SYNTHEtic Reversal of Epigenetic Silencing

Haynes, Karmella A. 1 and Pamela Silver 1,2

Department of Systems Biology 1, Harvard Medical School and The Wyss Institute of Biologically Inspired Engineering 2, Harvard University, Boston, Massachusetts 02115

Running head: Synthetic Reversal of Epigenetic Silencing

Address correspondence to: Pamela Silver, Ph.D., 200 Longwood Avenue WAB 536, Boston, MA 02115. Fax: 617-432-5012; E-mail: pamela_silver@hms.harvard.edu

Controlling cell fate-determining gene expression is key to stem cell differentiation, tissue regeneration, and cancer therapy. To date, custom-built transcription factors recognize the information encoded in specific DNA sequences. Chromatin proteins undergo covalent modifications and form complexes that encode a second layer of information that determines proximal gene activity. Here, we employ a novel gene targeting approach that exploits a specific chromatin modification to reactivate silenced loci in human cells. We used the human Polycomb chromatin protein and homologues from other species to construct modular synthetic transcription factors, called Pc-TFs, that recognize the repressive tri-methyl histone H3 lysine 27 (H3K27me3) signal and switch silenced genes to an active state. Pc-TF expression in U2OS osteosarcoma cells leads to increased transcription of the senescence locus CDKN2A (p16) and other loci in a chromodomain and activation module-dependent manner, a switch to a senescence phenotype, and reduced cell proliferation. These results indicate that silenced developmental regulators can be re-activated by a synthetic transcription factor that interacts with chromatin rather than DNA, resulting in an altered cell state. As such, our work extends the flexibility of transcription factor engineering and is the first example of chromatin-mediated synthetic transcription factor targeting.

Chromatin plays a pivotal role in regulating gene expression by establishing heritable cell states. Unlike transcription factors that bind specific DNA sequences, many chromatin-associated transcription regulators bind to the histone proteins that package DNA into nucleosomes. Post-translational modification of the residues along the unfolded N-terminal histone tails is a dynamic mechanism that allows a single genetic locus to assume different transcriptional states. Enzymes catalyze the covalent attachment and removal of methyl, acetyl, and other molecular groups. Effector proteins specifically recognize these marks and establish gene silencing or activation by altering the accessibility of the locus to RNA polymerase and other transcription factors (1, 2). One of the most well-studied chromatin modification systems is Polycomb, an evolutionarily conserved set of proteins that was first characterized in Drosophila (3). Tumor suppressors and cell differentiation genes are subject to Polycomb-mediated regulation in mammals (4, 5).

The Polycomb silencing mechanism involves two core complexes that generate and interact with histone methylation. Polycomb Repressive Complex 2 methylates histone H3 lysine 27 (Fig. 1A). The trimethylated lysine (H3K27me3) is bound by the Polycomb chromodomain of CBX, a member of Polycomb Repressive Complex 1 (Fig. 1B) (3, 6), which may create condensed chromatin that impedes RNA Polymerase activity (6). Transformations of tissue identity in developing animals upon mis-expression of Polycomb proteins (7-9) underscore the importance of this gene regulation system in cell fate determination.

Our knowledge of chromatin’s role in development and disease has spurred progress in artificial manipulation of epigenetic states. Several studies have revealed Polycomb factor mis-regulation and abnormal accumulation of silenced chromatin in cancer cells (10-12), prompting the development of anti-cancer drugs that globally inhibit repressive chromatin modifications (13). In order to advance the technology for manipulating transcriptional states, we have developed a modular Polycomb-
based artificial transcription factor (Pc-TF), consisting of the methyl-histone-binding Polycomb chromodomain and the VP64 activator (14), that targets and activates H3K27me3-associated Polycomb-repressed loci (Fig. 1C). Unlike small molecule drugs, synthetic transcription factors are modular, a characteristic that allows us to potentially tune Pc-TF specificity.

**EXPERIMENTAL PROCEDURES**

**Cell culture and transfection** - HEK293 and U2OS Flp-in T-REx cells were grown in Dulbecco’s modified Eagle’s medium (DMEM) or McCoy’s 5A medium (respectively) supplemented with 10% tetracycline-free fetal bovine serum and 1% penicillin and streptomycin. Cells were grown at 37°C in a humidified CO2 incubator. Transient transfections of HEK293 cells were carried out in 24-well culture dishes using 1.8 ul Fugene6 (Roche) in 1 ml medium. U2OS stable cell lines were generated by transfecting 5x 10^5 cells with 2 µg plasmid DNA plus 4 µl Lipofectamine 2000 (Invitrogen) in antibiotic-free medium for 6 h. Two days after transfection wash out, cells were grown under 1.0 µg/ml puromycin selection, screened for inducible mCh expression upon treatment with 1 µg/ml dox, and 2 – 3 clones were expanded and maintained as stable lines.

**Plasmid constructs** - DNA fragments with universal cloning sites (EcoRI, NotI, XbaI, SpeI, and PstI) were generated by PCR and assembled according to a modified BioBrick DNA assembly method (15). Human codon optimized human, fly, and zebrafish Polycomb chromodomains were synthesized by Genscript. The Gal4 transcription activator and deletion variants were cloned as EcoRI/ XbaI fragments into pcDNA3.1(+) (Invitrogen) in antibiotic-free medium for 6 h. Two days after transfection wash out, cells were grown under 1.0 µg/ml puromycin selection, screened for inducible mCh expression upon treatment with 1 µg/ml dox, and 2 – 3 clones were expanded and maintained as stable lines.

**Luciferase activity assays** - HEK293 23;4;9 cells (16) were treated with 1 µg/ml dox expression for 2, 4, 6, or 8 days to induce Gal4-EED expression. 4-day dox treated or untreated cells were plated at ~1x10^7/well in 24-well culture dishes, transfected with 200 ng of plasmid DNA, grown without dox for 48 h, and harvested for luciferase assays (Bright-Glo kit, Promega). Scepter digital counter (Millipore) was used for cell counting.

**Microscopy** – Transiently transfected HEK293 cells were imaged in 24-well plates at 200x using a Nikon fluorescence microscope (Model TE2000-E). U2OS cell lines were seeded at ~5x10^4/well on glass cover slips in 24-well culture dishes, treated with 1 µg/ml dox overnight, fixed with 4% paraformaldehyde/ 1x PBS for 10 min., and permeabilized with 0.5% Triton-X 100/ 1x PBS for 10 min. Slips were inverted into vinol mounting medium on glass slides and imaged at 600x. Microscopy of live colonies was carried out on cells plated at low density ( ~600 cells/cm^2) in 6-well plates and grown in the presence or absence of 1 µg/ml dox for 5 days.

**Antibodies for Western blot analysis** - Primary staining with rabbit anti-DsRed 632946 (Clontech) was followed by secondary staining with donkey anti-rabbit-HRP (Promega) and mouse anti-GAPDH-HRP conjugate G9295 (Sigma-Aldrich).

**Reverse transcription PCR** - SuperScript III (Invitrogen) was used to generate cDNA from 2 µg of total RNA isolated with either TRIzol (Invitrogen) or the RNeasy kit (QIAGEN). See Table S1 for primers. Real-time quantitative PCR reactions (15 µl each) contained SYBR Green master mix (Applied Biosystems), 2.25 pmol primers, and 2 µl of a 1:10 cDNA dilution (1:1000 dilution for GAPDH and mCh).

“Relative mRNA level” was calculated as 2^[C(t)GAPDH - C(t)experimental gene].

**Cross-linked chromatin immunoprecipitation** – ~1x10^8 cells were treated with 1 µg/ml dox for 5 days, collecting by scraping, and incubated 25 ml 1% formaldehyde/ 1x Dulbecco’s PBS for 10 min at room temperature. Cross-linking was quenched with 25 mM glycine for 5 min. Cross-linked cells were washed once with 1x PBS and twice in 175 µg/ml PMSF/ 1x PBS.
Cell lysis and chromatin preps were carried out in buffers supplemented with 175 µg/ml PMSF and 12.5 mg/ml PLAAC protease inhibitor cocktail. Cells were incubated for 10 min at 4°C in 10 ml lysis buffer (62.5 mM HEPES-KOH pH 7.6; 146 mM NaCl; 1 mM EDTA; 10.4% glycerol; 0.5% NP-40; 0.26% Triton-X 100), washed once in 8 ml wash buffer (200 mM NaCl; 1 mM EDTA; 0.5 mM EGTA; 20 mM Tris pH 8.0), and disrupted in 6 ml sonication buffer (10 mM EDTA; 5 mM EGTA; 204 mM Tris pH 8.0) using a Fischer Scientific Sonic Dismembrator (Model 550). Soluble chromatin was incubated with 5 mg/ml N-laurylsarcosine for 10 min and spun at ~10,000 rcf to remove impurities. DNA was purified from a chromatin sample by phenol/chloroform extraction and resolved via electrophoresis to confirm ~500 bp fragments.

For immunoprecipitations, each 6 ml sample was supplemented with 1.2% Triton X-100, 0.12% Na-deoxycholate, 0.8x TE pH 7.5, PLAAC, and PMSF. U2OS Flp-in T-REx chromatin was incubated with rabbit anti-H3K27me3 (Millipore) or rabbit IgG (sc-2027, Santa Cruz Biotech.) overnight, and Protein A sepharose beads for 3 h at 4°C. Chromatin from Pc-TF or TF expressing cells was precipitated with mouse anti-myc 3400 or mouse IgG 3420 bead conjugates (Cell Signaling Technology). Chromatin-antibody-bead complexes were washed essentially as described by Millipore (ChIP Assay MCPROTO407). DNA-protein complexes were eluted in 0.5 ml 1% SDS/TE. 100 µl of chromatin (input) was brought up to 500 µl with 1% SDS/TE. Samples were treated with Pronase for 1 h at 42°C, then incubated at 65°C for 48 h to reverse cross-linking. DNA was extracted twice with phenol/chloroform:isoamyl alcohol (25:24:1), once with chloroform, supplemented with 15 µg GlycoBlue (Ambion) and 200 mM NaCl, and ethOH-precipitated. DNA was resuspended in 0.5 ml 10 mM Tris pH 8.0.

**RESULTS**

**Pc-TF reactivates a Polycomb-repressed reporter gene.** In order to determine whether Pc-TF recognizes methylated histones and reactivates a silenced locus, we used an inducible Polycomb silencing reporter system in HEK293 cells (16). Polycomb chromatin formation at the GAL4TK-Luc transgene (loc) in HEK293 cells is mediated by Gal4-EED (silencer). Doxycycline (dox) treatment activates expression of the silencer, leading to Luc repression and accumulation of H3K27me3 and Polycomb proteins at the promoter (16). Expression of a synthetic activator that consists of a Gal4 DNA binding domain (Gal4DB), red fluorescent mCherry tag, VP64, and an SV40 nuclear localization signal leads to a ~10-fold increase in luc expression over basal expression levels (Fig. S1A). Conversely, 4 to 8 days of dox-induced silencer expression leads to a ~7-fold decrease of luc activity (Fig. S1B). Following luc silencing for 4 days, recruitment of VP64 directly upstream of the promoter leads to luc reactivation at levels ~10-fold greater than initial basal expression levels (Fig. S1C). Deletion of the Gal4DB or the VP64 domain abolished luc reactivation, indicating that VP64 stimulates transcription when it is recruited near the promoter and that the Gal4DB alone is not sufficient to reactivate luc.

The synthetic transcription activator was reengineered to recognize H3K27me3 by replacing the Gal4DB with a Polycomb...
chromodomain (PCD). hPCD-TF includes amino acids 1-62 from human CBX8, a peptide that preferentially binds H3K27me3 in vitro (17). We also built Pc-TFs using the chromodomains from a distant vertebrate (zebrafish) and flies (Drosophila) to investigate whether variations in amino acids around the conserved hydrophobic residues (Fig. 2A) might alter Pc-TF activity. fshPCD-TF includes amino acids 2-63 of zebrafish Pc1, a CBX homologue that acts during embryonic development (18). flyPCD-TF includes amino acids 16-77 of Drosophila Polycomb (Pc), a peptide that shows preferential H3K27me3 binding (19, 20) and confers Polycomb-like genomic localization when the peptide is tethered to a heterologous protein (19, 21).

Human Pc-TF-mediated reactivation of the Polycomb-silenced luc reporter relies upon a previously established silenced state. Mitotically stable luc silencing (16) was generated by transient expression of the repressor (4 days of dox treatment), followed by dox wash-out and growth in dox-free medium for 10 days. Cells expressing the human Pc-TF show a modest significant increase in luc expression compared to the unsilenced reporter (Fig. 2C). Reactivation by human Pc-TF requires previous Gal4-EED expression and the PCD module, suggesting that Pc-TF is recruited to the reporter gene through a PCD-H3K27me3 interaction.

Pc-TFs activate transcription at methyl-histone associated loci. We identified Pc-TF targets by looking for genes that were up-regulated upon Pc-TF expression. Transgene constructs were placed under the control of a dox-inducible CMV promoter (Fig. S2A) and integrated into a U2OS cell line that expresses a tetracycline repressor (Fip-in TREx, referred to herein as the “parent line”). Quantitative reverse transcription PCR (qRT-PCR) (Fig. S2B) and Western blots (Fig. S2C) confirmed dox-inducible expression. The mCh tag showed clear nuclear localization in all cell lines (Fig. 3A). In order to identify Pc-TF target genes, total RNA from cells grown with or without dox for 96 h was used for Nanostring quantitative RNA analysis of 500 development and differentiation-associated genes (22). We tested 14 positive hits by qRT-PCR of duplicate RNA samples (Fig. S3). CDKN2A (p16) senescence-associated gene, matrix metallopeptidase 12 (MMP12), natriuretic peptide (NPPA), and tumor necrosis factor (TNF) were up-regulated by one or more of the Pc-TF cell lines in a VP64- and PCD-dependent manner.

Cross-linked chromatin immunoprecipitation (ChIP) assays showed enrichment of histone methylation near the promoters of CDKN2A, MMP12, and NPPA (Fig. 3C, S4), but not at TNF (data not shown) suggesting that the first three genes are direct targets of Pc-TF. hPCD-TF was significantly enriched (p < 0.01) near the p16 promoter (probe 3) compared to a non-targeted protein in which the PCD was deleted. All three Pc-TF homologues showed significant enrichment near the MMP12 promoter (probe 6). Thus, the ChIP data are consistent with the qRT-PCR data for p16 and MMP12, indicating that these genes are direct targets of the Pc-TFs. We did not detect enrichment of Pc-TFs at promoters lacking H3K27me3 (TNF, ATOH1, ACTC1, and CASP10, data not shown) indicating that H3K27me3 is required for Pc-TF binding at endogenous loci. Interestingly, the H3K27me3-enriched promoters of NPPA and eomesodermin (EOMES), an additional gene included in our analysis (Fig. 3C, S4), show no Pc-TF enrichment, showing that Pc-TF may preferentially bind a subset of H3K27me3 sites.

Human Pc-TF expression is accompanied by a switch to cellular senescence. Consistent with p16 activation, hPCD-TF expressing cells showed senescence-associated phenotypes. Hallmarks of senescence include increased levels of β-galactosidase (β-gal) activity and decreased cell proliferation (23). Treatment with C12FDG followed by flow-cytometry has been used to measure senescence-associated β-gal activity in rotenone-stressed or aged senescent human fibroblasts (24). We optimized the assay for U2OS cells and found that 94 h of exposure to 0.2 μg/ml rotenone, followed by staining with 1.5 μg/ml C12FDG, was sufficient to detect an increase in SA-β-gal activity over background levels. We observed senescence associated SA-β-gal activity in Pc-TF-expressing cells at levels comparable to cells stressed with rotenone. We repeated these results in two independent hPCD-TF expressing cell lines, and not in cells expressing VP64 or PCD.
deletion proteins or in parental cells (Fig. 4A, B).

To measure changes in cell proliferation, we plated cells at low density and observed growth of isolated colonies after 5 days of dox treatment. Colony size is diminished in hPCD-TF-expressing cells compared to cells that do not express hPCD-TF (Fig. 4C). Our results demonstrate that the chromatin-targeted synthetic transcription activator induces a switch in cell state, from rapid proliferation to senescence, in the U2OS cancer cell line.

**DISCUSSION**

Synthetic biology presents opportunities to deepen our understanding of molecular biology while creating useful new tools (25). Using the Polycomb chromodomain (PCD) to target a transcription activator to methylated histones has revealed some complexities underlying chromatin interaction. The H3K27me3-associated luc reporter (Fig. 2C) and CDKN2A are targeted by the human Pc-TF but not by the zebrafish or fly homologues. Perhaps the human PCD participates in human protein-specific interactions with other chromatin-associated proteins in addition to H3K27me3. Interestingly, MMP12 is targeted by all three Pc-TFs in a PCD-dependent manner. Thus, at least one locus has a chromatin structure that permits general PCD binding. Finally, our data indicate that not all H3K27me3-associated regions are bound by Pc-TF. hPCD-TF is enriched at the region immediately upstream of the CDKN2A promoter, but not at the other H3K27me3-enriched regions. Two H3K27me3-associated promoters (NPPA and EOMES) showed no Pc-TF enrichment, indicating that certain regions may be more accessible to Pc-TF binding than others.

So far, we have confirmed p16 as a target of Pc-TF, which has implications for cancer therapy. Polycomb-mediated suppression of the senescence-associated p16 gene may augment cancer proliferation (23, 26). Recently, drugs that broadly disrupt chromatin-mediated gene silencing in cancer cells have entered clinical trials (13), but lack specificity at the target gene level and may cause off-target effects. Pc-TFs integrate the information from epigenetic modifications with transcriptional regulation, providing a way to target therapeutic genes in a disease state-specific manner.

Manipulating Polycomb-mediated gene regulation is also relevant for regenerative medicine. Human stem cells can be differentiated in vitro by depleting Polycomb factor levels and activating repressed genes (27). Using Pc-TFs to reverse epigenetic silencing at these loci has the possible advantage of enabling reconfiguration of the modular domains to adjust the range of activated target genes. Thus far, we have shown that changing the PCD module influences target gene specificity in the case of p16 and MMP12 (Fig. 3B, C). Changing the TATA motif-dependent VP64 module (14) to a different activator, such as SP1 or p65 (28, 29) could expand the range of Pc-TF-activated genes. Ongoing identification of chromatin recognition domains (30) suggests great potential for synthetic transcription factors that interact with a variety of chromatin modifications. Our development of synthetic Polycomb-based transcription activators establishes a novel strategy that integrates the information encoded in chromatin with targeted gene regulation.

**REFERENCES**


**FOOTNOTES**

For their generous gifts of cell lines, DNA, and other resources we thank K. Hansen (HEK293 23;4;9 cells), S. Blacklow (U2OS TREx Flp-in cells), and A. Meissner and C. Gifford (Nanostring assay). This work was supported grants GM057476 and GM36373 to P.A.S. and 1F32GM087860 to K.A.H from the NIH.

The abbreviations used are: H3K27me3, histone H3 trimethylated at lysine 27; Pc-TF, Polycomb-based artificial transcription factor; Gal4DB, Gal4 DNA binding domain; PCD, Polycomb chromodomain

**FIGURE LEGENDS**

Fig. 1. Using the Polycomb system to engineer a chromatin-targeted transcription factor. A. Model of H3K27 methylation by EZH1/2 of Polycomb Repressive Complex 2 (PRC2) (3, 6). B. Gene repression by
Polycomb Repressive Complex 1 (PRC1) (3, 6). Trimethylation of H3K27 acts as a docking site for the PRC1 complex. C. Gene activation by a Polycomb-based synthetic transcription factor (Pc-TF). A conserved 62 amino acid Polycomb chromodomain from CBX8 of the PRC1 complex, fused to a transcription activation domain, binds H3K27me3. As a result, the silencing mark is translated into gene activation.

Fig. 2. Pc-TF-mediated activation of a Polycomb-repressed reporter gene in HEK293 cells. A. Fusion protein gene constructs. ClustalW alignment of the Polycomb chromodomains from human CBX8 (hPCD), zebrafish CBX2 (fshPCD), and Drosophila Pc (flyPCD). Boxes mark conserved aromatic residues within the H3K27me3 binding pocket (19, 20). Pc-TF fusion proteins (h, fsh, and flyPCD-TF) contain a red fluorescent mCherry (mCh) protein, VP64 activation domain, SV40 nuclear localization signal (NLS), and myc epitope tag. VP64 deletions (h, fsh, and flyPCD-RFP) and a PCD deletion (TF) serve as negative controls for Pc-TF activity assays. B. Nuclear localization of fusion proteins in transiently transfected HEK293 cells. Dashed lines mark the periphery of representative mCherry-positive cells. Scale bar = 20 µm. C. hPCD-TF-mediated GAL4TK-Luc (luciferase) reporter gene activation requires prior exposure to Polycomb silencing. The schematic illustration (left) shows the predicted behavior of the synthetic Pc-TF in the absence or presence of histone methylation, controlled by doxycycline-induced expression of Gal4-EED. Pc-TF fusion proteins and deletion controls were expressed in cells before and after doxycycline (dox) treatment. Error bars represent standard deviation (s.d.), n = 3 measurements. A one-tailed t-test was used to compare the values under the horizontal bracket (*, p < 0.01).

Fig. 3. Endogenous genes targeted by synthetic Pc-TFs in U2OS cells. A. Expression of Pc-TFs and deletion variants in U2OS cells. Dotted lines mark the periphery of representative mCherry-positive cells. Scale bar = 20 µm. B. Reverse transcription PCR of cDNA synthesized from total RNA of untreated or dox-treated cells. Relative expression levels are qPCR values normalized to the untreated parental line (P.L., -dox). Error bars represent standard error of the mean, n = 2 RNA samples. C. ChIP analysis of Pc-TF target loci. Immunoprecipitated (IP) DNA was amplified using primers 1 – 16 (Table S1). Probes n1 and n2, located at the constitutively active GAPDH promoter, are negative controls. Mean enrichments of IP DNA are shown as percentages of input minus background (non-specific mock IP DNA). Error bars represent s.d., n = 3 quantitative PCR reactions. A one-tailed t-test was used to identify Pc-TF enrichments that were significantly higher (*, p < 0.01) than a randomly cross-linked protein that has no chromatin binding domain (TF).

Fig. 4. Senescence-associated phenotypes in Pc-TF-expressing U2OS cells. A. Histograms of senescence-associated β-gal activity detected by C12FDG staining and flow cytometry (24). Rotenone treatment of the parental cell line (P.L.) was used as a positive control. B. Average differences in median fluorescence. Error bars represent s.d., n = 2 independent induction trials. C. Proliferation of Pc-TF expressing cells (hPCD-TF, +dox) compared to untreated cells and the parental line.
Figure 1

A

PRC2
SU(N)2
EED
1-4
EED 1-2

H3K27 methylation

Gene

B

Natural Polycomb-mediated gene repression

PRC1
BMII1
PCGF (1-6)
MEI18
CBX
2-4.6.2-8

PhC
1-3
RING
1-4.8

Silenced gene

C

Engineered chromatin-targeted gene activation

Pc-TF fusion protein

Pc

Activated gene
Figure 2

A

![Diagram showing protein sequences and interactions]

B

![Images showing cell morphology with different treatments]

C

![Diagram showing luciferase activity with and without doxycycline]


![Graph showing luciferase activity with different treatments]

Gal4-EED (dox)
Figure 4