

AI-driven designed protein epigenomics

Shiri Levy^{1,2,*}, Hannele Ruohola-Baker^{1,2,*}

¹Institute for Stem Cell and Regenerative Medicine, University of Washington, School of Medicine, Seattle, WA 98109, USA

²Department of Biochemistry, University of Washington, School of Medicine, Seattle, WA 98195, USA

*Authors for correspondence:
Email: shiri007@uw.edu; hannele@uw.edu

Received date: January 24, 2023

Accepted date: February 03, 2023

Copyright: © 2023 Levy S, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Citation: Levy S, Ruohola-Baker H. AI-driven designed protein epigenomics. Clin Res Oncol. 2023;1(1):1-3.

Abstract

The biological revolutions of computationally designed proteins, induced pluripotent stem cells (iPSCs), and the CRISPR-Cas9 system finally enables modifications that can spur deep understanding of spatial requirement of epigenetic information. This commentary describes the utility of a computationally designed protein, EED Binder (EB), when fused to dCas9 (EBdCas9) for identifying critical sites of PRC2 dependent histone H3K27me3 marks in the chromatin. By using EBdCas9 and gRNA, PRC2 function can be inhibited at specific loci, resulting in precise reduction of EZH2 and H3K27me3 marks, and in some (but not all) locations, activation of the gene and functional outcomes (such as regulation of cell cycle or trophoblast transdifferentiation). Interestingly, a functional TATA box located more than 500bp upstream of a TBX18 TSS was found to be repressed by PRC2, supporting the theory that epigenetic regulators control the repression of transcriptional elements on the promoter region. The EBdCas9 technology may provide a useful tool for controlling gene regulation through epigenomic control.

Main body

While H3K27me3 marks are broadly decorating the promoter regions of developmental genes, and generally associate with gene repression [1], a central question in the field is which specific PRC2-dependent H3K27me3 marks (if any) are essential for gene repression. The development of a new tool [2,3], for the first time enables spatial dissection of specific H3K27me3 histone marks associated with individual genes.

PRC2 (Polycomb Repressive Complex 2) plays a crucial role in the process of determining the fate of cells during development by modifying chromatin through a specific histone modification called H3K27me3. Among the PRC2 proteins which facilitate new H3K27me3 modification on the chromatin are the reader EED which anchor the PRC2 complex to the chromatin and the writer/catalyzer, methyltransferase, EZH2 which methylates the lysine 27 sidechain in histone 3 (H3K27). To understand the requirement of the PRC2 in different developmental stages of pluripotency we generated a computer-designed protein that binds the EZH2 binding site in EED (EED binder; EB), thereby competing with EZH2 [2]. EB binding to EED inhibits EZH2 interactions with EED, resulting in EZH2 and JARID2 degradation, a global reduction of H3K27me3 marks, and ultimately, activation of thousands of PRC2-dependent developmental genes [2]. To answer the question of whether any specific H3K27me3 marked nucleosomes are critical alone for the control of gene expression we fused EB to catalytically dead Cas9 (EBdCas9). In the past decade, numerous groups utilized dCas9 as a versatile tool for epigenomic editing using effector and repressor molecules to shed light on gene regulation, cellular differentiation, and as potential therapeutic agents in pathological diseases and regenerative medicine [4-19]. However, EBdCas9 is the first computer designed protein used in this context.

Targeting EBdCas9 and appropriate single guide RNA to specific chromatin locus on a promoter region of a gene, such as the transcription factor in pacemaker cells, TBX18 results in PRC2 repression and activation of TBX18 gene expression. Moreover, this derepression of TBX18 gene was inherited to the daughter cells since Tbx18 was expressed in subsequent cell divisions after removal of EBdCas9 and the guide RNA. While the repression of EZH2, H3K27me3, and JARID2 by EBdCas9/gRNA were initially limited to a single gRNA determined locus, in specific locations epigenomic changes readily spread onto the promoter region towards the transcription start site,



Figure 1. Lady Histone's curly hair mimics heterochromatin (orange rolls) regions, and euchromatin (green rolls) regions following epigenomic remodeling using the novel designed protein that inhibits the repressive epigenetic modifications.

TSS. Additionally, the previously facultative heterochromatin was replaced with histone marks associated with gene activation, such as H3K27ac. The enzyme which catalyzes the transfer of acetyl marks to histone, histone acetyltransferase, p300 was also recruited to the locus. The generated euchromatin was characterized by the presence of RNA polymerase 2, and the hydroxylation of DNA methyl marks at CpG island proximal to TSS. Therefore, the targeted PRC2 inhibition results in epigenomic remodeling, native activation of the transcriptional machinery, gene expression and inherited epigenomic memory.

Interestingly, we identified two kinds of PRC2 responsive loci, the loci in which PRC2 function was essential for gene repression, and loci in which PRC2 did not function to repress transcription. We sought to understand what is/are the molecular mechanism/s responsible for such profoundly different effects.

In the case of TBX18, one unique identifier was detected on the promoter region of the gene, in close proximity to the impactful single guide RNA that resulted in substantial epigenomic and transcriptional changes. The Element prediction tool [20] identified a potential TATA box in close proximity to the guide RNA. However, this domain was atypically far, >500bp upstream of TSS (typically TATAbox is found -30bp from TSS). Deletion of this identified upstream TATA box eliminated the epigenomic remodeling and transcriptional activation of the TBX18 proving that the domain was a bona fide TATA box. Moreover, targeted PRC2 inhibition by EBdCas9 allowed the recruitment of TATA Binding Protein (TBP) proving that the distal TATA box normally masked by PRC2 repressive marks can be functional in regulating TBX18 expression.

Control of epigenomic regulation by AI-designed proteins holds immense therapeutic promise for human disease without using chemical drugs or manipulation of endogenous gene sequence. Diffuse Midline Glioma (DMG), is an aggressive form of brain cancer that develops in glial cells in the pons, primarily affecting children [21,22]. A lysine-to-methionine mutation in oncohistone H3.3 gene (H3.3K27M) defines DMG tumors [21,22]. In rapidly dividing cells, such as in DMG, the cell cycle regulator *p16* is repressed due

to hypermethylation at the promoter area, however, the capacity to activate *p16* for cell cycle arrest could lead to a breakthrough in DMG therapeutics. Transient introduction of EBdCas9 and appropriate guide RNA to *p16* gene promoter region of DMG patient cell line led to reduction of EZH2 and H3K27me3 epigenomic marks, *p16* transcript and protein activation. Importantly, impactful reduction of over 50% in DMG cell replication and viability was observed in these experiments [3], supporting the validity of AI-designed EpiBinders in future therapies.

Finally, regenerative medicine offers the potential to repair diseased cells and innovatively provide new gene and cell therapies for previously intractable conditions. One of the hallmarks of regenerative medicine is the ability to transdifferentiate one cell type to another beyond their natural bifurcation point. Likewise, the extraembryonic placental vs embryonic cellular fate decision (the first in eutherian development) is dependent on PRC2 [23]. While abundant H3K27me3 marks are associated with embryonic lineage, genome-wide depletion of H3K27me3 marks is associated with extraembryonic placental lineage [23-28]. Primed stage iPSCs are considered embryonic-like cells, which have already passed the bifurcation point and cannot readily generate extraembryonic lineage. Therefore, we used primed stage iPSC as a transdifferentiation model to directly reprogram primed iPSC to extraembryonic placental cells. Directing EBdCas9 with appropriate guide RNAs to multiple transcription factor genes at primed iPSC stage resulted in the transdifferentiation of iPSCs to placental progenitor cells (confirmed by RNAseq analysis, unique placental morphological changes, and placental markers).

Conclusions

The field of regenerative medicine is rapidly growing and offers a wide range of potential therapies for the treatment of various diseases. One promising area of research within regenerative medicine is the use of epigenetic modifiers towards epigenetic therapies. The absolute gem in epigenetic therapies is the ability to correct defects in gene regulation without altering the underlying genetic code. In this commentary we have discussed how a computer designed

mini-protein fused to dCas9 can specifically activate genes through releasing epigenetic repression [3]. In the future these novel AI-designed epigenetic competitor- or recruiter-proteins may join the forefront of the epigenetic modifiers towards effective and impactful therapeutics. Already a generation of AI-designed EpiBinder proteins that aim to target different epigenomic pathways is underway, suggesting a birth of an emerging revolution; AI-driven designed protein epigenomics.

Acknowledgments

We thank Dr. Julie Mathieu, Dr. Thelma Escobar, and Clara McCurdy for helpful comments and discussions. This work was supported in part by grants from Brotman Baty Institute (BBI) for Precision Medicine (Catalytic Collaboration Award) and National Institute of Health R01GM097372, R01GM97372-03S1, R01GM083867, 1P01GM081619 for HR-B.

References

- Margueron R, Reinberg D. The Polycomb complex PRC2 and its mark in life. *Nature*. 2011 Jan 20;469(7330):343-9.
- Moody JD, Levy S, Mathieu J, Xing Y, Kim W, Dong C, et al. First critical repressive H3K27me3 marks in embryonic stem cells identified using designed protein inhibitor. *Proceedings of the National Academy of Sciences*. 2017 Sep 19;114(38):10125-30.
- Levy S, Somasundaram L, Raj IX, Ic-Mex D, Phal A, Schmidt S, et al. dCas9 fusion to computer-designed PRC2 inhibitor reveals functional TATA box in distal promoter region. *Cell Reports*. 2022 Mar 1;38(9):110457.
- Qi LS, Larson MH, Gilbert LA, Doudna JA, Weissman JS, Arkin AP, et al. Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. *Cell*. 2013 Feb 28;152(5):1173-83.
- Pengue G, Lania L. Krüppel-associated box-mediated repression of RNA polymerase II promoters is influenced by the arrangement of basal promoter elements. *Proceedings of the National Academy of Sciences*. 1996 Feb 6;93(3):1015-20.
- Groner AC, Meylan S, Ciuffi A, Zangger N, Ambrosini G, Dénerveau N, et al. KRAB-zinc finger proteins and KAP1 can mediate long-range transcriptional repression through heterochromatin spreading. *PLoS Genetics*. 2010 Mar 5;6(3):e1000869.
- Gao R, Dong R, Du J, Ma P, Wang S, Fan Z. Depletion of histone demethylase KDM2A inhibited cell proliferation of stem cells from apical papilla by de-repression of p15 INK4B and p27 Kip1. *Molecular and Cellular Biochemistry*. 2013 Jul;379:115-22.
- Kearns NA, Pham H, Tabak B, Genga RM, Silverstein NJ, Garber M, et al. Functional annotation of native enhancers with a Cas9-histone demethylase fusion. *Nature Methods*. 2015 May;12(5):401-3.
- Shechner DM, Hacisuleyman E, Younger ST, Rinn JL. Multiplexable, locus-specific targeting of long RNAs with CRISPR-Display. *Nature Methods*. 2015 Jul;12(7):664-70.
- Thakore PI, D'ippolito AM, Song L, Safi A, Shivakumar NK, Kabadi AM, et al. Highly specific epigenome editing by CRISPR-Cas9 repressors for silencing of distal regulatory elements. *Nature Methods*. 2015 Dec;12(12):1143-9.
- Amabile A, Migliara A, Capasso P, Biffi M, Cittaro D, Naldini L, et al. Inheritable silencing of endogenous genes by hit-and-run targeted epigenetic editing. *Cell*. 2016 Sep 22;167(1):219-32.
- Pradeepa MM, Grimes GR, Kumar Y, Olley G, Taylor GC, Schneider R, et al. Histone H3 globular domain acetylation identifies a new class of enhancers. *Nature genetics*. 2016 Jun;48(6):681-6.
- Chavez A, Tuttle M, Pruitt BW, Ewen-Campen B, Chari R, Ter-Ovanesyan D, et al. Comparison of Cas9 activators in multiple species. *Nature Methods*. 2016 Jul;13(7):563-7.
- Gilbert LA, Horlbeck MA, Adamson B, Villalta JE, Chen Y, Whitehead EH, et al. Genome-scale CRISPR-mediated control of gene repression and activation. *Cell*. 2014 Oct 23;159(3):647-61.
- Chen LF, Lin YT, Gallegos DA, Hazlett MF, Gómez-Schiavon M, Yang MG, et al. Enhancer histone acetylation modulates transcriptional bursting dynamics of neuronal activity-inducible genes. *Cell Reports*. 2019 Jan 29;26(5):1174-88.
- Hilton IB, D'ippolito AM, Vockley CM, Thakore PI, Crawford GE, Reddy TE, et al. Epigenome editing by a CRISPR-Cas9-based acetyltransferase activates genes from promoters and enhancers. *Nature Biotechnology*. 2015 May;33(5):510-7.
- Adamo A, Sesé B, Boue S, Castaño J, Paramonov I, Barrero MJ, Belmonte JC. LSD1 regulates the balance between self-renewal and differentiation in human embryonic stem cells. *Nature Cell Biology*. 2011 Jun;13(6):652-9.
- Goodman RH, Smolik S. CBP/p300 in cell growth, transformation, and development. *Genes & Development*. 2000 Jul 1;14(13):1553-77.
- Yang JH, Hayano M, Griffin P, Amorim JA, Bonkowski MS, Apostolides JK, Blanchette M, Munding EM, Bhakta M, Salfati EL, Lu Y. Loss of Epigenetic Information as a Cause of Mammalian Aging. *Cell* 2023; 186, 305-326 e327.
- Sloutskin A, Danino YM, Orenstein Y, Zehavi Y, Doniger T, Shamir R, et al. ElemeNT: a computational tool for detecting core promoter elements. *Transcription*. 2015 May 27;6(3):41-50.
- Mohammad F, Weissmann S, Leblanc B, Pandey DP, Højfeldt JW, Comet I, et al. EZH2 is a potential therapeutic target for H3K27M-mutant pediatric gliomas. *Nature Medicine*. 2017 Apr;23(4):483-92.
- Stafford JM, Lee CH, Voigt P, Descostes N, Saldaña-Meyer R, Yu JR, et al. Multiple modes of PRC2 inhibition elicit global chromatin alterations in H3K27M pediatric glioma. *Science Advances*. 2018 Oct 31;4(10):eaau5935.
- Saha B, Home P, Ray S, Larson M, Paul A, Rajendran G, et al. EED and KDM6B coordinate the first mammalian cell lineage commitment to ensure embryo implantation. *Molecular and Cellular Biology*. 2013 Jul 15;33(14):2691-705.
- Banaszynski LA, Wen D, Dewell S, Whitcomb SJ, Lin M, Diaz N, et al. Hira-dependent histone H3.3 deposition facilitates PRC2 recruitment at developmental loci in ES cells. *Cell*. 2013 Sep 26;155(1):107-20.
- Liu X, Wang C, Liu W, Li J, Li C, Kou X, et al. Distinct features of H3K4me3 and H3K27me3 chromatin domains in pre-implantation embryos. *Nature*. 2016 Sep 22;537(7621):558-62.
- Yang J, Ryan DJ, Wang W, Tsang JC, Lan G, Masaki H, et al. Establishment of mouse expanded potential stem cells. *Nature*. 2017 Oct 19;550(7676):393-7.
- Yang Y, Liu B, Xu J, Wang J, Wu J, Shi C, et al. Derivation of pluripotent stem cells with in vivo embryonic and extraembryonic potency. *Cell*. 2017 Apr 6;169(2):243-57.
- Yan L, Yang M, Guo H, Yang L, Wu J, Li R, et al. Single-cell RNA-Seq profiling of human preimplantation embryos and embryonic stem cells. *Nature structural & molecular biology*. 2013 Sep;20(9):1131-9.