

UCSF

UC San Francisco Electronic Theses and Dissertations

Title

Activation of Histone 3 Lysine 9 methyl writing and reading capabilities within the G9a-GLP heterodimer

Permalink

<https://escholarship.org/uc/item/0532z155>

Author

Sanchez, Nicholas

Publication Date

2021

Peer reviewed|Thesis/dissertation

Activation of Histone 3 Lysine 9 Methyl Writing and Reading Capabilities Within the G9a-GLP Heterodimer

by
Nicholas Sanchez

THESIS

Submitted in partial satisfaction of the requirements for degree of
DOCTOR OF PHILOSOPHY

in

Biochemistry and Molecular Biology

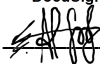
in the

GRADUATE DIVISION

of the

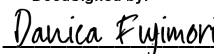
UNIVERSITY OF CALIFORNIA, SAN FRANCISCO

Approved:

DocuSigned by:

251AC7A581B94CA...

Bassem Al-Sady

Chair

DocuSigned by:

328308DD7E5546A...

Danica Fujimori

DocuSigned by:

328308DD7E5546A...

David Agard

Committee Members

Acknowledgements

My first and sincerest acknowledgement is for my parents, Diane and Javier Sanchez. To my dad who gave me my first job at age 7 picking up cigarette butts around our car dealership. Thank you for teaching me the value of hard work as well as the beauty of critical thinking. To my mom who taught me my first lessons in advocating for myself. Through you I've found the strength to live true to who I am. Thank you both for setting my foundation to grow into the person I am today.

To my family, both Sanchez and Davis who continue to inspire me. To my brother Shua. My big brother, my college roommate, my physics tutor. Without you I would never have the guts to do anything of this caliber. Tia Angelica who always reminds me to smile even when life is hard. Tia Rosario for being a constant role model in my life. Aunt Dawn for being a constant nerdy role model in my life.

To Eru who makes my life happier today than it has ever been.

To those from the beautiful state of Wisconsin. Angela Drake from whom I'm always challenged and learning new perspectives on life. Jordan Ho for still being that kid who kicked me out of his seat in art class. Tyler Lieberthal for upping my cooking game. Mina Shahlapour for being such a light in my life. Evan Flietner for all the family fun new years. PK for her spontaneity. Carmel Assa and Ethan Rosen for keeping the "doopid" alive in me. To advisors Gloria Hawkins and Yolanda Garza whose direction and guidance showed me how to pave my road to success. To Eric Strieter and lab for making research cool. Thank you Ellen Valkevich for taking me in with zero lab skills and continuing to be a lifelong friend and mentor.

To members of the Al-Sady and Redding labs old and new. Bassem for always pushing me to be better. RA Greenstein, I could not have done it without you. Dana Kennedy for the smiles and laughs. Luke Strauskulage, Madeline Keenen, Lena Kallweit, Cal Clemmer, Lucy Brennan, Eric Siemental, Henry Ng, Khanh Ngo, Liv Jensen, Christine Kugel, Sy Redding. I have grown so much having you in my life. Thank you for all the support both intellectually and emotionally.

To the UCSF tetrad entering class of 2014. Getting to know you all was the most worthwhile part of this experience. Thank you Justin Salat who was the best roommate anyone could ever have. Thank you for all the laughs, all the adventures and all the times you'd help me eat whatever monstrosity I cooked. Thank you Allison Cohen for being a damn good scientist and being down for anything. Thank you Efren Reyes for all the early morning workouts and afternoon happy hours. Over the years, my classmates have pushed me to run a marathon, climb mountains and make hit music videos. I've loved every second with you all.

To the broader UCSF community. Thank you for fostering such a caring, collaborative, stimulating environment. To those from the Hooper Foundation who made every day such a joy. Paola Soto and Bernarda Lopez for keeping my Spanish skills sharp. Yi Zhou for keeping my Mandarin skills sharp. Thank you David Agard and Danica Fujimori for advising this work and pushing me in all the right ways.

**Activation of Histone 3 Lysine 9 methyl writing and reading capabilities within the
G9a-GLP heterodimer**

Nicholas Sanchez

Abstract

Unique among metazoan repressive histone methyltransferases, G9a and GLP, which target histone 3 lysine 9 (H3K9), require dimerization for productive H3K9 mono (me1)- and dimethylation (me2) *in vivo*. Intriguingly, even though each enzyme can independently methylate H3K9, the predominant active form *in vivo* is a heterodimer of G9a and GLP. How dimerization influences the central H3K9 methyl binding (“reading”) and deposition (“writing”) activity of G9a and GLP, and why heterodimerization is essential *in vivo* remains opaque. Here, we examine the H3K9me “reading” and “writing” activities of defined, recombinantly produced homo- and heterodimers of G9a and GLP. We find that both reading and writing are significantly enhanced in the heterodimer. Compared to the homodimers, the heterodimer has higher recognition of H3K9me₂, and a striking ~ 10-fold increased k_{cat} for nucleosomal substrates under multiple turnover conditions, which is not evident on histone tail peptide substrates. This however is not encoded by altered nucleosome affinity, which is dominated by the G9a protomer and comparable across the homo- and heterodimer. Our results indicate that heterodimerization may be required to relieve autoinhibition of H3K9me reading and chromatin methylation evident in G9a and GLP homodimers. Relieving this inhibition may be particularly important in early differentiation when large tracts of H3K9me₂ are deposited by G9a-GLP, which may require a more active form of the enzyme.

Table of Contents

1: Introduction.....	1
1.1 Introduction.....	1
1.2 References.....	4
2: G9a and GLP form stable dimers at 1:1 stoichiometry.....	10
2.1 Introduction.....	10
2.2 Results.....	12
2.3 Discussion.....	21
2.4 Methods.....	22
2.5 References.....	25
3: Heterodimerization Stimulates G9a and GLP Catalytic Activity.....	28
3.1 Introduction.....	28
3.2 Results.....	30
3.3 Discussion.....	40
3.4 Methods.....	42
3.5 References.....	44
4: Probing G9a-GLP Interaction with Chromatin.....	49
4.1 Introduction.....	49
4.2 Results.....	52
4.3 Discussion.....	66
4.4 Methods.....	68
4.5 References.....	70

List of Figures

Figure 2.1 G9a and GLP form stable 1:1 homo- and heterodimers.....	15
Figure 2.2 Full-length G9a and GLP purify as a heterodimer from insect cells.....	16
Figure 2.3 S-200 10/300 SEC Profiles of Heteromeric and Homomeric G9a and GLP....	17
Figure 2.4 S200 16/600 SEC Profiles.....	18
Figure 2.5 Stability of G9a-G9a and G9a-GLP complexes through a dilution assay.....	19
Figure 2.6 Mass Photometry identification of G9a and GLP homo and heterodimers.....	20
Figure 3.1 Effect of pH on G9a-GLP methylation reaction.....	33
Figure 3.2 G9a-GLP Pre-incubation Test.....	34
Figure 3.3 G9a-G9a Full Length vs ANKSET truncation initial rate comparison.....	35
Figure 3.4 Initial Rate Comparisons of G9a in homodimeric and heterodimeric forms....	36
Figure 3.5 Heterodimerization stimulates G9a and GLP catalytic activity.....	37-38
Figure 3.6 S-adenosyl methionine K_M and G9a-G9a mononucleosome Michaelis Menten kinetics.....	39
Figure 4.1 Heterodimerization facilitates G9a and GLP binding to H3K9 methyl peptides.....	56
Figure 4.2 Binding to ANK repeat is measured with fluorescence polarization.....	57
Figure 4.3 How to measure product recognition stimulation.....	58
Figure 4.4 Product Recognition Stimulation Assay Setup.....	59
Figure 4.5 G9a-GLP does not bind DNA.....	60
Figure 4.6 Determining enzyme conditions for product recognition stimulation.....	61
Figure 4.7 Reaffirming G9a-G9a product recognition result.....	62

List of Figures contd.

Figure 4.8 Comparison of G9a product recognition stimulation to that of G9a-GLP.....63

Figure 4.9 Troubleshooting G9a product recognition stimulation.....64

Figure 4.10 Alternative way to measure Product recognition stimulation.....65

1: Introduction

The animal genome is partitioned into active and inactive regions by gene-repressive structures that restrict access to gene-activating factors (Allshire and Madhani, 2018). Differential patterning of heterochromatin throughout the genome is one mechanism cells adopt that allows genetically identical cells to express vastly different protein networks while still reading the same DNA template (Wen *et al.*, 2009). Heterochromatin is denoted by regions of the genome marked via di or tri methylation of histone 3 lysine 9 (H3K9me) or H3K27 tri methylation (Simon and Kingston, 2013; Holoch, Moazed and Avenue, 2015). These two histone modifications are considered true epigenetic marks in that once installed at a gene locus, they can be inherited through cell division and still retain their ability to encode gene silencing (Hansen *et al.*, 2008; Audergon *et al.*, 2015; Deans and Maggert, 2015; Rangunathan, Jih and Moazed, 2015). These heterochromatic histone methylation marks are recognized by reader proteins with gene silencing functions such as HP1 α , which binds H3K9me and can physically compact DNA, presumably to restrict access by transcriptional machinery (Torres and Fujimori, 2015; Larson *et al.*, 2017; Keenen *et al.*, 2021). Other reader domain proteins associate with K9 or 27me sites, initiating cascades that translocate gene loci to different regions of the nucleus, reposition nucleosomes, and associate tightly with DNA all in an effort to create the barrier to transcription that is heterochromatin.

Expansions of H3K9me that invade genic regions are carried out by the methyltransferases SETDB1 and SETDB2 along with G9a and GLP (Wen *et al.*, 2009; Nicetto and Zaret, 2019; Nicetto *et al.*, 2019). The PRDM family H3K9 methyltransferases functions distinctly at oncogenes (Mzoughi *et al.*, 2016) while SUV39H1 and SUV39H2, which catalyze H3K9me₃, direct their activities to form heterochromatin at the gene-poor centromere (Peters *et*

al., 2001; Maison *et al.*, 2011; Jehanno *et al.*, 2017). Despite catalyzing the same post-translational modification, each H3K9 methyltransferase directs their activity to distinct regions of the genome (Rea *et al.*, 2000).

In early mammalian development, the genome of embryonic stem cells is highly transcriptionally active and features little heterochromatin (Percharde, Bulut-Karslioglu and Ramalho-Santos, 2016). The overall lack of heterochromatin is a mechanism by which cells of the embryo achieve pluripotency (Wen *et al.*, 2009; Percharde, Bulut-Karslioglu and Ramalho-Santos, 2016). As differentiation begins, H3K9me heterochromatin expands, re-sculpting the genome architecture and restricting fate (Nicetto and Zaret, 2019). Once ascribed to a certain fate, a cell maintains H3K9me domains through division in order to maintain its transcriptional profile. During embryogenesis, G9a, GLP and the SETDB1 methyltransferases direct the expansion of large tracks of H3K9me₂, which can adopt a lineage-specific pattern (Wen *et al.*, 2009; Zyllicz *et al.*, 2015). These methyltransferases have the capacity to methylate mono-, di-, and tri- H3K9 as well as mono- and di- K27. However, their activity represents the bulk of K9 mono and di methylation of the cell (Rea *et al.*, 2000; Tachibana *et al.*, 2005). Their heterochromatin expansion directly represses lineage inappropriate genes or silences enhancers that in turn direct the transcription of several genes.

Like several heterochromatic methyltransferases, G9a and GLP have the capacity to both read and write H3K9me (Tachibana *et al.*, 2002, 2005; Collins *et al.*, 2008; Liu *et al.*, 2015). G9a and GLP both contain an Ankyrin (ANK) repeat domain, which confers methyl-histone binding (reading) activity, and a SET domain, which confers methyltransferase (writer) activity. The capacity to both read and write their H3K9me product allows a methyltransferase to stay engaged at a specific locus and promotes the expansion of H3K9me at that site (Erdel and

Greene, 2016). Additionally, the presence of methylated substrates appears to stimulate G9a or GLP catalytic activity in an ANK domain-dependent manner (Liu *et al.*, 2015). Other H3K9 methylases require such positive feedback for either lateral spreading (Al-Sady, Madhani and Narlikar, 2013; Müller *et al.*, 2016) or epigenetic maintenance (Audergon *et al.*, 2015; Ragunathan, Jih and Moazed, 2015). How the ANK and SET domains of G9a and GLP regulate one another is opaque but may be central to understanding their function in cell fate control.

Unlike other metazoan H3K9 methyltransferases, G9a and GLP must associate with each other to carry out H3K9 methylation – when the interaction between the two enzymes is broken *in vivo*, the bulk of H3K9me1 and me2 is lost (Tachibana *et al.*, 2005, 2008). This is despite the observation that each enzyme is capable of methylating histones *in vitro* in the absence of its binding partner (Tachibana *et al.*, 2002, 2005; Chin *et al.*, 2005, 2006; Liu *et al.*, 2015). The interaction interface of G9a and GLP occurs between each enzyme's SET domains. We hypothesized that forming the G9a-GLP complex (G9a-GLP) has a direct regulatory effect on the H3K9 methylation reaction. However, the biochemical and biophysical nature of homo- and heterotypic associations between G9a and GLP are poorly understood, as are any effects on H3K9me writing and reading that these associations may have. In this study, we investigated the regulation heterodimerization imposes on G9a and GLP's ability to read and write H3K9me.

1.2 References

Al-Sady, B., Madhani, H. D. and Narlikar, G. J. (2013) ‘Division of labor between the chromodomains of HP1 and Suv39 methylase enables coordination of heterochromatin spread’, *Molecular Cell*. Elsevier Inc., 51(1), pp. 80–91. doi: 10.1016/j.molcel.2013.06.013.

Allshire, R. C. and Madhani, H. D. (2018) ‘Ten principles of heterochromatin formation and function’, *Nature Reviews Molecular Cell Biology*. Nature Publishing Group, 19(4), pp. 229–244. doi: 10.1038/nrm.2017.119.

Audergon, P. N. C. B. *et al.* (2015) ‘Restricted epigenetic inheritance of H3K9 methylation.’, *Science (New York, NY)*, 348(6230), pp. 132–135. doi: 10.1126/science.1260638.

Chin, H. G. *et al.* (2005) ‘Sequence specificity and role of proximal amino acids of the histone H3 tail on catalysis of murine G9a lysine 9 histone H3 methyltransferase’, *Biochemistry*, 44(39), pp. 12998–13006. doi: 10.1021/bi0509907.

Chin, H. G. *et al.* (2006) ‘Catalytic properties and kinetic mechanism of human recombinant Lys-9 histone H3 methyltransferase SUV39H1: Participation of the chromodomain in enzymatic catalysis’, *Biochemistry*, 45(10), pp. 3272–3284. doi: 10.1021/bi051997r.

Collins, R. E. *et al.* (2008) ‘The ankyrin repeats of G9a and GLP histone methyltransferases are mono- and dimethyllysine binding modules.’, *Nature structural & molecular biology*, 15(3), pp. 245–50. doi: 10.1038/nsmb.1384.

Deans, C. and Maggert, K. A. (2015) ‘What do you mean, “Epigenetic”?’ , *Genetics*, 199(4), pp. 887–896. doi: 10.1534/genetics.114.173492.

Erdel, F. and Greene, E. C. (2016) ‘Generalized nucleation and looping model for epigenetic memory of histone modifications’. doi: 10.1073/pnas.1605862113.

Hansen, K. H. *et al.* (2008) ‘A model for transmission of the H3K27me3 epigenetic mark’, *Nature Cell Biology*, 10(11), pp. 1291–1300. doi: 10.1038/ncb1787.

Holoch, D., Moazed, D. and Avenue, L. (2015) ‘RNA-mediated epigenetic regulation of gene expression’, *Nature genetics*, 16(2), pp. 71–84. doi: 10.1038/nrg3863.RNA-mediated.

Jehanno, C. *et al.* (2017) ‘Biochimica et Biophysica Acta A model of dynamic stability of H3K9me3 heterochromatin to explain the resistance to reprogramming of differentiated cells’, *BBA - Gene Regulatory Mechanisms*. Elsevier B.V., 1860(2), pp. 184–195. doi: 10.1016/j.bbagr.2016.11.006.

Keenen, M. M. *et al.* (2021) ‘HP1 proteins compact dna into mechanically and positionally stable phase separated domains’, *eLife*, 10, pp. 1–38. doi: 10.7554/eLife.64563.

Larson, A. G. *et al.* (2017) ‘Liquid droplet formation by HP1 α suggests a role for phase separation in heterochromatin’, *Nature*. Nature Publishing Group, 547(7662), pp. 236–240. doi: 10.1038/nature22822.

Liu, N. *et al.* (2015) ‘Recognition of H3K9 methylation by GLP is required for efficient establishment of H3K9 methylation, rapid target gene repression, and mouse viability’, *Genes and Development*, 29(4), pp. 379–393. doi: 10.1101/gad.254425.114.

Maison, C. *et al.* (2011) ‘SUMOylation promotes de novo targeting of HP1 α to pericentric heterochromatin.’, *Nature genetics*. Nature Publishing Group, 43(3), pp. 220–227. doi: 10.1038/ng.765.

Müller, M. M. *et al.* (2016) ‘A two-state activation mechanism controls the histone methyltransferase Suv39h1’, *Nature Chemical Biology*, 12(3), pp. 188–193. doi: 10.1038/nchembio.2008.

Mzoughi, S. *et al.* (2016) ‘The role of PRDMs in cancer: One family, two sides’, *Current Opinion in Genetics and Development*. Elsevier Ltd, 36, pp. 83–91. doi: 10.1016/j.gde.2016.03.009.

Nicetto, D. *et al.* (2019) ‘H3K9me3-heterochromatin loss at protein-coding genes enables developmental lineage specification’, *Science*, 363(6424), pp. 294–297. doi: 10.1126/science.aau0583.

Nicetto, D. and Zaret, K. S. (2019) 'Role of H3K9me3 heterochromatin in cell identity establishment and maintenance', *Current Opinion in Genetics and Development*, 55, pp. 1–10. doi: 10.1016/j.gde.2019.04.013.

Percharde, M., Bulut-Karslioglu, A. and Ramalho-Santos, M. (2016) 'Hypertranscription in Development, Stem Cells, and Regeneration', *Developmental Cell*. Elsevier Inc., 40(1), pp. 9–21. doi: 10.1016/j.devcel.2016.11.010.

Peters, A. H. F. M. *et al.* (2001) 'Loss of the Suv39h histone methyltransferases impairs mammalian heterochromatin and genome stability', *Cell*, 107(3), pp. 323–337. doi: 10.1016/S0092-8674(01)00542-6.

Ragunathan, K., Jih, G. and Moazed, D. (2015) 'Epigenetic inheritance uncoupled from sequence-specific recruitment', *Science*, 348(6230), p. science.1258699-. doi: 10.1126/science.1258699.

Rea, S. *et al.* (2000) 'Regulation of chromatin structure by site-specific histone H3 methyltransferases', *Nature*, 406(6796), pp. 593–599. doi: 10.1038/35020506.

Simon, J. A. and Kingston, R. E. (2013) ‘Occupying Chromatin: Polycomb Mechanisms for Getting to Genomic Targets, Stopping Transcriptional Traffic, and Staying Put’, *Molecular Cell*. Elsevier Inc., 49(5), pp. 808–824. doi: 10.1016/j.molcel.2013.02.013.

Tachibana, M. *et al.* (2002) ‘G9a histone methyltransferase plays a dominant role in euchromatic histone H3 lysine 9 methylation and is essential for early embryogenesis’, *Genes and Development*, 16(14), pp. 1779–1791. doi: 10.1101/gad.989402.

Tachibana, M. *et al.* (2005) ‘Histone methyltransferases G9a and GLP form heteromeric complexes and are both crucial for methylation of euchromatin at H3-K9’, *Genes and Development*, 19(7), pp. 815–826. doi: 10.1101/gad.1284005.

Tachibana, M. *et al.* (2008) ‘G9a/GLP complexes independently mediate H3K9 and DNA methylation to silence transcription.’, *The EMBO journal*, 27(20), pp. 2681–90. doi: 10.1038/emboj.2008.192.

Torres, I. O. and Fujimori, D. G. (2015) ‘Functional coupling between writers, erasers and readers of histone and DNA methylation’, *Current Opinion in Structural Biology*, 35, pp. 68–75. doi: 10.1016/j.sbi.2015.09.007.

Wen, B. *et al.* (2009) ‘Large histone H3 lysine 9 dimethylated chromatin blocks distinguish differentiated from embryonic stem cells.’, *Nature genetics*, 41(2), pp. 246–250. doi: 10.1038/ng.297.

Zylicz, J. J. *et al.* (2015) ‘Chromatin dynamics and the role of G9a in gene regulation and enhancer silencing during early mouse development’, *eLife*, 4(NOVEMBER2015), pp. 1–25.
doi: 10.7554/eLife.09571.

2: G9a and GLP form stable dimers at 1:1 stoichiometry

2.1 Introduction

Having multiple modes of interaction and valency is a property by which heterochromatin associated proteins are able to establish and maintain a heterochromatic domain (Dodd and Sneppen, 2011; Zhang *et al.*, 2014; Sneppen and Dodd, 2015). As such, heterochromatic methyltransferases typically associate into higher order, multivalent, complexes in order to carry out their function *in vivo*. Many of these complexes contain additional readers of histone methylation, as in the case of H3K9 methyltransferase Suv39H1 which directly interacts with H3K9me reader HP1 (Maison *et al.*, 2016). These complexes may also contain additional writers of histone methylation as in the case of H3K27 methyltransferase EZH2, which along with binding readers of histone methylation and other chromatin interacting proteins, is able to dimerize with another copy of EZH2, doubling the multivalency already present in a single EZH2 complex (Margueron *et al.*, 2009; Yuan *et al.*, 2012; Davidovich *et al.*, 2014).

G9a and GLP are a unique example of two methyltransferases that must associate in order to carry out their function *in vivo* (Tachibana *et al.*, 2008). Though 1:1 stoichiometric heteromeric association appears to be the preferred mode of interaction of G9a and GLP, G9a homomeric complexes are observed in the absence of GLP and *vice versa* (Tachibana *et al.*, 2005). While heteromeric G9a-GLP complexes are necessary for H3K9me to occur *in vivo*, the extent and role to which G9a and GLP homomeric complexes function *in vivo* remains elusive. A 1:1 stoichiometric complex may represent a dimeric form of G9a and GLP, however, like HP1,

G9a and GLP may have multiple modes of interaction, allowing them to form higher order stoichiometric complexes (Canzio *et al.*, 2011). In this section, I sought to characterize the biophysical nature of G9a and GLP homo and heteromeric complexes. I investigated the number of molecules represented by a homo and heteromeric G9a-GLP complex as well as characterized the stability of each complex type.

2.2 Results

To directly assay the interplay between reading, writing, and dimerization, we expressed and purified G9a and GLP truncated to the C-terminal ANK and SET domains (ANK-SET), consistent with prior studies (Tachibana *et al.*, 2005) (**Figure 2.1 A**). We separately expressed G9a and GLP in *E. coli* each fused to an N-terminal MBP tag to isolate each methyltransferase individually. To isolate the G9a-GLP heterodimer, we adopted an *E. coli* co-expression strategy whereby G9a was N-terminally tagged with a 6xHIS extension and GLP tagged with an N-terminal Maltose Binding Protein (MBP). Upon expression of each methyltransferase, we performed a sequential affinity purification of cobalt affinity resin to isolate 6xHIS:G9a and associated MBP:GLP followed by amylose resin to isolate MBP-GLP and associated 6xHIS:G9a (**Figure 2.1 B**). We assessed our HIS:G9a-MBP:GLP heterodimers to have 1:1 stoichiometry via Quantitative SYPRO Red Gel Staining. Our *E. coli* expression system limited us to using G9a and GLP truncated to the c-terminal ANK and SET domains, as we could not express or purify full length GLP. We were, however, able to express and purify full length G9a and GLP using an Sf9 expression system. Here we adopted a similar sequential affinity purification strategy using 6xHIS:G9a (full length) and N-terminally tagged strep:GLP (full length) (**Figure 2.2**). The Sf9 expression system, though able to produce full-length versions of each methyltransferase, did not produce protein yields sufficient for our assay needs. As such, we continued with our *E. coli* expression system using truncated G9a and GLP.

We applied Size Exclusion Chromatography (SEC) to our purified MBP:G9a, MBP:GLP, and HIS:G9a-MBP:GLP purified constructs and were surprised to find the retention profiles of each construct to be relatively similar (**Figure 2.3**). Given that G9a and GLP homodimerization had only been characterized in mammalian cells (Tachibana *et al.*, 2005), our initial expectation

of expressing and purifying G9a and GLP individually was that they would purify as monomers. The similar SEC retention profiles of G9a, GLP homo and heterotypic expressions led us to hypothesize our G9a and GLP constructs when expressed individually were forming homodimers in *E. coli*. We further explored this a higher resolution SEC column (s-200 16/600) that would give better separation between potential monomer and dimer species of G9a and GLP. Again, we observed similar retention profiles of each homotypic and heterotypic construct, suggesting dimerization or multimerization was occurring in each expression context (**Figure 2.4**).

To explore homo and heterodimerization of G9a and GLP in more depth, we re-adopted our co-expression strategy to express two alleles of G9a, 6xHIS:G9a and MBP:G9a both tagged at the N-terminus. Similar to above, we assessed dimerization by a sequential cobalt then amylose resin purification (**Figure 2.1 B**). Quantification of SyPRO Red stained bands indicated dimerization at 1:1 stoichiometry for both 6xHIS:G9a-MBP:G9a and 6xHIS:G9a-MBP:GLP complexes. We examined the stability of G9a homo- and G9a-GLP heterodimers using a dilution-based assay, which assesses the off-rate of the complex. Each complex was diluted to 40nM and allowed to dissociate for 1-2 hours at room temperature. We assessed dimer association by precipitating the His-tagged protein with cobalt resin and determining the fraction of MBP protein that remained bound. We observed little dissociation in either G9a homo or heterodimers (**Figure 2.5**).

For further characterization of the G9a-GLP heterodimer, we removed both the His and MBP tags by TEV-mediated proteolysis (**Figure 2.1 C**). Using this “tagless” G9a-GLP molecule, we next ascertained the number of enzymes per complex. To do so, we determined its molecular weight using size exclusion chromatography (SEC) followed by multi-angle light

scattering (MALS). We observed one major peak in our SEC-MALS measurement (**Figure 2.1 D**) and determined it to have a molecular weight of ~135 kDa, roughly the theoretical molecular weight of our truncated G9a/GLP heterodimer (142 kDa). These data suggest that G9a-GLP is limited to a heterodimeric complex with one G9a and one GLP molecule under the concentration regimes used in the assay (0.4 – 4 μ M). We further confirmed G9a and GLP form a heterodimer via mass photometry measurement where we were also able to identify G9a and GLP homodimers from our homotypic expression strains (**Figure 2.6**). Interestingly, with the mass photometry single molecule approach, we observed the presence of G9a and GLP monomers from homotypic purifications. Perhaps at the concentrations of this assay (~10nM) G9a and GLP homodimers are unstable, unlike the G9a-GLP heterodimer, which did not appreciably dissociate in our mass photometry measurements. Taken together these data indicate we were able to express and isolate stable homo and heterodimers of G9a and GLP.

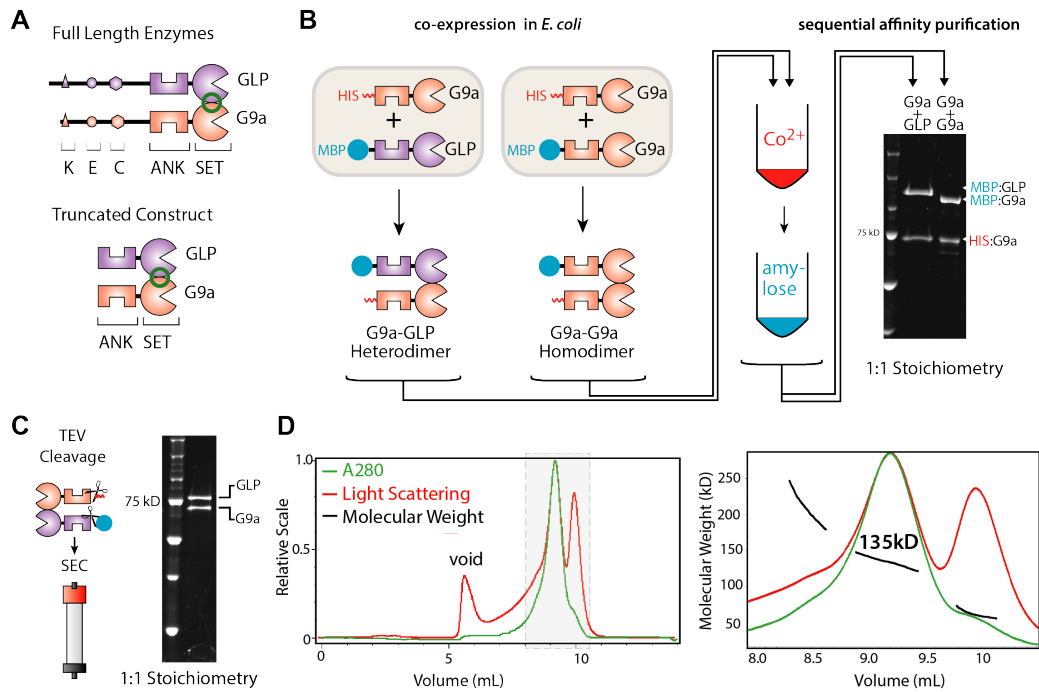


Figure 2.1: G9a and GLP form stable 1:1 homo- and heterodimers. **A.** Domain architecture of G9a-GLP. TOP: full-length enzymes. The C-terminus of G9a and GLP feature an automethylation residue (K), an acidic patch (E), and a cysteine-rich region (C). The N-termini contain ankyrin repeats (ANK) that bind H3K9me and a SET domain (SET) which is both the methylation catalytic domain and the dimerization interface (green circle). BOTTOM: Truncation ANK-SET construct used in this study. **B.** *E. coli* coexpression and purification strategy for identification of G9a homo and heterodimers. **C.** TEV cleavage of G9a-GLP heterodimers and purification via size exclusion chromatography **D.** SEC-MALS trace of G9a-GLP. LEFT: Full A280 (green) and Light Scattering (Red) traces. RIGHT: magnification of the main peaks with molecular weight determination (black). The measured molecular weight of the complex is 135kD (theoretical MW 142kD).

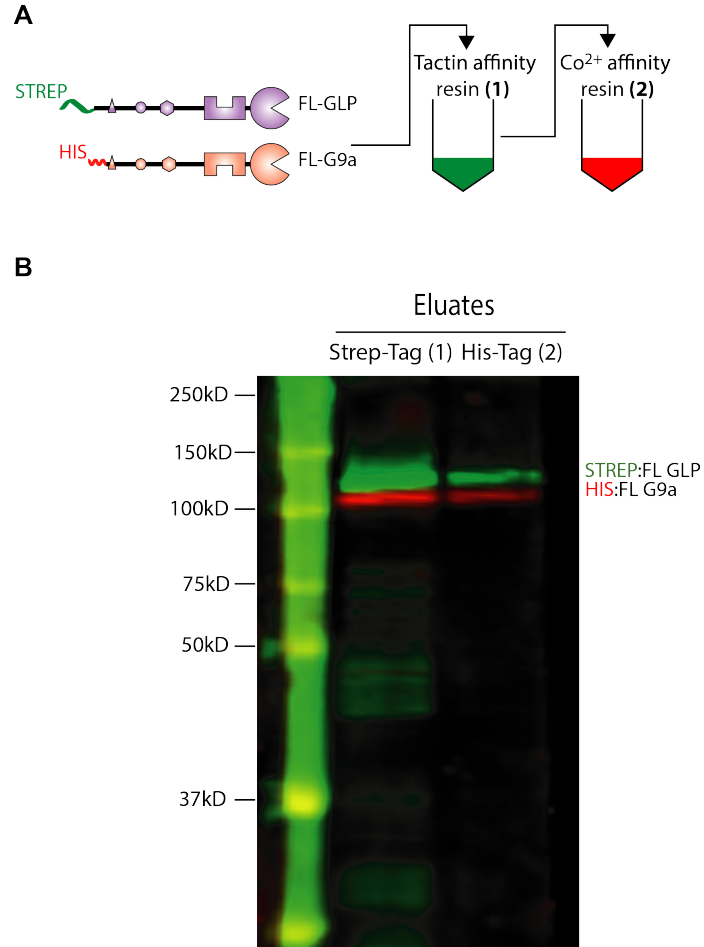


Figure 2.2: Full-length G9a and GLP purify as a heterodimer from insect cells. **A.** Full-length GLP and G9a were N-terminally tagged with a STREP or 6XHis tag, respectively, co-expressed from a baculoviral construct in Sf9 insect cells, and isolated by sequential STREPTACTIN (Tactin) and Cobalt (Co²⁺) affinity resins. **B.** Western blots of baculovirus-infected or uninfected Sf9 lysates (LEFT) or affinity resin eluates (RIGHT) with anti-His tag or anti-STREP tag antisera. The single-channel and merged images are shown.

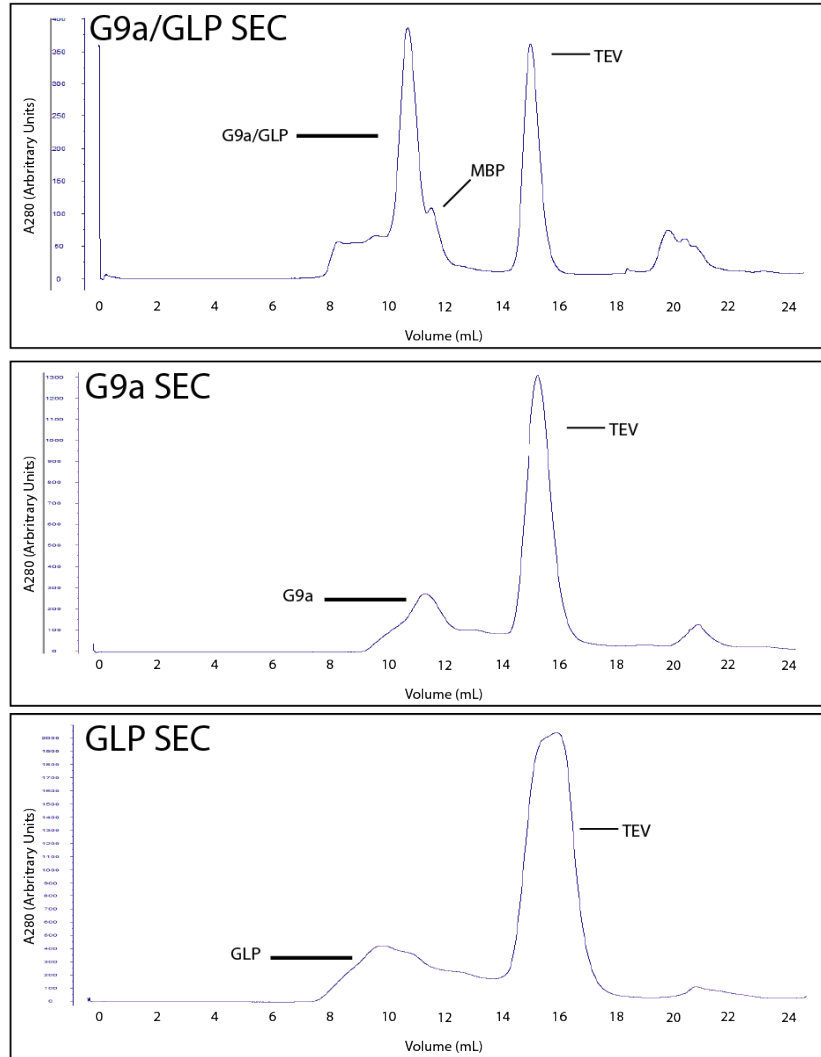


Figure 2.3 S-200 10/300 SEC Profiles of Heteromeric and Homomeric G9a and GLP
 Size Exclusion A280 profiles (blue) of G9a-GLP heterodimer (top), G9a-G9a homodimer (middle) and GLP-GLP homodimer (bottom). TEV protease is present in each of these runs as this is the routine run I would use to prep the enzymes. Retention profiles of all methyltransferase constructs on this column are all similar, suggesting species of similar molecular weight.

G9a-GLPΔN

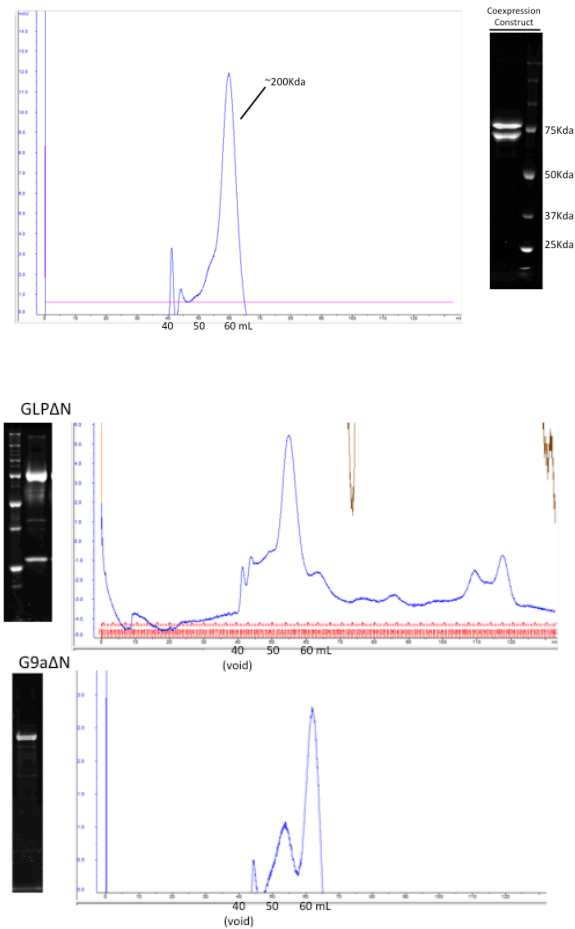


Figure 2.4 S200 16/600 SEC Profiles Size Exclusion A280 profiles (blue) of G9a-GLP heterodimer (top), GLP-GLP homodimer (middle) and G9a-G9a homodimer (bottom). SDS-PAGE gels of each construct is run alongside each trace. Retention profiles of all methyltransferase constructs on this column are all similar suggesting species of similar molecular weight.

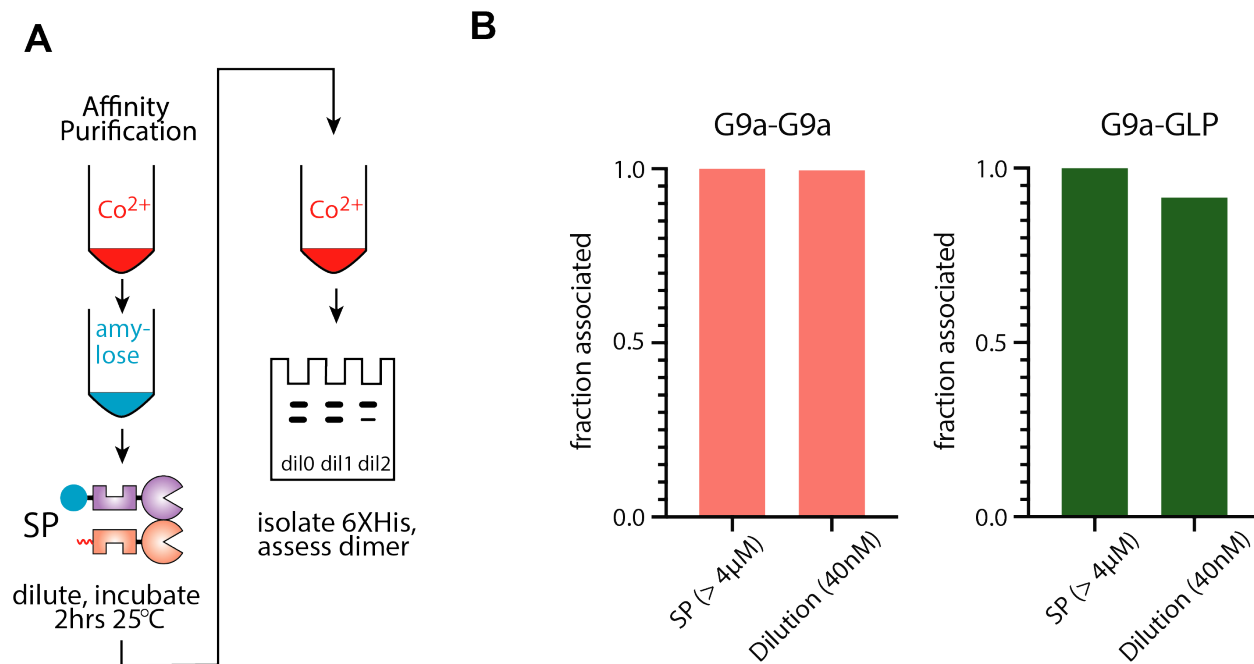


Figure 2.5: Stability of G9a-G9a and G9a-GLP complexes through a dilution assay. A. Experiment scheme. His:G9a coexpressed with either MBP:G9a or MBP:GLP was purified by sequential affinity purification as in Figure 1. The complexes were diluted and kept at 25°C for 1-2hrs. After this incubation, His:G9a was isolated via cobalt resin precipitation. **B.** The relative amount of His and MBP tagged proteins retained after cobalt resin precipitation was quantified by SyPRO Red staining and normalized to a stock protein (SP, > 4mM) of purified undiluted protein (fraction associated). The highest dilution for G9a-G9a and G9a-GLP, 40nM or > 100 times dilution, is shown.

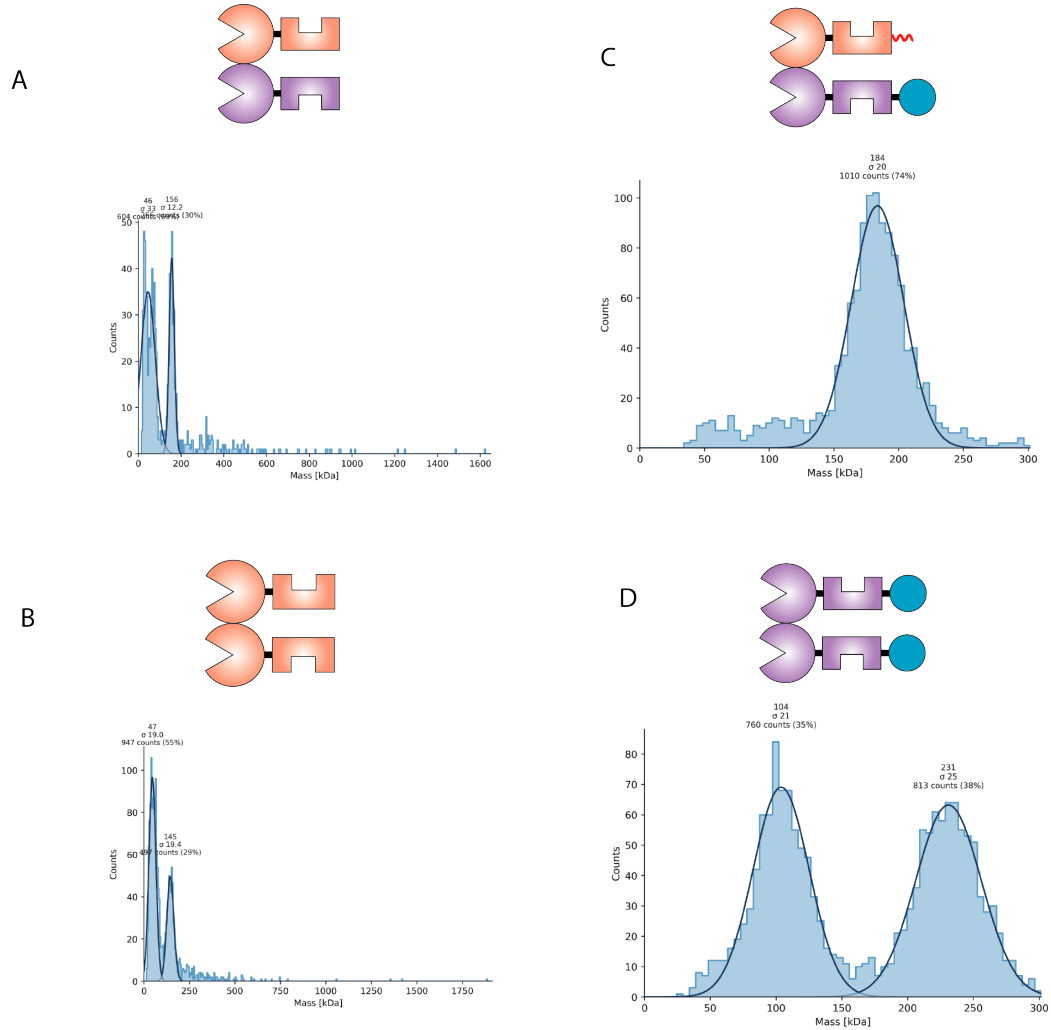


Figure 2.6: Mass Photometry identification of G9a and GLP homo and heterodimers A) G9a-GLP theoretical MW 142kda, empirical 156 +/- 12.2, B) G9a-G9a theoretical MW 137kda, empirical monomer 47 +/- 19, empirical dimer 145 +/- 19.4 C) HIS:G9a-MBP:GLP theoretical MW186 Kda, empirical 184 +/- 20Kda D) MBP:GLP-MBP:GLP theoretical MW 230kda, empirical monomer 104 +/- 21, empirical dimer 231 +/- 25

2.3 Discussion

The degree to which dimerization, or the stability of dimers, is regulated by homo- and heteromeric associations is unknown. Our data argue that the ANK-SET portion of G9a and GLP homo- and heteromeric species form 1:1 dimers with low apparent dissociation and that efficient dimerization is not limited in the homomeric context. However, *in vivo* experiments indicate that when G9a and GLP are present at equal cellular stoichiometry, the heterodimer is the preferred form and that homodimers can form when G9a or GLP is in excess (Tachibana *et al.*, 2005). This *in vivo* preference may be due to regulation outside the ANK-SET domain. Given our mass photometry measurements it is possible homodimers are less stable than heterodimers (**Figure 2.6**). Taken together with our results, a picture emerges where G9a and GLP constitutively form homo- or heterodimers with relative pools determined by the steady-state accumulation of either protein. Homodimers, when formed, may then execute unique functions, for example in DNA repair for GLP (Lu *et al.*, 2019), and in terminal muscle differentiation, where G9a and GLP control non-overlapping genes sets (Battisti *et al.*, 2016). We speculate that the heterodimer, with increased catalysis on chromatin and product recognition (see later sections), is required when large domains of H3K9me1/2 are first formed and then maintained through cell division (Wen *et al.*, 2009; Zyllicz *et al.*, 2015), and that specialized and local chromatin methylation, or methylation of non-histone targets, may be carried out by homodimers.

2.4 Methods

Purification of G9a/GLP Homo and Heterodimers.

To isolate the ANK-SET G9a-GLP heterodimer, we coexpressed N-terminally tagged His:G9a and MBP:GLP from a single plasmid (QB3 Berkeley Macrolab expression vectors) in *E. coli* DE3 Rosetta cells and performed sequential cobalt- and amylose-charged resin affinity chromatography purification. Cells co-expressing His and MBP constructs were lysed on ice via sonication in lysis buffer (100mM Tris pH 8, 300mM NaCl, 10% glycerol (v/v), 0.1% Tween-20, with freshly added 1mM β -mercaptoethanol (BME), 1mM PMSF, 5mM Benzamidine, 200 μ M Leupeptin, Aprotinin, Pepstatin, Phenantroline). The clarified lysate was then bound to cobalt-charged resin (Takara) for 1hr and washed twice with lysis buffer. His tagged proteins were eluted with lysis buffer containing 400mM imidazole and bound immediately to amylose resin (NEB) for 1hr. MBP tagged proteins were eluted with lysis buffer + 20mM maltose. Affinity tags were then removed by incubation with 12mg TEV protease for 1hr at 25°C. TEV protease was absorbed to cobalt resin and the cleaved heterodimer was further purified by size exclusion chromatography (Superdex 200 Increase 10/300 column) and buffer exchanged into storage buffer (100mM Tris pH 8, 100mM KCl, 10% glycerol, 1mM MgCl₂, 20mM ZnSO₄, 10mM BME). Homodimeric MBP:G9a or MBP:GLP ANK-SET constructs were purified as above, omitting the cobalt resin purification. All protein constructs were quantified using SDS page with BSA standards and Sypro Red stain.

Dilution Experiment.

Purified His:G9a::MBP:GLP or His:G9a::MBP:G9a complexes were diluted to 40nM in lysis buffer and allowed to dissociate for 1-2hr at room temperature in a volume of 20 μ L. 7 μ L cobalt-charged resin was added to each sample and incubated for 30min to allow resin binding. The resin was washed twice with lysis buffer and 20 μ L 1x Laemmli Buffer was then added to resin. The ratio of MBP tagged protein to his tagged protein was assessed via SDS page with BSA standards and Sypro Red stain. Fraction assembled was =1 for a “stock protein” (SP) control that was not put through this assay and had a concentration of $>4\mu$ M.

Insect Cell Expression.

Baculovirus containing full-length His:G9a and STREP:GLP co-expression cassettes (QB3 Berkeley Macrolab) were used to infect Sf9 cells (Expression Systems, Davis, California) grown in ESF921 media. Cells were infected for 72hr at an MOI of 0.1. Infected cells were flash frozen and stored at -80°C until thawed for purification. Cells were lysed on ice via sonication in lysis buffer (100mM Tris pH 8, 300mM NaCl, 10% glycerol (v/v), 0.1% Tween-20, with freshly added 1mM β -mercaptoethanol (BME), 1mM PMSF, 5mM Benzamidine, 200 μ M Leupeptin, Aprotinin, Pepstatin, Phenantroline). The clarified lysate was then bound to streptactin superflow resin (IBA) for 1hr and washed twice with 100mM Tris pH 8, 750mM NaCl, 10% glycerol (v/v), 0.1% Tween-20, freshly added 1mM BME. Strep tagged proteins were eluted with lysis buffer containing 5mM desthiobiotin and bound immediately to cobalt-charged resin (Takara) for 1hr. His tagged proteins were eluted with lysis buffer + 400mM imidazole. Complex formation was assessed via western blot.

Size-Exclusion Chromatography coupled to Multi Angle Light Scattering.

For size-exclusion, protein samples were injected at 10 μ M into a silica gel KW804 chromatography column (Shodex; Shanghai, China). For MALS, the chromatography system was coupled to an 18-angle light-scattering detector (DAWN HELEOS II, Wyatt Technology; Santa Barbara, CA) and a differential refractometer (Optilab-rEX, Wyatt Technology). Data collection was done at a flow rate of 0.4mL per minute. SEC MALS data were collected and analyzed using Astra 7 software (Wyatt technology; Santa Barbara, CA).

2.5 References

Battisti, V. *et al.* (2016) ‘Unexpected Distinct Roles of the Related Histone H3 Lysine 9 Methyltransferases G9a and G9a-Like Protein in Myoblasts’, *Journal of Molecular Biology*. Elsevier B.V., 428(11), pp. 2329–2343. doi: 10.1016/j.jmb.2016.03.029.

Bian, C., Chen, Q. and Yu, X. (2015) ‘The zinc finger proteins ZNF644 and WIZ regulate the G9A/GLP complex for gene repression’, *eLife*, 2015(4), pp. 1–17. doi: 10.7554/eLife.05606.

Canzio, D. *et al.* (2011) ‘Chromodomain-mediated oligomerization of HP1 suggests a nucleosome-bridging mechanism for heterochromatin assembly’, *Molecular Cell*. Elsevier Inc., 41(1), pp. 67–81. doi: 10.1016/j.molcel.2010.12.016.

Davidovich, C. *et al.* (2014) ‘NAR Breakthrough Article: A dimeric state for PRC2’, *Nucleic Acids Research*, 42(14), pp. 9236–9248. doi: 10.1093/nar/gku540.

Dodd, I. B. and Sneppen, K. (2011) ‘Barriers and silencers: A theoretical toolkit for control and containment of nucleosome-based epigenetic states’, *Journal of Molecular Biology*. Elsevier Ltd, 414(4), pp. 624–637. doi: 10.1016/j.jmb.2011.10.027.

Erdel, F. *et al.* (2020) ‘Mouse Heterochromatin Adopts Digital Compaction States without Showing Hallmarks of HP1-Driven Liquid-Liquid Phase Separation.’, *Molecular cell*, pp. 1–14. doi: 10.1016/j.molcel.2020.02.005.

Larson, A. G. *et al.* (2017) ‘Liquid droplet formation by HP1 α suggests a role for phase separation in heterochromatin’, *Nature*. Nature Publishing Group, 547(7662), pp. 236–240. doi: 10.1038/nature22822.

Lu, X. *et al.* (2019) ‘GLP-catalyzed H4K16me1 promotes 53BP1 recruitment to permit DNA damage repair and cell survival’, *Nucleic Acids Research*. Oxford University Press, 47(21), pp. 10977–10993. doi: 10.1093/nar/gkz897.

Maison, C. *et al.* (2016) ‘The methyltransferase Suv39h1 links the SUMO pathway to HP1 α marking at pericentric heterochromatin’, *Nature Communications*, 7, p. 12224. doi: 10.1038/ncomms12224.

Margueron, R. *et al.* (2009) ‘Role of the polycomb protein EED in the propagation of repressive histone marks.’, *Nature*, 461(7265), pp. 762–767. doi: 10.1038/nature08398.

Sneppen, K. and Dodd, I. B. (2015) ‘Cooperative stabilization of the SIR complex provides robust epigenetic memory in a model of SIR silencing in *Saccharomyces cerevisiae*’, *Epigenetics*, 10(4), pp. 293–302. doi: 10.1080/15592294.2015.1017200.

Tachibana, M. *et al.* (2005) ‘Histone methyltransferases G9a and GLP form heteromeric complexes and are both crucial for methylation of euchromatin at H3-K9’, *Genes and Development*, 19(7), pp. 815–826. doi: 10.1101/gad.1284005.

Tachibana, M. *et al.* (2008) ‘G9a/GLP complexes independently mediate H3K9 and DNA methylation to silence transcription.’, *The EMBO journal*, 27(20), pp. 2681–90. doi: 10.1038/emboj.2008.192.

Wen, B. *et al.* (2009) ‘Large histone H3 lysine 9 dimethylated chromatin blocks distinguish differentiated from embryonic stem cells.’, *Nature genetics*, 41(2), pp. 246–250. doi: 10.1038/ng.297.

Yuan, W. *et al.* (2012) ‘Dense chromatin activates polycomb repressive complex 2 to regulate H3 lysine 27 methylation’, *Science*, 337(6097), pp. 971–975. doi: 10.1126/science.1225237.

Zhang, H. *et al.* (2014) ‘Statistical mechanics model for the dynamics of collective epigenetic histone modification’, *Physical Review Letters*, 112(6). doi: 10.1103/PhysRevLett.112.068101.

Zylicz, J. J. *et al.* (2015) ‘Chromatin dynamics and the role of G9a in gene regulation and enhancer silencing during early mouse development’, *eLife*, 4(NOVEMBER2015), pp. 1–25. doi: 10.7554/eLife.09571.

3: Heterodimerization Stimulates G9a and GLP Catalytic Activity

3.1 Introduction

The SET domain is a conserved feature of histone methyltransferases giving them the ability to write histone methylation (Dillon *et al.*, 2005). SET domains contain two binding pockets: one for binding the substrate lysine peptide and another for binding the co-factor S-adenosyl methionine. Bringing these two substrates into proximity, the SET domain allows methylation to be transferred from SAM to the substrate lysine and hence a methyltransferase reaction occurs (Tachibana *et al.*, 2002). SET domains rely on tyrosine residues in their binding pockets to carry out methylation. In the SET domain active site, one key tyrosine positions SAM and the substrate lysine for appropriate transfer, and up to three tyrosines restrict the rotation of the lysine epsilon amino terminus, biasing the final methyl state of that residue (Collins *et al.*, 2005; Wu *et al.*, 2010). A lysine residue can be methylated up to three times, yielding mono-, di- and tri- methylation. Restricting the rotation of the substrate lysine epsilon amino group in the active site restricts the number of methylation states it can achieve by preventing the geometries necessary to access different sides of the lysine residue.

Repressive SET domain proteins require a consensus ARK sequence in order to perform methylation (Chin *et al.*, 2005). ARK sequences are not specific to histone substrates, which in fact, vastly opens up the substrate pool of SET domain containing methyltransferases beyond that of methylating histones. G9a and GLP have a wide range of substrates they can methylate, from the proteins that recruit them to genomic sites (e.g. WIZ1), to various transcription factors (e.g. HIF1a, FOXO1), to their very own N-termini (Chin *et al.*, 2007; Rathert *et al.*, 2008; Poulard *et al.*, 2017; Bao *et al.*, 2018; Chae *et al.*, 2019; Chopra *et al.*, 2020). This promiscuous activity links G9a and GLP activity to broad areas of cellular regulation.

G9a and GLP are not the only SET domain methyltransferases able to form dimers, however, they are unique in that 1) their dimerization interface is linked to the SET domain and 2) they dimerize heterotypically (Tachibana *et al.*, 2005). *Drosophila* SU(VAR)3-9 dimerizes via N-terminal interactions while human EZH1 of the Polycomb Repressive Complex was also observed to dimerize though it is unclear exactly which domains are necessary for dimerization (Eskeland *et al.*, 2004; Davidovich *et al.*, 2014). A viral SET domain protein (vSET) homodimerizes via SET domain interactions (Manzur *et al.*, 2003; Wei and Zhou, 2010). It is yet unclear how dimerization via SET domains may regulate methyltransferase activity. In this section, we examined the differences between G9a and GLP homo or heterodimeric forms in their ability to methylate various substrates in order to elucidate any regulation imposed by heterodimerization.

3.2 Results

Because dimerization occurs within the catalytic SET domain, we asked whether dimerization constitutively affects the methyltransferase activity of G9a-GLP. We determined Michaelis Menten kinetic parameters of G9a and GLP homodimers as well as the G9a-GLP heterodimer using an H3 histone tail peptide substrate (H3₁₋₂₀, (Al-Sady, Madhani and Narlikar, 2013)) (**Figure 3.5 A, G**). In preparation for these experiments, we found a pH dependence on the linearity of the initial rate curves we measured with G9a-GLP (**Figure 3.1**). Initial rates measured at pH = 8 proceeded linearly in our time courses while measurements at pH = 8.5 produced an exponential shape. Similar results have been previously been observed with the G9a homodimer (Patnaik *et al.*, 2004). This pH dependence may reflect low stability in these enzymes at higher pH's whereby prolonged exposure inactivate the enzymes. Additionally, we found no difference in methyltransferase reaction behavior upon preincubating G9a-GLP with the peptide substrate or SAM and then initiating methylation with the respective other substrate (**Figure 3.2**). Subsequent kinetic assays involved mixing peptide and SAM substrates and initiating the reaction via addition of methyltransferase. Finally, we observed no difference in activity between full length G9a-G9a and our ANKSET G9a-G9a truncation construct, validating our use of the truncation (**Figure 3.3**).

Comparing the G9a-GLP heterodimer to their respective homodimers, while the K_M of all enzyme species were similar (e.g. **Figure 3.5 B**), we observed changes of the k_{cat} within the heterodimer beyond those expected from equal contributions from G9a and GLP: The G9a and GLP k_{cats} on H3₁₋₂₀ are $\sim 33\text{min}^{-1}$ and $\sim 14\text{min}^{-1}$, respectively. Rather than an intermediate value, G9a-GLP's k_{cat} is $\sim 33\text{min}^{-1}$, suggesting one or both enzymes are more active in the heterodimer than their respective homodimers. To determine which enzyme is stimulated in the heterodimer

we made point mutations in either G9a (G9aS, Y1120V, Y1207F) or GLP (GLPS, Y1240F) abrogating their catalytic activity while still allowing dimerization with a wild type allele of their binding partner (Estève *et al.*, 2005; Tachibana *et al.*, 2008). Comparing G9aS-GLP to GLP-GLP we observe that GLP's k_{cat} increased ~ 2 times upon heterodimerization (**Figure 3.5C**). Activity comparisons of G9a-GLPS and G9a-G9a suggested no significant increase in activity of G9a in the heterodimer (**Figure 3.4**), indicating that only GLP's activity is meaningfully enhanced in the heterodimer. Thus, these data imply a modest catalytic enhancement in G9a-GLP versus the homodimeric forms.

Next, we asked if the G9a-GLP heterodimer exhibits altered methylation kinetics on a mononucleosome, which mimics its cellular chromatin substrate (**Figure 3.5D**). We first measured Michaelis Menten parameters under multiple turnover conditions, just like for H3₁₋₂₀, for G9a-GLP and G9a-G9a on a reconstituted mononucleosome. In striking contrast to our observations on H3₁₋₂₀, G9a-GLP's k_{cat} is 10 times higher than that of G9a-G9a (**Figure 3.5E,G** and **3.6B**). To further confirm this observation, and to additionally compare G9a-GLP to GLP-GLP, we performed methylation timecourses under near-saturating conditions for the mononucleosome, (5 μ M mononucleosome), as inferred from Michaelis-Menten curves in Figure 3E. Under these conditions, the heterodimer also exhibits higher activity (~ 8 times higher k_{cat}) than G9a-G9a. Similar to the H3₁₋₂₀ substrate, the GLP homodimer is the slowest enzyme construct of the three, preventing us from producing a multiple turnover Michaelis-Menten curve. (**Figure 3.5F**). We conclude that the nucleosome substrate brings to fore the intrinsic catalytic differences between homo and heterodimers, revealing a dramatically increased k_{cat} for G9a-GLP. However, concomitant with this k_{cat} increase, we also noticed an increase in the nucleosome K_{M} for G9a-GLP (**Figure 3.5E,G**). This may indicate a decrease in binding affinity

for the nucleosome or may also suggest a difference in product release. In a differential product release model, G9a and GLP homodimers stay associated with the nucleosome after methylation has occurred, essentially trapping the enzymes on the product nucleosome, disabling their ability to attach to unmethylated substrates. The G9a-GLP heterodimer may have a faster product release rate, allowing it to overcome entrapment by the methyl nucleosome product.

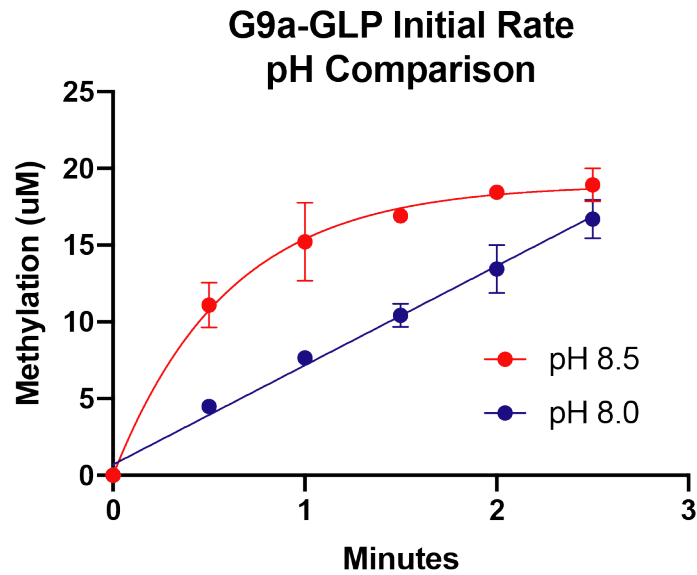


Figure 3.1 Effect of pH on G9a-GLP methylation reaction. G9a-GLP reaction kinetics were measured with a saturating concentration of SAM and peptide in two separate buffers containing Tris pH 8.5 or pH 8.0. Reactions run in pH 8.5 buffer (red) followed a single exponential fit, while pH 8 buffer caused the reaction to run linearly. Measurement taken in duplicate.

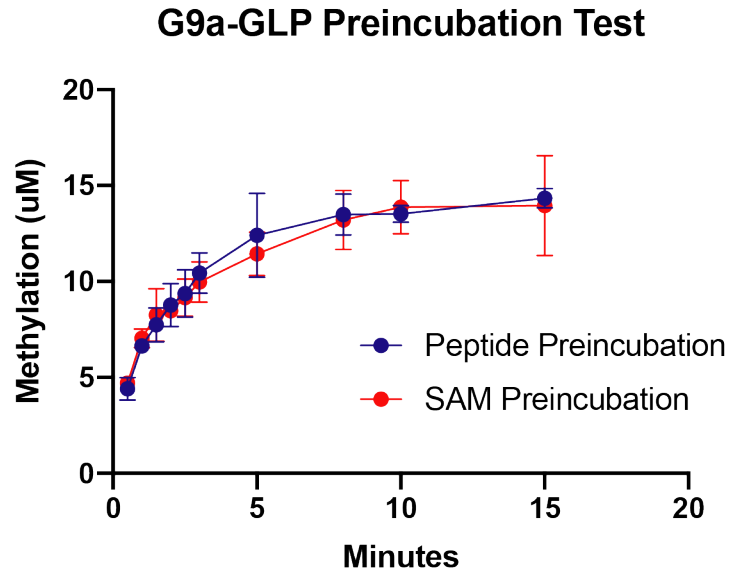


Figure 3.2 G9a-GLP Pre-incubation Test. G9a-GLP was incubated with H3 tail peptide substrate (blue) or SAM (red) for 10 minutes prior to addition of the second substrate to initiate the reaction. We found no difference in reaction kinetics upon preincubation with either substrate. Measurement taken in duplicate.

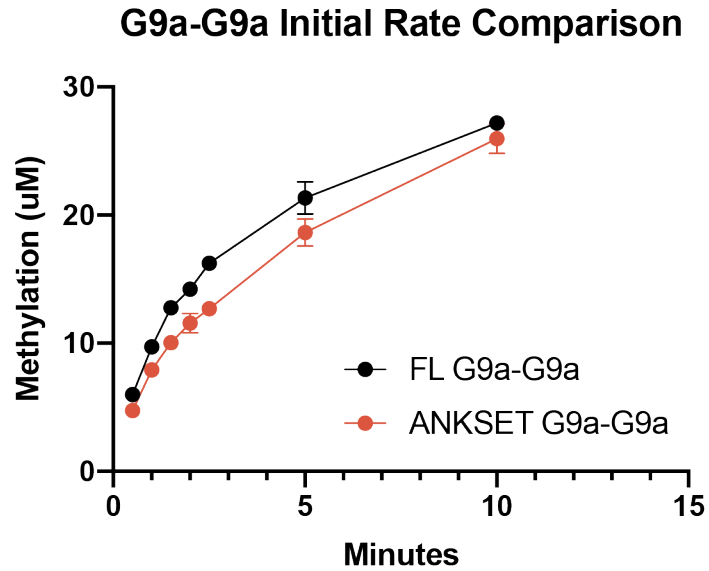


Figure 3.3 G9a-G9a Full Length vs ANKSET truncation initial rate comparison Compared activities of Full Length G9a (Black) and ANKSET truncation (red) used in this paper. Activities were nearly identical. Saturating concentrations of H3 tail peptide and SAM were used. Measurements taken in duplicate. Buffer containing tris pH = 8.5 was used in this experiment.

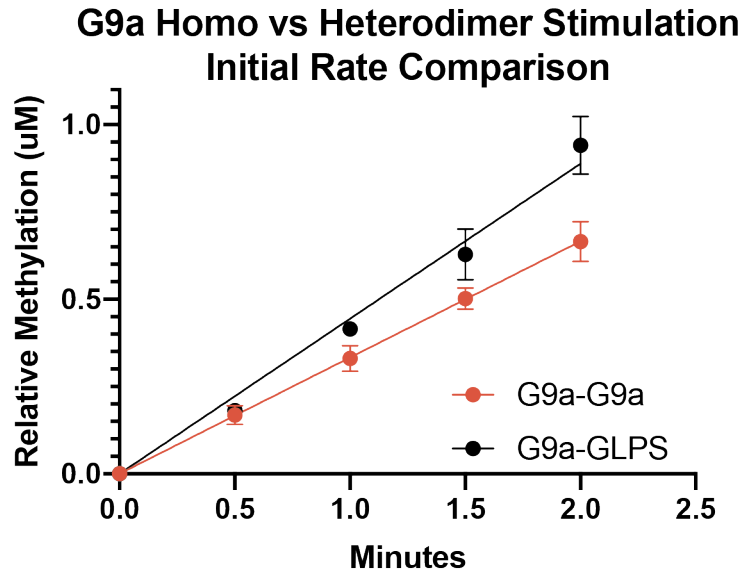
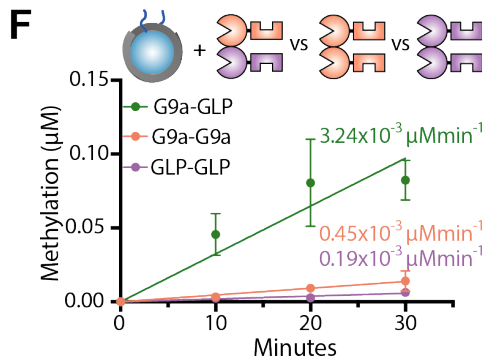
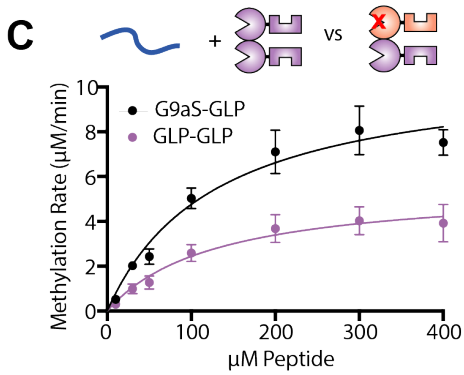
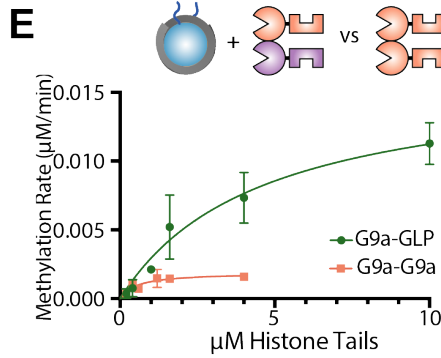
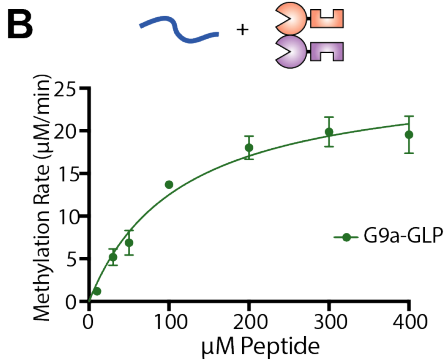
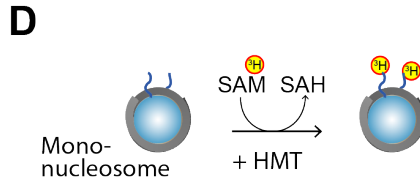
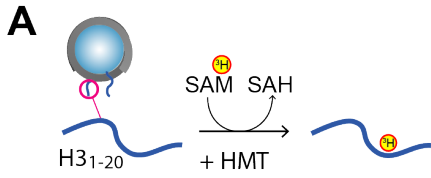


Figure 3.4 Initial Rate Comparisons of G9a in homodimeric and heterodimeric forms. G9a-G9a homodimer (red) and G9a-GLPS (black) kinetics were measured with saturating concentrations of H3 tail peptides and SAM. Relative methylation signal is plotted here. Relative rates - G9a-G9a 0.75, G9a-GLPS 1.



G

	H3 ₁₋₂₀ peptide		core nucleosome	
	k _{cat} (min ⁻¹)	K _M (μM)	k _{cat} * (min ⁻¹)	K _M (μM)
dimer				
G9a-G9a	32.6 24.6-34.6	133.6 58.4-165.8	0.0047 0.0034-0.0067	0.54 0.18-1.4
GLP-GLP	14.0 9.5-15.5	128.6 45.5-199.7	N/A	N/A
G9a-GLP	33.5 26.0-34.5	114.1 57.7-141.5	0.041 0.029-0.072	4.6 2.0-13.3
G9aS-GLP	27.2 20.5-28.3	128.1 61.1-163.8	N/A	N/A

Figure 3.5: Heterodimerization stimulates G9a and GLP catalytic activity. **A.** Reaction scheme of H3 tail peptide methylation. SAM cofactor was labeled with tritium (^3H). **B.** Michaelis Menten fit of G9a-GLP heterodimer methylating H3 tail peptide. **C.** Michaelis Menten fit of GLP homodimer (purple) vs G9aS-GLP heterodimer (black). **D.** Reaction scheme of mononucleosome methylation reactions. **E.** Michaelis Menten fit of G9a-GLP heterodimer (green) and G9a homodimer (pink) methylating mononucleosomes. **F.** Initial rate comparison of G9a-GLP (green), G9a homodimer (pink), and GLP homodimer (purple) methylating mononucleosomes under saturating conditions. Reactions contained 10mM histone tails (5mM mononucleosome). **G.** Kinetic parameters for histone peptide and mononucleosome methylation. Values reflect kinetic parameters (upper value) and 95% confidence intervals (lower values) extracted from duplicate measurements.

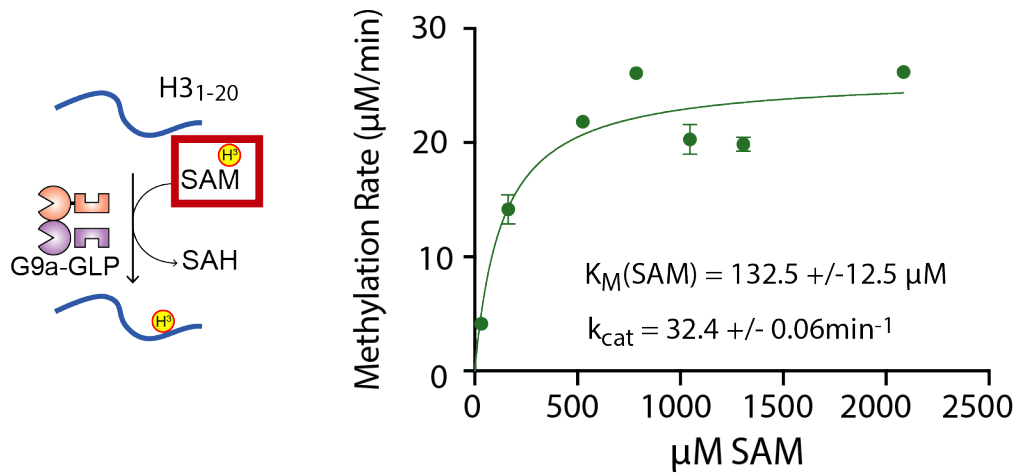
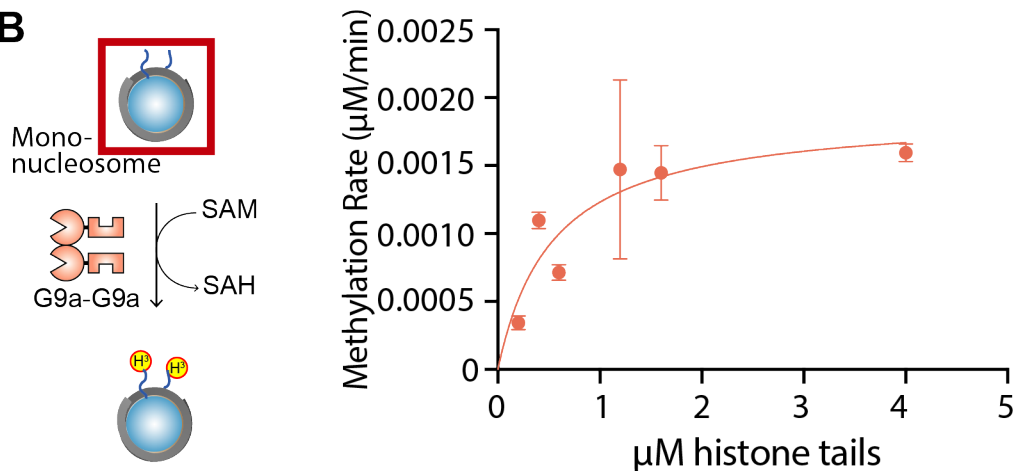
A**B**

Figure 3.6 S-adenosyl methionine K_M and G9a-G9a mononucleosome Michaelis Menten kinetics. **A.** The K_M for SAM was determined with G9a-GLP and H3₁₋₂₀ peptide substrates. Initial rates are plotted at indicated total SAM concentrations and the curve fit to $V = V_{\text{max}} \cdot [S] / (K_M + [S])$. **B.** A magnification of Michaels-Menten fit for G9a-G9a Multiple Turnover kinetics as shown in Figure 3E. All measurements taken in duplicate.

3.3 Discussion

We observe both modest, constitutive activation of GLP on H3 tail peptides, and a dramatic increase in the k_{cat} towards nucleosomes compared to the homodimers. Based on our peptide methylation analysis, we observed GLP to be constitutively activated upon heterodimerization (**Figure 3.5 B**). One interpretation of this result is a general release of an autoinhibited state of GLP. Indeed, autoinhibition loops have been identified in a number of SET domain containing methyltransferases. The fission yeast Clr4 enzyme contains an autoinhibitory loop that is relieved by automethylation (Iglesias *et al.*, 2018) and similar inhibitory loops have been shown for PRC2 and NSD1 (Graham, Tweedy and Carlson, 2016; Bratkowski, Yang and Liu, 2017). The crystal structure of the GLP SET domain revealed an alpha helix in the c-terminal post-SET region that could act as a potential steric block to the SAM binding pocket (Wu *et al.*, 2010). In vivo analysis of GLP suggested it's catalytic activity is not necessary to recover loss of H3K9me (Tachibana *et al.*, 2008). It's possible that the slight increase of GLP methyltransferase activity is an artifact of a larger regulatory phenomenon brought on by its heterodimerization with G9a.

We observed a much larger (8-10x) effect of heterodimerization affecting methyltransferase activity when we measured activities on nucleosomes (**Figure 3.5 C**). In this context, we observed increased activity of both G9a and GLP. With eight full-length histones and 147bp of DNA, the nucleosomes used in this study represent a more complex substrate than the H3 tail peptide substrate. Forming a heterodimer may improve each enzyme's ability to survey the full nucleosome landscape and interact appropriately with the H3 tail. The observed increase in K_m and k_{cat} comparing G9a-GLP to G9a-G9a is consistent with a potential non-productive binding mode of G9a-G9a to chromatin that is avoided when G9a-GLP engages with

the nucleosome. Another model is that the G9a-GLP interaction with chromatin allosterically stimulates each enzyme's ability to methylate the H3 tail. The subunit, SUZ12, of the PRC1 complex is able to sense dense chromatin substrates and stimulate methyltransferase EZH2 (Yuan *et al.*, 2012). In this way it's possible a site exists on the G9a and GLP that when in the heterodimer is able to engage with chromatin in a stimulatory manner, while as a homodimer it does not stimulate activity. Further kinetic and structural analysis must be done to inform and distinguish these models.

3.4 Methods

Preparation of mononucleosome substrates.

Histone proteins and nucleosomes were purified as described (Canzio *et al.*, 2011) with the following modifications: Following assembly nucleosomes were dialyzed overnight into storage buffer (above) or FP storage buffer (50mM HEPES pH 7.5, 100mM KCl, 10% glycerol) depending on the application. Dialyzed samples were concentrated using 10kda Millipore Sigma Amicon Ultra Centrifugal Filter Units. For fluorescence polarization, nucleosomes were reconstituted with a 5' fluorescened DNA template.

Methylation Assays.

All methylation reactions used a tritium-based assay to detect the methylated product. Substrate peptides or nucleosomes were mixed in a solution containing 9mM tritiated S-Adenosyl Methionine (SAM, Perkin Elmer) cofactor, and reactions were initiated upon addition of 0.4-0.8mM enzyme. Before adding enzyme, peptide reactions were supplemented with 1000mM cold SAM to fully saturate the SAM binding pocket (**Figure 3.6**). Due to signal limitations, nucleosome reactions were not supplemented with SAM (see below). Reactions were run in 100mM Tris pH 8, 100mM KCl, 1mM MgCl₂, 20mM ZnSO₄, 10mM BME and quenched with laemmli buffer. Peptide reactions were performed as described (Al-Sady, Madhani and Narlikar, 2013). Mononucleosome methylation reactions were read out via autoradiography. Proteins were separated on a 18% SDS PAGE gel which was dried and exposed to a GE Phosphor screen for 72 hours along with a standard curve of tritiated SAM spotted on Whatman paper, and imaged on a STORM imager. Images were quantified using Image Quant software (Cytiva).

Kinetics.

All methylation reactions were performed under Multiple Turnover conditions (Fersht, 1998). G9a and GLP each catalyze a two-substrate bi-bi reaction using the cofactor SAM and H3K9 as substrates (Patnaik *et al.*, 2004). To determine initial rates, methylation time courses were traced at various concentrations of H3K9-containing substrate while keeping the concentration of SAM and enzyme constant. Plots of initial rate vs. concentration H3K9 substrate measured in duplicate or triplicate were fit to $V = V_{\max} * [S] / (K_M + [S])$ using Prism software to extract K_M and k_{cat} values as well as the 95% confidence interval bounds of the fit. For peptide reactions, we saturated the SAM binding pocket (see above) with a mixture of tritiated and non-tritiated SAM to measure kinetic parameters under pseudo first-order conditions. For nucleosome experiments, we could not fully saturate the SAM pocket, due to signal limitations. The k_{cat} values reported here are thus limited to first-order rate constants with a subsaturating second substrate.

3.5 References

Al-Sady, B., Madhani, H. D. and Narlikar, G. J. (2013) 'Division of labor between the chromodomains of HP1 and Suv39 methylase enables coordination of heterochromatin spread', *Molecular Cell*. Elsevier Inc., 51(1), pp. 80–91. doi: 10.1016/j.molcel.2013.06.013.

Bao, L. *et al.* (2018) 'Methylation of hypoxia-inducible factor (HIF)-1 α by G9a/GLP inhibits HIF-1 transcriptional activity and cell migration', *Nucleic Acids Research*. Oxford University Press, 46(13), pp. 6576–6591. doi: 10.1093/nar/gky449.

Bratkowski, M., Yang, X. and Liu, X. (2017) 'crosstalk Polycomb repressive complex 2 in an autoinhibited state', 292, pp. 13323–13332. doi: 10.1074/jbc.M117.787572.

Canzio, D. *et al.* (2011) 'Chromodomain-mediated oligomerization of HP1 suggests a nucleosome-bridging mechanism for heterochromatin assembly', *Molecular Cell*. Elsevier Inc., 41(1), pp. 67–81. doi: 10.1016/j.molcel.2010.12.016.

Chae, Y. C. *et al.* (2019) 'FOXO1 degradation via G9a-mediated methylation promotes cell proliferation in colon cancer', *Nucleic Acids Research*. Oxford University Press, 47(4), pp. 1692–1705. doi: 10.1093/nar/gky1230.

Chin, H. G. *et al.* (2005) 'Sequence specificity and role of proximal amino acids of the histone H3 tail on catalysis of murine G9a lysine 9 histone H3 methyltransferase', *Biochemistry*, 44(39), pp. 12998–13006. doi: 10.1021/bi0509907.

Chin, H. G. *et al.* (2007) ‘Automethylation of G9a and its implication in wider substrate specificity and HP1 binding’, *Nucleic Acids Research*, 35(21), pp. 7313–7323. doi: 10.1093/nar/gkm726.

Chopra, A. *et al.* (2020) ‘Hypoxia-Inducible Lysine Methyltransferases: G9a and GLP Hypoxic Regulation, Non-histone Substrate Modification, and Pathological Relevance’, *Frontiers in Genetics*, 11(September), pp. 1–15. doi: 10.3389/fgene.2020.579636.

Collins, R. E. *et al.* (2005) ‘In Vitro and in Vivo Analyses of a Phe / Tyr Switch Controlling Product Specificity of Histone Lysine Methyltransferases *’, 280(7), pp. 5563–5570. doi: 10.1074/jbc.M410483200.

Davidovich, C. *et al.* (2014) ‘NAR Breakthrough Article: A dimeric state for PRC2’, *Nucleic Acids Research*, 42(14), pp. 9236–9248. doi: 10.1093/nar/gku540.

Dillon, S. C. *et al.* (2005) ‘The SET-domain protein superfamily: Protein lysine methyltransferases’, *Genome Biology*, 6(8). doi: 10.1186/gb-2005-6-8-227.

Eskeland, R. *et al.* (2004) ‘The N-Terminus of Drosophila SU(VAR)3-9 Mediates Dimerization and Regulates Its Methyltransferase Activity’, *Biochemistry*, 43(12), pp. 3740–3749. doi: 10.1021/bi035964s.

Estève, P. O. *et al.* (2005) ‘Functional analysis of the N- and C-terminus of mammalian G9a histone H3 methyltransferase’, *Nucleic Acids Research*, 33(10), pp. 3211–3223. doi: 10.1093/nar/gki635.

Fersht, A. (1998) *Structure and Mechanism in Protein Science: A Guide to Enzyme Catalysis and Protein Folding*. 1st edn. W. H. Freeman, New York.

Graham, S. E., Tweedy, S. E. and Carlson, H. A. (2016) ‘Dynamic behavior of the post-SET loop region of NSD1: Implications for histone binding and drug development’, *Protein Science*, 25(5), pp. 1021–1029. doi: 10.1002/pro.2912.

Iglesias, N. *et al.* (2018) ‘Automethylation-induced conformational switch in Ctr4 (Suv39h) maintains epigenetic stability’, *Nature*. Springer US, 560(7719), pp. 504–508. doi: 10.1038/s41586-018-0398-2.

Manzur, K. L. *et al.* (2003) ‘A dimeric viral SET domain methyltransferase specific to Lys27 of histone H3’, *Nature Structural Biology*, 10(3), pp. 187–196. doi: 10.1038/nsb898.

Patnaik, D. *et al.* (2004) ‘Substrate Specificity and Kinetic Mechanism of Mammalian G9a Histone H3 Methyltransferase *’, 279(51), pp. 53248–53258. doi: 10.1074/jbc.M409604200.

Poulard, C. *et al.* (2017) ‘A post-translational modification switch controls coactivator function of histone methyltransferases G9a and GLP’, *EMBO reports*, 18(8), pp. 1442–1459. doi: 10.15252/embr.201744060.

Rathert, P. *et al.* (2008) ‘Protein lysine methyltransferase G9a acts on non-histone targets’, *Nature Chemical Biology*, 4(6), pp. 344–346. doi: 10.1038/nchembio.88.

Tachibana, M. *et al.* (2002) ‘G9a histone methyltransferase plays a dominant role in euchromatic histone H3 lysine 9 methylation and is essential for early embryogenesis’, *Genes and Development*, 16(14), pp. 1779–1791. doi: 10.1101/gad.989402.

Tachibana, M. *et al.* (2005) ‘Histone methyltransferases G9a and GLP form heteromeric complexes and are both crucial for methylation of euchromatin at H3-K9’, *Genes and Development*, 19(7), pp. 815–826. doi: 10.1101/gad.1284005.

Tachibana, M. *et al.* (2008) ‘G9a/GLP complexes independently mediate H3K9 and DNA methylation to silence transcription.’, *The EMBO journal*, 27(20), pp. 2681–90. doi: 10.1038/emboj.2008.192.

Wei, H. and Zhou, M.-M. (2010) ‘Dimerization of a viral SET protein endows its function’, *Proceedings of the National Academy of Sciences*, 107(43), pp. 18433–18438. doi: 10.1073/pnas.1009911107.

Wu, H. *et al.* (2010) 'Structural biology of human H3K9 methyltransferases', *PLoS ONE*, 5(1).
doi: 10.1371/journal.pone.0008570.

Yuan, W. *et al.* (2012) 'Dense chromatin activates polycomb repressive complex 2 to regulate H3 lysine 27 methylation', *Science*, 337(6097), pp. 971–975. doi: 10.1126/science.1225237.

4: Probing G9a-GLP Interaction with Chromatin

4.1 Introduction

A feature conserved from yeast to humans of heterochromatic methyltransferases is their ability to stimulate methyltransferase activity upon recognition of existing H3K9 methylation on nucleosomes proximal to their substrate nucleosome (Hall *et al.*, 2002; Liu *et al.*, 2015; Müller *et al.*, 2016; Lee *et al.*, 2018). I term this recognition-stimulation event *product recognition stimulation* and allows existing H3K9 methylation (H3K9me), or H3K27me in the case of PRC2 heterochromatin, at a gene locus to create a positive feedback loop that promotes the installation of more methylation at that gene region. Product recognition stimulation is fundamental to both establishment and maintenance of a heterochromatic domain.

To establish a heterochromatic domain, histone methyltransferases are often recruited to sites in the genome in a DNA sequence dependent manner ((Simon and Kingston, 2013; Bian, Chen and Yu, 2015; Holoch, Moazed and Avenue, 2015). Upon installing, or establishing, H3K9me at their recruitment site, H3K9 methyltransferases rely on product recognition stimulation to propagate, or spread, their modifications beyond that site in a DNA sequence independent manner (Hall *et al.*, 2002). With the help of product recognition stimulation, H3K9 methylated gene regions can expand to megabase sizes (Wen *et al.*, 2009). Product recognition stimulation also works to maintain the methylation status of a gene region through cell division (Audergon *et al.*, 2015; Rangunathan, Jih and Moazed, 2015), allowing it to recover from 1) dilution of methyl marks through DNA replication (Dodd *et al.*, 2007; Dodd and Sneppen, 2011; Hathaway *et al.*, 2012; Hodges and Crabtree, 2012) and 2) spurious invasion by H3K9 demethylases (Erdel and Greene, 2016; Jehanno *et al.*, 2017). Product recognition stimulation by

H3K9 methyltransferases is considered to be a fundamental mechanism enabling the epigenetic inheritance of H3K9me, one of very few histone modifications to demonstrate this behavior.

G9a and GLP have been reported to perform product recognition stimulation. In a study by Nan Liu and colleagues, G9a was found to only stimulate activity upon recognition of H3K9me₂ nucleosomes while GLP stimulated only upon reading H3K9me₁ (Liu *et al.*, 2015). This is consistent with fluorescence polarization measurements demonstrating G9a ANK has the highest affinity for H3K9me₂ peptides compared to other H3K9me states, while GLP ANK prefers binding H3K9me₁ (Collins *et al.*, 2008). These measurements were taken with truncations of each enzyme that only contain the ANK repeats. Given that the SET domain was determined to be required for G9a-GLP dimerization, measurements of truncations only containing the ANK repeat suggest monomeric forms of G9a and GLP were used. It is unclear how the presence of the SET domain and thus dimerization of G9a and GLP may affect reading in the ANK repeat.

Product recognition stimulation of G9a and GLP occurs only on chromatin in a manner that is restricted *in cis*, and cannot be induced or propagated *in trans* (Liu *et al.*, 2015). *In cis* product recognition stimulation describes stimulation that occurs on nucleosomes connected on the same DNA strand to methylated nucleosomes, most likely the immediate proximal nucleosome. *In trans* stimulations suggests a methyl nucleosome from distal DNA sites or from other chromatin DNA can trigger product recognition stimulation. G9a/GLP product recognition stimulation is reminiscent of the *S. pombe* HMT, Clr4, which is stimulated only when an activating H3K9me nucleosome is proximal to an unmethylated substrate nucleosome (Al-Sady, Madhani and Narlikar, 2013). This is unlike the heterochromatic HMTs Suv39H1 and PRC2, which can be stimulated *in trans* by a soluble methylated peptide that induces allosteric changes

communicated to each HMTs SET domain (Müller *et al.*, 2016; Poepsel, Kasinath and Nogales, 2018a). The available data thus suggests a “guided state” model for G9a-GLP, which predicts unique structural orientation of the G9a-GLP dimer on the substrate and high dependency on the nucleosomal arrangement. However, it remains possible that G9a-GLP is not constrained to unique structural orientations (flexible orientation model) where interaction with H3K9me instead serves to either tightly retain the enzyme on chromatin, lowering its K_M as in the case of Suv39H1, or allosterically activate its active site, as in the case of PRC2, with the observed *in cis* restriction deriving from other constraints (Müller *et al.*, 2016; Poepsel, Kasinath and Nogales, 2018a). Though the current data supports a guided state model for G9a-GLP, more rigorous structural and enzymological studies need to be carried out to test this model and explore other alternatives or variants of this model.

In this section, I sought to better understand G9a and GLP recognition of H3K9me by 1) understanding how homodimerization and heterodimerization of these enzymes may affect reading of these marks and 2) reconstituting product recognition stimulation. I envisioned to build nucleosome arrays with varying geometries and methyl states to explore the guided state model as well as learn if as a heterodimer G9a-GLP can overcome the limitations of reading only me1 or me2 nucleosomes to induce product recognition stimulation. In the end, I was not able to faithfully reconstitute product recognition stimulation, however, I found that G9a-G9a recognition of the H3K9me2 state is highly dependent on dimerization with GLP-GLP.

4.2 Results

The presumably monomeric ANK domains of G9a and GLP have been shown to interact with both H3K9me1 and H3K9me2 histone tails, to largely similar extents (Collins *et al.*, 2008; Liu *et al.*, 2015). Using fluorescence polarization, I asked whether the ability to engage H3K9me1/me2 is preserved in the dimeric ANK-SET molecule and whether heterodimerization affected this reading function (**Figure 4.1**). Interestingly, except for GLP ANK-SET binding to me1 (**Figure 4.1 A**), the presence of the SET domain in the context of homodimers is broadly inhibitory to H3K9me binding compared to published ANK alone data (Collins *et al.*, 2008; Liu *et al.*, 2015). GLP ANK-SET has reduced binding to me2 peptides (**Figure 4.1 B**), while G9a ANK-SET has no discernable affinity for me1 and only slight affinity for me2 peptides (**Figure 4.1 C,D**). Heterodimerization abrogated the negative regulation imposed by the SET domain, allowing interaction with me1 and me2 peptides in a range comparable to the ANK domain alone (**Figure 4.1 E,F**). To discern whether our fluorescence polarization assays were indeed measuring binding in the ANK domain as opposed to the SET domain, I made point mutations in GLP ANK repeat (GLPA, W877A/W882A/E885A) or GLP SET domain (GLPS, Y1240F) abrogating binding in either domain (Collins *et al.*, 2008). Low affinity binding of an H3K9me1 peptide was observed in GLPA condition while affinity similar to WT GLP was observed with GLPS condition (**Figure 4.2**). I conclude that peptide-binding measurements of the WT GLP are reflective of binding in the ANK domain. Taken together, these results suggest that within the ANK-SET homodimeric context, the ANK domains of both G9a and GLP are partially compromised in their ability to bind methyl peptides and that this inhibition is overcome upon heterodimerization.

To measure product recognition stimulation, the methylation rate on wild type nucleosomes flanked by methyl nucleosomes is compared to the methylation rate of wild type nucleosomes flanked by control, non-methylatable nucleosomes (Margueron *et al.*, 2009; Al-Sady, Madhani and Narlikar, 2013; Müller *et al.*, 2016) (**Figure 4.3**). I constructed arrays containing twelve nucleosomes with a random assortment of wild type and either di-methylated or H3K9R control nucleosomes. In order to guarantee faithful positioning of histone octamers onto our array DNA, I relied on the 601 positioning sequence placed at twelve sites on the array DNA template (Lowary and Widom, 1998). Each 601 positioning sequence was separated by 40-50bp DNA, also known as the linker regions of the array. Array DNA was amplified in a plasmid engineered to use a single restriction enzyme, EcoRV, to excise the array DNA and reduce the plasmid backbone to small >500bp fragments (Gibson *et al.*, 2019). Arrays were assembled the same as mononucleosomes except that wild type octamers were mixed at an equal ratio to di-methyl or K9R octamers. As such, each array had a random assortment of wild type and non wild type nucleosomes (**Figure 4.4 A**). I found the ANKSET G9a-GLP constructs used in this study had no measureable affinity to DNA (**Figure 4.5**). As such, in order to increase my yield of assembled arrays, I did not perform a final clean up of the unassembled DNA, assuming it would not affect the methylation reaction.

G9a and GLP are able to methylate lysine 9 di-methyl nucleosomes to tri-methyl as well as lysine 27 to mono and di (Collins *et al.*, 2005; Wu *et al.*, 2011). In order to faithfully measure a true product recognition event (i.e. methylation specific to the wild type nucleosome), I had to design the assay to distinguish methylation of the wild type nucleosome from any side reaction that occurred on the K9R or premethylated nucleosomes. To this end I used a 6xHIS extension with additional Glycine-Serine linker fused to the c-terminus of histone H3 (6xHIS:H3) which

effectively increased the molecular weight of H3 by ~1Kda. This molecular weight addition was enough to distinguish 6xHIS:H3 migration from untagged H3 on an 18% SDS-PAGE gel (**Figure 4.4 B**). I measured methylation via autoradiography, using the tritiated cofactor S-adenosyl methionine to track methylation. Coupling separation of tagged and 6xHIS:H3 with autoradiograph gave very reliable signal and distinction between methylation on wild type and di-methyl or K9R nucleosomes (**Figure 4.4 C**). For these assays, I tagged the wild type histone H3 with 6xHIS and left di-methyl or K9R H3 untagged (**Figure 4.4 A**).

G9a was previously shown to stimulate methylation activity upon reading H3K9me2 (Liu *et al.*, 2015). In my initial experiments I tried to replicate this result by comparing methylation with array substrates containing HIS:WTH3 and H3K9me2 nucleosomes. Product recognition stimulation could be concentration dependent whereby enzyme or substrates below the reaction K_m would show stimulation and saturating the K_m would not show stimulation (Müller *et al.*, 2016). If product recognition stimulation was not dependent on enzyme concentration it would suggest a k_{cat} effect, indicating a structural change may occur upon recognition of H3K9me2 (Poepsel, Kasinath and Nogales, 2018b). As such, I measured reaction kinetics on array substrates with two concentrations of G9a-G9a (**Figure 4.6**). Differences in methylation rate were observed at 0.35uM G9a-G9a and not at 10uM G9a-G9a. Observing product recognition stimulation at the low enzyme concentration led us to hypothesize stimulation is dependent on the K_m more than the k_{cat} of G9a-G9a.

However, this result remained difficult to reproduce. Attempts were made to reproduce the above result with identical experimental conditions. In my first attempt, I observed a higher relative activity on the H3K9me2 array over H3K9R array than previously observed (**Figure 4.7**). Observing this result, I compared again replicated these experimental conditions now

comparing G9a-G9a activity to that of G9a-GLP. For this experiment I used an endpoint assay with a 60 minute time point determined from my previous experiments (**Figure 4.8**).

Unfortunately, the system did not behave as expected, as G9a-G9a was unable to methylate the H3K9me2 array faster than the H3K9R array. With every iteration of this assay, I needed to prep new array substrates. To determine if the variability I observed in this assay was due to differences in array preps, I replicated the experimental conditions, measuring G9a-G9a methylation with two different preps of arrays. In these experiments I also examined any differences between preps of G9a-G9a (**Figure 4.9**). These experiments also failed to demonstrate an increased activity of G9a-G9a on H3K9me2 arrays.

Finally, I attempted to measure stimulation in an alternative way (**Figure 4.10**), comparing methylation time traces of G9a-G9a and a G9aA-G9aA homodimer containing point mutations in its ANK repeat abrogating its reading function (G9a W791A/ W796A/E799A, G9aA-G9aA). If product recognition stimulation occurred, I would have observed a higher activity in G9a-G9a than G9aA-G9aA. Instead, I observed a higher activity in the G9aA-G9aA compared to G9a-G9a which may suggest the G9aA-G9aA prep had a higher specific activity than G9a-G9a. Overall, this study suggested product recognition stimulation could not be robustly performed under these assay conditions.

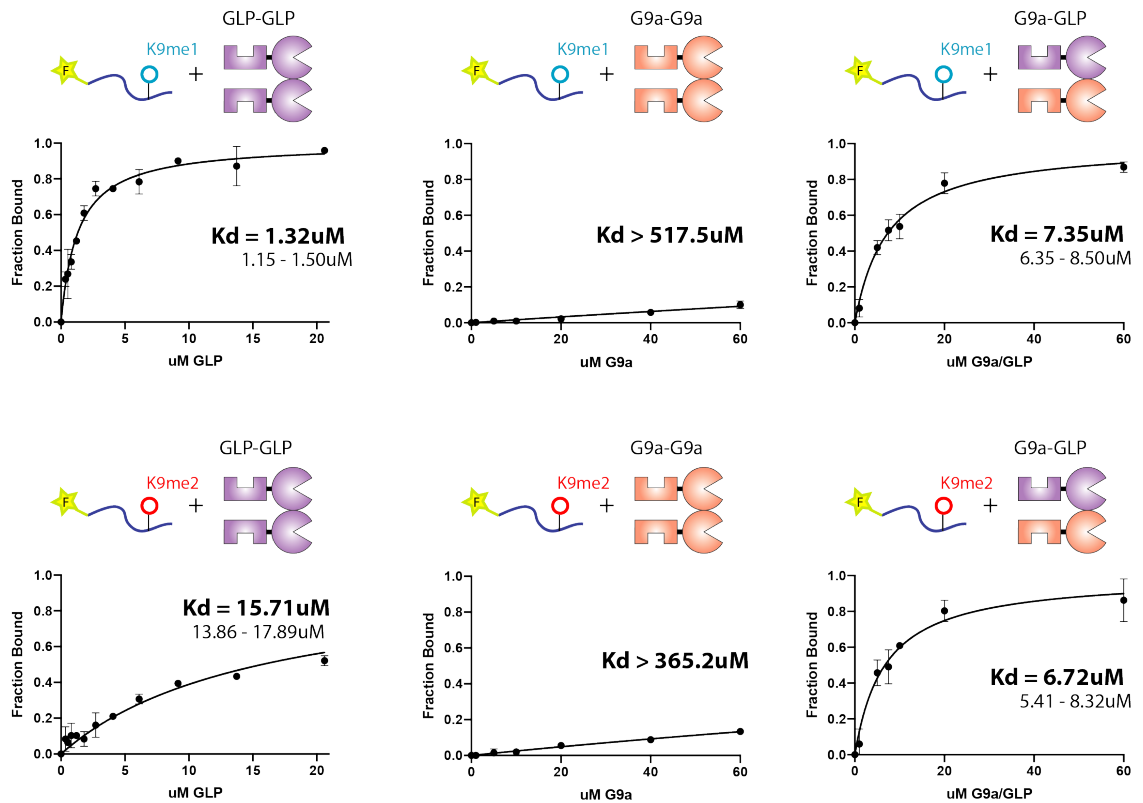


Figure 4.1: Heterodimerization facilitates G9a and GLP binding to H3K9 methyl peptides. Fluorescence Polarization measuring binding of H3K9 mono and dimethyl peptides to **A.&B.** GLP homodimers; **C.&D.** G9a homodimers; **E.&F.** G9a-GLP heterodimers. K_d values are indicated on plots. The 95% confidence interval (CI) is shown for fits with significant binding saturation (A., E., F). For curves with limited saturation and fits with wide range of K_d values, the lower bound of the 95% CI is shown.

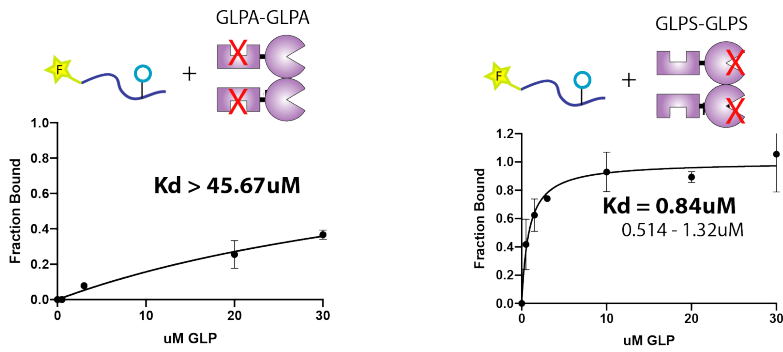


Figure 4.2 Binding to ANK repeat is measured with fluorescence polarization Point mutations were made in GLP ANK repeat (GLPA, Left) or GLP SET domain (GLPS, Right) abrogating binding in either domain. Low affinity binding of a H3K9me1 peptide was observed in GLPA condition while affinity similar to WT GLP was observed with GLPS condition. I conclude that peptide-binding measurements of the WT GLP are reflective of binding in the ANK domain.

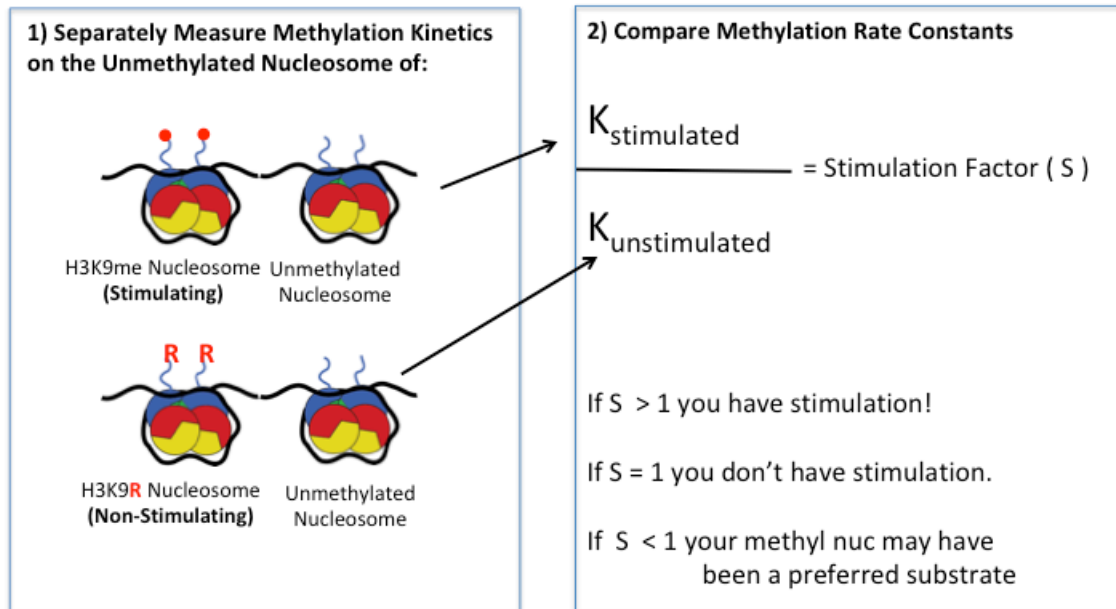


Figure 4.3 How to measure product recognition stimulation To measure product recognition stimulation *in cis*, first assemble two chromatin types (left): 1) one that contains a mixture of methylated and unmethylated nucleosomes and 2) one that contains unmethylatable (non-stimulating, e.g. H3K9R) and unmethylated nucleosomes. These chromatin types can be dinucleosomes or nucleosome arrays. Measure methylation kinetics of both substrates in solution with your methyltransferase of choice and extract a rate constant, K . This K value could be a k_{cat} value, the rate constant of a single time trace, or simply the signal from an end point assay. Compare K values from methylating the stimulating substrate and non-stimulating substrate (right). If the quotient of K_{stim} over K_{unstim} is greater than 1, you've observed product recognition stimulation.

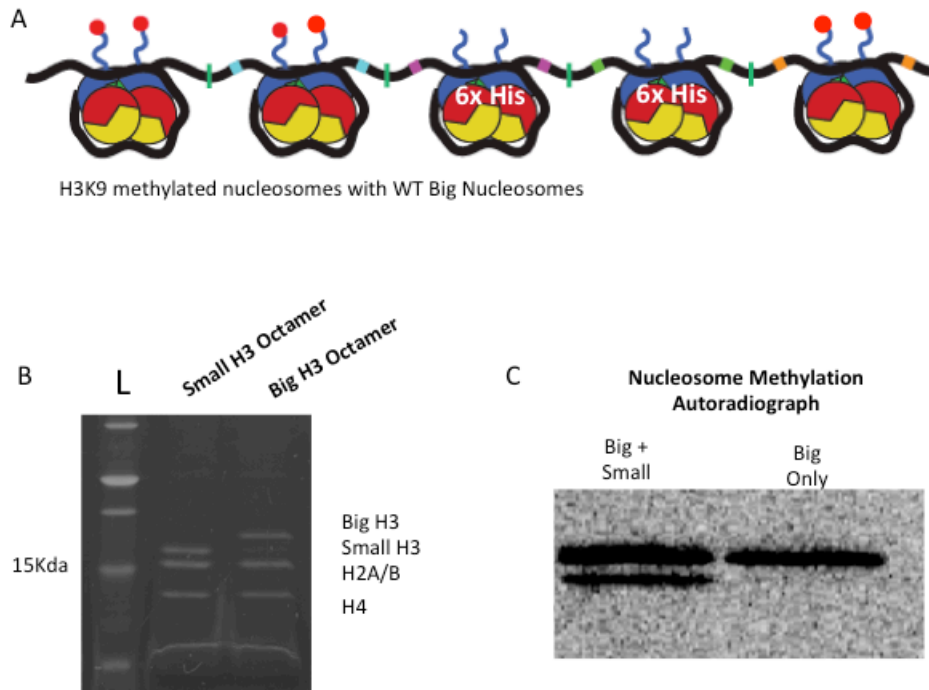


Figure 4.4 Product Recognition Stimulation Assay Setup: A) Example array used in this study. Stimulating arrays contained a random assortment of H3K9me2 and HIS:H3K9me0 nucleosomes while control arrays contained random assortments of H3K9R and HIS:H3K9me0 nucleosomes. B) 18% SDS PAGE distinction of histone octamers containing WT H3 (small H3 Octamer) and HIS:H3 (big H3 Octamer). C) Example Autoradiograph showing methylation of an array containing Big and Small nucleosomes (left) or only Big nucleosomes (right).

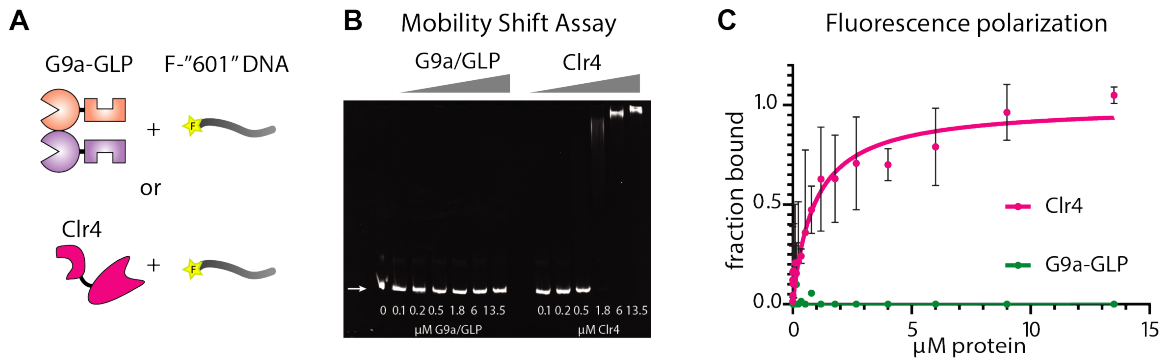


Figure 4.5: G9a-GLP does not bind DNA. **A.** G9a-GLP heterodimer or the monomeric fission yeast Clr4 H3K9 methyltransferase were assessed for binding the fluorescein-labeled 147bp "601" mononucleosome DNA template. **B.** Electro-Mobility shift Assay with G9a-GLP or Clr4 at indicated concentrations. **C.** Fluorescence polarization assay with G9a-GLP or Clr4 at indicated concentrations.

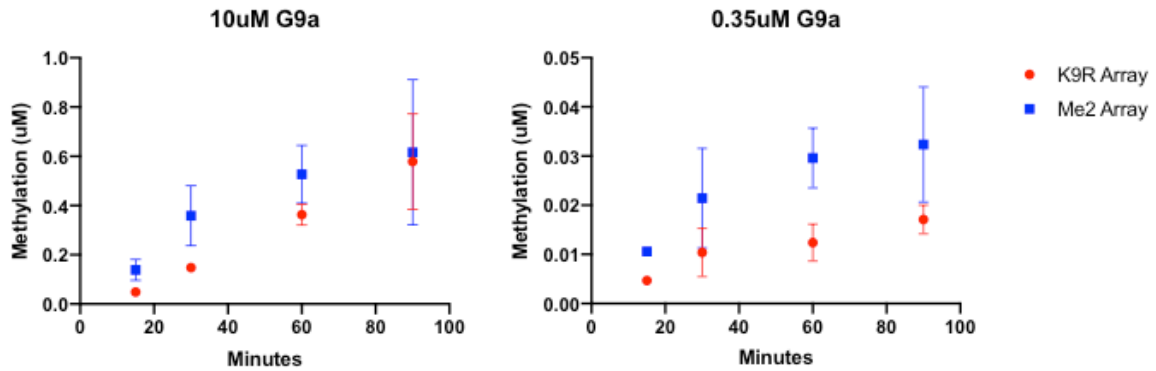


Figure 4.6 Determining enzyme conditions for product recognition stimulation: Mixed HIS:WTH3 arrays containing either H3K9me2 (blue) or H3K9R (red) were methylated by G9a-G9a at two concentrations G9a-G9a. Differences in methylation rate were observed at 0.35uM G9a-G9a (right) and not at 10uM G9a-G9a (left). Observing product recognition stimulation at the low enzyme concentration led us to hypothesize stimulation is dependent on the K_m more than the K_{cat} of G9a-G9a.

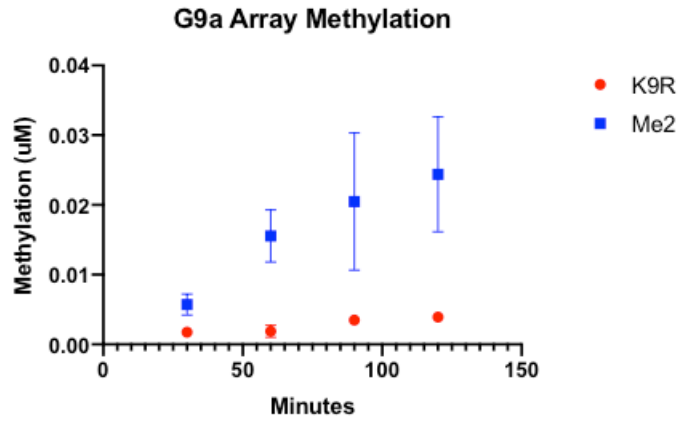


Figure 4.7 Reaffirming G9a-G9a product recognition result Mixed HIS:WTH3 with H3K9me2 (blue) and H3K9R arrays (red) were methylated with G9a-G9a at 0.35uM concentration. Product recognition was observed again in this condition.

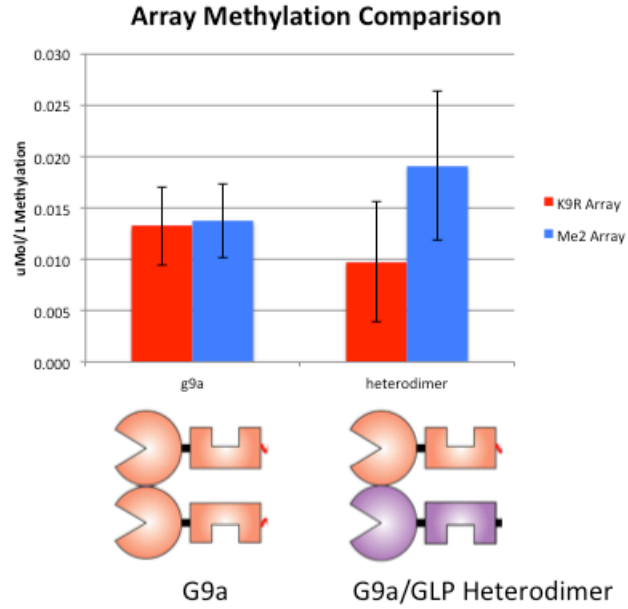


Figure 4.8 Comparison of G9a product recognition stimulation to that of G9a-GLP: Mixed HIS:WTH3 with H3K9me2 (blue) and H3K9R arrays (red) were methylated with G9a-G9a at 0.35uM concentration. Signal was measured at an endpoint of 60 min. Product Recognition Stimulation was observed in the G9a-GLP context, but not G9a-G9a.

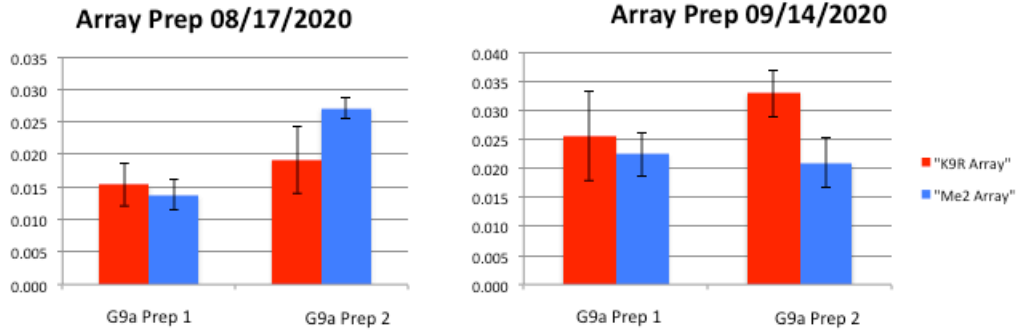


Figure 4.9 Troubleshooting G9a product recognition stimulation: Compared two fresh preps of G9a and two preps of arrays. Measured methylation at a 60 minute time point. Stimulation was not appreciably measured in any context.

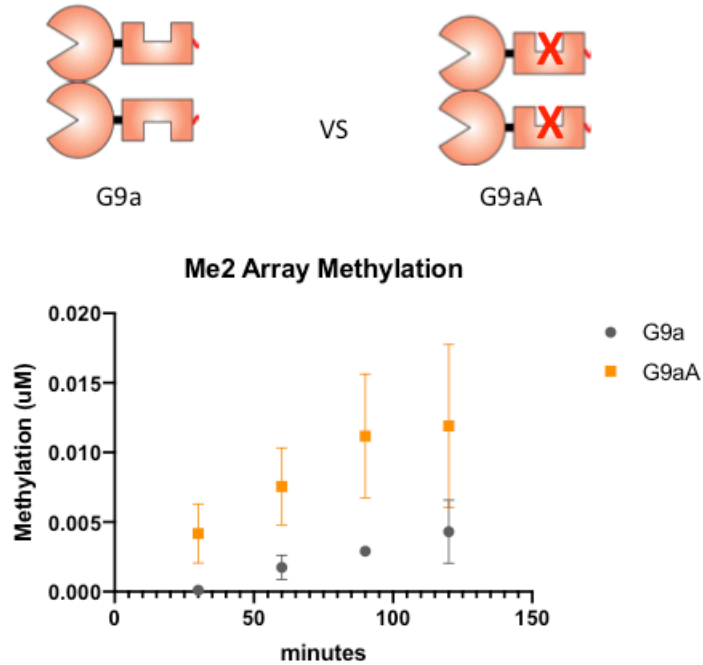


Figure 4.10 Alternative way to measure Product recognition stimulation: Methylation activity was compared between G9a-G9a (grey) and G9aA-G9aA (yellow) on an array containing HIS:WTH3 and H3K9me2 nucleosomes. Higher methylation activity was observed in G9aA-G9aA context. This may have been due to specific activity differences between preps.

4.3 Discussion

Our data suggest that H3K9me reading, in particular H3K9me2 is inhibited within the ANK-SET homodimer. As far as H3K9me reading, only the heterodimer appears capable of significant H3K9me2 recognition (**Figure 2.1**), while both GLP-GLP and G9a-GLP can recognize H3K9me1. This differs from findings on the presumably monomeric G9a and GLP ANK domains alone as measured by fluorescence polarization (Collins *et al.*, 2008) or ITC (Liu *et al.*, 2015). Because the affinities I measure for H3K9me1/2 binding, where observed, are within 2-fold of these published affinities for ANK alone, I do not believe that the ANK domains in our hands have lower specific activity. Instead, I interpret our data to indicate that either 1. inhibitory contacts between ANK and SET, or ANK and ANK across the two dimers are alleviated in G9a-GLP or 2. The heterodimerization interface induces a conformational change within the ANK. Beyond these interactions in the “naïve” state where the protein has not engaged substrate, H3K9me ANK binding sites may be initially unavailable in the naïve protein but become fully available following initial nucleosome engagement. This could account for the observation of H3K9me1 or me2 specific stimulation of methylation by G9a and GLP homodimers on chromatin circles (Liu *et al.*, 2015).

In the end, I was not able to reliably reproduce product recognition stimulation with G9a-G9a. The conditions I used reflect those used in the previous study where stimulation was first observed (Liu *et al.*, 2015) with the following differences 1) my chromatin substrate was a linear array instead of a circular plasmid array, 2) my buffer contained 100mM KCl, where 20mM KCl was used previously, 3) I had stoichiometric amounts of free DNA in solution with my array substrates that were remnants of the array DNA prep. When I designed this assay, I assumed these factors would not affect whether or not stimulation would occur given that 1) other

methyltransferases stimulate with linear arrays (Müller *et al.*, 2016), 2) my buffer is compatible with mononucleosome methylation (section 3 of this work), and 3) G9a-GLP does not interact with DNA (**Figure 4.5**). The ANKSET truncated G9a-G9a used in this study is the same construct used previously to measure product recognition stimulation. In my study, I observed G9a-G9a to have no measureable affinity for H3K9me2 (**Figure 4.1 E**). In some contexts, however, I was still able to see it's increased activity on H3K9me2 arrays over H3K9R arrays (**Figure 4.6, 4.7**). It's possible that G9a-G9a interacts with H3K9me2 differently on the peptide than in the chromatin context, though further studies are required to address this possibility. Overall, product recognition stimulation in G9a-GLP does not appear to be a very robust process and more experiments are needed to exactly identify the intricacies and nuances of this process.

4.4 Methods

Fluorescence polarization.

Polarization assays with fluoresceinated H3₁₋₂₀ K9me1 or K9me2 peptides, 147bp 601 DNA template, and core nucleosomes were performed and K_d values estimated as described (Canzio *et al.*, 2011) with the following modifications: The reaction buffer was 100mM Tris pH 8, 100mM KCl, 1mM MgCl₂, 20uM ZnSO₄, 10mM BME, 0.1% NP-40. The concentration of peptides and nucleosomes was 150nM and 200nM for DNA. Polarization measurements were conducted on a Biotek Cytation 5 (peptide and nucleosome) or Molecular Devices Spectramax (DNA) plate reader with low volume plates (Corning) in either 10 μ l (peptide, nucleosome) or 40 μ l (DNA) total volume.

Electromobility Shift Assay.

To assess binding to 601 DNA via EMSA, G9a-GLP or Clr4 were incubation with 200nM of a 147bp 601 DNA template in 50mM HEPES pH 7.5, 100mM KCl, 10% glycerol, 0.1% NP-40 in a 10 μ L volume. Upon mixing, 50% glycerol was added to 5% final and samples were run on a 5% Tris-Glycine Native Gel. Samples were imaged using Chemidoc MP Imaging System (Biorad).

Preparation of nucleosome array substrates.

Array DNA was prepared via EcoRV digest of a plasmid containing 12 tandem 601 sites. Histone proteins and nucleosome arrays were purified as described (Canzio *et al.*, 2011) with the following modifications: Assemblies contained equal stoichiometric amounts of HIS:H3K9me0 octamers and either H3K9me or H3K9R octamers. Following assembly

nucleosomes were dialyzed overnight into G9a/GLP reaction buffer (below). Dialyzed samples were concentrated using 10kda Millipore Sigma Amicon Ultra Centrifugal Filter Units.

Methylation Assays.

All methylation reactions used a tritium-based assay to detect the methylated product. Substrate peptides or nucleosomes were mixed in a solution containing 9mM tritiated S-Adenosyl Methionine (SAM, Perkin Elmer) cofactor, and reactions were initiated upon addition of enzyme. Due to signal limitations, reactions were not supplemented with SAM. Reactions were run in 100mM Tris pH 8, 100mM KCl, 1mM MgCl₂, 20mM ZnSO₄, 10mM BME and quenched with laemmli buffer. Array methylation reactions were read out via autoradiography. Proteins were separated on a 18% SDS PAGE gel which was dried and exposed to a GE Phosphor screen for 72 hours along with a standard curve of tritiated SAM spotted on Whatman paper, and imaged on a STORM imager. Images were quantified using Image Quant software (Cytiva).

4.5 References

Al-Sady, B., Madhani, H. D. and Narlikar, G. J. (2013) 'Division of labor between the chromodomains of HP1 and Suv39 methylase enables coordination of heterochromatin spread', *Molecular Cell*. Elsevier Inc., 51(1), pp. 80–91. doi: 10.1016/j.molcel.2013.06.013.

Audergon, P. N. C. B. *et al.* (2015) 'Restricted epigenetic inheritance of H3K9 methylation.', *Science (New York, NY)*, 348(6230), pp. 132–135. doi: 10.1126/science.1260638.

Bian, C., Chen, Q. and Yu, X. (2015) 'The zinc finger proteins ZNF644 and WIZ regulate the G9A/GLP complex for gene repression', *eLife*, 2015(4), pp. 1–17. doi: 10.7554/eLife.05606.

Canzio, D. *et al.* (2011) 'Chromodomain-mediated oligomerization of HP1 suggests a nucleosome-bridging mechanism for heterochromatin assembly', *Molecular Cell*. Elsevier Inc., 41(1), pp. 67–81. doi: 10.1016/j.molcel.2010.12.016.

Collins, R. E. *et al.* (2005) 'In Vitro and in Vivo Analyses of a Phe / Tyr Switch Controlling Product Specificity of Histone Lysine Methyltransferases *', 280(7), pp. 5563–5570. doi: 10.1074/jbc.M410483200.

Collins, R. E. *et al.* (2008) ‘The ankyrin repeats of G9a and GLP histone methyltransferases are mono- and dimethyllysine binding modules.’, *Nature structural & molecular biology*, 15(3), pp. 245–50. doi: 10.1038/nsmb.1384.

Dodd, I. B. *et al.* (2007) ‘Theoretical Analysis of Epigenetic Cell Memory by Nucleosome Modification’, *Cell*, 129(4), pp. 813–822. doi: 10.1016/j.cell.2007.02.053.

Dodd, I. B. and Sneppen, K. (2011) ‘Barriers and silencers: A theoretical toolkit for control and containment of nucleosome-based epigenetic states’, *Journal of Molecular Biology*. Elsevier Ltd, 414(4), pp. 624–637. doi: 10.1016/j.jmb.2011.10.027.

Erdel, F. and Greene, E. C. (2016) ‘Generalized nucleation and looping model for epigenetic memory of histone modifications’. doi: 10.1073/pnas.1605862113.

Gibson, B. A. *et al.* (2019) ‘Organization of Chromatin by Intrinsic and Regulated Phase Separation’, *Cell*. Elsevier Inc., 179(2), pp. 470-484.e21. doi: 10.1016/j.cell.2019.08.037.

Hall, I. M. *et al.* (2002) ‘Establishment and maintenance of a heterochromatin domain.’, *Science (New York, N.Y.)*, 297(5590), pp. 2232–7. doi: 10.1126/science.1076466.

Hathaway, N. A. *et al.* (2012) ‘Dynamics and memory of heterochromatin in living cells’, *Cell*. Elsevier Inc., 149(7), pp. 1447–1460. doi: 10.1016/j.cell.2012.03.052.

Hodges, C. and Crabtree, G. R. (2012) ‘Dynamics of inherently bounded histone modification domains’, *Proceedings of the National Academy of Sciences*, 109(33), pp. 13296–13301. doi: 10.1073/pnas.1211172109.

Holoch, D., Moazed, D. and Avenue, L. (2015) ‘RNA-mediated epigenetic regulation of gene expression’, *Nature genetics*, 16(2), pp. 71–84. doi: 10.1038/nrg3863.RNA-mediated.

Jehanno, C. *et al.* (2017) ‘Biochimica et Biophysica Acta A model of dynamic stability of H3K9me3 heterochromatin to explain the resistance to reprogramming of differentiated cells’, *BBA - Gene Regulatory Mechanisms*. Elsevier B.V., 1860(2), pp. 184–195. doi: 10.1016/j.bbagr.2016.11.006.

Lee, C. *et al.* (2018) ‘Allosteric Activation Dictates PRC2 Activity Independent of Its Recruitment to Chromatin Article Allosteric Activation Dictates PRC2 Activity Independent of Its Recruitment to Chromatin’, *Molecular Cell*. Elsevier Inc., 70(3), pp. 422-434.e6. doi: 10.1016/j.molcel.2018.03.020.

Liu, N. *et al.* (2015) ‘Recognition of H3K9 methylation by GLP is required for efficient establishment of H3K9 methylation, rapid target gene repression, and mouse viability’, *Genes and Development*, 29(4), pp. 379–393. doi: 10.1101/gad.254425.114.

Lowary, P. T. and Widom, J. (1998) 'New DNA sequence rules for high affinity binding to histone octamer and sequence-directed nucleosome positioning', *Journal of Molecular Biology*, 276(1), pp. 19–42. doi: 10.1006/jmbi.1997.1494.

Margueron, R. *et al.* (2009) 'Role of the polycomb protein EED in the propagation of repressive histone marks.', *Nature*, 461(7265), pp. 762–767. doi: 10.1038/nature08398.

Müller, M. M. *et al.* (2016) 'A two-state activation mechanism controls the histone methyltransferase Suv39h1', *Nature Chemical Biology*, 12(3), pp. 188–193. doi: 10.1038/nchembio.2008.

Poepsel, S., Kasinath, V. and Nogales, E. (2018a) 'Cryo-EM structures of PRC2 simultaneously engaged with two functionally distinct nucleosomes', *Nature Structural and Molecular Biology*. Springer US, 25(2), pp. 154–162. doi: 10.1038/s41594-018-0023-y.

Poepsel, S., Kasinath, V. and Nogales, E. (2018b) 'Cryo-EM structures of PRC2 simultaneously engaged with two functionally distinct nucleosomes', *Nature Structural and Molecular Biology*, 25(2), pp. 154–162. doi: 10.1038/s41594-018-0023-y.

Ragunathan, K., Jih, G. and Moazed, D. (2015) 'Epigenetic inheritance uncoupled from sequence-specific recruitment', *Science*, 348(6230), p. science.1258699-. doi: 10.1126/science.1258699.

Simon, J. A. and Kingston, R. E. (2013) ‘Occupying Chromatin: Polycomb Mechanisms for Getting to Genomic Targets, Stopping Transcriptional Traffic, and Staying Put’, *Molecular Cell*. Elsevier Inc., 49(5), pp. 808–824. doi: 10.1016/j.molcel.2013.02.013.

Wen, B. *et al.* (2009) ‘Large histone H3 lysine 9 dimethylated chromatin blocks distinguish differentiated from embryonic stem cells.’, *Nature genetics*, 41(2), pp. 246–250. doi: 10.1038/ng.297.

Wu, H. *et al.* (2011) ‘Histone methyltransferase G9a contributes to H3K27 methylation in vivo’, *Cell Research*, 21(2), pp. 365–367. doi: 10.1038/cr.2010.157.

Publishing Agreement

It is the policy of the University to encourage open access and broad distribution of all theses, dissertations, and manuscripts. The Graduate Division will facilitate the distribution of UCSF theses, dissertations, and manuscripts to the UCSF Library for open access and distribution. UCSF will make such theses, dissertations, and manuscripts accessible to the public and will take reasonable steps to preserve these works in perpetuity.

I hereby grant the non-exclusive, perpetual right to The Regents of the University of California to reproduce, publicly display, distribute, preserve, and publish copies of my thesis, dissertation, or manuscript in any form or media, now existing or later derived, including access online for teaching, research, and public service purposes.

DocuSigned by:

Nicholas Sanchez

200CC290A8B94B4...

Author Signature

9/3/2021

Date