Proceedings of the 2nd International Conference on Modern Medicine and Global Health
DOI: 10.54254/2753-8818/45/20240661

Zelda regulation of *slam* transcription in early Drosophila embryo

Yujun Li

Shenzhen College of International Education, 518043, China

s21064.li@stu.scie.com.cn

Abstract. The transcription activation of the zygotic genome is meaningful in the starting stage of development of embryos, it is regulated by a complicated network of protein-DNA interactions and protein-protein interactions including transcription factors and signaling pathways, that follows a time sequence precisely to result in patterned expression and cell fate determination. Zelda as a pioneering transcription factor, modifies expression by binding to highly conserved TAGteam sites. In the present study, we focus on the regulation specific to the gene *slow as molasses(slam*). The results from *in situ* hybridization and ChIP-seq assays support the vital influence of Zelda on the expression of *slam*, yet a TAGteam knockout in vitro mutagenesis experiment is proposed to validate and provide insights to the precise manner of the binding of Zelda. The paper provides analysis of the cis-regulatory event, while more investigations shall be done on the effect of the micro-three-dimensional dynamics of transcription that may result in PolII binding ChIP peaks that appear scattered contiguous crowd, or into the flanking region.

Keywords: genome activation, transcription factors, early embryogenesis, *Drosophila*

1. Introduction

During cleavage phase (0-2hr) of the developing embryo of *Drosophila melanogaster* after fertilization, zygotic genome transcription can be activated under the effect of transcription factors, prior to cellularization at mitotic cycle 14 (2-3hr). Patterning genes are activated timely and precisely to enable the later differentiation of cells that form tissues and imaginal discs in a cascade of regulatory events. One of the essential regulators that is dominantly functional on the global activation of at least 120 genes is a C2H2 zinc finger transcription factor Zelda (Zld), which retains function in the *Drosophila* genus [1,2]. Zld is a pioneering transcription factor as it depletes nucleosomes locally to counteract intrinsic barrier at enhancers, and the extent of barrier-lowering effect depending on number and positions of Zld binding motifs.[3] Zelda is both maternally deposited and zygotically translated with different functionalities regarding the required zinc-fingers. The second zinc finger (ZnF2) showed maternal-specific function. Pre-blastoderm unspliced product ZLD-PB is of a size of 1596 amino acids, with 6 zinc fingers, an N-terminal acidic patch, yet no enzymatic domains have been identified. In contrast with the chromatin level regulatory mechanisms in maternal genomes that include non-sequence-specific recognition of insulator functioning proteins and boundaries between topologically associating domains

DOI: 10.54254/2753-8818/45/20240661

(TADs), Zelda, as a zygotic genome transcription factor, was observed to bind specifically to heptamer motifs called TAG teams (CAGGTAG, TAGGTAG and CAGGTAA). The binding of Zld exhibits a tendency to be near transcription start site (TSS) of target genes, though Zld-ChIP assay shows that it interacts with over two thousand domains [4]. Later studies also model that Zld participates in the formation of hubs involving multiple combinations of interactions between promoters and enhancers, not determined by activations of transcription, and thus in the organization of the spatial cluster of enhancer–promoter cis-regulatory chromatin loops before the formation of zygotic TADs.[5].

Slow as molasses (slam) is one of the ubiquitous genes in Drosophila during superficial cleavage with an expression score of 89.06. Despite germline cells, slam is expressed in all somatic cells of the blastoderm embryo. It is associated with the formation of basal junction for germ cell migration during interface of cycle 14, and might function parallelly with *hmgcr*[6]. Little *slam* is translated in nuclear cycle 1-9, but 10-fold up-regulation of RNA and protein occurs in the 2-3hr window of mid-blastula transition (MBT). [7] The transcript of *slam* appears through the periods of cycle 13 and the slow phase of cellularization [8]. *Slam* is crucial in forming metaphase furrow [9], and its mRNA is suggested to form a complex mediated by some other factors with its encoding protein accumulated at the basal region in a positive feedback loop, performing subcellular localization. The gene is also expressed in adult abdomen (scored 73.98), as well as integumentary system, somatic cell, and mouth part with lower expression scores [10]. Maternal deposited RNA and protein of the gene exist in the embryo.

In the present study we analyze data obtained from in situ hybridized Drosophila early embryos and ChIP binding profiles of Polymerase II (PolII) and Zelda (Zld), mapped to the dm6 release gene model, and interpret the regulatory role of Zld on slam as well as other ubiquitously expressed genes in D. melanogaster through discussion. We design a cas9 knock out experiment to confirm the vital role of TAGteam site binding of Zld and come up with assumptions regarding special structure in the vicinity of *slam*.

2. Methods

Fly strain

Drosophila melanogaster wild type and zld^- mutant.

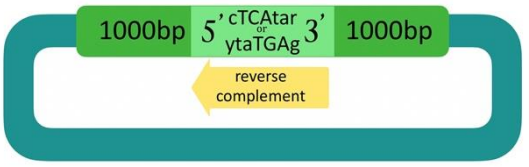
ISH

In situ Hybridization was performed by Y. Ma of Rushlow lab. Probes incorporated with UTP targeting mRNA of slam region were used in antibody staining to visualize locations where expression of the gene takes place.

In vitro mutagenesis

We designed knock-out experiments of TAGteams suggested to bind Zld by performing cas9 double stranded cut with gRNA designed with the support of Chopchop website, and mutated TAGteams manually through homology directed repair. In this oligonucleotide-directed mutagenesis, the site-directed homology arms are 1000bp upstream and downstream of the mutated TAGteam sites. We created primers for PCR of 3 donor plasmids targeted at TAGteam sites in the suggested Zld binding region. The donor plasmids and the guide RNA plasmids will be injected into *Drosophila* germ line at the end of mitosis. However it might be necessary to eliminate the potential influence of gRNA staying bound to DNA or remains close that prevents repair, or of collisions between inactive PolII and cas9.

Schematic representation of donor plasmid:



Sequences of the three pairs of primers:

DOI: 10.54254/2753-8818/45/20240661

Original location on 2L: 6373263-6373258

5’CTAATATTAC(CTATGAG)TTTTCTTTG3’

5’CAAAGAAA(**CTCATAG)**GTAATATTAG3’

Original location on 2L: 6373386-6372295

5’GCTGTTAATC(TTATGAG)ATCGCTGCTA3’

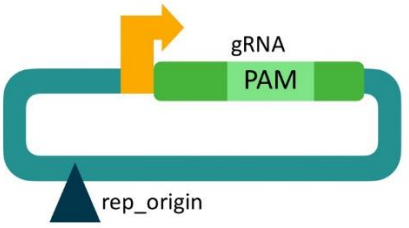
5’TAGCAGCGAT(**CTCATAA)**GATTAACAGC3’

Original location on 2L: 6371125-6373489

5’ATCCATATT(TTATGAG)ATCATTGTA3’ tata

5’TACAATGAT(**CTCATAA)**AATATGGAT3’

Schematic representation of guide RNA plasmid:



Sequences of gRNA:

CTGAAACAAAGAAAACAGGT

ATTAAATATCGGTTAGCGCG

TTAAGGAAAATCTGAGGCAG

Pol II-chip

The data of Polymerase II-binding Chromatin Immunoprecipitation profile is from the study by Blythe et. al.

Zld-chip

The data of Zelda-binding Chromatin Immunoprecipitation profile is from Rushlow Lab. The two binding profiles are mapped to the dm6 gene model from Flybase, generated graph with Integrated Genome Browser.

3. Results

3.1. Expression of Slam in early embryogenesis

In situ hybridization (ISH) assay was performed to visualize the expression of *slam* in *Drosophila* embryo before the completion of nucleic cycle 14. Wild type and *zld*- *Drosophila* embryos were hybridized with RNA probe. [Fig.1] In the early wild type embryo, at the stage when cytokinesis hasn’t begun, *slam* is expressed ubiquitously at even levels. Then the expression shows segmental pattern. However, later during cellularization, the mRNA signal fades and gradually disappears with no apparent spatial order.

DOI: 10.54254/2753-8818/45/20240661

In zld- embryos, slam expression is concentrated at the anterior end, and some nuclei at the sites of furrow ingression display weak dispersive signals. The evident anomaly supports the crucial activation of slam by Zld, consistent with loss-of function results in paper by Liang[1], in which *zld- slam+* embryos didn't succeed in full ingression in furrow canals.

The segmental pattern is absent however the anterior end expression is still activated, this may suggest the activation role of concentration gradient is still functional if the possibly undegraded signal is large enough, and the nucleosome barrier is probably high in the chromosomes. Thus the maternal Zelda breaks the constraints of concentration gradient and act much stronger than or outpace the effect of gradient. The Bicoid target sequences TAATC (5 in the vicinity of *slam*),[11](Markstein et al. in 2002); Hunchback (Hb) target TAATCC(1 among the Zld binding peak region) is likely to function cooperatively on expression of *slam*. [Fig.2]

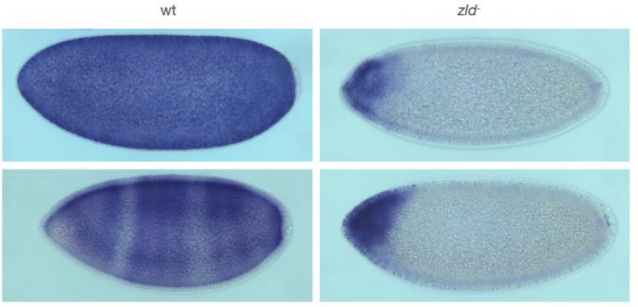


Figure 1. Difference in transcription between wild-type and *Zelda* knockout embryos before gastrulation.

In wild-type embryo (left column), slam is first expressed evenly in the whole embryo, whereas in *zld-* embryo(right column), expression is concentrated upon the anterior end.

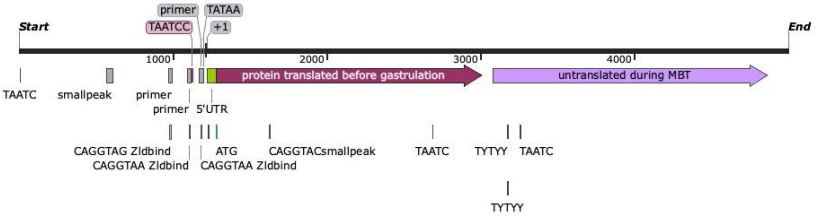


Figure 2. Annotations of *slam* region.

The view created by SnapGene starts from 6372295bp on the 2L chromosome.

3.2. Analysis of genome region

It can be noticed that cyc-8 Zld binding peaks on the profile lays on the upstream cyc-12 Pol II binding peaks, indicating that the transcription of slam occurs after the heptamer binding of Zld in a characterized manner: TAGteam that correlates to the Zld-ChIP peak is 93bp upstream of the TATA box TSS. The ChIP profiles are consistent with the in-situ observations by supporting the Zelda activation of *slam*. The three TAGteams facilitate Zld binding of each other and have led to a broader peak.

Another medium Zld binding peak is in the downstream of *slam*, located in the intron of lower strand gene DIP-epsilon, however, the peak only contains 1 CAGGTAA site. The site may or may not interact with Zld binding on *slam*, and the two binding regions may interact spatially.

The binding effect of the four identified CAGGTAR motif can be further investigated by deleting identifiability to Zld through KO mutation. In a suggested *in vitro* mutagenesis experiment, we use CRISPR-cas9 to facilitate homology directed repair. The deletion of each of the three TAGteam sites is predicted to cause weaker Zld binding, and the absence of the motif located in the remote peak may result in less obvious effects in the expression multiple genes.

DOI: 10.54254/2753-8818/45/20240661

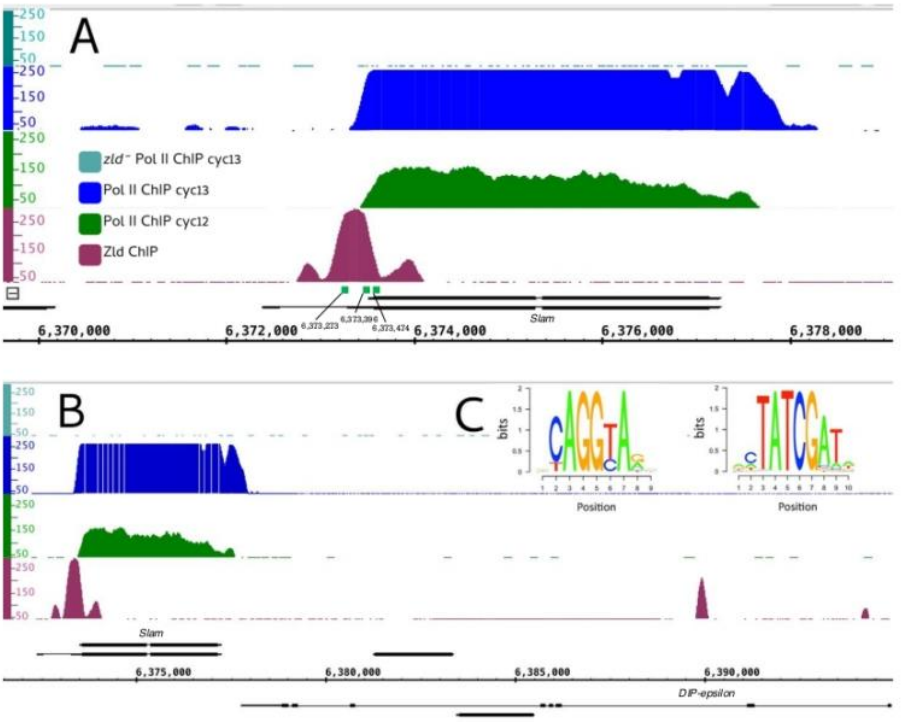


Figure 3. Integrated genome browser view of the slam gene.

(A)The top track represents PolII-ChIP pile up profile of zld⁻ expression [12], the second and third track represents PolII-ChIP profile of nucleic cycle 12 and 13 respectively, and the bottom track represents Zelda-binding ChIP-seq profile, with locations of TAGteams in the peak binding of nucleic cycle 8 marked below as green blocks. (B)Zoom out view of the ChIP-seq profiles, a smaller Zld binding peak is present in the untranslated region of a downstream gene *DIP-epsilon*(C)Position weight matrix of the enriched sites [4].

4. Discussion

In general, the expression of *slow as molasses* follow the characteristics summarized in previous studies, cluster of TAGteams upstream of cds are bound to Zld forming a major peak in the ChIP profile. There are also two smaller peaks coupling the major peak, one of which downstream extending into the coding region. This small peak contains two overlapping AGGTA motifs (AGGTACCT) that may display unknown function in weak interactions of biological molecules. Chromatin immunoprecipitation followed by sequencing (ChIP-seq) can be employed to identify specific proteins or factors that may interact with AGGTACCT or similar motifs, to investigate the potential biological heterogeneity, such as cell subpopulations or dynamic changes in chromatin structure, that may contribute to scattered peaks. In addition, Pol II binding peaks of the genes portraying scattered contiguous small peaks seem to have more complex extensions in the downstream flanking region. Promoter proximal pausing[13], co-transcriptional processing[14], and enhancer-promoter interactions[15] might explain the issue. Moreover, other nearby gene pairs such as *elba2* and *insv* shows intriguing peak shapes, yet the peaks of pairing *elba1* and *elba3*, as well as *CG14014* and *CG 18269* are relatively smooth, indicating possible transcriptional interference[16] in the former genes.

Considering the observation of only the first exon being translated before cellularization, another TSS *Inr*(Long Vo Ngoc et al. 2019)[17] (TYTYY)[Fig.2] has been spotted on the second exon, though it probably would not be functional.

DOI: 10.54254/2753-8818/45/20240661

In future studies, the coordination between transcription factors, the influential forces acted by the zinc fingers examined *in vivo*, the navigation of transcription factors in chromatin landscape and cross-species comparisons on zinc finger transcription factors as well as pioneering transcription factors can be further explored.

5. Conclusion

In conclusion, this study has shed light on the intricate regulatory network governing the transcription activation of the zygotic genome during the early stages of Drosophila embryonic development, with a particular focus on the gene slow as molasses (slam). The in situ hybridization results and ChIP-seq assays have reaffirmed the significant influence of Zelda on slam's expression. The observed differences in slam expression between wild-type and zld- embryos further validate the essential role of Zld in the regulatory cascade. The concentration gradient and the presence of other regulatory elements such as Bicoid and Hunchback may interact to fine-tune slam expression.

Analysis of the genome region has shown the specific binding patterns of Zld and Polymerase II (PolII) with respect to slam, supporting the idea that Zld facilitates slam transcription by binding to TAGteams, which are strategically positioned upstream of the transcription start site (TSS). Furthermore, the study has revealed potential additional binding sites and the possibility of spatial interactions between different genomic regions.

The study also raises questions about the scattered contiguous small peaks observed in the PolII binding profiles, hinting at the complexities of promoter proximal pausing, co-transcriptional processing, and enhancer-promoter interactions that warrant further investigation.

References

[1] Liang, H.-L., Nien, C.-Y., Liu, H.-Y., Metzstein, M. M., Kirov, N., & Rushlow, C. (2009). *The zinc-finger protein Zelda* is a key activator of the early zygotic genome in Drosophila.

[2] THE MODENCODE CONSORTIUM, Roy, S., Ernst, J., Kharchenko, P. V., Kheradpour, P., Negre, N., Eaton, M. L., Landolin, J. M., Bristow, C. A., Ma, L., Lin, M. F., Washietl, S., Arshinoff, B. I., Ay, F., Meyer, P. E., Robine, N., Washington, N. L., Di Stefano, L., Berezikov, E., … Zieba, J. (2010). Identification of Functional Elements and Regulatory Circuits by Drosophila modENCODE. *Science*, *330*(6012), 1787–1797. https://doi.org/10.1126/science.1198374

[3] Sun, Y., Nien, C.-Y., Chen, K., Liu, H.-Y., Johnston, J., Zeitlinger, J., & Rushlow, C. (n.d.). *Zelda overcomes the high intrinsic nucleosome barrier at enhancers during Drosophila zygotic genome activation*.

[4] Nien, C.-Y., Liang, H.-L., Butcher, S., Sun, Y., Fu, S., Gocha, T., Manak, J. R., & Rushlow, C. (2011). Temporal Coordination of Gene Networks by Zelda in the Early Drosophila Embryo. *PLoS Genetics*, *7*(10).

[5] Espinola, S. M. (2021). Cis-regulatory chromatin loops arise before TADs and gene activation, and are independent of cell fate during early Drosophila development. *Nature Genetics*, *53*.

[6] Müller, H.-A. J. (2002). Germ Cell Migration: As Slow as Molasses. *Current Biology*, *12*(18), R612–R614. https://doi.org/10.1016/S0960-9822(02)01131-4

[7] Yan, S. (n.d.). *Localization and translation control of slam in Drosophila cellularization*.

[8] Lecuit, T., Samanta, R., & Wieschaus, E. (2002). Slam Encodes a Developmental Regulator of Polarized Membrane Growth during Cleavage of the Drosophila Embryo. *Developmental Cell*, *2*(4), 425–436. https://doi.org/10.1016/S1534-5807(02)00141-7

[9] Acharya, S., Laupsien, P., Wenzl, C., Yan, S., & Großhans, J. (2014). Function and dynamics of slam in furrow formation in early Drosophila embryo. *Developmental Biology*, *386*(2), 371–384. https://doi.org/10.1016/j.ydbio.2013.12.022

[10] *Slam expression in fruit fly*. (n.d.). Retrieved August 6, 2023, from https://www.bgee.org/gene/FBgn0043854

DOI: 10.54254/2753-8818/45/20240661

[11] Markstein M, Markstein P, Markstein V, Levine MS. Genome-wide analysis of clustered Dorsal binding sites identifies putative target genes in the Drosophila embryo. Proc Natl Acad Sci U S A. 2002 Jan 22;99(2):763-8. doi: 10.1073/pnas.012591199. Epub 2001 Dec 18. PMID: 11752406; PMCID: PMC117379.

[12] Blythe, S. A., & Wieschaus, E. F. (2015). Zygotic Genome Activation Triggers the DNA Replication Checkpoint at the Midblastula Transition. *Cell*, *160*(6), 1169–1181. https://doi.org/10.1016/j.cell.2015.01.050

[13] Dollinger, R., & Gilmour, D. S. (2021). Regulation of Promoter Proximal Pausing of RNA Polymerase II in Metazoans. *Journal of Molecular Biology*, *433*(14), 166897. https://doi.org/10.1016/j.jmb.2021.166897

[14] Neugebauer, K. M. (2002). On the importance of being co-transcriptional. *Journal of Cell Science*, *115*(Pt 20), 3865–3871. https://doi.org/10.1242/jcs.00073

[15] Richter, W. F., Nayak, S., Iwasa, J., & Taatjes, D. J. (2022). The Mediator complex as a master regulator of transcription by RNA polymerase II. *Nature Reviews. Molecular Cell Biology*, *23*(11), 732–749. https://doi.org/10.1038/s41580-022-00498-3

[16] A, M., Jw, H., S, P., Y, S., & Hw, B. (2007). Transcriptional interference: An unexpected layer of complexity in gene regulation. *Journal of Cell Science*, *120*(Pt 16). https://doi.org/10.1242/jcs.007633

[17] Vo Ngoc L, Kassavetis GA, Kadonaga JT. The RNA Polymerase II Core Promoter in Drosophila. Genetics. 2019 May;212(1):13-24. doi: 10.1534/genetics.119.302021. PMID: 31053615; PMCID: PMC6499525. [6]Müller, H.-A. J. (2002). Germ Cell Migration: As Slow as Molasses. Current Biology, 12(18), R612–R614. https://doi.org/10.1016/S0960-9822(02)01131-4