>
<tr>
<td>-Sex comb on midleg</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Scm</td>
<td>drosophila</td>
<td></td>
</tr>
<tr>
<td>SCML1</td>
<td>human</td>
<td></td>
</tr>
<tr>
<td>-Enhancer of Zeste</td>
<td></td>
<td></td>
</tr>
<tr>
<td>E (Z)</td>
<td>drosophila</td>
<td></td>
</tr>
<tr>
<td>mes-2</td>
<td>c.elegans</td>
<td>[35,42,46]</td>
</tr>
<tr>
<td>Enxl/Ezh2</td>
<td>mouse</td>
<td>[34]</td>
</tr>
<tr>
<td>Enx2/Ezh1</td>
<td>mouse</td>
<td>[58]</td>
</tr>
<tr>
<td>EZH1</td>
<td>human</td>
<td>[1]</td>
</tr>
<tr>
<td>EZH2</td>
<td>human</td>
<td>[16]</td>
</tr>
<tr>
<td>-Extra sex combs</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Esc</td>
<td>drosophila</td>
<td></td>
</tr>
<tr>
<td>mes-6</td>
<td>c.elegans</td>
<td></td>
</tr>
<tr>
<td>eed</td>
<td>mouse</td>
<td>[20,73]</td>
</tr>
<tr>
<td>EED</td>
<td>human</td>
<td>[74]</td>
</tr>
<tr>
<td>-Pleiohomeotic</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pho</td>
<td>drosophila</td>
<td></td>
</tr>
<tr>
<td>YY1</td>
<td>human</td>
<td></td>
</tr>
<tr>
<td>-Enhancer of polycomb</td>
<td></td>
<td></td>
</tr>
<tr>
<td>E(Pc)</td>
<td>drosophila</td>
<td></td>
</tr>
<tr>
<td>Epc1;Epc2</td>
<td>mouse</td>
<td></td>
</tr>
<tr>
<td>EPC1;EPC2</td>
<td>human</td>
<td></td>
</tr>
</tbody>
</table>

**References**: [62], [69], [91], [63], [2], [27], [ch.4], [52], [69], [32,88], [84], [3], [57], [33], [29], [29], [7], [86], [19], [38], [35,42,46], [34], [58], [1], [16], [26,30,70], [35], [20,73], [74], [10], [78], [79], [79], [79].
1. General introduction

both for PcG and TrG proteins and is also considered a unique PcG protein.

In conclusion, PcG proteins constitute a class of structurally unrelated proteins, which are genetically defined by the occurrence of similar homeotic transformations when they are mutated, in Drosophila. PcG proteins are involved in the stable maintenance of gene repression of, at least, the homeotic genes over many cell generations, thereby preserving the cell’s identity. Genetic studies in mice has shown that PcG proteins, beside being involved in proper development of the bodyplan, are also involved in the axial pattern formation during early development, hematopoiesis, oncogenesis and neural development.

3. Polycomb group protein complexes.

PcG genes have been charcterized by the similar homeotic transformation which occur upon the mutation of a PcG gene in Drosophila (21, 22, 39, 50, 80). Mutation of Pc, Ph and Polycomblike (Pcl) display severe homeotic transformations, whereas deletion of the majority of the PcG genes show only a weak homeotic phenotype (50). Interestingly, the homeotic transformations are strongly enhanced when more than one PcG gene is mutated. Mutation of PcG genes in mouse results in similar phenotypic changes as observed in Drosophila (12, 43, 50, 62).

The synergistic effects of double PcG mutants have been interpreted as such that PcG proteins form multimeric protein complexes. Further, it has been found that different PcG proteins, including Pc, Ph and Psc, bind in overlapping patterns on polyteny chromosomes in Drosophila salivary gland cells (68, 92). Biochemical evidence for the existence of multimeric PcG protein complexes came from co-immunoprecipitation and co-fractionation experiments performed in lysates from Drosophila embryos. Using specific antibodies against the Drosophila PcG proteins Polycomb (Pc) and Polyhomeotic (Ph), it was shown that both proteins co-immunoprecipitate.

Recently, a PcG complex has been purified from Drosophila embryos, using affinity chromatography (75). To facilitate the purification two Drosophila PcG proteins were flag-tagged and used to generate transgenic lines. Complexes purified via flag-tagged Posterior sex combs (FPsc) or Polyhomeotic (FPh) contained the same PcG proteins, showing that Psc, Ph, Pc and Sex combs on midleg (Scm) are part of one PcG complex, designated PRC1. The PcG proteins E(Z) and Pcl do not copurify with the PRC1 complex. Size fractionation of the PRC1 complex showed the heterogenous nature of the PRC1 complex, since Psc, Ph, Scm and Pc ran in the void volume and migrated at approximately 200-600 kDa. In addition, gel filtration analyses of PcG complexes from mouse embryos and mouse F9 cells have been done (31). The elution profile indicate that the murine PcG complex can be resolved into at least 3 forms: a very high molecular weight fraction eluting in the void volume, a high molecular weight fraction and a peak fraction of approximately 200-1000 kDa. The fractions contained the murine PcG proteins M33, Rae28 and Bmi1. The relative absence of these proteins in the lower molecular weight fractions, suggest that very few of these murine PcG proteins exist as monomers. All these data indicated the existence of multimeric PcG protein complexes. It is, however, not known to what extent the PcG proteins are able to multimerixe. Various models (64, 65) predict that PcG complexes spread over the chromatin fibre. This may indicate that the complexes polymerize. However, no solid experimental data support polymerization of PcG complexes.

In spite of the widespread existence of PcG proteins, very little is known about the underlying molecular mechanism by which PcG proteins silence genes. Hardly anything is known about the identity of target genes beside the homeotic genes, or how the PcG proteins
interact with the DNA of their target genes. However, one important observation is that PcG proteins form multimeric protein complexes. At present many more PcG proteins and protein-protein interacts have been characterized, which will be discussed in the conclusive chapter.

4. DNA connection

4.1. Polycomb response elements (PREs).

Increasing evidence is accumulating about the interaction between the different PcG proteins and their capacity to form multimeric protein complexes. How these PcG complexes are associated to their target genes is still unclear. Genetic studies have characterized cis-regulatory sequences that maintain transcriptionally inactive states of homeotic, in transgenic flies. The silencing mediated by these cis-regulatory sequences is dependent on the presence of functional PcG proteins, and hence these sequences have been termed PcG Response Elements (PREs) (14, 15, 55, 76).

PREs are cis-acting DNA elements which mediate PcG silencing. Two characteristics classify them as PREs. i) In reporter gene constructs, PREs are able to reproduce the boundaries of homeotic gene expression domains throughout development. This property is dependent on the presence of functional PcG proteins. ii) Transposon mediated insertion of an PRE element into the genome is able to create an ectopic PRE site to which PcG proteins are recruited. As yet, no PREs that concede with the characteristics stated above, have been identified in vertebrates.

4.2. The PcG protein Pleiohomeotic binds to a specific DNA sequence in PREs.

Recently, Pleiohomeotic (Pho) has been identified as the first DNA binding PcG protein. A search for proteins that interact with a PRE from the segment polarity gene engrailed resulted in the identification of Pho. Pho is a homolog of the vertebrate protein Yin-Yang (YY1), a ubiquitously expressed zinc finger containing DNA binding protein that is able to act as either a transcriptional repressor or activator (10). YY1 contains two acidic domains at the N-terminus, which provide the transcriptional activating activity. The C-terminal region encompassing the four zinc fingers of YY1 have been identified as a transcriptional repression domain. In addition, the spacer region N-terminal of the zinc fingers has an accessory transactivation function (4). However, no known transactivation domains are detected in Pho, and only the zinc finger region and a small region of the spacer of mammalian YY1 are found to be conserved (10). Mammalian YY1 has also been shown to interact with a histone deacetylase, through a domain not present in Pho. Thus, whereas repression mediated by YY1 is likely to involve histone deacetylase, this is more unlikely in the case of Pho. Therefore, although Pho and YY1 are very homologous in the DNA binding zinc finger motif it is not certain whether YY1 is a bona fide mammalian homolog of Pho and even whether Pho functions in a similar way as YY1. One possibility is that the function of Pho is limited to repression or mainly acts as a recruiter and anchor for PcG protein complexes.

5. PcG complex mediated gene silencing

5.1 PcG proteins are gene repressors

PcG proteins are known to be involved in the repression of gene activity in Drosophila and in vertebrates (50, 61). There are indirect lines of evidence that suggest that PcG proteins lead to formation of a repressed chromatin state (see 5.2). However there is no direct evidence that PcG proteins function via alterations in chromatin structure. In addition to genetic and structural studies, a first step, towards the
elucidation of the mechanism by which PcG proteins repress gene activity, has been done with the design of a repression assay in cell lines (9, 56). Not until recently a PcG protein, Pho, has been identified with DNA binding properties. No other PcG protein has been found to bind directly to DNA. To investigate the ability of PcG proteins to repress gene activity they have been targeted to reporter genes as LexA or GAL4 fusion proteins. Reporter genes have been activated by either the transcription factor HEB or HSF. It has been shown that targeting of bmi-1 represses the HEB/HSF induced gene activity more than ten fold. Also the Drosophila PcG proteins Psc, Su(z)2 and Pc are able to repress gene activity in mammalian cell lines (9, 56). The finding that Drosophila PcG proteins function as repressors in mammalian cells suggests that the mechanism of repression is conserved between mammals and Drosophila.

Pc homologs contain a conserved protein motif, the chromodomain. The chromodomain is essential for binding to chromatin. It was shown that, using the repression assay, mutations in the chromodomain, which abolishes chromatin binding, do not impair the gene repression function of the Pc protein. This shows that a Pc protein without functional chromodomain is capable of silencing (56). It is also detected that Pc homologs share an evolutionary conserved C-terminus. Although several lethal Pc alleles in Drosophila, are characterized by C-terminal deletions, this domain does not affect the binding of the protein to chromatin (25). It is found that deletion of the C-terminus abolishes gene repression (56). Therefore, the conserved C-terminus of Pc homologs is essential for the gene repression function of Pc proteins

The in vivo repression assay is used to assess the ability of PcG proteins to repress gene activity and to analyze which protein motifs are responsible for gene repression. The mechanism how targeted gene repression is achieved by PcG proteins in this assay is still unclear.

5.2 Chromatin structure and gene silencing.

Many aspects of the molecular mechanism underlying the role of PcG proteins in the stable transmission of gene activity are enigmatic. The notion that PcG proteins operate through the formation of multimeric complexes does not contribute much to the question how these PcG complexes keep genes in a repressed state. An important clue about the molecular mechanism came from the identification of a protein motif in Pc homologs. Pc homologs have a domain which is shared with HP1, a heterochromatin-associated protein. The shared motif between Pc and HP1 has been termed the chromodomain (chromatin organization modifier) (62, 77). The chromodomain is a conserved region of about 60 amino acids which was originally found in Drosophila modifiers of variegation, Su(var) 205, also known as heterochromatin protein 1, HP1. Modifiers of variegation modify the structure of chromatin to the condensed morphology of heterochromatin, a cytologically visible condition in which gene expression is repressed. This phenomenon is termed position-effect variegation (PEV). The discovery of the chromodomain in Pc provided an important, direct link between the regulation of gene activity and chromatin structure. It suggests that Pc and HP1 operate through common mechanisms, which may involve heterochromatinization. Is there a functional connection between PEV and PcG function? If so, mutations in PcG genes should also affect PEV. The Enhancer of Polycomb (E(Pc)) mutation may represent a protein that functions both in PEV and in PcG function. E(Pc) deletion does not cause a homeotic transformation but does enhance the phenotypes of other PcG mutants and also strongly suppresses PEV. The E(Pc) gene might interact with the
chromodomain in order to enhance activity of both the Pc and HP1 proteins (44, 79).

Pc homologs, like many other PcG proteins, are unable to bind DNA. However, the chromodomain is essential for the binding of the Pc proteins to chromatin. Carboxyl-terminal truncations of Pc homologs do not affect chromosomal binding. Mutations affecting only the chromodomain abolish the binding to chromatin (54, 67). Furthermore, it has been shown that the chromodomain is important for targeting the protein to specific chromosomal sites. HP1 and Pc localize to distinct loci on polytene chromosomes in Drosophila. A chimeric HP1-Pc protein, consisting of the chromodomain of Pc in the context of HP1 binds both to heterochromatin, as HP1 does, and polycomb binding sites on polytene chromosomes. Expression of the chimeric HP1-Pc protein also results in the mislocalisation of endogenous HP1 to Polycomb binding sites and endogenous Pc to the heterochromatic chromocenter. Particularly interesting is the finding that the chimeric HP1-Pc protein is able to promote PEV. This supports the notion that heterochromatin-mediated silencing and PcG mediated gene silencing share mechanistic similarities (67).

Although the biochemical role of the chromodomain is not known, the altered distribution of the Pc proteins in chromodomain mutants suggest a role in protein-protein interaction. Indeed, structure analysis using nuclear magnetic resonance spectroscopy on a murine HP1-like protein, MoMOD1, revealed that the chromodomain may act as a protein interaction motif. The MoMOD1 is suggested to act as an adaptor mediating interactions between different proteins (5). However, until this moment no proteins have been identified to interact directly with Pc homologs through the chromodomain of the protein. In the light of the formation of multimeric complexes the chromodomain may be involved in recruiting proteins to the complex. Whether the chromodomain itself or proteins that can associate with the chromodomain is functionally important for chromatin organization and thereby influencing gene activity is unclear.

5.3 Chromatin remodelling.

The general idea is that PcG proteins repress gene activity by the induction of conformational changes of the chromatin structure. The way how this is achieved is still unclear. Although PcG gene silencing has been compared with PEV both phenomena do not employ the same mechanism. Both at the level of chromatin configuration and at the level of the covered distance of silencing, PcG mediated gene repression and PEV show unique features. PEV induces changes in chromatin structure, which results in a heterochromatinisation of the DNA, which results in an increased nucleosome ordering and decreased accessibility of restriction enzymes (90). In contrast, the PcG induced gene silencing state of chromatin does not prevent the action of restriction enzymes (53, 72). Secondly, PEV induced gene silencing results form spreading of the heterochromatin configuration into regions which used to have a chromatin structure that allows gene transcription. This spreading is able to cover several hundreds of kilobases in distance. PcG proteins do not spread over large distances over the chromatin, but they only cover the PRE and flanking regions, promoters and enhancers, and are only cover several kilobases (60, 83).

Recent findings indicate that PcG function may involve local changes in chromatin structure rather than conformational changes over long distances. In Drosophila, dMi-2 protein has been found to participate in PcG functioning and, at the same time, it has become clear that Xenopus Mi-2 is a component of a histone deacetylase complex (41, 89). This complex contains nucleosome remodelling activity and it has been demonstrated that several transcription factors
repress gene activity by recruiting histone deacetylases. It is suggested that nucleosome remodelling and deacetylase activities of a Mi-2 complex, recruited by the interaction with the transcription factor Hunchback (Hb), may result in local chromatin changes that allow binding of PcG proteins. Alternatively, the proposed Hb-Mi-2 complex may directly bind to PcG proteins and recruit the complex to DNA. A role in chromatin remodelling is also found for TrG proteins, a group of proteins that counteract the activities of PcG proteins. Several members of the TrG, GAGA factor and brahma (45, 85) are involved in nucleosome remodelling processes. Interestingly, several proteins of the TrG, GAGA factor and Trithorax (TrX, ) associate with PREs which also bind PcG proteins, in *Drosophila* (60, 83). Binding sites for the DNA binding protein GAGA factor have been found in several PREs. GAGA factor is involved in disruption of nucleosome positioning together with the nucleosome remodelling complex NURF. It has been proposed that GAGA factor functions as an anti-repressor protein, facilitating the formation and maintenance of nucleosome free regions (18, 85). As such GAGA factor could be constitutively bound to PREs either to recruit the TrG complex or to facilitate the binding of the PcG complex to a nucleosome free region of a PRE.

Although several findings indirectly suggest that PcG mediated silencing involves nucleosome remodelling, no direct interactions between proteins of the PcG complex and proteins involved in chromatin remodelling have been found yet.

**5.4 Models.**

At present, it is common knowledge that PcG proteins maintain the repressive state of genes and that changes in chromatin structure may be involved to achieve this. What are the chromatin, architectural structures which can provide stable repression of gene activity. Hitherto, mainly two different models can be distinguished, the “spreading” model and the “looping” model. The spreading model involves a general and global association of PcG complexes covering a large part of the chromatin fibre. The looping model, on the other hand, focusses on specific, local association of PcG complexes to the DNA.

Immunoprecipitation studies of PcG proteins from *in vivo* cross linked chromatin, in *Drosophila*, initially showed that the Pc protein is associated with a stretch of approximately 240 kilobases. It had been found that the PcG protein covered the entire inactive *Ubx* and *Abd-A* genes and regulatory regions of the Bithorax cluster (59). This finding lead to the idea that PcG proteins bind to the chromatin and spread over the chromatin fibre, covering hundreds of kilobases. The “spreading” model argues that DNA has become inaccessible for factors of the transcription machinery, thereby repressing gene activity (59, 61). This model is similar to telomere silencing by the SIR proteins in yeast. In this model a specific DNA binding protein binds to a target sequence at the telomeres and recruits the SIR proteins, which can interact with each other, thereby covering the DNA.

The “spreading” began to lose its’ supporters with the finding that the PcG proteins actually do not spread over many hundreds of kilobases but bind only locally at the PREs and the regulatory region of a gene, covering not more than several kilobases (60, 82, 83). The “looping” model argues that individual PcG proteins bind to several sites, which may be several kilobases away from each other, surrounding a gene (including a PRE). The individual PcG proteins interact with each other resulting in a stabilized PcG complex and looping of the DNA. The looping of the DNA will provide a higher chromatin density and a larger PcG complex (64, 65). This idea is in agreement with the findings that two PREs inserted at homologues sites or even at distant sites enhance one another’s silencing
activity. The enhanced silencing may result from chromosome pairing by the interaction of PcG proteins which are located at different PREs, forming large PcG complexes.

A different option is that no significant structural chromatin changes are induced at all and that PcG mediated silencing is merely a result of transcription interference. PcG mediated gene repression, then, could function through interference of the PcG complex with the basal transcription machinery, prevent the binding of transcription factors or prevent the interaction of enhancers and promoter (61). In the light of the possible nucleosome remodelling activities of PcG mediated silencing (outlined in 5.3), a local chromatin remodelling and interference with the basal transcription machinery is feasible. In this model no drastic chromosomal changes occur.


PcG and TrG proteins function in the maintenance of the gene transcription states which are established early in embryogenesis. They are part of a cellular memory system, which is responsible for the stable inheritance of gene activity. Interestingly, the active and inactive chromatin states are submitted through mitotic and, in some cases, meiotic cell divisions, as reported in Drosophila, yeast and plants (13, 28, 36). Meiotic inheritance of determined states is striking in the light of the fact that the developing embryo must be able to erase the epigenetic information of its parental gametes in order to allow differentiation of a variety of cell lineages. PcG and TrG proteins probably act through similar chromosomal elements, acting as a molecular switch either to direct the adjacent chromatin into an active or inactive state. To sustain cellular memory, such chromatin states need to be heritable. How is this cellular memory achieved?

To achieve stable repression over many cell generations PcG proteins have to reassemble at the right places during each cell division or after transcriptional displacement. It has been proposed that PREs of active genes are marked by proteins or chromatin modifications that prevent the formation of PcG silencing complexes (66). The other possibility is that PcG proteins remain associated on silenced. Recently it has been demonstrated that different human PcG proteins remain chromatin-associated near centromeric regions throughout all stages of mitosis, in several human cell lines. In developing Drosophila, it was shown that between 5 and 7% of several PcG proteins are bound to metaphase chromosomes (8, 71). It is suggested that by maintaining chromosome-associated PcG proteins/complexes during mitosis the PcG mediated inactive gene state is tagged. In this way the sites of PcG action are ensured to be inherited by successive cell generations, thus providing the daughter cells with the necessary components for maintaining gene expression patterns.

7. Scope of the thesis

In order to study the role of PcG proteins in chromatin mediated gene regulation we aimed to unravel the composition of the human PcG protein complex and to functionally characterize its components. In chapter 2 and 3 the identification of proteins that are part of the human PcG complex is described. HPH1 and HPH2, two human homologs of the Drosophila PcG protein Polyhomeotic are identified to interact with the vertebrate PcG protein Bmi1 (chapter 2). RING1, a protein that hitherto has not been recognized as a PcG protein, is being found to interact with vertebrate Polycomb homologs (chapter 3). In chapter 4 a human homolog of Drosophila Pc, HPC2 is described. Interference with HPC2 function results in misregulation of the c-myc proto-oncogene, cellular transformations of human cell lines and apoptosis. Chapter 5
describes multiple interactions between RING1 and other human PcG proteins, indicating that RING1 is an integral component of human PcG complexes. Interference with RING1 function results in the misregulation of the \textit{c-jun} and \textit{c-fos} oncogenes and the \textit{enlarged} homeotic gene, accompanied with cellular transformations and the induction of tumors in nude mice. Finally, in chapter 6 the role of distinct PcG complexes in the regulation of different target genes is discussed.

References


1. General introduction

group-related mammalian gene, encodes a transcriptional negative regulator with tumor suppressive activity. EMBO J., 14: 5672-5678.


1. General Introduction

