

Structural and biochemical insights into the assembly of the DPY-30/Ash2L heterotrimer

John Haddad

Thesis submitted to the Faculty of Graduate and
Postdoctoral Studies in partial fulfillment of the
requirements for the Masters degree in Biochemistry

Department of Biochemistry, Microbiology and
Immunology
Faculty of Medicine
University of Ottawa

© John Haddad, Ottawa, Canada, 2017

ABSTRACT

In eukaryotes, the SET1 family of methyltransferases carry out the methylation of Lysine 4 on Histone H3. Alone, these enzymes exhibit low enzymatic activity and require the presence of additional regulatory proteins, which include RbBP5, Ash2L, WDR5 and DPY-30, to stimulate their catalytic activity. While previous structural studies established the structural basis underlying the interaction between RbBP5, Ash2L and WDR5, the formation of the Ash2L/DPY-30 complex remains elusive. Here we report the crystal structure of the Ash2L/DPY-30 complex solved at 2.2Å. Our results show that a C-terminal amphipathic α -helix on Ash2L makes several hydrophobic interactions with the DPY-30 homodimer. Moreover, the structure reveals that a tryptophan residue on Ash2L, which directly precedes its C-terminal amphipathic α -helix, makes key interactions with one of DPY-30 α -helix. Finally, biochemical studies of Ash2L revealed a hitherto unknown ability of this protein to bind anionic lipids.

ACKNOWLEDGEMENTS

I would like to thank Dr Jean-François Couture for allowing me to do my masters project in his lab and immerse myself in the fascinating world of structural biology. I am thankful to Dr. Couture for helping me develop my inquisitive thinking, problem-solving skills, and numerous other invaluable skills. Our many discussions in his office during the course of my project always kept the science interesting and stimulated my interest from beginning until end. I would also like to thank Véronique Tremblay who mentored me from the start of my masters and was always there to assist me on a technical level with my experiments. I owe it to Dr. Couture and to Véronique Tremblay for making me the scientist I am today!

I am also very grateful to have had the chance to do my studies alongside the students of the Couture lab. Sabina Sarvan and Pamela Zhang were great role models, who inspired me to push myself to do great science and offered so much guidance and friendship. I would like to thank Elisa Bergamin, my favorite scientist of all times, for continuously encouraging me, for seeing the best in me, for the numerous times she's helped me in the lab, and for her friendship. She has truly been a great inspiration to me and has shown me the true meaning of hard work, sacrifice, and passion!

I would like to thank my TAC members Dr. Jeffrey Dilworth and Dr. John Baenziger for their help and support throughout my master's degree. Also, I feel lucky to have been surrounded by many people on the fourth floor who have been there for me during the different stages of my degree, especially the COOP and honors students from

the Couture lab, the members of the Downey lab, and Kateland Simmons (whose late arrival in my masters timeline made all the difference!). I also appreciate the hard work that the BMIGSA has done to organize the many social events (ice cream and barbecues) that allowed me to unwind during my incubation times. Lastly, I am grateful to my family members who have helped me in so many ways during my degree and facilitated the whole struggle of being a graduate student.

TABLE OF CONTENTS

Abstract.....	ii
Acknowledgements	iv
List of figures.....	viii
List of tables.....	ix
Abbreviations	x
Introduction.....	1
1. Structure of the nucleosome and chromatin	1
2. Regulation of chromatin compaction	2
2.i. ATP-dependent chromatin remodelers	2
2.ii. Histone variant exchange	3
2. iii. Post-translational modification of histones	4
3. Methylation of histones.....	5
4. Methylation of characterized arginine and lysine residues on histones.....	6
5. Methylation of Lysine 4 on histone H3	9
6. H3K4 methyltransferases.....	11
7. H3K4 demethylases.....	12
8. Effectors of H3K4 methylation	13
9. SET1 family of methyltransferases from yeast to human	14
10. Links between the SET1 family of methyltransferases, the WRAD complex and diseases	15
11. The 6 members of the SET1 methyltransferase family	18
12. The WRAD regulatory subcomplex	19
13. The WRAD complex and cancer.....	
14. The role of the subunits of the WRAD regulatory subcomplex.....	22
14.i. Role for WDR5 in SET1 methyltransferases	22
14.ii. Role for the Ash2L-RbBP5 heterodimer in SET1 methyltransferases.....	23
14.iii. Role for DPY-30 in SET1 methyltransferases	24
15. The DPY-30/Ash2L heterotrimer.....	28
16. Research goal 1: Gain structural insights into the formation of the Ash2L-DPY-30 heterotrimer.....	33
17. Functions of WRAD in the nucleus	34
18. Non-nuclear functions of WRAD	35

19. Cardiolipin.....	36
20. Structural characterization of the interactions of cardiolipin with proteins	39
21. Research goal 2: Determine whether Ash2L binds lipids in vitro.....	40
Materials and Methods.....	41
1. Cloning, Expression, and Purification of the Ash2L/DPY-30 heterotrimer.....	41
2. Crystallization of the Ash2L/DPY-30 heterotrimer.....	42
3. Antibodies	43
4. <i>In vitro</i> pull-down experiment	43
5. Tandem Flag- and GFP-co-Immunoprecipitation	44
6. Protein-lipid overlay assay	45
7. Biotin-cardiolipin pull-down.....	45
8. Lipids.....	46
9. Liposome preparation	46
10. Liposome binding assay.....	47
11. Lipid competition assay.....	48
Results	49
1. Purification of the DPY-30/Ash2L ^{SPRY+SDI} heterotrimer	49
2. Crystallization and structure determination of the DPY-30/Ash2L ^{SPRY+SDI}	50
3. Overall view of the DPY-30/Ash2L ^{SPRY+SDI} heterotrimer	53
4. Structural analysis of the interactions between DPY-30 and Ash2L	54
5. Role of Ash2L tryptophan 505 in DPY-30 binding.....	69
6. Role of the DPY-30 hydrophobic pocket in Ash2L binding.....	60
7. Ash2L/DPY-30 bind to anionic lipids.....	60
8. Ash2L has lipid-binding properties and binding to lipids and DPY-30 is mutually exclusive	69
9. A construct encompassing the SPRY and SDI domains of Ash2L is sufficient for CL-binding	73
Discussion	76
1. Ash2L/DPY-30 Heterotrimer: Insights into the orientation and binding of Ash2L to DPY-30.....	76
2. Structural and biochemical characterization of the lipid-binding function of Ash2L and its impact on the assembly of the Ash2L/DPY-30 heterotrimer	86
Conclusion	93
References.....	94

Appendix.....	116
----------------------	------------

LIST OF FIGURES

Figure 1. Structure of a SET domain	7
Figure 2. Domain organization of the members of the SET1 methyltransferase family and of the WRAD regulatory subunits	20
Figure 3. Stimulatory effect of the regulatory WRAD subunits on SET1 methyltransferase activity	26
Figure 4. The RbBP5-Ash2L heterodimer rigidifies the iSET motif on the SET domain of MLL3.	29
Figure 5. Structures of the DPY-30 D/D homodimer and the DPY-30D/D/Ash2LSDI complex.	31
Figure 6. Structure of a cardiolipin molecule.....	37
Figure 7. Purification and crystallization of the Ash2L/DPY-30 heterotrimer.	51
Figure 8. The N- and C-terminal residues on the Ash2L SDI position the SDI helix within the DPY-30 hydrophobic cleft.....	55
Figure 9. Alignment of the Ash2L sequence of different species.....	57
Figure 10. W505 is dispensable for the interaction between Ash2L and DPY-30. ...	61
Figure 11. Mutating the hydrophobic pocket on DPY-30 composed of Ala70 and Ala73 completely abrogates binding of DPY-30 to Ash2L.....	64
Figure 12. The DPY-30/Ash2L heterotrimer binds to anionic lipids in vitro.....	67
Figure 13. Ash2L binding to cardiolipin and DPY-30 is mutually exclusive.	71
Figure 14. The SPRY and SDI domains of Ash2L are important for binding to cardiolipin.....	74
Figure 15. DPY-30 is physically distant from MLL3 SET domain.	79
Figure 16. DPY-30 interactors have an aromatic/proline residue preceding their SDI-like helices.....	84

List of tables

Table 1. Antibodies used in this study.....	43
Table 2. Lipids used in this study	46
Table 3. List of p-Values for the binding of Ash2L constructs to various liposomes	115

ABBREVIATIONS

A

A	Alanine
ABM	Ash2L binding motif
ADP	Adenosine Diphosphate
AF4	ALL1-fused gene from chromosome 4 protein
AF9	ALL1-Fused Gene From Chromosome 9 Protein
Ala	Alanine
AOD	Amine oxidase domain
aP2	Adipocyte protein 2
AS	Activating segment
ASC-2	Activating Signal Cointegrator 2
ASCOM	ASC-2 complex
Ash1	Absent, Small, Homeotic 1
Ash2L	Absent, Small, Homeotic-like 2
ATP	Adenosine Triphosphate
ATXR5	Arabidopsis Trithorax-Related Protein5

B

BAP18	BPTF-associated Protein of 18 kDa
BF-1	Brain factor-1
BIG-1	Brefeldin A-inhibited guanine nucleotide-exchange Protein1
BIR	Baculovirus inhibitor of apoptosis protein repeat
BME	Beta-mercaptoethanol
BMP	Bis(Monoacylglycero)Phosphate
BPTF	Bromodomain PHD Finger Transcription Factor
BRCA	Breast Cancer Gene
BRCT	BRCA1 C Terminus
BRD	Bromodomain

C

CDP	Cytidine 5'-(trihydrogen diphosphate)
C/EBP-alpha	CCAAT/enhancer-binding protein alpha
c. Elegans	Caenorhabditis elegans
CENP-A	Centromere protein A
CFP1	CxxC finger protein 1
CH	Cholesterol
CHD	Chromodomain-helicase-DNA-binding protein
ChIP-seq	Chromatin immunoprecipitation followed by high-throughput sequencing
CIMPR	Cation-independent mannose 6-phosphate receptor
CL	Cardiolipin
COMPASS	Complex Proteins Associated with Set1
Cps	COMPASS

CUL4-DDB1 Cullin 4-Damage Specific DNA Binding Protein 1

D

D Aspartic acid
D. Danio
DAG Diacylglycerol
D-AKAP2 Dual-specific A kinase anchoring protein 2
D/D Docking and dimerization
DExx DExx box domain
DNA Deoxyribonucleic acid
DNMT3A DNA Methyltransferase 3 Alpha
Dot1 Disruptor of Telomeric Silencing-1
DPY-30 Dumpy-30

E

ECL Enhanced Chemiluminescence
E. Escherichia
EF-Tu Elongation factor thermo unstable
eGFP Enhanced Green fluorescent protein
E Glutamic acid
EM Electron microscopy
ENL Eleven-nineteen-leukemia
ER Endoplasmic reticulum
ESC Embryonic stem cells

F

FL Full-length
FXR Farnesyl X Receptor
FYRC F/Y-rich C-terminal
FYRN F/Y-rich N-terminal

G

G. Gallus
GFP Green fluorescent protein
Gly Glycine
GST Glutathione-S-transferase

H

H Histidine
H. Homo
H2AR3 Arginine 3 on histone H2A
H3R2 Arginine 2 on histone H3
H3T3 Threonine 3 on histone H3
H3K4 Lysine 4 on histone H3
H3K4me1 Mono-methylated H3K4
H3K4me2 Di-methylated H3K4

H3K4me3	Tri-methylated H3K4
H3R8	Arginine 8 on histone H3
H3K9	Lysine 9 on histone H3
H3K9me3	Tri-methylated H3K9
H3S10	Serine 10 on histone H3
H3R17	Arginine 17 on histone H3
H3K27	Lysine 27 on histone H3
H3S28	Serine 28 on histone H3
H3K36	Lysine 36 on histone H3
H3K36me3	Tri-methylated H3K36
H3K79	Lysine 79 on histone H3
H3K79me2	Di-methylated H3K79
H4R3	Arginine 3 on histone H4
H4K20	Lysine 20 on histone H4
HBO1	Histone acetyltransferase binding to ORC 1
HDAC	Histone deacetylase
HEK293	Human Embryonic Kidney 293
HeLa	Henrietta Lacks
HELICc	Helicase superfamily c-terminal domain
HMG	High-mobility group
HOTTIP	HOXA transcript at the distal tip
HOX	Homeobox
HRP	Horseradish peroxidase
HSP90	Heat shock protein 90
HSPA1A	Heat Shock Protein A1A

I

ING	Inhibitor of Growth
INO80	Inositol requiring 80
ISWI	Imitation SWI

J

JARID	Jumonji AT-rich interactive domain
JMJD2A	Jumonji domain 2A

K

KANSL	KAT8 Regulatory NSL Complex Subunit
KIF4	Kinesin Family Member 4
KLF10	Kruppel Like Factor 10

L

L	Leucine
LB	Luria-Bertani
Leu	Leucine
LSD	Lysine-specific demethylase
LXR	Liver X Receptor

Lys	Lysine
M	
MDC1	Mediator of DNA damage checkpoint protein 1
Mef2d	Myocyte Enhancer Factor 2D
Meisetz	Meiosis-induced factor containing PR/SET domain and zinc-finger
MKLP1	Mitotic kinesin-like protein 1
MLL	Mixed-lineage leukemia
MPD	2-Methyl-2,4-pentanediol
M.	Mus
MYC	Avian myelocytomatosis virus oncogene
N	
NCOA6	Nuclear Receptor Coactivator 6
NFY	Nuclear transcription factor Y
Nlrp3	NLR Family Pyrin Domain Containing 3
NURF	Nucleosome Remodelling Factor
NT2-D1	NTERA-2 clone D1
O	
Oct4	Octamer-binding transcription factor 4
P	
PA	Phosphatidic acid
PAF	Polymerase-associated factor
PAGE	Polyacrylamide gel electrophoresis
PAX	Paired box gene
PBS	Phosphate buffer saline
PC	Phosphatidylcholine
PEG	Polyethylene glycol
PG	Phosphatidylglycerol
PHD	Plant homeodomain
PHF20	PHD Finger Protein 20
PI	Phosphatidylinositol
PtdIns(4)P	Phosphatidylinositol-4-phosphate
PtdIns(4,5)P2	Phosphatidylinositol-4,5-biphosphate
PtdIns(3,4,5)P3	Phosphatidylinositol-3,4,5-triphosphate
PPAR γ	Peroxisome proliferator-activated receptor gamma
PRC	Polycomb repressive complex
PRDM	PRDI-BF1 and RIZ homology domain containing
PS	Phosphatidylserine
PTIP	PAX transcription activation domain interacting protein
PTM	Post-translational modification
Q	
Q	Glutamine

R

RI α	Protein Kinase A regulatory subunit I alpha
RII α	Protein Kinase A regulatory subunit II alpha
RARE	Retinoic acid response element
RbBP5	Retinoblastoma-Binding Protein 5
RIAM	Rap1-GTP-interacting adaptor molecule
RIPA	Radioimmunoprecipitation assay
ROS	Reactive oxygen species
RNA	Ribonucleic acid
RRM	RNA recognition motif

S

S.(cerevisiae)	Saccharomyces
S.(scrofa)	Sus
SAM	S-Adenosyl Methionine
SAH	S-Adenosyl Homocysteine
SDS	Sodium dodecyl sulfate
SDI	Sdc1 interacting
SET	Suppressor of variegation 3-9, Enhancer of zeste, Trithorax
SETMAR	SET Domain And Mariner Transposase Fusion
Sin3A	Swi-independent 3a
SLB	Sample-loading buffer
SM	Sphingomyelin
SMT3	Suppressor of Mif Two 3
SMYD	SET and MYND domain-containing
SOX2	SRY (sex determining region Y)-box 2
SPRY	SPia and Ryanodine
SU	Sulfatide
Suv4-20	Suppressor of variegation 4-20
SWI/SNF	SWItch/Sucrose Non-Fermentable

T

TEV	Tobacco etch virus
TG	Triglycerides
TGN	Trans-golgi network
Thr	Threonine
TIEG1	TGF β -inducible early gene-1
Trx	Trithorax
Trr	Trithorax-related

U

ULP1	Ubiquitin-like specific protease 1
UTX	Ubiquitously transcribed tetratricopeptide repeat, X chromosome
UV	Ultraviolet

V	
VISA	Virus-induced signaling adapter
Vs	Versus
W	
W	Tryptophan
WBM	WDR5 binding motif
WD40	Tryptophan-aspartate repeat
WDR5	WD40-containing Repeats 5
WDR82	WD Repeat Domain 82
WHW	Wing-Helix-Wing
Win	WDR5 interacting
WPPW	Tryptophan-proline-proline-tryptophan
WRAD	WDR5-RbBP5-Ash2L-DPY30
Wt/WT	Wild-type
Y	
Y	Tyrosine

INTRODUCTION:

1. Structure of the nucleosome and chromatin:

The genetic material of a typical eukaryotic cell approximately measures two meters and must be restricted to the confines of the nucleus. The cell employs four α -helical basic proteins to create a scaffold around which DNA can be compacted: histones H2A, H2B, H3, and H4. First, two histone H3-H4 heterodimers dimerize to form a heterotetramer, upon which two H2A-H2B heterodimers will bind. The H2A protomers contact H3 and H4 at the extremities of the heterotetramer, meanwhile the H2B protomers form an extensive dimerization interface (Arents et al., 1991) to form a symmetrical disk-shaped histone octamer. A DNA fragment of approximately 150bp will then wrap twice around the histone octamer of basic histone proteins to form a repetitive structure known as the nucleosome (Noll, 1977, Luger et al., 1997). This first level of compaction yields a “beads on a string” type structure in which nucleosomes are interspersed with segments of DNA (Olins and Olins, 1974). A linker histone, namely histone H1, binds to DNA at the entry or exit of every nucleosome (Zhou et al., 1998). Nucleosomes can then be positioned in a zigzag fashion to further compact chromatin into a 30 nm fiber (Oudet et al., 1975, Bednar et al., 1998). This 30nm fiber can undergo additional condensation steps to generate more complex chromatin structures to ultimately form large macromolecular structures termed chromosomes (Zhu and Li, 2016).

2. Regulation of chromatin compaction:

A variety of nuclear proteins, such as DNA damage repair factors, transcription factors, and the transcriptional and replication machinery must access regions of DNA during various cellular processes. At other times, these regions must be concealed in order to prevent the inappropriate transcription of genes or non-coding RNAs. Accordingly, DNA can be unwound from nucleosomes or nucleosomal arrays can be condensed to modulate gene expression (Luger et al., 2012). Three major mechanisms are employed by cells to control the openness of chromatin: ATP-dependent chromatin remodeling, exchange of canonical histones for histone variants, and post-translational modification of histones (PTM).

i. ATP-dependent chromatin remodelers:

Chromatin remodelers utilize the energy from the hydrolysis of an ATP molecule to modulate the nucleosomal architecture. Four major families of chromatin remodelers are known: SWI/SNF, ISWI, CHD, and INO80 (Längst et al., 2015). They harbor a shared ATPase domain, which is formed by the tandem Dextx and HELICc folds (Eisen et al., 1995). These enzymes carry additional domains conferring them a nucleosome/internucleosomal DNA-binding capacity and domains that recognize precise epigenetics marks (Tang et al., 2010). Moreover, a variety of complexes can be assembled around these catalytic subunits, which attributes them with a specialized chromatin-remodeling function (Längst et al., 2015). Chromatin remodelers may loosen

the DNA around the histone octamer to either slide nucleosomes in a precise direction, or to insert/evict nucleosomes from a genomic region (Logie et al., 1999, Tsukiyama and Wu, 1995, Varga-Weisz et al., 1997). Collectively, these actions locally impact the structure of chromatin and can act globally to orchestrate the assembly of chromosomes.

ii. Histone variant exchange

In the cases of histones H2A, H2B, and H3, specific variants, which differ in amino acid sequence or domain architecture (Bernstein et al., 2006) impinge a specific activity to chromatin by altering the structure of the histone octamer. For example, the insertion of histone H2A.Z into the nucleosome disrupts the interactions between the H2A/H2B dimers and the H3/H4 tetramer (Suto et al. 2000). Exchanging canonical histones for histone variants can also affect gene expression and mark transcriptional compartments. For example, histone H3 variants H3.1/H3.2 and H3.3 are associated with silenced chromatin, (heterochromatin) (Hake et al., 2006) and active chromatin (euchromatin) (Jin and Felsenfeld 2006; Wirbelauer et al. 2005), respectively. Histone variants can also serve as markers of chromosomal structures/regions. This is exemplified by the predominant localization of the histone H3 variant CENP-A at the centromeres of chromosomes (Shelby et al. 1997, Blower and Karpen 2001) and the deposition of the H2A variant H2A.X near DNA damaged sites (Rogakou et al., 1998, Rogakou et al., 1999). Similarly, the histone variant macroH2A is deposited on the silenced X chromosome and is important for dosage compensation (Chadwick and Willard 2001). Alternatively, deposition of histone variants may directly impact the post-transcriptional

modification of histones. For example, due to a single amino acid difference, ATXR5 methylates H3K27 on histone H3.1 but not H3.3 (Jacob et al., 2014).

iii. Post-translational modification of histones:

Histones can be modified on their N-terminal protruding tails, in their structured core fold, or at their C-termini (Cosgrove et al., 2004). These modifications include acetylation, methylation, phosphorylation, and ubiquitination and less studied modifications, such as ADP-ribosylation, glycosylation, and formylation (Rothbart et al., 2014). Acetylation of lysine residues and phosphorylation of threonine and serine residues can weaken interactions between positively charged histone proteins and negatively charged DNA. The acetyl group neutralizes the positive charge of lysine residues and abolishes electrostatic histone-DNA contacts (Hong et al., 1993); meanwhile phosphate groups create electrostatic repulsion with the phosphate groups on DNA (Banerjee et al., 2011). Specialized domains, referred to as bromodomains, can recognize acetylated lysine residues and are recruited to acetylated chromatin to accomplish a specific function (Dhalluin et al., 1999). Akin to bromodomains, histone phosphorylation can serve as a binding epitope for a protein. The BRCT domain of MDC1 recognizes H2A.X phosphorylated at residue Ser139 (Stucki et al., 2005) while the BIR domain of Survivin can bind to phosphorylated H3T3 (Jeyaprabakash et al., 2011). Similarly, 14-3-3 ζ interacts with the phosphorylated forms of both H3S10 and H3S28 (MacDonald et al., 2005). Phosphorylation of specific histone residues has also been shown to interfere with the recognition of adjacent modified residues by their epigenetic readers (Zentner et al.,

2013). For instance, the phosphorylation of Ser10 on histone H3 during mitosis prevents H3K9 methylation by the methyltransferase G9a (Duan et al., 2008).

3. Methylation of Histones:

Methylation of histone proteins was initially identified by Murray and colleagues after observing the accumulation of radioactive material on a family of basic proteins following the treatment of various tissues of rabbits with C¹⁴-labeled methionine (Murray et al., 1964). Since then, many other proteins were found to be methylated, the first ones being EF-Tu, Rubisco, citrate synthase, cytochrome c, calmodulin (Lanouette et al., 2014). The ε-amine of lysine residues on histones can be methylated up to three times (Paik and Kim, 1967, Takemoto et al., 1964), whereas arginine residues on histones can be mono- or di-methylated, in a symmetrical or an asymmetrical fashion (Kakimoto and Akazawa, 1970). Following Murray et al.'s groundbreaking discovery, Paik et al. used various biochemical approaches to identify protein methylase I (Paik and Kim, 1968) and protein methylase III (Paik and Kim, 1970); two proteins that were ultimately shown to have arginine and lysine methyltransferase activity, respectively. In the case of lysine methylation, the reaction is primarily carried out by a family of methyltransferases harboring a catalytic domain referred to as **Suppressor of variegation, Enhancer of zeste, Trithorax (SET)** (Jenuwein, 2006). Exceptionally, the seven β-strand methyltransferase Dot1, an H3K79-specific methyltransferase, and the Prdm family, which have a catalytic PR domain at their N-termini, do not possess a SET domain (van Leeuwen et al., 2002, Hohenauer and Moore, 2012). Regardless of the structural differences between enzymes,

they all catalyze the transfer of an electrophilic methyl group from S-Adenosyl Methionine (SAM) cofactor to the ϵ -amine of a lysine residue (Dillon et al., 2005). Structural studies of SET domain histone lysine methyltransferases showed that these enzymes have a unique fold wherein the SET domain consists of 12 β -strands interspersed with large loops. In most SET domain enzymes, the SET domain is interrupted by a regulatory α -helix, termed the iSET region. A highly variable N-terminal portion (nSET) precedes the SET domain of all methyltransferases and is believed to participate in the structural stability of the SET domain (Wilson et al., 2002, Trievel et al., 2002). The SET domain is succeeded by a C-terminal region, referred to as cSET, that makes important interactions with the histone substrate and the cofactor (Couture et al., 2006) (**figure 1**).

4. Methylation of arginine and lysine residues on histone proteins:

Several methylated lysine and arginine residues on histone proteins have been studied in the context of their links to transcriptional regulation. The asymmetrical di-methylation of H3R17, the symmetrical di-methylation of H3R2, and the asymmetrical di-methylation of H4R3me2 and H2AR3me2 are found in the promoter region of highly transcribed genes. Conversely, asymmetrically di-methylated H3R2, symmetrically di-

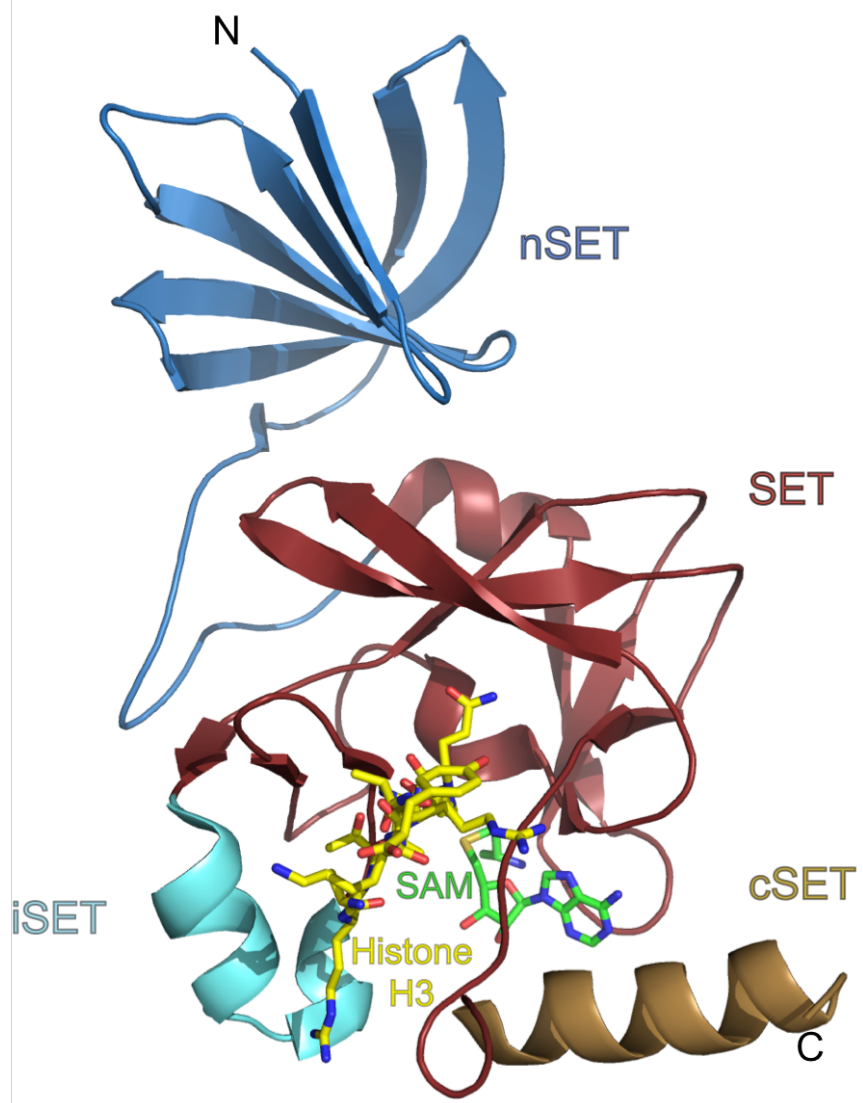


Figure 1. Structure of a SET domain. Crystal structure of the SET domain of the SET7/9 methyltransferase represented in cartoon diagram showing the nSET (blue), the conserved SET fold (red), the iSET motif (cyan), and cSET (sand) regions. The N-terminus and C-terminus are labeled.

methyated H3R8, and symmetrically di-methyated H4R3 and H2AR3 occupy promoter regions of transcriptionally repressed genes (Lorenzo and Bedford, 2011). The methylation of H3K9 and H4K20 residues is linked to the formation of constitutive heterochromatin while H3K9me3, a mark deposited by Suv39H1 and G9a enzymes, recruits the heterochromatin protein 1 (Lachner et al., 2001) which in turn is bound by the H4K20 di- and tri-methyltransferases Suv20H1 and Suv20H2 (Hahn et al., 2013) and other heterochromatin promoting factors. H3K27 methylation has a repressive effect on gene transcription related to X-chromosome inactivation as well as imprinted genes and linked to the regulation of poised genes during development (Bernstein et al., 2006, Cao and Zhang, 2004). On the other hand, methylation of H3K4, H3K36, and H3K79 is found within euchromatin or regions with high transcriptional activity (Schübeler et al., 2004, Wagner et al., 2012). Consistent with their occurrence in actively transcribed regions, the H3K36 methyltransferase Set2 and the H3K79 methyltransferases Dot1 are able to associate with the elongating form of RNA polymerase II, and these PTMs are enriched in gene bodies (Li et al., 2002, Krogan et al., 2003). Overall, these observations demonstrate that histone methylation is linked to a myriad of nuclear transactions.

5. Methylation of Lysine 4 on Histone H3:

Histone H3K4 methylation was initially characterized in *tetrahymena thermophila*. Isolation of the transcriptionally active macronuclei from the inactive micronuclei revealed that methylation of lysine 4 on histone H3 (H3K4) was positively enriched in the macronuclei. (Strahl et al., 1999). The positive correlation between H3K4 methylation and actively transcribed genes was further documented by chromatin

immunoprecipitation followed by high-throughput sequencing (ChIP-seq) showing that H3K4 tri-methylation (H3K4me3) is enriched at the promoter regions of actively transcribed genes (Santos-Rosa et al., 2002, Bernstein et al., 2005). Similar studies further showed that H3K4 mono-methylation (H3K4me1) marks enhancer regions of such genes along with other histone modifications (Heintzman et al., 2007, Visel et al., 2009) while H3K4 di-methylation (H3K4me2) is distributed more broadly, and often co-localizes with H3K4me3 but extends further in the body of genes (Schneider et al., 2004; Bernstein et al., 2005). Particularly, the HOX cluster of genes, which contains genes important for development, is decorated with H3K4me3 along its entire length (Bernstein et al., 2005). The methylation of H3K4 may mediate effects on transcription by working with other methylation events on histone proteins. Actively transcribed genes marked with H3K4me3 at their promoters, are often accompanied by H3K36me3 and H3K79me2 in the gene bodies, which are also found within actively transcribed genes (Guenther et al., 2007). The methylation of lysine or arginine residues can also dampen the effects of H3K4 methylation. Tri-methylation of H3K4 antagonizes the heterochromatic H3K9me3 modification, and counters gene expression silencing (Barski et al., 2007, Binda et al., 2010). Similarly, the tri-methylation of H3K27 localizes to the promoter regions of developmental genes along with H3K4me3, and silences these genes until cells undergo differentiation (Bernstein et al., 2006). The asymmetrical di-methylation of H3R2, a neighboring residue to H3K4, prevents effectors from binding to the methylated H3K4 PTM (Iberg et al., 2008). Further, the symmetric di-methylation of H4R3 interferes with the ability of the H3K4-specific methyltransferase MLL4 from binding to histone H4

(Dhar et al., 2012). Given its important links with transcription, the deposition of this mark onto transcriptional units is highly organized and controlled.

6. H3K4 methyltransferases:

In mammals, several enzymes methylate H3K4. These enzymes include the members of the SET1 family, SMYD3, SET7/9, Ash1, MEIS1, and SETMAR (Gu et al., 2013). Ash1 was first identified in *Drosophila* as an indispensable methyltransferase for the regulation of the *Hox* gene cluster (Byrd et al., 2003). The mammalian homologue shows a parallel function, and specifically methylates H3K4 (Gregory et al., 2007). SET7/9 is a H3K4 mono-methyltransferase that plays an active role in myogenesis (Tao et al., 2011), in the β -cells of the pancreatic islets (Ogihara et al., 2009), and is expressed during ROS-associated and genotoxic stresses (Chen et al., 2014, He et al., 2015, Lezina et al., 2014). SMYD3, from the SMYD family of methyltransferases, methylates H3K4 in a chaperone HSP90-dependent fashion (Hamamoto et al., 2004) to regulate muscle development (Proserpio et al., 2013), early embryonic development (Bai et al., 2016), early lineage commitment (Suzuki et al., 2015) and cardiac development (Fujii et al., 2011). Given that gene expression is finely tuned during the life cycle of a cell, especially for genes involved during developmental processes, mechanisms must be in place to limit the potential stimulatory effects of H3K4 methylation on transcription and other nuclear processes.

7. H3K4 demethylase:

The methylation of H3K4 is a reversible PTM. A first class of demethylases named LSD1 and LSD2, which harbor a catalytic Amine Oxidase Domain (OAD), is essential to demethylate the mono- and di-methylated forms of H3K4 (Shi et al., 2004, Karytinis et al., 2009). LSD1 allows the reprogramming of the epigenetic landscape by modulating H3K4 di-methylation levels (Katz et al., 2009) to regulate adipogenesis (Musri et al., 2010), skeletal muscle differentiation (Choi et al., 2010), and neural development (Sun et al., 2010). LSD2 shows structural homology to LSD1, and accordingly both demethylases share a common function for gene imprinting at maternal loci (Karytinis et al., 2009, Holmes et al., 2012, Ciccone et al., 2009). Additionally, LSD2 demethylase activity is largely targeted to gene bodies and may play a role in transcriptional elongation (Fang et al., 2010). The JARID1 family of demethylases, which harbor a catalytic iron-dependent dioxygenases containing Jumonji C domain, is responsible for the demethylation of the di- and tri-methylated forms of H3K4 (Christensen et al., 2007). JARID1A, a well-studied member of this family, regulates the expression of cell cycle, mitochondrial (Lopez-Bigas et al., 2008) and cytokine-related genes (Klose et al., 2007) as well as the progesterone receptor (Stratmann et al., 2007). JARID1B is implicated in senescence (Nijwening et al., 2011), the differentiation of neural and hematopoietic progenitor cells (Dey et al., 2008, Cellot et al., 2013), and in the formation of the neural lineage (Schmitz et al., 2011). JARID1A associates to specific G9a complexes and aids H3K9 methyltransferase in the formation of heterochromatin (Chaturvedi et al., 2012). JARID1A and its paralogue JARID1B also associate to PCR2 complexes to silence

developmental genes and counter the activity of H3K4-specific methyltransferases (Pasini et al., 2008, Zhou et al., 2009). Moreover, JARID1B acts as a co-repressor for the transcription factors BF-1, PAX9 (Tan et al., 2003), and TIEG1/KLF10 (Kim et al., 2010). JARID1C localizes to promoters and enhancers in ESC and in neural progenitor cells, where its demethylase activity impacts gene transcription (Outchkourov et al., 2013) such as neuronal genes (Tahiliani et al., 2007). Finally, JARID1D is indispensable for spermatogenesis (Akimoto et al., 2008). These studies demonstrated that an intricate interplay between H3K4 methyltransferases and demethylases is critical for the regulation of differentiation programs and developmental processes. However, mechanisms by which the methylated H3K4 elicits these transcriptional outcomes remain elusive.

8. Effectors of H3K4 methylation:

Methylation of lysine residues does not lead to a drastic change in the bulkiness of the lysine ϵ -amine, impact its charge and, in contrast to ubiquitination and acetylation, this PTM does not directly impact the structure of chromatin. However, many effector proteins harboring chromo, PHD, PWWP, and tudor domains (Patel and Wang, 2013) are able to recognize different states of H3K4 methylation and bring upon a specific biological outcome. For example, several members of the ING family of tumor suppressors recognize the tri-methylated form of H3K4. The PHD finger of ING2 binds to H3K4me3 and allows the association of the mSin3a–HDAC1 complex to the promoter of genes linked to cell proliferation (Shi et al., 2006) and ING4 recognizes the same PTM

to recruit the acetyltransferase HBO1 (Hung et al., 2009) to chromatin. A similar activity by ING1 is important to regulate DNA repair and apoptotic activities of the protein (Peña et al., 2008). The double chromodomain of CHD1 and the PHD domain of NURF's subunit BPTF, two chromatin remodelers, recognize H3K4me3 (Sims et al., 2005, Wysocka et al., 2006). The H3K9me3- and H3K36me3-specific demethylase JMJD2A binds to the tri-methylated form of H3K4 using one of its tudor domains (Lee et al., 2008). The PHD finger of PHF20, a subunit of the MOF16 acetyltransferase complex, recognizes chromatin marked with the di-methylated form of H3K4 (Klein et al., 2016). The association of Spindlin1 to H3K4me3 within ribosomal DNA regions is important for its stimulatory effect on transcription (Wang et al., 2011). In addition of serving as a binding platform for different effector proteins, H3K4 methylation can also impede on the ability of specific proteins to bind to histones. For example, H3K4 methylation blocks the binding of the de novo DNA methyltransferase DNMT3a to chromatin (Hu et al., 2009). Overall, these observations highlight that a functionally diverse set of proteins requires the methylation of H3K4 to operate on chromatin and that regulation of the predominant family of H3K4 methyltransferases, the SET1 methyltransferases, is critical for regulating a myriad of nuclear functions.

9. SET1 family of methyltransferases from yeast to humans:

Yeast SET1 (ySET1) was initially discovered as a H3K4-specific methyltransferase regulating telomeric (Nislow et al., 1997) and rDNA silencing (Briggs et al., 2001) as well as DNA damage repair (Corda et al., 1999). Subsequent studies assigned ySET1 as

an important regulator of gene expression (Boa et al., 2003). In *drosophila melanogaster*, three orthologues of the SET1 family are expressed which include drosophila SET1 (dSET1), trithorax (Trx), and Trithorax-related (Trr), which have non-redundant yet essential biological functions in fly development (Nislow et al., 1997). In mammals, six predominant SET1 methyltransferases are known. The paralogues SET1A and SET1B share homology with dSET1 and ySET1 while Mixed Lineage Leukemia 1 (MLL1) and MLL2 are similar to drosophila Trx. Finally, the paralogues MLL3 and MLL4 share similarities, both in sequence homology and organization of functional domains, with drosophila Trr (Mohan et al., 2011) (**figure 2A**).

10. The 6 members of the SET1 methyltransferase family:

SET1A/B are global regulators of all three forms of H3K4 methylation throughout the genome (Lee et al., 2005, Lee et al., 2007). These enzymes bind to WDR82, which in turn allows this complex to bind to the elongating form of RNA polymerase II (phosphorylated on its 5th serine of the heptad repeat) (Wu et al., 2008). SET1A/B also uniquely associates with CFP1 to recruit the enzyme to non-methylated CpG islands in the genome thereby contributing to the punctate localization of the PTM to the promoter region of actively transcribed genes. (Lee et al., 2005).

MLL1/2 predominantly display a mono- and di-methyltransferase activity towards H3K4 (Shinsky et al., 2015) *in vitro* and MLL1 tri-methylates approximately 5% of gene promoters (Wang et al., 2009) in HEK293 cells. Both MLL1 and MLL2 are

proteolytically cleaved by the caspase protease into N-terminal and C-terminal fragments, which reassemble into a functional complex (Yokoyama et al., 2002). Additional factors then bind to the various domains of MLL1 and MLL2 to assist them in mediating transcriptional regulation. One of the unique factors associating with MLL1 and helping the recruitment of the methyltransferase to its target genes are the tumour suppressor menin and the chromatin-anchoring lens epithelium growth factor (Huang et al., 2012, Yokoyama et al., 2008) (Wang et al., 2009). The MOF H4K16-acetyltransferase complex associates with the C-terminus of MLL1, via its PHD finger and promotes gene expression alongside MLL1 (Dou et al., 2005). The CXXC domain of MLL1 binds to the PAF-bound RNA polymerase II and participates in HOXA9 expression (Milne et al., 2010). Consistently, MLL1 colocalizes with the RNA polymerase II at the 5' end of coding genes (Guenther et al., 2005). The MLL1 and MLL2 methylate H3K4 on a specific subset of developmental and differentiation-related genes. In fact, MLL1 is required for erythropoiesis and regulates more specifically the expression of HoxA7, HoxC8, and HoxC9 genes (Yu et al., 1995, Hanson et al., 1999, Yokoyama et al., 2004). Moreover, MLL1 is indispensable for the proper segment identity in mice (Yu et al., 1995) and for neurogenesis (Lim et al., 2009). MLL1 promotes the proliferation and expansion of hematopoietic progenitor cells (Jude et al., 2007, Ernst et al., 2004) and regulates blood homeostasis in mice (Yu et al., 1995). Knockout of MLL2 leads to cellular apoptosis, delayed development of mice (Glaser et al., 2006) and skewed differentiation of ESCs (Lubitz et al., 2007). MLL2 has also been implicated in spermatogenesis, oogenesis (Glaser et al., 2009, Andreu-Vieyra et al., 2010) and in the formation of the germ cell and cardiac lineages (Glaser et al., 2009, Wan et al., 2014).

MLL3 and MLL4 are nuclear receptor-bound mono-methyltransferases that specifically mono-methylate enhancer regions (Hu et al., 2013). MLL3 and MLL4 associate with the transcriptional coactivator ASC-2 in a complex termed ASCOM, and collectively associate with a variety of nuclear receptors and their target genes (Goo et al. 2003). For example, the ligand-activated retinoic acid receptor binds to RARE DNA elements and recruits the ASCOM complex (Goo et al., 2003). The same complex is also recruited to the Liver X Receptor (LXR) and the Farnesyl X Receptor (FXR) target genes when these nuclear receptors are bound to the T317 agonist and bile acids, respectively. Disruption of MLL3 expression impairs LXR and FXR transactivation, reduces lipid droplet formation in the liver and disrupts bile acid homeostasis (Lee et al., 2008, Kim et al., 2009). Akin to LXR and FXR, ASCOM interplays with PPAR γ to control *aP2* gene expression and adipogenesis (Lee et al., 2008*). MLL3 and MLL4 share common binding partners that influence their function in H3K4 methylation. PTIP is found within certain MLL3/MLL4 complexes (Cho et al., 2007), and is required by the Pax2 transcription factor to recruit MLL3 to particular DNA elements (Patel et al., 2007). MLL3 and MLL4 associate with UTX (Lee et al., 2007, Van Nuland et al., 2013), an H3K27me₃-specific demethylase, which removes this repressive mark from various promoters and enhancers to allow gene expression via the remaining H3K4 methylation mark (Herz et al., 2012, Kim et al., 2014). MLL3 and MLL4 also have non-overlapping roles. For example, MLL3's role in adipogenesis is more prominent than its orthologue MLL4 (Son et al., 2016), while the latter one is more important for the differentiation of NT2-D1 cells into neuronal cells (Dhar et al., 2012). Accordingly, MLL4 knock-out mice

display unique phenotypes in adipogenesis and myogenesis, suggesting non-redundancy between the two enzymes (Lee et al., 2013). These examples highlight the important and diverse biological functions of members of the SET1 family of methyltransferases.

11. Links between the SET1 family of methyltransferases and diseases:

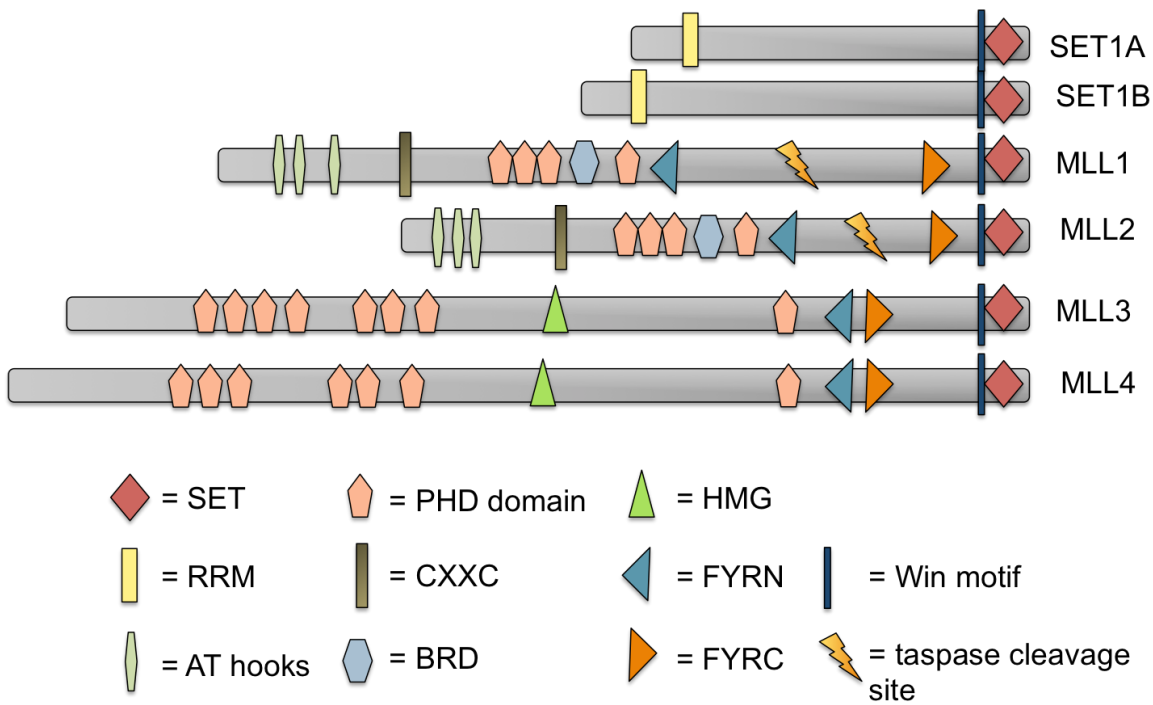
The members of the SET1 family of methyltransferases are mutated in a variety of cancers. The most studied cases are myeloid and myeloid leukemias arising in 5-10% of infants with leukemias (Cao et al., 2015). The chromosomal region spanning the N-terminal portion of MLL1 translocates with several elongation factors including, but not limited to, AF4, AF9, and ENL. These fusion proteins cause aberrant expression of hematopoietic self-renewal genes, which enable the pathologies to develop (Ayton and Cleary, 2001.). Intriguingly, the MLL fusion proteins and the wild-type complex cooperate to drive the high expression of MLL1 target genes. Mechanistically, the third PHD domain of MLL and the CXXC finger on the wild-type protein bind to the Hox9 locus to drive gene expression (Milne et al., 2010). In another example, the *MLL2* gene is overexpressed in colon and breast cancers (Natarajan et al., 2010) and amplified in pancreatic cancers and glioblastomas (Huntsman et al., 1999). Insertions of the Hepatitis B virus DNA into the *MLL2* gene are linked to hepatocellular *carcinoma* (Tamori et al., 2005) and somatic mutations of the same gene are frequently detected in non-Hodgkin lymphomas (Pasqualucci et al., 2011, Morin et al., 2011). Analogously, the *MLL3* gene is mutated in a large variety of cancers including glioblastomas and pancreatic (Balakrishnan et al., 2007), breast (Wang et al., 2011), colorectal (Sjöblom et al., 2006,

Watanabe et al., 2011) and bile duct cancers (Ong et al., 2012). MLL4 mutations were found in patients suffering of ovarian cancer (Ow et al., 2014) and hepatocellular carcinoma (Cleary et al., 2013). Similarly, other studies demonstrated that MLL4 contributes to tumor growth and angiogenesis in colon cancer (Ansari et al., 2012). Finally, mutations of *MLL3* and *MLL4* genes were identified in patients with mental diseases such as the kleeftstra and the kabuki syndromes (Ford et al., 2015). These examples show that MLL enzymes play indispensable roles in gene regulation and mutations of the SET1 family of methyltransferases cause an array of cancers and diseases.

12. The WRAD regulatory subcomplex:

Interestingly, despite playing different biological roles, all members of the SET1 family of methyltransferases reside in large mega complexes that vary in subunit composition (Van Nuland et al., 2013). For example, yeast SET1 associates with Cps30, Cps50, Cps60, and Cps25, a complex that is now referred to as Complex of Proteins Associated with SET1 (COMPASS) (Miller et al., 2001). Akin to yeast, members of the SET1 family of methyltransferases associate to a protein complex, also referred to as the WRAD complex, that includes WD40-containing Repeats 5 (WDR5), Retinoblastoma-Binding Protein 5 (RbBP5), Absent, Small, Homeotics 2-Like (Ash2L), and Dumpy-30 (DPY-30) (Dou et al., 2006) (**figure 2B**). *In vitro* reconstitution of the complex and biochemical studies of WRAD demonstrated that this regulatory subcomplex plays a critical role in

A.



B.

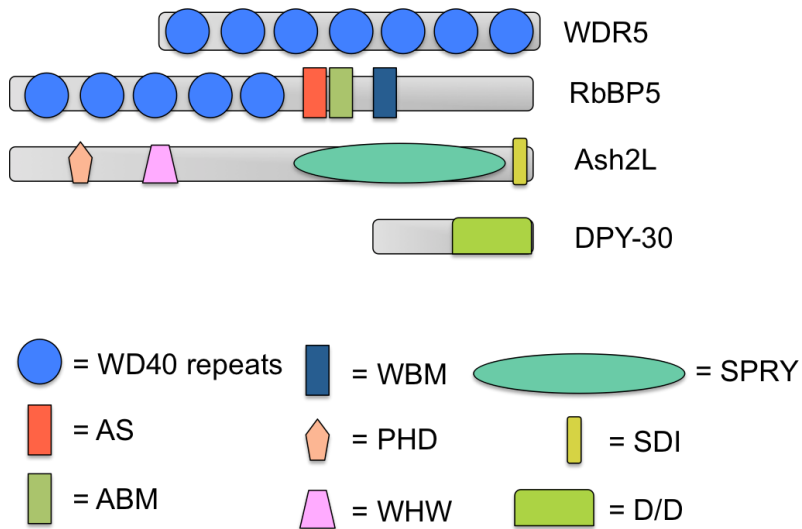


Figure 2. Domain organization of the members of the SET1 methyltransferase family and of the WRAD regulatory subunits. A) The domains composing SET1A, SET1B, MLL1, MLL2, MLL3, and MLL4, as well as WDR5, RbBP5, Ash2L, and DPY-30 are illustrated schematically with the main protein chain shown as a grey rectangle, and the various domains are depicted using various shapes and colours. The legend key for all shapes used with the corresponding name of the domain is indicated below. Abbreviations: SET = **Su(var)3-9, Enhancer of zeste, and Trithorax**, RRM= **RNA Recognition Motif**, PHD = **Plant Homeodomain**, BRD = **Bromodomain**, HMG = **High-mobility Group**, FYRN = **F/Y-rich N-terminus**, FYRC = **F/Y-rich C-terminus**, Win = **WDR5 interacting**. AS = **Activating segment**, ABM = **Ash2L Binding Motif**, WBM = **WDR5 Binding Motif**, WHW = **Wing-Helix-Wing**, SPRY = **SPla and Ryanodine**, SDI = **Sdc1 Interacting**, D/D = **Docking and dimerization domain**.

allosterically regulating the methyltransferase activity of MLL1 as well as the other members of the SET1 family (Dou et al., 2006).

13. The WRAD complex and cancer:

Similar to mutations found in the body of genes encoding H3K4 methyltransferases, genetic aberrations in the WRAD complex were identified. Overexpression of WDR5 is linked to the progression of several types of cancers including breast cancer (Dai et al., 2015), bladder cancer (Chen et al., 2015), leukemia (Ge et al., 2016), and neuroblastomas (Sun et al., 2015). RbBP5 is important for the epithelial to mesenchymal transition of prostate cancer cells, which is required for their metastasis (Li et al., 2016). Ash2L is highly expressed in different tumors and tumor cell lines (Lüscher-Firzlaff et al., 2008). Moreover, Ash2L cooperates with LSD1 in hepatocellular carcinomas to alter H3K4 dimethylation levels at different gene loci (Magerl et al., 2010). Finally, DPY-30 plays an important role in the proliferation, migration, and invasion of gastric cancer cells (Lee et al., 2015). The link between the misregulation of these proteins and cancers highlights the important role of properly regulating the SET1 family of methyltransferases.

14. Role of the subunits of the WRAD regulatory subcomplex:

In an attempt to further understand the biochemical basis underlying the formation of the WRAD complex, protein reconstitution studies revealed that the WRAD subunits associate in a pairwise fashion wherein direct interactions are formed between MLL1 and

WDR5, WDR5 and RbBP5, RbBP5 and Ash2L, and Ash2L and DPY-30 (Miller et al., 2001, Dou et al., 2006).

i. Role for WDR5 in the regulation of SET1 methyltransferases:

Initial characterization of WDR5 revealed that the protein nucleates the formation of the WRAD complex by directly binding to RbBP5 and MLL1 (Dou et al., 2006). Further studies revealed that WDR5 folds as a seven-bladed b-propeller structure (Schuetz et al., 2006) wherein a short stretch of residues on MLL1 that is found approximately 30 residues upstream of the catalytic SET domain, referred to as the Win motif, binds in the central pore in WDR5 (Patel et al., 2008). Subsequent studies revealed that all SET1 enzymes contain a unique win motif that can bind to the MLL1 binding pocket on WDR5 (Zhang et al., 2012). Binding to RbBP5 is achieved by a pocket located on the other side of MLL1 binding pocket of WDR5 (Avdic et al., 2011) and therefore bridges methyltransferases to the RAD subunits (RbBP5, Ash2L, DPY-30). When the interaction between MLL1 and WDR5 is abolished, the other members of the complex are unable to bind to the SET domain of MLL1 compromising its methyltransferase activity (Patel et al., 2008, Krogan et al., 2002, Steward et al., 2006, Dehe et al., 2006).

ii. Role of the Ash2L-RbBP5 complex in the activation of SET1 methyltransferase activity:

A short region on RbBP5, referred to as the Ash2L Binding Motif (ABM) consisting of acidic residues, preceding the WDR5 Binding Motif (WBM), binds to a basic patch located on top of Ash2L's SPIA and RYanodine (SPRY) domain (Zhang et al., 2015). Several lines of evidence points to a model in which the Ash2L/RbBP5 complex stimulates the enzymatic activity of members of the SET1 family of methyltransferases. First, the addition of Ash2L to an MLL1-WDR5-RbBP5 complex leads to an increase of 130-fold in the turnover number for H3K4 methylation (Patel et al., 2009). Second, the addition of an RbBP5-Ash2L heterodimer to the SET domain of MLL1 promotes H3K4 methylation to the same extent as an M-WRA complex (Cao et al., 2010). Third, similar to the results obtained for MLL1, the absence of RbBP5 and Ash2L results in a low turnover for H3K4 methylation by the other members of the SET1 family of methyltransferases. A cryo EM structure of the SET domain of yeast SET1 and MLL1 in complex with the four regulatory subunits depicts the complex as forming a Y shaped structure wherein the central SET domain is sandwiched between the WDR5-RbBP5 heterodimer and the Ash2L-DPY30 heterotrimer. The Ash2L-DPY-30 heterotrimer forms a compact triangular base at the bottom of the structure, whereas RbBP5 and WDR5 form two detached lobes of the Y shaped structure. RbBP5 extends behind the SET domain and along with part of the Ash2L-DPY-30 heterotrimer, contacts the SET domain, further supporting a direct role of Ash2L and RbBP5 in the allosteric regulation of SET1 methyltransferase activity. (Takahashi et al., 2013). Such architecture was further

supported by a high-resolution structure of the MLL3/Ash2L/RbBP5 complex. In this seminal study, residues 336-354 of RbBP5 (residues 336-344 are referred to as the AS motif while residues 344-354 are referred to as Ash2L binding Motif (ABM)) was co-crystallized in complex with the SET domain of MLL3 and the SPRY domain of Ash2L. The structure shows that a crucial arginine residue on the i-SET region of MLL3, located between the 4th beta strand and the α B α -helix, forms multiple electrostatic interactions with Ash2L and RbBP5, simultaneously. This residue is crucial for the formation of the complex, and thus for the stimulatory effect of the Ash2L-RbBP5 heterodimer on the methyltransferase activity of MLL3. The RbBP5 AS-ABM and Ash2L SPRY collectively form a crevasse that nestles the iSET α B α -helix, opposite to where the substrate histone H3 binds (**figure 3**). Further, investigations on the dynamics of the MLL3 SET domain reveal that the Ash2L-RbBP5 reduces the mobility of the i-SET motif, especially of the α B α -helix and thus enhances the binding of the SET domain to the histone H3 substrate (Li et al., 2016)

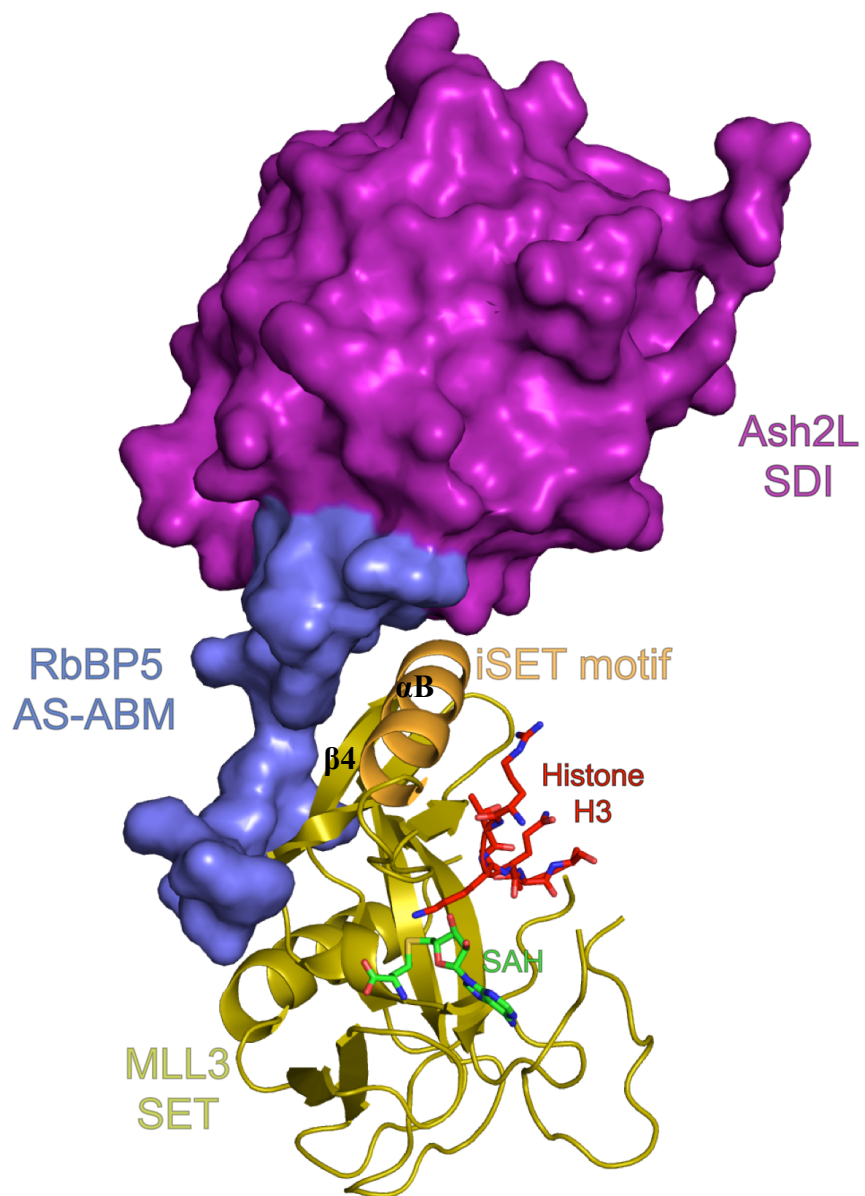


Figure 3. The RbBP5-Ash2L heterodimer rigidifies the iSET motif on the SET domain of MLL3. Crystal structure of a reconstituted active MLL3 complex composed of the SPRY domain of Ash2L (purple), the AS-ABM region of RbBP5 (blue), and the SET domain of MLL3 (olive) bound to the histone H3 substrate (red) and the S-Adenosyl Homocysteine (SAH) cofactor (green). The RbBP5-Ash2L heterodimer is represented in surface diagram showing how these proteins collectively create a groove that shelters the iSET α -helix (orange), shown in cartoon diagram, on the SET domain of MLL3. Important secondary structural elements are labeled.

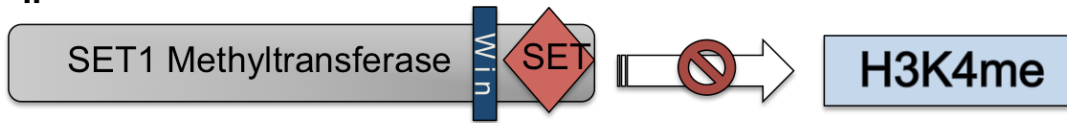
iii. Role for DPY-30 in SET1 methyltransferases:

DPY-30 is an evolutionarily conserved protein that was initially discovered in *C. elegans* as a key player in X-Chromosome dosage compensation (Hsu et al., 1994). Subsequent biochemical studies showed that DPY-30 co-purifies with SET1 enzymes (Roguev et al., 2001, Cho et al., 2007). Specifically, the knockdown of DPY-30 results in a global loss of H3K4 tri-methylation in various model systems (Jiang et al., 2011, Krogan et al., 2002, Schneider et al., 2005, Dehe et al., 2006). Moreover, depletion of DPY-30 prevents ESCs differentiation into the neural lineage (Jiang et al., 2011) and the proliferation as well as the differentiation of hematopoietic progenitor cells (Yang et al., 2014). Binding of DPY-30 to SET1 complexes is mediated via a direct interaction with Ash2L (South et al., 2010, Dehe et al., 2006). Detailed biochemical studies revealed that the interaction is mediated by a region corresponding to a 44-residue helical bundle on DPY-30, termed the docking and dimerization domain (D/D), and the last 30 residues of Ash2L (also referred to as the Sdc1-DPY-30 Interacting (SDI) domain) (South et al., 2010) (**figure 4**).

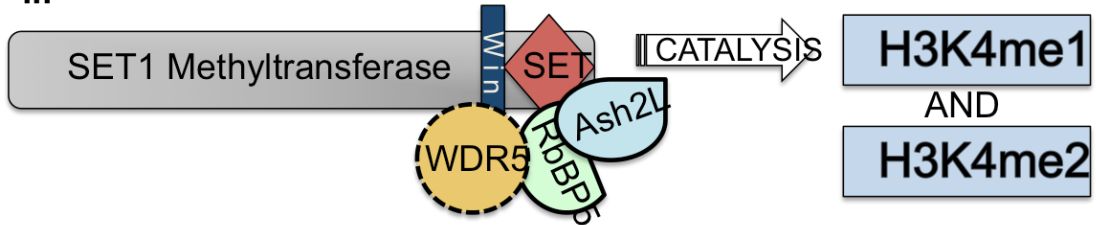
15. The DPY-30/Ash2L Heterotrimer:

The DPY-30 D/D was crystallized in complex with the Ash2L SDI. Two DPY-30 D/D domains dimerize in an antiparallel fashion to form a typical x-type four-helix bundle. The two helices of each protomer intersect at about 45° and the first and second α -helices of both protomers pack against each other via hydrophobic

i.



ii.



iii.

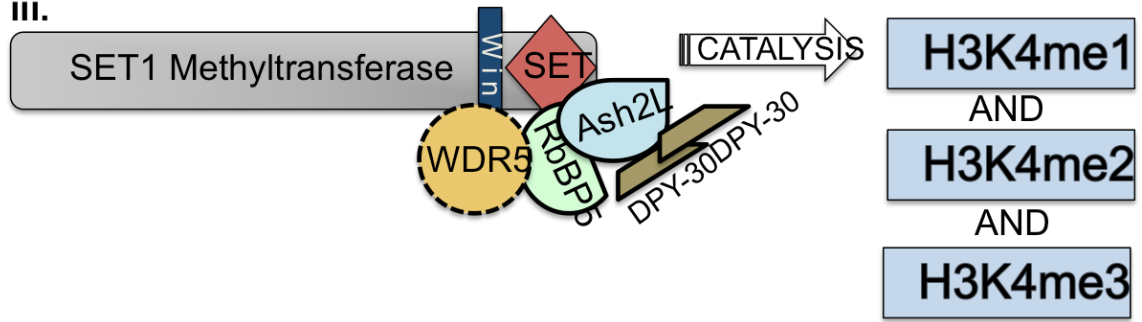
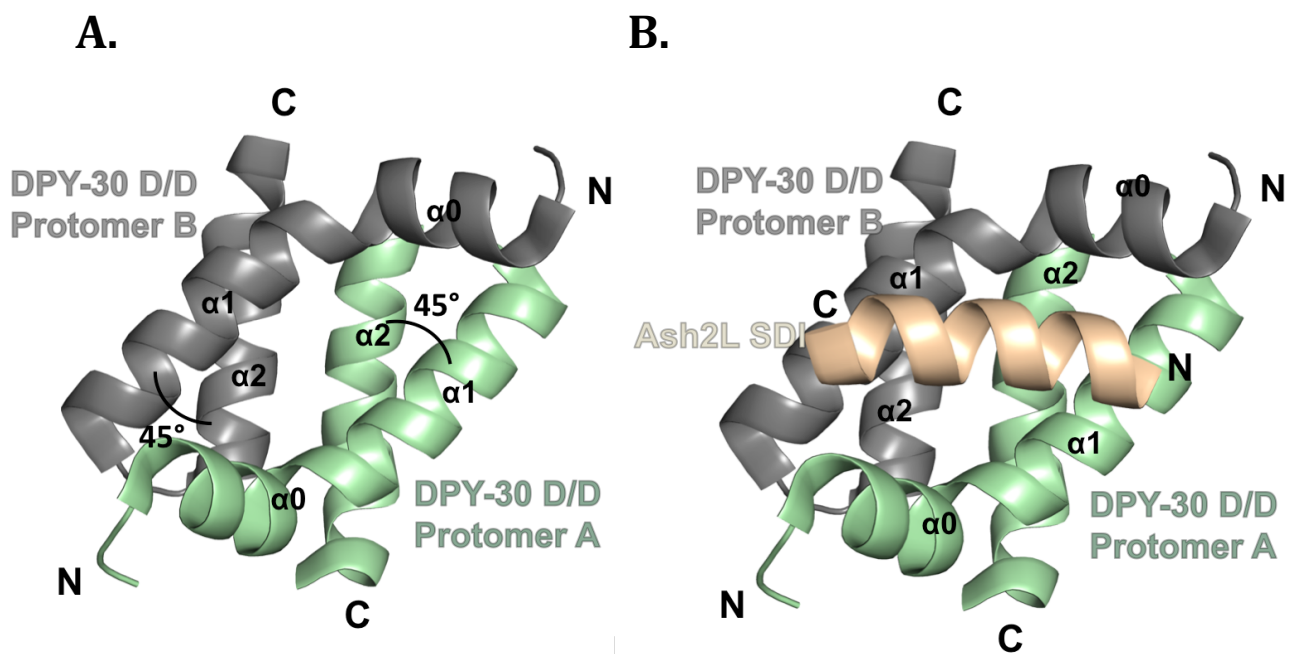


Figure 4. Stimulatory effect of the regulatory WRAD subunits on SET1 methyltransferase activity. The stimulatory effect of the WRAD subunits depicted by various shapes/colours, on SET1 methyltransferases, which are all represented by a generic gray construct, is schematically described. i. SET1 methyltransferases alone (the apo-enzymes) are unable to efficiently methylate H3K4. ii. WDR5 binds to the win motif of SET1/MLL and recruits the RbBP5-Ash2L heterodimer to the SET domain to stimulate the mono- and di-methyltransferase activity of SET1 enzymes on H3K4. Depending on the SET1 member, WDR5 may not be required to target RbBP5-Ash2L to the SET domain, which is represented by the discontinuous line on the WDR5 protein. iii. Ash2L recruits DPY-30 into SET1 complexes, enabling the catalysis of the tri-methylation of H3K4.



C.

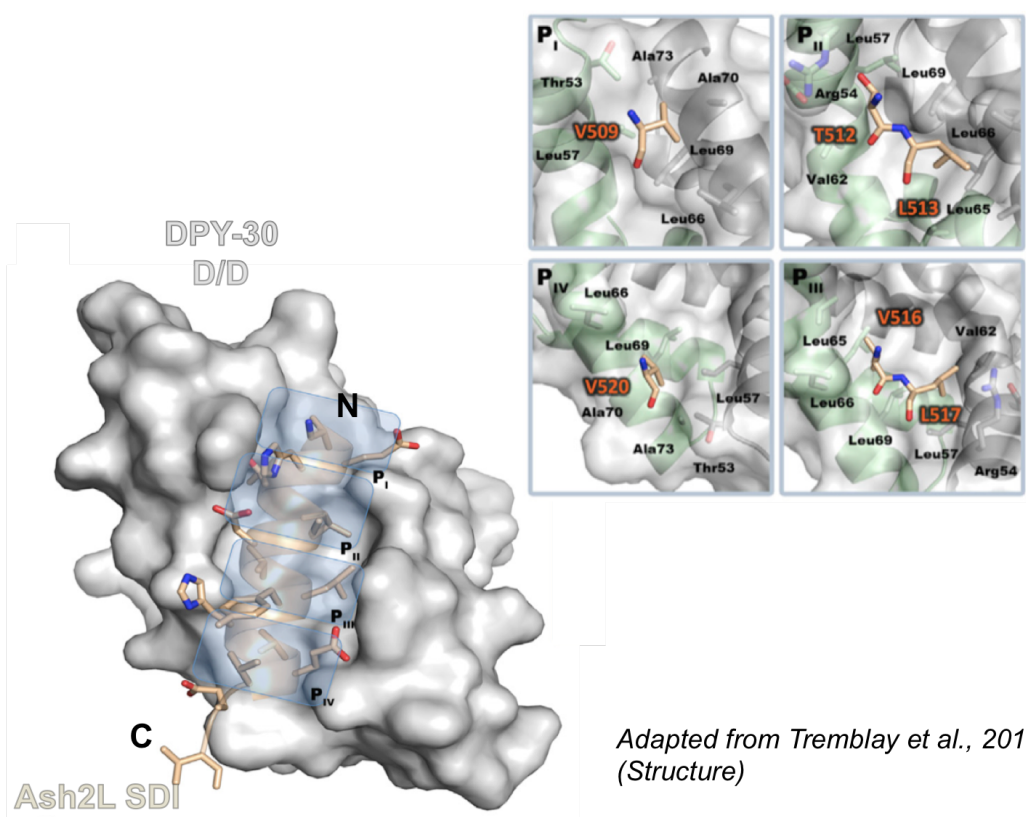


Figure 5. Structures of the DPY-30 D/D homodimer and the DPY-30D/D/Ash2LSDI complex. A. Structure of a homodimer of the DPY-30 D/D domain forming a typical X-type four-helix bundle with crossing over angles of 45° between helices $\alpha 0/\alpha 1$ and $\alpha 2$ of each protomer. The N-termini and C-termini and each protomer of DPY-30 are labeled. B. Structure of the DPY-30^{D/D}/Ash2L^{SDI} complex shown in cartoon diagram with each protein and the C- and N-termini labeled. C. Structure of the DPY-30^{D/D}/Ash2L^{SDI} complex with the DPY-30 homodimer in surface diagram (gray), emphasizing the hydrophobic crevasse that shelters the SDI helix on Ash2L, shown in cartoon diagram (beige). Four binding pockets on DPY-30 that contact hydrophobic residues on the SDI α -helix are highlighted. The specific interactions between the four binding pockets on DPY-30 and Ash2L are displayed in the corner and residues that are important are represented in sticks and labeled.

interactions (Wang et al., 2009). The non-polar face of the SDI amphipathic α -helix lies in a hydrophobic crevasse formed by residues on the two first helices of both protomers. The hydrophobic interactions between the Ash2L SDI and the DPY-30 D/D exhibit a pseudo two-fold symmetry around a central residue on the SDI helix. The non-polar face of the SDI helix of Ash2L is composed of residues **VxxTLxxVLxxV**, which fill four binding pockets on DPY-30 (**Figure 5**). On either side of the center, a pair of aliphatic chain-containing residues fit into two large pockets. These pairs of hydrophobic residues are flanked by a pair of valine residues on each side that contact two shallow pockets. Mutational analysis of DPY-30 and Ash2L SDI, *in vitro* and in cells demonstrated that two key leucine residues, Leu513 on Ash2L and Leu69 on DPY-30, are crucial for the formation and function of the heterotrimer (Tremblay et al., 2014).

16 . Research goal 1: Gain structural insights into the formation of the Ash2L-DPY-30 heterotrimer

The Briggs laboratory demonstrated that a region comprising the last 55 residues of DPY-30 and a C-terminal amphipathic α -helix on Ash2L are the minimal constituents required for the interaction of both proteins (South et al., 2010). In addition, the same authors also demonstrated that a short segment located between the SPRY and SDI domains of Ash2L is critical for its interaction to DPY-30 (South et al., 2010). However, the initial structure solved by Tremblay et al. failed to provide the molecular insights underlying the function of this segment in the formation of the Ash2L/DPY-30 complex (Tremblay et al., 2014). To gain a better understanding of the formation of the DPY-30/Ash2L heterotrimer, I used x-ray crystallography and different binding approaches to

gain a better knowledge of the full extent of interactions occurring between DPY-30 and Ash2L.

17. Functions of WRAD in the nucleus:

In addition to stimulating the methyltransferase activity of the SET1 enzymes, the WRAD subcomplex plays a key role in integrating other biological cues to control the methylation of H3K4. WDR5 associates with the KANSL1 and KANSL2 subunits of the MOF16 acetyltransferase and together, activate gene transcription (Dias et al., 2014). In another instance, WDR5 associates with the HDAC3 H3K27-deacetylase during hypoxia, and promotes the H3K4 methylation of mesenchymal genes by the SET1 complexes, while repressing epithelial gene expression (Wu et al., 2011). WDR5 binds to the transcription factor C/EBP-alpha (Grebien et al., 2015) and to the SMC-selective pituitary homeobox 2 (Gan et al., 2011) and is recruited to their target genes for methylation and activation of gene expression. Ash2L contains a DNA-binding domain, called the wing-helix-wing domain (Sarvan et al., 2011), and similarly interacts with numerous transcription factors to mediate gene expression. For example, Ash2L is recruited by the NFY transcription factor to tri-methylate CCAAT promoters (Fossati et al., 2011). Second, Ash2L binds to the GATA-Binding Protein-3 and enhances the expression of the estrogen receptor α gene (Qi et al., 2014). Third, during muscular differentiation, the phosphorylated form of Mef2d interacts with Ash2L to promote the expression of muscle-specific genes (Rampali et al., 2007). The WDR5-RbBP5 heterodimer binds to the CUL4-DDB1 ubiquitin E3 ligase to influence nucleosomal H3K4 methylation (Higa et al., 2006). WDR5 and Ash2L bind to the Myc oncogene and

regulate the epigenetic status of bivalent promoters (Thomas et al., 2015, Ullius et al., 2014). These regulatory subunits cooperate with the Yamanaka pluripotency factors to regulate cell development programs. In fact, WDR5 and Ash2L interact directly with Oct4 (Ang et al., 2011) and Sox2, respectively (Yang et al., 2015), while DPY-30 plays an important role in the recruitment of Oct4 to its target genes (Yang et al., 2015). Moreover, WDR5 binds to the HOTTIP long non-coding RNA (Wang et al., 2011) and several ESC-specific RNAs, which are indispensable for its accumulation on chromatin, and for the regulation of ESC self-renewal (Yang et al., 2014). These few examples show that the WRAD subunits can interact and cooperate with a variety of factors that influence the dynamic of H3K4 methylation on chromatin thereby adding a layer of complexity to the regulation of this post-translational modification. This ability of WRAD to bind to proteins other than the subunits of SET1 methyltransferases also extends to cellular compartments outside of the nucleus, and attributes them with novel functions.

18. Non-nuclear functions of WRAD:

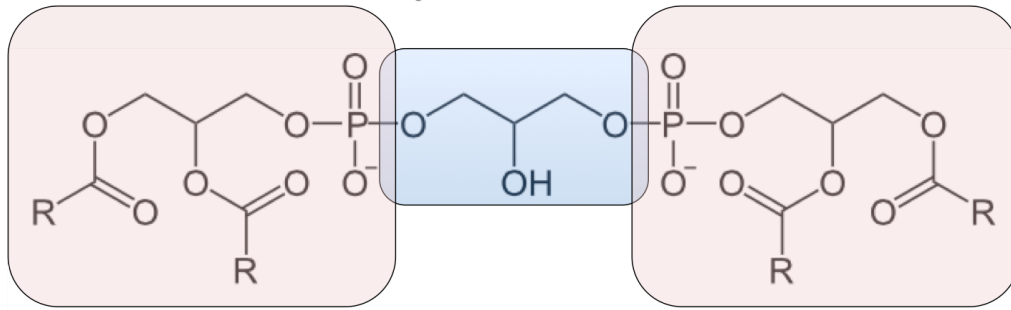
The WRAD subcomplex is known to play roles outside of the nucleus. During cytokinesis, WDR5 binds to PRC1, KIF4, and MKLP1 at the midbody of dividing cells, and regulates the disassembly of the constrictive microtubules structure (Bailey et al., 2015). DPY-30 localizes to the trans-Golgi network and binds to a resident guanine nucleotide exchange factor named BIG-1. In this compartment, DPY-30, along with Ash2L and RbBP5, plays an important role in endosomal trafficking. Accordingly, knockdown of any of these RAD subunits leads to an accumulation of cellular cargo at

cell protrusions (Xu et al., 2009). Lastly, in cells infected with the Sendai virus, WDR5 translocates from the nucleus to mitochondria, and interacts with the VISA complex to trigger the host anti-virus innate response (Wang et al., 2010). To achieve these functions, the WRAD subunits must be able to shuttle out of the nucleus and localize to other organelles such as the endoplasmic reticulum and mitochondria. The targeting of the WRAD subunits to these non-nuclear cellular compartments may be accomplished through the binding of membrane-specific lipids, such as the mitochondrial form of cardiolipin.

19. Cardiolipin:

Cardiolipin is a primitive lipid that is almost exclusively located in the inner membrane of the mitochondria of eukaryotes (Stoffel and Schiefer, 1968) and small amounts are found in the endoplasmic reticulum (Donaldson et al., 1977). The predominant structure of cardiolipin consists of two phosphatidic acid groups connected to carbons C1 and C3 of a glycerol backbone, to form a symmetrical molecule (MacFarlane et al., 1957) **(Figure 6)**. The synthesis of this molecule is catalyzed by cardiolipin synthase, which combines a molecule of phosphatidylglycerol with a molecule of CDP-diacylglycerol in the inner membrane of mitochondria. This basic cardiolipin molecule can be remodeled to give rise to a variety of forms that vary in their acyl chains (Schlame et al., 2016). Its four acyl chains occupy a large

Glycerol Backbone



Phosphatidic acid moiety

Phosphatidic acid moiety

Figure 6. Structure of a cardiolipin molecule. The chemical structure of a generic cardiolipin molecule is shown with the main glycerol chain and the two phosphatidic acid moieties connected to carbons C1 and C3 of the glycerol backbone highlighted in blue and red, respectively. The conventional letters are used to symbolize all atoms and the “R” symbol represents acyl chains of variable nature.

proportion of the molecule relative to its small polar head group, giving cardiolipin a conical shape, which is ideal to form monolayers with negative curvature (Renner et al., 2011). At physiological pH, the two phosphate groups on cardiolipin are deprotonated to give this lipid a negative two charge, which classifies cardiolipin as an anionic lipid (Olofsson et al., 2013). Given its localization in the mitochondria, cardiolipin is indispensable for the function of numerous proteins that regulate the cell's bioenergetics. For instance, cardiolipin stabilizes the super complexes III and IV of the electron transport chain (Yu et al., 1980; Robinson et al., 1990) and is essential for the activity of the ADP/ATP carrier protein (Drees et al., 1988). Moreover, its translocation to the outer membrane of mitochondria signals apoptosis, mitophagy, (Li et al., 2015) and inflammation (O'Neill et al., 2013). Finally, cardiolipin is linked to several cardiovascular diseases, like the Barth syndrome, and autoimmune diseases (Shen et al., 2015).

20. Structural characterization of the interactions of cardiolipin with proteins:

The co-crystal structures of cardiolipin bound to diverse proteins have been studied thoroughly and compared to infer structural elements that dictate the specific binding of proteins to cardiolipin. A cardiolipin molecule could be broken down into two pieces: the polar head group containing a hydroxyl group, and two phosphate groups, and a second part containing non-polar acyl chains. Polar residues (asparagine, tyrosine, threonine) or charged residues such as a lysine and arginine on proteins form electrostatic interactions with the polar head group. Interestingly, tryptophan residues have a large occurrence in

regions of proteins that especially bind to the hydroxyl group of cardiolipin. Regions on the protein that bind to the acyl chains of cardiolipin are rich in the hydrophobic residues valine, leucine, and isoleucine, which form hydrophobic interactions with these non-polar tails. Regions binding the polar head group tend to be more flexible; meanwhile acyl chain-binding elements are composed of rigidifying amino acids. Generally, cardiolipin binding patches have a high helical content, and a reoccurring supersecondary structure is two α -helices of variable lengths separated by a short loop. However, the head group also binds to random coil, whereas acyl-binding regions contact a β -strand in some instances (Planas-Iglesias et al., 2015).

21. Research goal 2: Determine whether Ash2L binds lipids *in vitro*.

The terminal SDI α -helix of Ash2L is amphipathic showing polar residues on one side and non-polar residues on the other side. In metazoans, the amino-acid sequence of Ash2L indicates that a tryptophan residue is located after the SPRY domain and only four residues before Ash2L SDI α -helix. Taken together, these features may constitute a potential lipid-binding platform wherein the polar residues of the SDI helix could form electrostatic interactions with the lipid head-group while the non-polar face of the helix can form hydrophobic interactions with the acyl chains of lipids. Further, computational thermodynamics studies testing the movement of hydrophobic peptides in lipid bilayers revealed that tryptophan residues are important to anchor transmembrane domains in a lipid bilayer (De Jesus et al., 2012). These structural insights helped us suggest that Ash2L binds to lipids.

MATERIAL AND METHODS

1. Cloning, Expression, and Purification of the Ash2L/DPY-30 heterotrimer:

The cDNA of the DPY-30 human gene was cloned in the parallel vector pGST2 (Sheffield et al., 1999) downstream of the glutathione sulfotransferase tag and a Tobacco Etch Virus (TEV) cleavage site. The cDNA of an Ash2L construct containing the SPRY and SDI domains was cloned in a homemade vector allowing the expression of proteins in fusion with a hexahistidine-tagged Small Ubiquitin-like Modifier (SMT3). Rosetta *E. coli* cells were transformed with GST-DPY-30 and His-SMT3-Ash2L^{SPRY+SDI} and the resulting colonies were grown in LB media supplemented with ampicillin (100mg/ml) until the cell density reached an OD⁶⁰⁰ of 0.4-0.5. Expression of proteins was then induced with 0.1mM of Isopropyl β -D-1-thiogalactopyranoside, at 18°C, during 16 hours. Cells expressing GST-DPY-30 were harvested and resuspended in PBS1X pH 7.4 buffer and sonicated 3 times for 1 minute to lyse the cells. The cell lysate was clarified by high-speed centrifugation at 16 000rpm, during 30 minutes, at 4°C. After filtration of the clarified lysate with a 0.22micron filter, the supernatant was incubated with a glutathione-sepharose resin (GE Healthcare) during 1 hour. The resin was extensively washed with PBS1X pH 7.4 buffer, resuspended in 20mM Tris pH 7.0, 150mM NaCl, 5mM β -Mercaptoethanol (BME), and incubated with 500 μ g TEV protease during 16 hours. Cells expressing His-SMT3-Ash2L^{SPRY+SDI} were harvested and resuspended in buffer A (50mM NaPi pH 7.0, 500mM NaCl, 5mM BME) and sonicated 3 times for 1 minute to lyse the cells. The cell lysate was clarified by high-speed centrifugation at 16 000rpm, during 30 minutes, at 4°C (Avanti J25, Beckman Coulter). After filtration of clarified

lysate with a 0.22micron filter, the supernatant was applied to a Talon cobalt affinity resin (Clontech) by gravity and the resin was extensively washed with Buffer A. The resin was resuspended in Buffer A supplemented with 5mM imidazole pH 7.0 and incubated with 500µg of the protease ULP1. The resin was gently centrifuged and the supernatant was concentrated using a 10KDa cutoff Amicon concentrator. The proteins were then quantified using OD²⁸⁰ (DPY-30 ϵ = 0.397 and Ash2L^{SPRY+SDI} ϵ = 1.842). DPY-30 and Ash2L^{SPRY+SDI} were mixed in a stoichiometric ratio of 2:1 respectively and incubated during 1 hour. The DPY-30/ Ash2L^{SPRY+SDI} heterotrimer was further purified by size-exclusion chromatography on a Superdex S200 preparative column pre-equilibrated in 20mM Tris pH 7.0, 150mM NaCl, 5mM BME. Fractions corresponding to the major peak of the heterotrimer, which eluted at approximately 75mL, were collected, pooled and concentrated to around 60mg/mL.

2. Crystallization of the Ash2L/DPY-30 heterotrimer:

Crystallization trials were carried out using various sparse matrix screens. Crystals were obtained using the vapor diffusion technique with a mother liquor composed of 20mM citrate pH 4.7 and polyethylene glycol (PEG) 6000. Crystals were harvested and transferred into the mother liquor supplemented with 20% MPD40 and flash frozen in liquid nitrogen. A full dataset was collected at Argonne National Laboratories on the 17-ID beamline at LS-CAT (the dataset was integrated and scaled using Aimless by Dr. Joseph Brunzelle). The structure of Ash2L^{SPRY+SDI}/DPY-30 heterotrimer was solved by molecular replacement using Phaser (Zwart et al., 2008) and Ash2L^{SDI}/DPY-30 heterotrimer and Ash2L^{SPRY} domain as search models. The model was completed using

iterative rounds of refinement and manual building using Phenix (Adams et al., 2010) and Coot (Emsley et al., 2010), respectively (Dr. Jean-Francois Couture built and refined the structure of the Ash2L^{SPRY+SDI}/DPY-30 heterotrimer). Default COOT settings were used to observe the calculated and observed fourier maps, and to identify the positive fourier hypothesized to be a lipid head group.

3. Antibodies:

Table 1. Antibodies used in this study.

Primary antibody name (company, catalogue number)	Dilution of primary antibody (BSA or milk)	Secondary antibody used	Dilution of secondary antibody (BSA or milk)
Anti-Ash2L (Bethyl laboratories, A300-112A)	1:10 000 (milk 2%)	Anti-Rabbit	1:10 000 (milk 2%)
Anti-FLAG (Sigma, F3165)	1:5 000 (milk 2%)	Anti-Mouse	1:10 000 (milk 2%)
Anti-GFP (Living colours, 632381)	1:2 000 (milk 2%)	Anti-Mouse	1:10 000 (milk 2%)
Anti-His (Santa Cruz Biotechnology, sc-803)	1:500 (BSA 2%)	Anti-Rabbit	1:10 000 (BSA 2%)
Anti-GST-HRP (Abcam, ab3416)	1:10 000 (BSA 2%)	N/A	N/A
Anti-DPY-30 (made in-house)	1:50 000 (milk 2%)	Anti-Rabbit	1:10 000 (milk 2%)

4, In vitro pull-down experiment:

Mutations of DPY-30 and Ash2L^{SPRY+SDI} were generated using a site-directed mutagenesis kit (Stratagene) and the mutants were overexpressed and purified as for the wild-type proteins. To perform the pull-down, the supernatant of wild-type DPY-30 was diluted in PBS1X supplemented with 0.1% Triton (PBS-T) and applied to glutathione

sepharose during 1 hour at 4°C. After three washes with PBS-T, Ash2L^{SPRY+SDI} wild-type or mutants were incubated with the GST-bound DPY-30, during 2 hours and the resin was washed three times with PBS-T. The proteins were eluted off the beads with SDS-sample loading buffer (SLB) 2X and boiled at 92°C before separating them by SDS-PAGE. Coomassie staining was used to visualize the proteins. Inputs of the His-SMT3-Ash2L^{SPRY+SDI} constructs were separated on SDS-PAGE and probed using an anti-His antibody. The same protocol was used to study the interaction between GST-DPY-30 mutants and wild type His-SMT3-Ash2L^{SPRY+SDI} on glutathione-sepharose beads.

5. Tandem Flag- and GFP-co-Immunoprecipitation:

Human Embryonic Kidney 293 (HEK293) cells were co-transfected with plasmids expressing a flag-tagged and eGFP-tagged copy of wild-type DPY-30 or DPY-30 mutants, using polyethylenimine as a transfection reagent. Two days following transfection, cells were harvested and lysed in radioimmunoprecipitation assay (RIPA) buffer (Wright, 1989). Cell lysates were centrifuged and the supernatant was incubated during 2 hours to anti-FLAG M2 magnetic beads (Invitrogen) pre-equilibrated with RIPA buffer at 4°C. Beads were washed three times with RIPA buffer, and flag-DPY-30 was eluted with 200µg of flag peptide. The eluted protein was then applied to a GFP trap (Bulldog bio) during 2 hours. Proteins were eluted with SLB 2X and separated by SDS-PAGE. Ash2L was detected using an anti-Ash2L antibody (Bethyl laboratories) while flag-tagged DPY-30 was detected using an anti-flag antibody (sigma). eGFP-DPY-30 was detected using an anti-GFP antibody (Living colours). The appropriate secondary

antibodies conjugated with HRP were added and ECL was used as a substrate for HRP. Detection was done via chemiluminescent signal on the Li-Cor Odyssey Fc imaging system.

6. Protein-lipid overlay assay:

Commercial phospholipid strips (Echelon, catalogue no. P-6002) were blocked in 5% BSA during 1 hour. Strips were then incubated with 100nM of GST-DPY-30/Ash2L^{SPRY+SDI} complex during 1 hour, at 4°C in binding buffer composed of 10mM citrate pH 4.7, 150mM NaCl, 5mM BME, 0.1% Tween-20. The strips were washed with the binding buffer and bound proteins were blotted using an HRP coupled anti-GST antibody. Detection of the bound proteins was performed on the Li-Cor Odyssey Fc imaging system.

7. Biotin-cardiolipin pull-down:

Biotinylated cardiolipin was solubilized in methanol at 1µg/µL. Either recombinantly purified SMT3-Ash2L^{SPRY+SDI}, DPY-30, or a DPY-30/SMT3-Ash2L^{SPRY+SDI} complex (216nM) was incubated with 2µg of sn-1-Biotin-Labeled Cardiolipin (Echelon, catalogue no. L-C16B) in binding buffer composed of 10mM citrate pH 4.7, 150mM NaCl, 5mM BME, over streptavidin-agarose beads (GE Healthcare) pre-equilibrated with binding buffer during 16 hours at 4°C. Streptavidin resin was washed in the binding buffer and the proteins were eluted from the resin using SLB2X and boiled at 92°C. Proteins were separated by SDS-PAGE, and DPY-30 was blotted using a homemade anti-DPY-30 antibody. SMT3-Ash2L^{SPRY+SDI} was detected using an anti-His antibody.

8. Lipids:

Table 2. Lipids used in this study. A list of all lipids used in this study with their specific chemical name, the method of preparation, and information on their purchase are indicated.

Generic phospholipid name (chemical name)	Source of lipid	Company name (catalogue number)
Phosphatidylcholine (18:1 (Δ 9-Cis) Phosphatidylcholine)	Synthetic	Avanti Polar Lipids (850375C)
Phosphatidylethanolamine (1-palmitoyl-2-oleoyl-sn-glycero-3-phosphoethanolamine)	Synthetic	Avanti Polar Lipids (850757C)
Phosphatidylserine (L- α -phosphatidylserine)	Purified from porcine brain	Avanti Polar Lipids (840032C)
Phosphatidic acid (L- α -phosphatidic acid)	Purified from chicken eggs	Avanti Polar Lipids (840101C)
Cardiolipin (cardiolipin)	Purified from bovine heart	Avanti Polar Lipids (840012C)

9. Liposome preparation:

Stoichiometric amounts of PC backbone lipid and the lipid of interest, both dissolved in chloroform, were mixed in ratios reflecting the content of biological membranes (PC:PE liposomes, molar ratio of 70:30, PC:CL, molar ratio of 85:15, PC:PA, molar ratio of 90:10, PC:PS, molar ratio of 80:20). The chloroform was completely evaporated from the mixtures using nitrogen gas, until a film was formed on the vial. The film was resuspended in a buffer composed of 10mM citrate pH 4.7, 150mM NaCl, 5mM BME, 0.5M sucrose. The solution of lipid mixes was sonicated in a water bath sonicator (Covaris S220) until the solution was optically clear and the liposomes were stored at 4°C. The time of sonication was varied depending on the lipid mixes (PC:PE for 1.5 hours, and PC:CL, PC:PA, and PC:PS for 30 minutes).

10. Liposome binding assay:

A recombinantly purified Ash2L variant containing residues 1-627 (400nM) was incubated with 400μM of liposomes in a binding buffer composed of 100mM citrate 4.7, 150mM NaCl, 5mM BME, 0,5M sucrose during 15 minutes at 4°C. To remove the precipitate that may have formed during the binding reaction, the solution was clarified by centrifugation at 13 300rpm during 15 minutes. To pellet the liposomes, the protein-lipid mixtures were then centrifuged (Beckman Optima TL 100 Ultracentrifuge) at 62 000rpm during 2 hours at 4°C. The supernatant was removed and the pellet was resuspended in binding buffer. The fractions corresponding to the supernatant (25μL removed/230μL total) and pellet (resuspended in binding buffer) were separated on an SDS-PAGE gel and the proteins were detected by Coomassie staining. Bands on the gel were quantified in pixels using Image J, and the percentage of proteins bound to liposomes were calculated using the following formula:

$$\% \text{ Protein Bound} = \frac{\text{pixels in pellet band (a.u.)}}{\text{pixels in pellet band (a.u.)} + 9.2 \times \text{pixels in supernatant band(a.u.)}}$$

This experiment was done in triplicates, except when performed with full-length Ash2L for which the experiment was repeated four times. A two-tailed t-test assuming equal variances conducted in excel was used to evaluate the statistical differences between samples.

11. Lipid competition assay:

Recombinantly purified GST-DPY-30 and an Ash2L variant containing residues 1-627 were incubated in a binding buffer composed of 10mM citrate pH 4.7, 150mM NaCl, 5mM BME, with increasing amounts of PC:PE liposomes or PC:CL liposomes. Protein-lipid mixes were incubated with glutathione sepharose pre-equilibrated with the binding buffer during 1 hour. The beads were washed with the binding buffer and the proteins were eluted with SLB2X and boiled at 92°C. The proteins were separated by SDS-PAGE and stained with Coomassie.

RESULTS

Our group recently reported the crystal structure of a minimal construct corresponding to DPY-30 dimerization docking module in complex with Ash2L SDI helix (Tremblay et al., 2014). While this structure highlighted the importance of aliphatic residues in mediating the formation of the heterotrimer, it failed to capture a short segment between the Ash2L SPRY domain and the SDI.

1. Purification of the DPY-30/Ash2L^{SPRY+SDI} heterotrimer:

To produce a large quantity of human DPY-30, the cDNA corresponding to the full-length protein was cloned into the pGST2 parallel vector (Sheffield et al., 1998). This construction enables the expression of DPY-30 as a Tobacco Etch Virus (TEV) cleavable GST fusion protein. A fragment encoding Ash2L^{SPRY_A+SDI} was ordered from Genscript and cloned in a homemade plasmid allowing the expression of the protein in fusion with a hexahistidine-tagged SUMO. More details about these constructs are available in the materials and methods section. The Ash2L^{SPRY_A+SDI} construct contains the SDI and SPRY domains linked by residues that contribute to DPY-30 binding according to recent peptide overlay assays. GST-DPY-30 and His-SMT3-Ash2L^{SPRY_A+SDI} were individually overexpressed in *E. coli*. Application of GST-DPY-30 and His-SMT3-Ash2L^{SPRY_A+SDI} to the appropriate affinity resins served as a first step to purify both proteins. An enrichment of both proteins could be observed with traces of contaminants still present (**figure 7A**). The N-terminal affinity tags, no longer being necessary, were cleaved off to yield more

native-like proteins. The DPY-30/Ash2L heterotrimer was reconstituted using partially purified DPY-30 and Ash2L^{SPRY_A+SDI} at a 2:1 ratio and the sample was applied to a size-exclusion chromatography column to assess the integrity of the complex by evaluating its molecular weight. The sample eluted at a volume of approximately 74.5 mL, near the elution volume of a molecular weight standard of 44KDa, which is in agreement with the molecular weight of a DPY-30/Ash2L^{SPRY_A+SDI} complex of 46.3 KDa. The purity of the fractions forming the major peak were collected and before pooling, were analyzed by polyacrylamide gel electrophoresis and stained by coomassie. As shown in **figure 7B**, both proteins co-elute and DPY-30 migrates near the 50KDa marker, despite its molecular weight being 11.2KDa, while Ash2L^{SPRY_A+SDI} migrates below the 25KDa marker. This discrepancy in the molecular weight of DPY-30 can be explained by the low percentage of SDS in the gel electrophoresis running buffer, which would be unable to disrupt the DPY-30 homodimer. DPY-30 would thus migrate according to its molecular weight as a dimer or oligomers. While the PAGE gel shows that the preparation of the DPY-30/ Ash2L^{SPRY_A+SDI} still contains traces of contaminants, as seen in **figure 7B;D**, the purity was deemed satisfactory to carry out crystallization trials.

2. Crystallization and structure determination of DPY-30/Ash2L^{SPRY+SDI}:

Triangular shaped crystals of the DPY-30/ Ash2L^{SPRY_A+SDI} complex (**figure 7C**) were obtained by vapor diffusion in a solution composed of 10mM citrate pH4.7 and PEG 6000 that diffract at 2.1Å. The structure of the DPY-30/ Ash2L^{SPRY_A+SDI} heterotrimer was solved by molecular replacement using the crystal structures of DPY-30/ Ash2L^{SDI} and Ash2L^{SPRY} as search models. The results of the molecular replacement

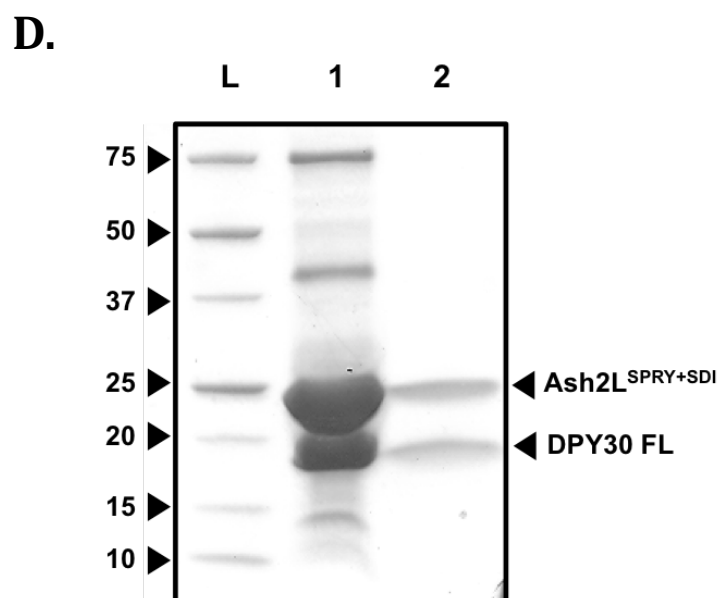
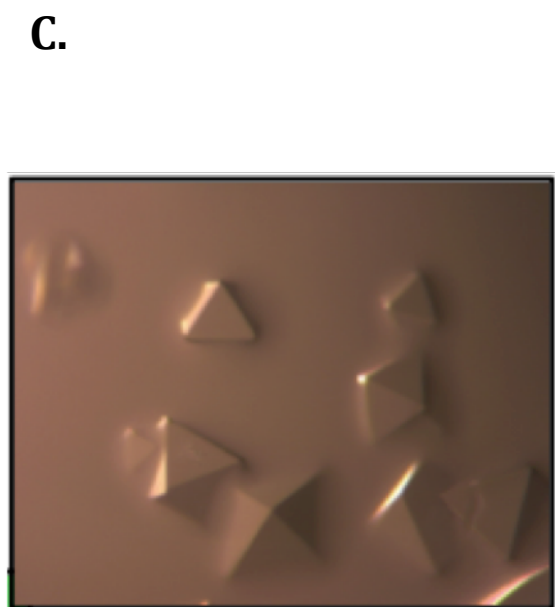
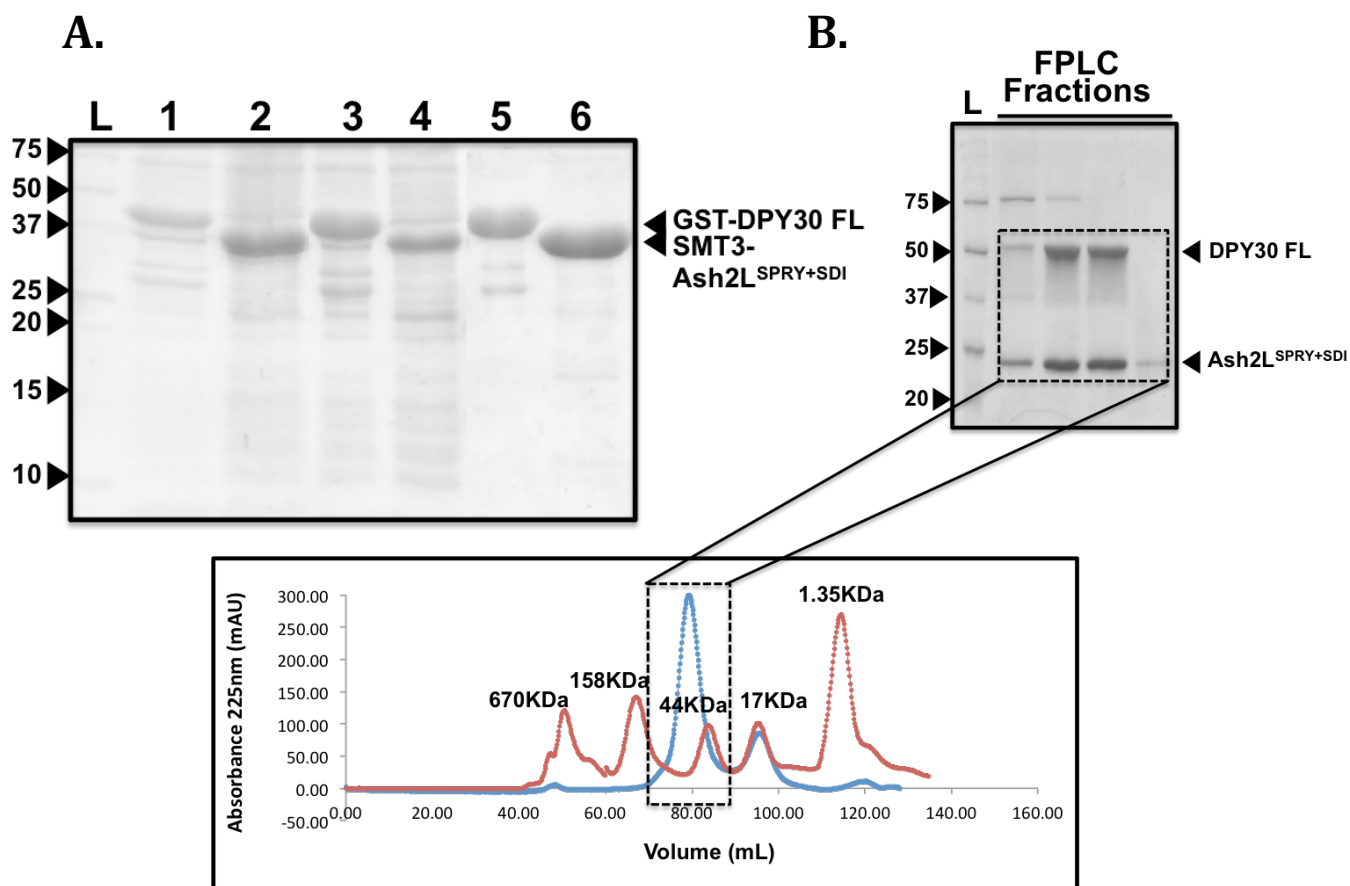


Figure 7. Purification and crystallization of the Ash2L/DPY-30 heterotrimer. A). Ash2LSPRY+SDI and FL DPY-30 were overexpressed in Rosetta cells using the pET expression system. The cells were then sonicated and the crude extract was centrifuged at high speed. The total extract for cells expressing DPY-30 FL (1) and Ash2L^{SPRY+SDI} (2) was separated on an SDS-PAGE gel as well as the supernatant after centrifugation, containing soluble DPY-30 FL (3) and Ash2L^{SPRY+SDI} (4). As a first purification step, the supernatants containing DPY-30 FL and Ash2LSPRY+SDI were applied to a GST resin and a talon cobalt resin respectively. Beads containing purified DPY-30 FL (5) and purified Ash2L^{SPRY+SDI} were also separated on an SDS-PAGE gel. B). Stoichiometric amount of affinity-purified proteins were combined in a 1:2 ratio to form the Ash2L/DPY-30 heterotrimer. This complex was purified by size-exclusion chromatography. The fractions corresponding to the major peak of the FPLC chromatogram were retrieved, concentrated, and used to prepare crystals of the heterotrimer. The curve in blue represents the sample for the heterotrimer, while the curve in red represents standards with known molecular weights that were separated separately. The molecular weights of these markers are indicated. The fractions forming the major peak were separated on an SDS-PAGE gel. The chromatogram is represented on the lower panel displaying the absorbance of protein peptide bonds (225nm) with respect to the volume of buffer eluted through the column. C) Crystals of the DPY-30/Ash2L^{SPRY+SDI} complex. D) Crystals of the Ash2L/DPY-30 heterotrimer were extensively washed in mother liquor, dissolved in water, and separated on a gradient SDS-PAGE gel (2). A sample of purified protein used to prepare the crystals (1) was also separated on gel for comparison.

show that the asymmetric unit contains a single copy of the DPY-30/ Ash2L^{SPRY_A+SDI} complex. The structure was completed with iterative rounds of model building using Coot (Emsley et al., 2010) and refinement using Phaser (Zwart et al., 2008). The final model is refined to an R/Rfree of 17.16%/19.74% and contains one molecule of Ash2L^{SPRY} and a dimer of DPY-30.

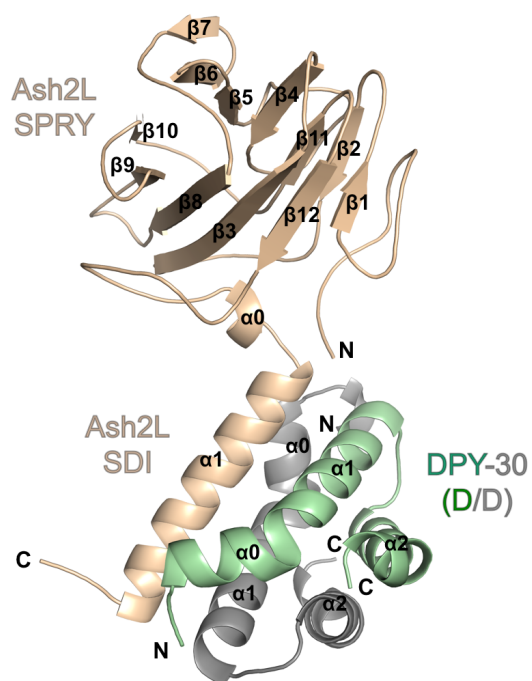
3. Overall view of the DPY-30/Ash2L^{SPRY+SDI} heterotrimer:

The SPRY domain of Ash2L is a globular structure mainly composed of two stacked β -sheets, each containing 6 anti-parallel β -strands, and is followed by a C-terminal SDI α -helix. These two domains fold as two distinct entities, separated by a 4-residue GWGA segment. As previously reported, the D/D domain of a DPY-30 protomer is composed of two α -helices. DPY-30 protomers homodimerize in an anti-parallel arrangement to form a typical X-type four-helix bundle with a hydrophobic crevasse in its center (**figure 8A**). In the refined model, the electronic density for the first 44 residues of DPY-30 is missing and as a consequence these residues are absent in the final model (**Figure 8A**). The DPY-30/ Ash2L^{SPRY_A+SDI} crystals were dissolved in water and separated on an SDS-PAGE gel, and a band corresponding to the molecular weight of full-length DPY-30 was identified, indicating that DPY-30 didn't degrade over time (**Figure 7D**). The N-terminal portion of DPY-30 presumably adopts a random coil conformation and as such are not homogeneously found on the same Bragg's planes. This hypothesis is further supported by secondary structure predictions of DPY-30 N-terminus showing that it is predicted to form a random coil (Dong et al., 2009).

4. Structural analysis of the interactions between DPY-30 and Ash2L:

For the structural analysis of the DPY-30/Ash2L^{SPRY+SDI} structure, residues on DPY-30 will be represented with the 3-letter code, while residues on Ash2L will be referred to with their one-letter code. Hydrophobic residues lining the non-polar face of the SDI helix are nestled into the hydrophobic groove formed by DPY-30. These protein-protein contacts are arranged in such a way that they form a pseudo two-fold symmetry from the central leucine residue (L513) of the SDI helix and onwards. In fact, an axis of rotation could be drawn around this residue, and the 180° rotation of the SDI α -helix could generate a quasi-identical heterotrimeric structure. The alignment of the DPY-30^{D/D}/Ash2L^{SDI} and DPY-30/Ash2L^{SPRY+SDI} structures reveals additional residues prolonging the C-terminus and N-terminus of the SDI helix (**figure 8B**). More specifically, in this new structure, two key residues located on the C- and N-terminus of the SDI α -helix form novel interactions with DPY-30. A C-terminal aspartic acid residue (D525) forms an electrostatic interaction with Thr73 on protomer A of DPY-30. At the N-terminal end of the SDI, a tryptophan residue belonging to the GWGA segment on Ash2L (W505) is contorted in a way that allows its side chain to contact a hydrophobic pocket located on protomer A of DPY-30 formed by residues Ala70, Ala73, and the aliphatic portion of Lys74 (**figure 8C**). Alignment of the amino acid sequence of different species ranging from fish to human shows that this 4-residue segment is conserved throughout evolution, and may thus have an important function (**figure 9**). More importantly, D525 and W505 represent the first asymmetrical elements reported on this otherwise symmetrical structure.

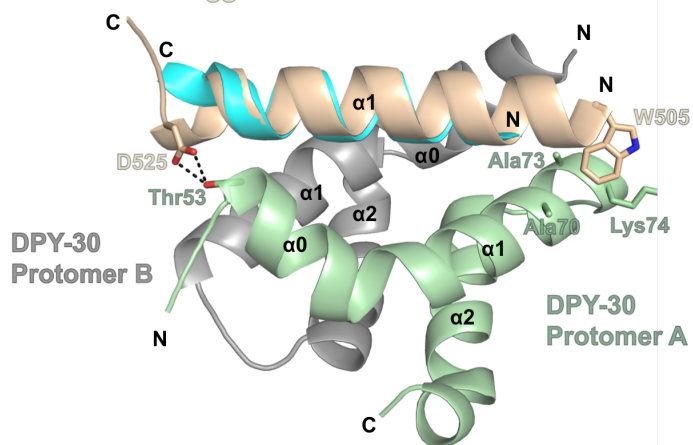
A.



B.

Ash2L SDI tagged V508-V524 (Tremblay et al., 2014)

Ash2L SDI untagged W505-R528



C.

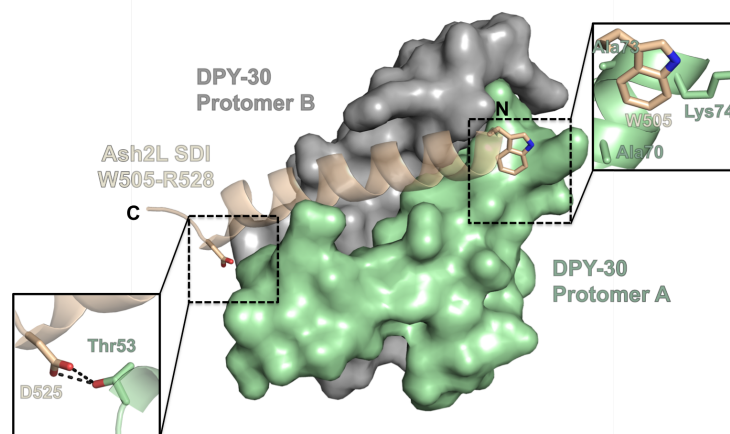


Figure 8. The N- and C-terminal residues on the Ash2L SDI position the SDI helix within the DPY-30 hydrophobic cleft. A) Global view of the DPY-30/Ash2L^{SPRY+SDI} heterotrimer where the protomers of the DPY-30 homodimer (green and gray) and Ash2L (beige) are represented in cartoon. The domains composing Ash2L and DPY-30 are labeled and the α -helices and β -strands are also highlighted. B) Structural alignment of the DPY-30^{D/D}/Ash2L^{SDI} and DPY-30/Ash2L^{SPRY+SDI} in cartoon diagram. The D/D domain is shown in the same colours as previously, while the SDI segment of both structures are represented in different colours and labeled, with the SDI from the Tremblay et al. structure in cyan, and the SDI from the novel structure in beige. The W505 and D525 residues are represented in sticks along with their interacting residues on DPY-30 to highlight the additional interactions formed in the novel crystal structure. C) Zoom-view of Ash2L SDI α -helix bound to the DPY-30 hydrophobic cleft. Proteins and residues are colored as in panel A and B. The inset figures show the interactions between the N-terminal capping SDI residue W505 with an hydrophobic pocket on DPY-30 composed of Ala70, Ala73, and Lys74, and electrostatic interactions between Ash2L D525 residue and Thr73 on protomer A of DPY-30.

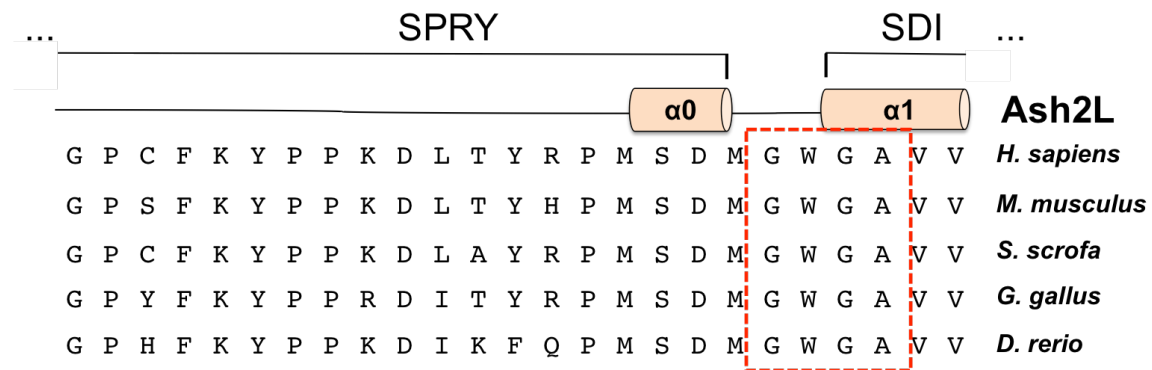


Figure 9. Alignment of the Ash2L sequence of different species. The amino acid sequences of Ash2L belonging to species *H. sapiens*, *M. musculus*, *S. scrofa*, *G. gallus*, and *D. rerio* were aligned and a region encompassing the terminal part of the SPRY domain, the GWGA segment, as well as a few residues at the beginning of the SDI α -helix was shown for all species. The name of the species is indicated to the right of its respective sequence. The red box highlights the GWGA segment that is conserved from fish to human.

5. Role of Ash2L tryptophan 505 in DPY-30 binding:

Tryptophan residues possess many physico-chemical properties. This residue has a bulky side chain, it harbors both hydrophobic and aromatic properties and it may form hydrogen bonds using the amine on its indole ring. To test the role of W505 in the interaction of Ash2L to DPY-30, the residue was mutated to several different residues of varying physico-chemical properties. A W505Y mutant retains the aromaticity and size of a tryptophan residue all the while increasing its polarity with its hydroxyl group. The W505Q mutant conserves, to some level, the bulkiness of a tryptophan residue and adds an amide group that can engage in electrostatic interactions. Mutations of the tryptophan residue to a leucine or alanine confer a hydrophobic character, and while a leucine residue is considered large in size, an alanine residue is the second smallest amino acid. Mutations to a negatively or positively charged residue, such as a glutamic acid or a lysine respectively, were used to completely oppose the non-polar character of tryptophan residues. A mutant was also designed in which the 4-residue GWGA segment was deleted. These mutants were all tested for their ability to bind to GST-DPY-30 using *in vitro* GST pull-down experiments. GST-DPY-30 was immobilized on glutathione beads while different non-GST tagged Ash2L^{SPRY_A+SDI} mutants were added to a resin bound by DPY-30. As shown in **figure 10A**, wild-type His-SMT3-Ash2L^{SPRY_A+SDI} did not bind to the resin coupled with GST tag alone while enrichment was detected by the resin bound DPY-30 confirming that in our buffer conditions, Ash2L^{SPRY_A+SDI} specifically interacts with DPY-30 *in vitro*. Unexpectedly, mutants W505Y, W505Q, W505L, and W505A of Ash2L^{SPRY_A+SDI} bound to DPY-30 as efficiently as the wild-type protein. Complete removal of the GWGA segment or a mutation of the tryptophan to a lysine, however,

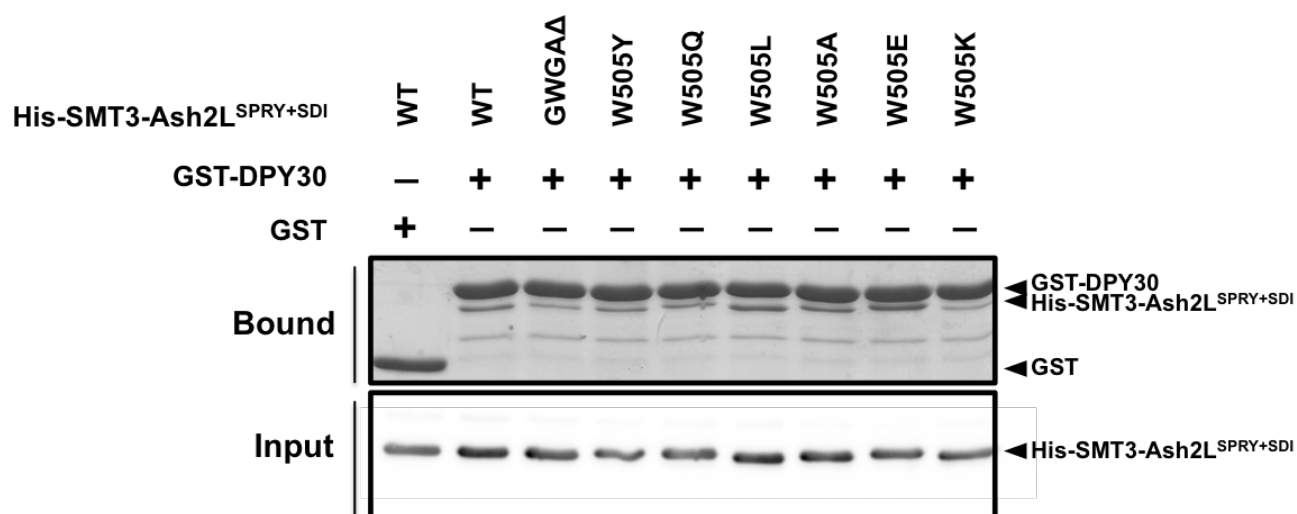
partially reduced binding of Ash2L^{SPRY_A+SDI} to DPY-30 by approximately 50%.

Surprisingly, the substitution of W505 with a negatively charged residue did not abolish the formation of the heterotrimer. The structure of the heterotrimer reveals that a lysine residue (Lys74) on a protomer of DPY-30 is in close proximity to W505. When W505 is mutated to a glutamic acid, this residue could possibly orient itself in such a manner that it could form a salt bridge with Lys74, and compensate for the entropic penalty of losing a hydrophobic contact. To test this possibility, *in vitro* GST pull-down experiments were performed between Ash2L^{SPRY_A+SDI} W505E and GST-DPY-30 K74G. As expected, these mutations partially reduced the binding of DPY-30 to Ash2L^{SPRY_A+SDI} by 50% approximately (**figure 10B**). Overall, our mutational analysis shows that W505 partially contributes to the formation of the Ash2L/DPY-30 complex, *in vitro*.

6. Role of the DPY-30 hydrophobic pocket in Ash2L binding:

After investigating the role of W505, we sought to determine the role of the hydrophobic pocket binding to the tryptophan residue. The methyl groups of two nearby alanine residues (Ala70 and Ala73) on a protomer of DPY-30 form this pocket, along with the aliphatic chain of Lys74. DPY-30's Ala70 and Ala73 residues were mutated to an aspartic acid residue, which is much larger in size, and more importantly carries a negative charge. Lys74 on DPY-30 was mutated to a leucine, which still carries an aliphatic chain and should bind to the same extent as wild-type DPY-30 to Ash2L. A K74G mutant, that only carries a hydrogen atom as side chain, was generated and should reduce binding to Ash2L. *In vitro* GST pull-downs were

A.



B.

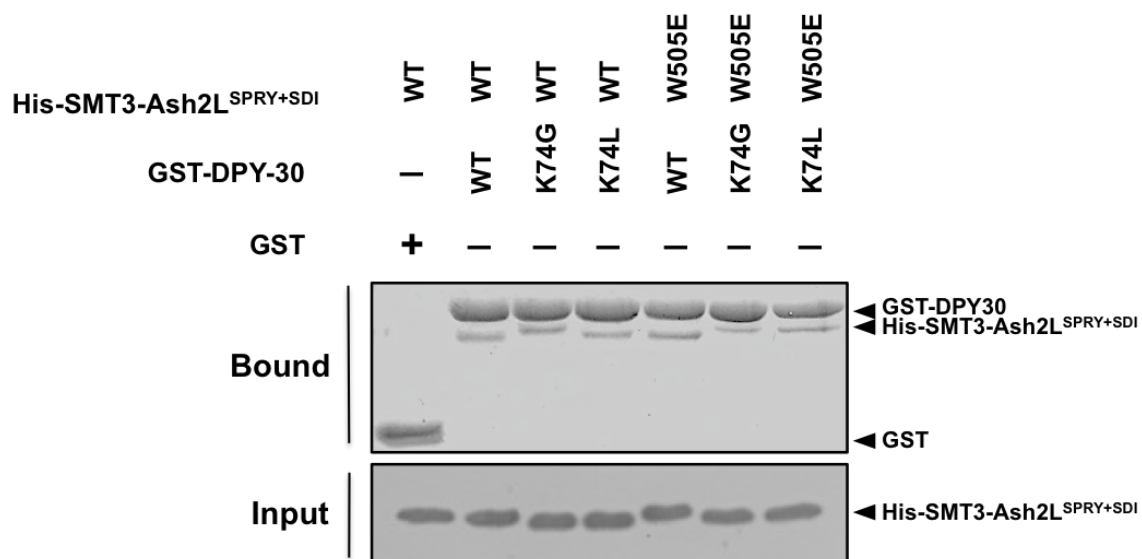
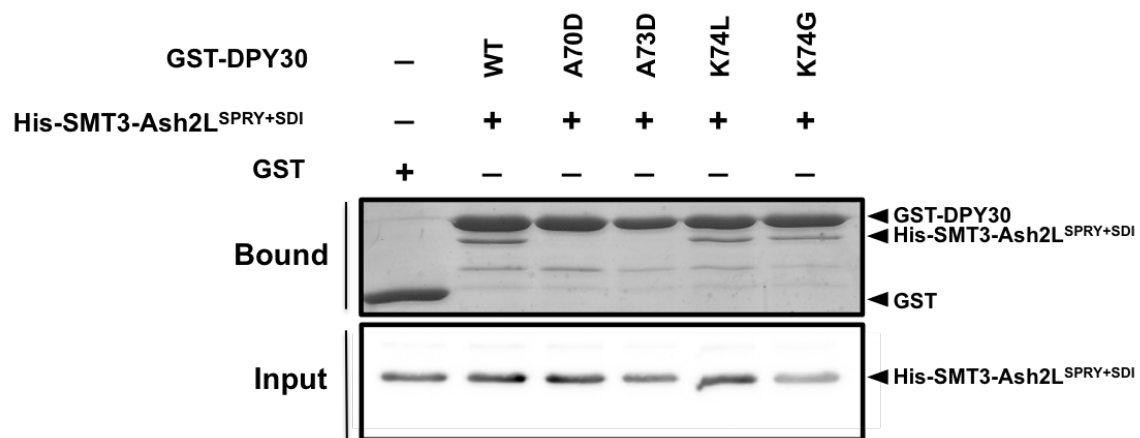


Figure 10. W505 slightly impairs the interaction between Ash2L and DPY-30. A) Binding studies performed on the DPY-30/Ash2L^{SPRY+SDI} complex *in vitro*. *In vitro* GST pull-down assays were performed between various point mutants of W505 on His-SMT3-Ash2L^{SPRY+SDI} and wild-type GST-DPY-30. The GST tag alone was used as a negative control, while wild-type Ash2L was used as a positive control. Coomassie staining was used to visualize proteins and inputs of Ash2L are shown below and detection using an anti-His antibody. B) Mutation of W505 to a negatively charged residue reduces Ash2L-DPY-30 binding when Lys74 is mutated. *In vitro* GST pull-down experiments were performed between the Ash2L^{SPRY+SDI} W505E mutant and the GST-DPY-30 K74G or K74L mutants. Binding experiments between Ash2L W505E and wild-type GST-DPY-30 and between wild-type Ash2L^{SPRY+SDI} and GST-DPY-30 K74G were used as negative controls. Coomassie staining was used to detect the proteins. Inputs are shown below and the anti-His antibody was used for detection.

performed between these GST-tagged DPY-30 mutants and wild-type Ash2L. As shown in **figure 11A**, a complete loss of interaction was observed between both DPY-30 A70D and A73D mutants and Ash2L while the mutations of Lys74 had no effect on the binding between DPY-30 and Ash2L. Overall, these results suggest that A70 and A73, residues found at both extremities of the DPY-30 helical bundle, have an important role in anchoring the SDI α -helix into DPY-30.

To further probe the role of A70 and A73 in the interaction between DPY-30 and Ash2L, we co-transfected human embryonic kidney 293 cell lines (HEK293) with plasmids coding for a flag-tagged copy and an eGFP-tagged copy of DPY-30 corresponding to DPY-30 wild-type or mutants. A tandem immunopurification was performed on the HEK293 cell extracts to isolate DPY-30 homodimers that are uniquely composed of two exogenous copies of DPY-30. Specifically, the cell extracts were applied to beads that are conjugated with anti-flag antibody in order to capture all flag-DPY-30, while a subsequent GFP trap purification of the eluted sample (from anti-flag resin), retains all homodimers containing both flag- and eGFP- tagged DPY-30. As shown in **figure 11B**, mutating Ala70 and Ala73 to aspartic acid lead to a drastic loss in the levels of Ash2L co-immunoprecipitated when compared to wild-type DPY-30 (**figure 11B**). Similar amounts of flag-tagged wild type and mutant DPY-30 were purified by the tandem purification process, which validates the accuracy of the comparisons made between all DPY-30 constructs. Moreover, the similar recovery of all constructs after the two-step purification rules out the possibility that the mutations impeded on the ability of DPY-30 to homodimerize.

A.



B.

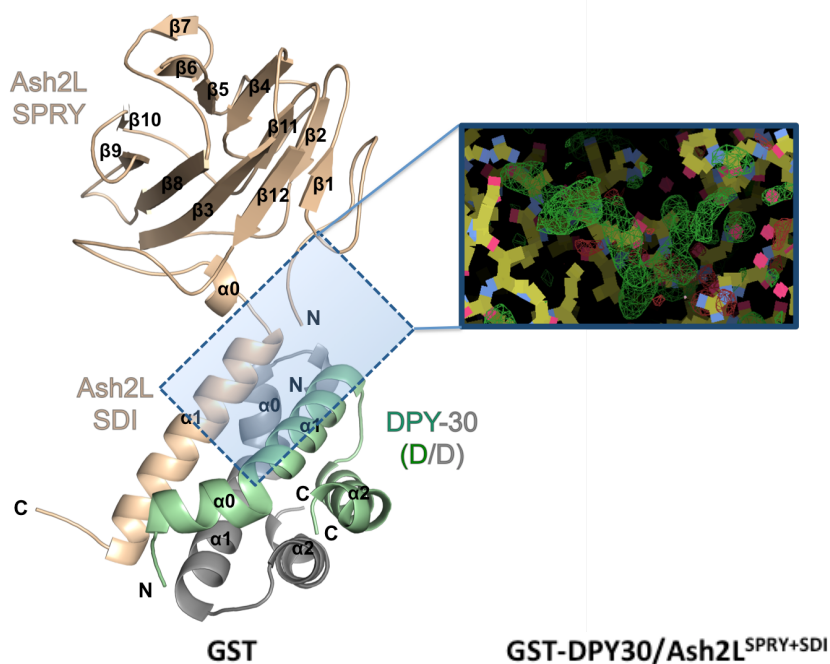
Figure 11. Mutating the hydrophobic pocket on DPY-30 composed of Ala70 and Ala73 completely abrogates binding of DPY-30 to Ash2L. A) Binding studies performed on the DPY-30/Ash2L^{SPRY+SDI} complex *in vitro*. *In vitro* GST pull-down assays were performed between point mutants of GST-DPY-30 and wild-type His-SMT3-Ash2L^{SPRY+SDI}. The GST tag alone was used as a negative control, while wild-type DPY-30 was used as a positive control. Coomassie staining was used to visualize proteins. Inputs of Ash2L are shown below and detection was done using an anti-His antibody. B) Mutating Ala70 and Ala73 on DPY-30 disrupt binding of DPY-30 to Ash2L in cells. HEK293 cells were co-transfected with eGFP- and flag-tagged copies of wild-type or mutant DPY-30. A tandem FLAG and eGFP co-immunoprecipitation was performed on the cell extracts, and bound Ash2L was immunoblotted with an anti-Ash2L antibody. Immunoprecipitated FLAG-DPY-30 and eGFP DPY-30 were detected using an anti-Flag and anti-GFP antibody, respectively. Inputs of DPY-30 and endogenous Ash2L are also shown.

Altogether, our mutational analysis show that Ala70 and Ala73 are crucial for the assembly of the DPY-30/Ash2L heterotrimer in HEK293 cells and likely in other cellular context.

7. Ash2L/DPY-30 binds to anionic lipids:

Upon closer inspection of the electronic density map of the DPY-30/Ash2L^{SPRY_A+SDI} heterotrimer, we identified a positive fourier map at the interface between DPY-30 and Ash2L that did not belong to either the Ash2L or DPY-30 polypeptides. In fact, we attempted to fit a generic peptide into this fourier map and it failed to refine. The shape of this fourier, however, was reminiscent of the structure of a lipid head group (**figure 12A**). Considering this hypothesis, it seemed, then, that DPY-30 and Ash2L could form a joint lipid-binding patch upon the formation of the heterotrimer. To test the possibility that the heterotrimer binds to lipids, the GST-DPY-30/ Ash2L^{SPRY_A+SDI} heterotrimer was incubated with strips coated with a variety of phospholipids and phosphatidylinositols at varying pHs. The heterotrimer bound to a few anionic phospholipids, including phosphatidylinositols (PI), phosphatidylserine (PS), phosphatidic acid (PA), and phosphatidylglycerol (PG), at a pH of 4.7. As shown in **figure 12B**, the heterotrimer interacts preferentially with cardiolipin (CL). Binding to cardiolipin would suggest that during the protein purification process, cardiolipin released from the outer membrane of *E. coli* bound to Ash2L to ultimately co-crystallized with it. Binding assays performed at different pHs also revealed that the binding of GST-DPY-30/ Ash2L^{SPRY_A+SDI} is optimal at lower pH. For this reason, all subsequent experiments were performed at a pH of 4.7. While performing binding studies with lipid arrays can help identify novel lipid binding activity, this approach relies on the use of lipids that are tethered to a solid matrix when in reality; the majority of lipids form

A.



B.

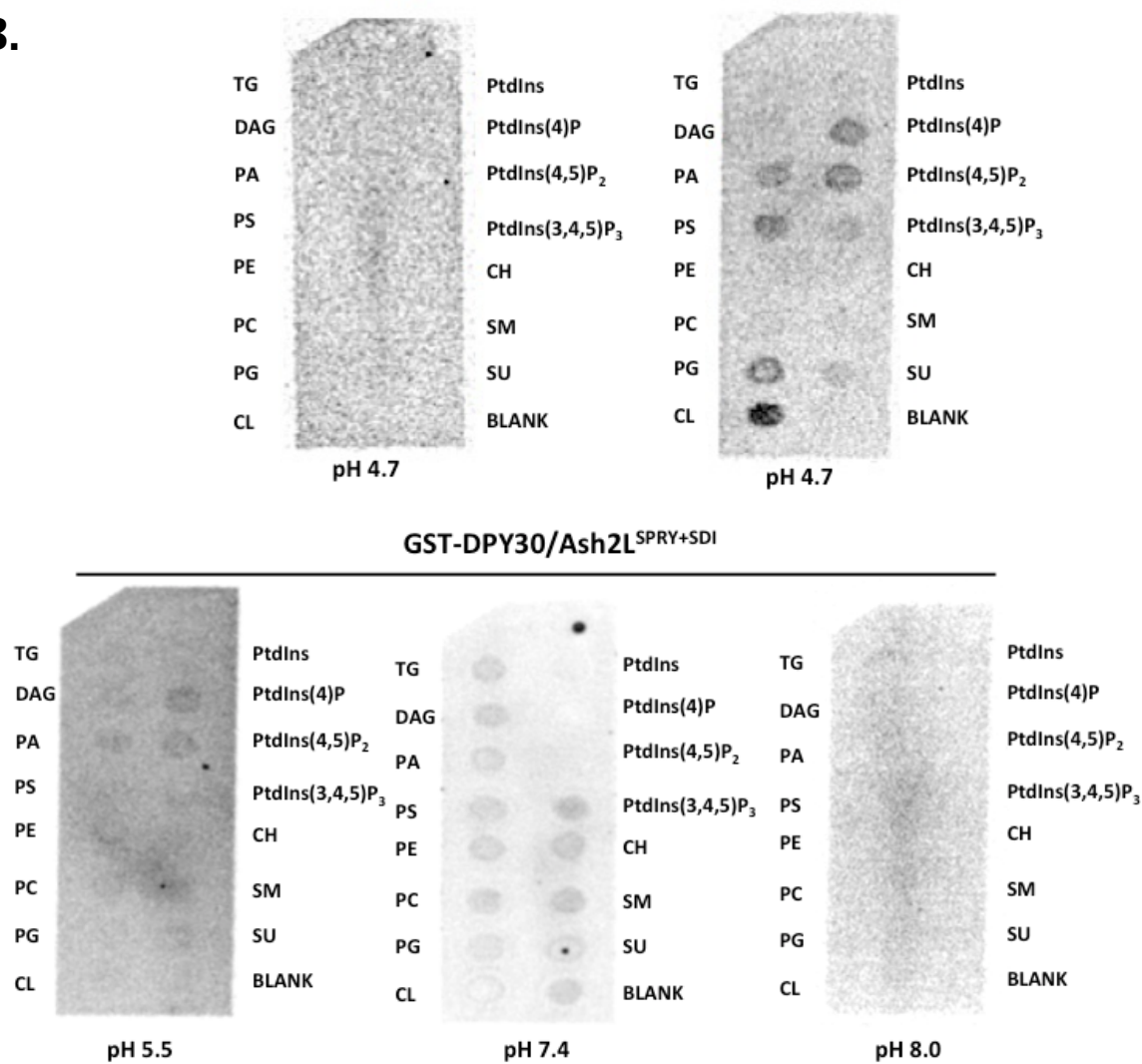


Figure 12. The DPY-30/Ash2L heterotrimer binds to anionic lipids in vitro. A) Structure of the DPY-30/Ash2L^{SPRY+SDI} heterotrimer highlighting the region in the DPY-30/Ash2L interface wherein a positive fourier resembling a lipid is found. The insert projecting from this area shows the positive fourier obtained from COOT. B) Lipid-protein overlay assay in which purified GST-DPY-30/Ash2L^{SPRY+SDI} heterotrimer was incubated with phospholipid coated strips at pHs of 4.7, 5.5, 7.0, and 8.0. pHs are indicated below the lipid strips. GST alone was used as a negative control. Binding was detected using an HRP coupled anti-GST antibody. Abbreviations: TG = triglycerides, DAG = diacylglycerol, PA = phosphatidic acid, PS = phosphatidylserine, PE = phosphatidylethanolamine, PC = phosphatidylcholine, PG = phosphatidylglycerol, CL = cardiolipin, PI = phosphatidylinositol, PtdIns(4)P = phosphatidylinositol-4-phosphate, PtdIns(4,5)P₂ = phosphatidylinositol-4,5-biphosphate, PtdIns(3,4,5)P₃ = phosphatidylinositol-3,4,5-triphosphate, CH = cholesterol, SM = sphingomyelin, SU = sulfatide.

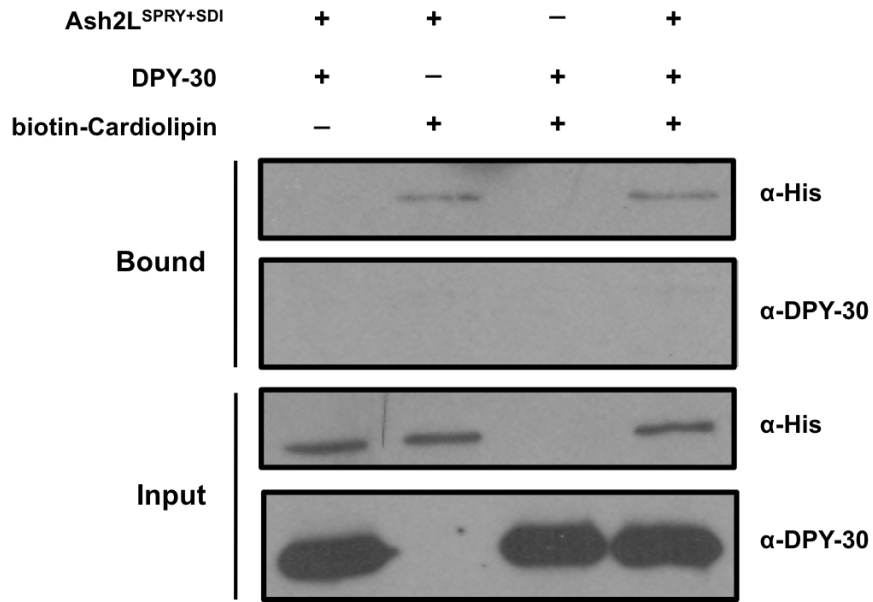
micellar structures. Moreover, these overlay assays did not clarify whether both DPY-30 and Ash2L obligatorily need to form a complex to bind to lipids, or if they could accomplish this task individually.

8. Ash2L has lipid-binding properties and binding to lipids and DPY-30 is mutually exclusive:

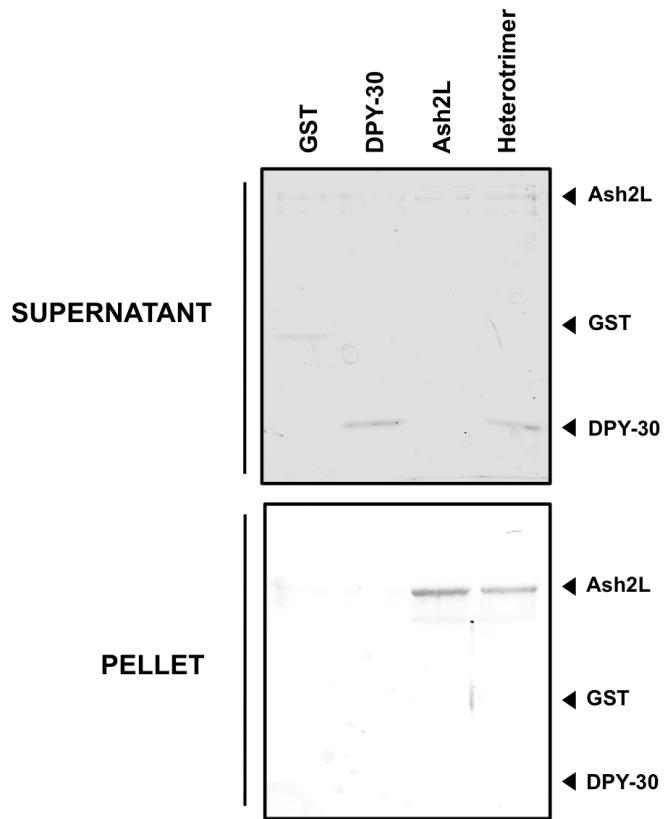
CL is the strongest hit in the lipid-binding screen, therefore a water-soluble version of CL, biotin-CL, was used to answer these preliminary questions. Biotinylated-CL was either combined with a DPY-30/His-SMT3-Ash2L^{SPRY_A+SDI} complex, the DPY-30 homodimer, or with His-SMT3-Ash2L^{SPRY_A+SDI}. Biotin-CL was then captured on streptavidin-coated beads, and protein binding to CL was probed by western blot. As shown in **figure 13A**, His-SMT3-Ash2L^{SPRY_A+SDI} alone was able to bind to CL to the same extent as it does in complex with DPY-30, while DPY-30 alone did not bind to CL. Interestingly, when the His-SMT3-Ash2L^{SPRY_A+SDI}/DPY-30 heterotrimer was incubated with biotinylated CL, a loss of interaction between both proteins was observed. The dissociation of DPY-30 from the complex indicates that CL can compete with DPY-30 for Ash2L binding. These observations suggest that DPY-30 and CL have overlapping binding sites on the Ash2L SDI α -helix and the interaction of either molecule to Ash2L may preclude the binding of the other. Interestingly, lipid-binding proteins often bind lipids that are embedded into phospholipid membranes. For this reason, liposomes were produced *in vitro* that are composed of a phosphatidylcholine (PC) backbone and a given lipid of interest at a ratio that reflects the lipid content of biological membranes. The binding of recombinant full-length Ash2L isoform 1 (residues 1-627), DPY-30, and the

Ash2L-DPY-30 complex to CL-containing liposomes was tested *in vitro*. As a negative control, liposomes containing phosphatidylethanolamine (PE) were used, which according to the results presented in **figure 13B**, does not interact with the Ash2L/DPY-30 heterotrimer. As previously observed, Ash2L alone, but not DPY-30 was able to interact with CL-containing liposomes. Intriguingly, Ash2L showed binding to the PE-containing liposomes in this assay, possibly due to long incubation times (**figure 13B**). Overall, these experiments suggest that DPY-30 and CL compete with each other for binding to Ash2L. To test this hypothesis, GST-DPY-30 and Ash2L were subjected to *in vitro* GST pull-down assays, with an increasing concentration of CL-containing liposomes. The experiment was repeated with increasing quantities of PE-containing liposomes as a negative control. As the concentration of CL-containing liposomes was increased, GST-DPY-30 decreasingly pulled-down Ash2L until no binding could be detected by coomassie staining. This effect was not observed with PE-containing liposomes, showing specificity in the Ash2L-CL interaction (**figure 13C**). The loss of Ash2L binding to DPY-30 with increasing CL concentration supports the mutually exclusive binding of CL and DPY-30 to Ash2L.

A.



B.



C.

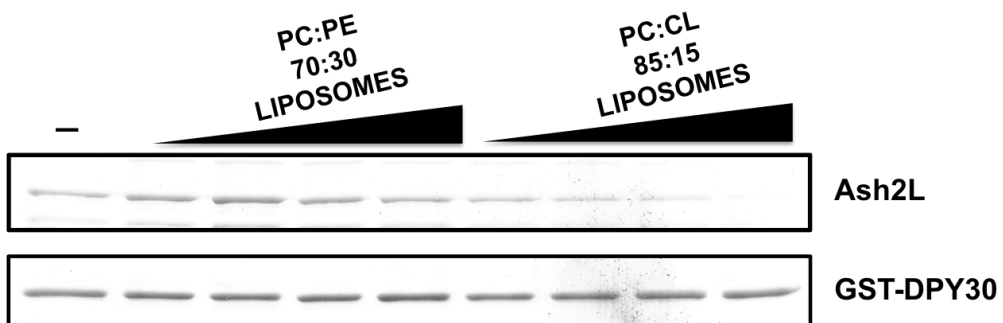


Figure 13. Ash2L binding to cardiolipin and DPY-30 is mutually exclusive. A) *In vitro* pull-down experiments with biotinylated cardiolipin (CL) were conducted in solution with purified DPY-30, His-SMT3-Ash2L^{SPRY+SDI}, or the heterotrimer. A negative control was also included in which GST-tagged heterotrimer was incubated with streptavidin beads without biotinylated cardiolipin. Inputs are indicated below. Bound Ash2L was detected with the anti-His antibody while bound DPY-30 was blotted with the anti-DPY-30 antibody. B) Liposome binding assays were performed using CL-containing liposomes with purified DPY-30, Ash2L, or the DPY-30/Ash2L heterotrimer. The protein in the supernatant (unbound) and pellet (bound) fractions following liposome pelleting were separated on a gel and visualized with coomassie solution. C) *In vitro* GST-pull-down assay performed between GST-DPY-30 and Ash2L in the presence of an increasing amount of phosphatidylethanolamine (PE)-containing liposomes (negative control) or CL-containing liposomes. A binding reaction between Ash2L and DPY-30 without liposomes was also performed as a negative control and is represented by the (-) symbol. Coomassie staining was used to visualize proteins.

9. A construct encompassing the SPRY and SDI domains of Ash2L is sufficient for CL-binding:

To test whether Ash2L displays the same lipid binding activities in the context of liposomes, the binding of Ash2L to CL-, PA- and PS-containing liposomes was tested. Ash2L was incubated with these different liposomes, separated by centrifugation and the supernatant/pellet fractions were separated on an SDS-PAGE gel and stained with coomassie. The quantity of Ash2L bound to the liposomes, i.e. pelleted along with the liposomes, versus the quantity of Ash2L remaining in the supernatant was measured and the percentage of Ash2L bound to the liposomes was then calculated, as described in materials and methods. As shown in **figure 14**, nearly 60% of Ash2L binds to CL-containing liposomes while only 9% of binding can be detected between the protein and PS-containing liposomes. Similarly, a small enrichment (3%) was observed when binding reactions were performed with PA- and PE-containing liposomes. To delineate the CL binding site on Ash2L, liposome-binding assays were performed with Ash2L constructs corresponding to Ash2L^{SPRY_Δ+SDI} or containing only the SPRY domain (Ash2L^{SPRY}). Akin to full-length Ash2L, approximately 60% of Ash2L^{SPRY_Δ+SDI} co-sediment with liposomes reconstituted with CL and show minimal interaction with PE-, PA-, and PS-containing liposomes. However, in contrast to full-length Ash2L, only approximately 25% of Ash2L^{SPRY} was retained by CL-containing liposomes suggesting that the SPRY domain and the SDI helix of Ash2L are both important for optimal binding to CL (**figure 14**).

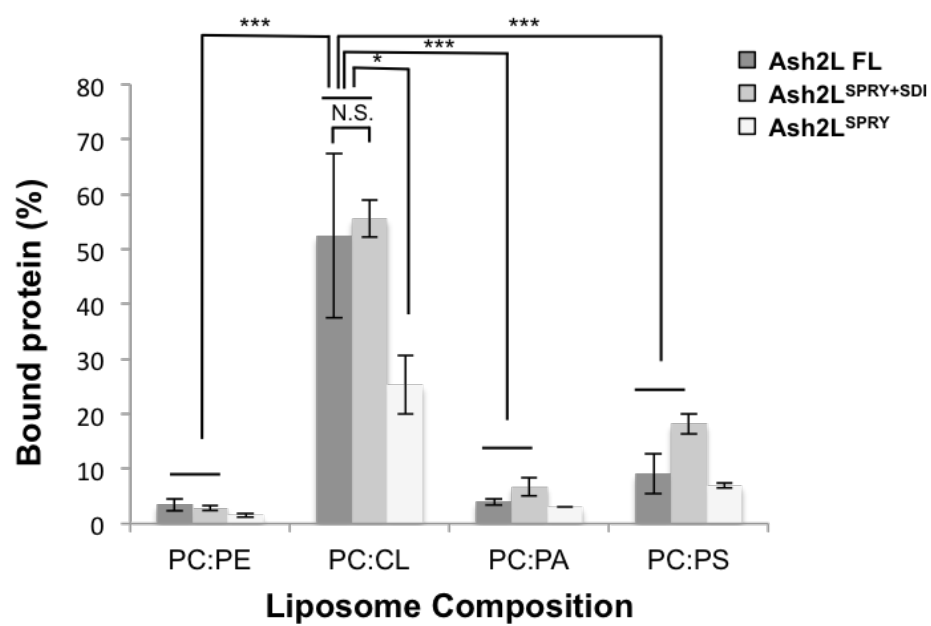
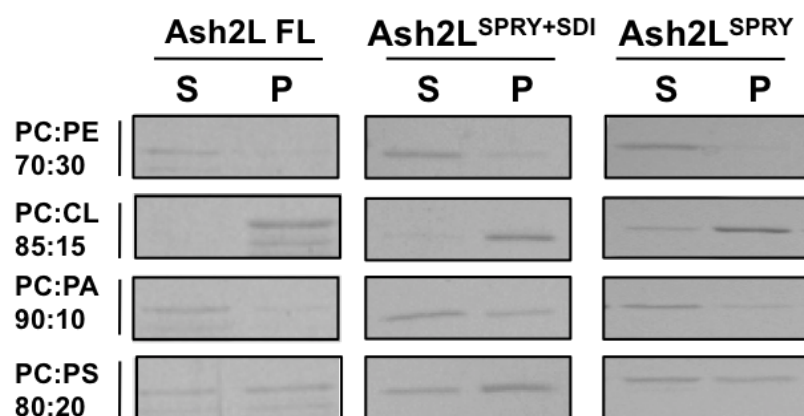


Figure 14. The SPRY and SDI domains of Ash2L are important for binding to cardiolipin. Liposome binding assays were performed between purified full-length Ash2L, Ash2L^{SPRY+SDI}, or Ash2L^{SPRY} and PE-, CL-, PA- or PS-containing liposomes. Bound (pellet) and unbound (supernatant) fractions were separated on a gel and stained with coomassie solution (Upper panel). Experiments were performed in triplicate for Ash2L^{SPRY} and Ash2L^{SPRY+SDI}, and in quadruplicate for full-length Ash2L. The percentage of bound protein was calculated as described in materials and methods, and represented in a bar graph, where Ash2L FL is in dark grey, Ash2L^{SPRY+SDI} in light grey, and Ash2L^{SPRY} in white. Vertical bars represent the standard deviation (lower panel). For statistical significance: * = p<0.05, ** = p< 0.01, *** = p< 0.001 (see appendix).

DISCUSSION

1. Ash2L/DPY-30 Heterotrimer: Insights into the orientation and binding of Ash2L to DPY-30.

Initial structural studies of Ash2L^{SDI} (residues V509-V524) bound to the DPY-30 complex highlighted several structural features important for the formation of the Ash2L/DPY-30 complex. However, this structure presented important shortcomings. The authors crystallized only fragments of both proteins which included only DPY-30 X-helix bundle and residues V509-V524 of Ash2L. Intriguingly, the same study (Tremblay et al., 2014), along with others (South et al., 2010), had shown that residues before V509 are important for high affinity binding. Therefore, the initial crystal structure published by Tremblay et al. did not capture all the interactions between Ash2L and DPY-30. Moreover, the Ash2L^{SDI} crystalized with DPY-30, was flanked by tag remnants located at its N- and C-termini which could alter the positioning of Ash2L^{SDI} into the DPY-30 X-bundle or affect the folding of the helix. The new structure presented in my thesis addresses these two issues as the construct used in my study is not tagged at its C-terminus and is preceded by the SPRY domain at its N-terminus thereby preserving the residues connecting the SPRY and SDI domains as well as residues C-terminal of the SDI. The structure is thus representative of the correct fold and length of the SDI helix and how it inserts into the D/D of DPY-30, and likely captures all possible interactions formed between both proteins. Overall, the DPY-30/Ash2L^{SDI} and DPY-30/Ash2L^{SPRY+SDI} structures align well, especially the DPY-30 X-helix bundles which superpose perfectly. The SDI helix from the DPY-30/Ash2L^{SDI} structure however is slightly closer to the DPY-30 X-bundle than the SDI helix from the DPY-

30/Ash2L^{SPRY+SDI} structure. The SPRY and SDI domains do not contact each other but are positioned in such a way that they collectively form a semi-circle, at the base of which DPY-30 binds. Rather than DPY-30 and Ash2L creating an alternative compact platform, as can be inferred from the cryo-EM structure of the reconstituted MLL1 complex (Takahashi et al., 2013), Ash2L utilizes its SDI domain as a recruitment platform to incorporate DPY-30 into SET1 complexes.

Despite setting up crystal trays with full-length DPY-30, we failed to observe electronic density for the first 44 residues of DPY-30. Based on this structural data, this region, predicted to be unstructured (Dong et al., 2005), does not seem to make any interactions with the SDI or the SPRY domains of Ash2L. These observations are further supported by pull-down assays showing that DPY-30 D/D domain alone recapitulates the binding of full-length DPY-30 to Ash2L and its yeast homolog Bre2 (South et al., 2010). Intriguingly, isothermal titration calorimetry experiments performed between DPY-30 and Ash2L, or DPY-30 D/D and Ash2L, showed that full-length DPY-30 binds to Ash2L with 10-fold more affinity compared to DPY-30 D/D (Zhang et al., 2015). While this data suggests that few residues in the first 44 residues of DPY-30 contribute to binding to Ash2L, the impact is relatively minor as a difference of 10-fold in the K_a corresponds to a $\Delta\Delta G$ of less than 1kJ/mol; a value that can be attributed to the contribution of few van der Waals. Collectively, combined with our observations that the first 44 amino acids are not conserved throughout evolution, these data suggest that the first 44 residues of DPY-30 are dispensable for its function in the biology of the SET1 family of methyltransferases *in vitro*. Knocking out DPY-30, using the CRISPR/Cas9 system,

followed by the re-expression of a DPY-30 construct lacking the first 44 residues would help evaluating the impact of the N-terminus on yet to discover novel non-nuclear functions of DPY-30.

Our crystal structure of DPY-30/Ash2L^{SPRY-SDI}, provided several novel insights into the binding of DPY-30 to Ash2L. Two new interactions, provided by residues that cap the SDI at its N- and C-terminus were found to interact with DPY-30 which include W505 and D525, respectively. The tryptophan residue places its indole ring in a hydrophobic pocket composed of Ala70, Ala73, and the aliphatic portion of Lys74, found on the longest helix of a single protomer of DPY-30. On this same helix, but at the other extremity, Thr53 forms an electrostatic interaction with D525 located at the C-terminus of the SDI α -helix. The importance of W505 was confirmed by mutational analysis showing that its substitution to a lysine residue leads to a partial reduction in the binding of DPY-30 to Ash2L. Moreover, this residue is conserved throughout evolution in different metazoans and a long segment at the junction between the SPRY and SDI domains of yeast Ash2L (Bre2) has also been shown to be crucial for DPY-30 (Sdc1) binding (South et al., 2010). Analysis of the sequences of predicted DPY-30 interactors (Tremblay et al., 2014) showed that these proteins have a cyclic/aromatic residue, 4 residues upstream of their putative “SDI” amphipathic α -helix. In fact, BAP18 and HDAC1, two known interactors of DPY-30, harbor a tryptophan and a tyrosine residue at this position, respectively; meanwhile BIG-1 and NCOA6 have a proline residue at this position (**figure 15**). The presence of an aromatic residue or a proline residue may thus be a key signature for the interaction of an

Putative SDI-like helix		
W	G A V V E H T L A D V L Y H V	Ash2L
W	T E T E I E M L R A A V K R F	BAP-18
P	G G A L P T S V R S I V T T L	NCOA6
Y	E A I F K P V M S K V M E M F	HDAC1
P	Q D I V Q N I V E E M V N I V	BIG1

Figure 15. DPY-30 interactors have an aromatic/proline residue preceding their SDI-like α -helices. The amino acid sequences of putative SDI-like α -helices from predicted DPY-30 interactors were aligned with the sequence of Ash2L. A box was traced around the residue preceding the putative SDI-like α -helix showing that this residue either an aromatic residue or a proline. The names of the proteins are indicated to the right of their respective sequences.

amphipathic α -helix to DPY-30. When residues Ala70 and Ala73 were mutated, a loss of interaction with Ash2L was observed *in vitro* and in HEK293 cells. While these results may suggest that the two alanine residues contribute to W505 binding, the impact of the mutation may have more pervasive effects. In fact, Ala70 and Ala73 are found in proximity of Val509 and Val524 of Ash2L; residues found near the N-terminus and the C-terminal end of the SDI helix, respectively. These observations suggest that the A $\rightarrow$ D mutations may also negatively impact the binding of these hydrophobic residues within DPY-30 hydrophobic cleft. This idea is supported by a previous study showing that the substitution of Leu69 in DPY-30, a residue that contributes to the stabilization of Ash2L SDI hydrophobic ridge within its dimerization docking domain, also leads to a complete loss of binding between Ash2L and DPY-30.

In contrast to the substitution of the Ala70 and Ala73 of DPY-30, the substitution of W505 to a positively charged residue only causes a mild reduction in the interaction between DPY-30 and Ash2L. One possible explanation for this discrepancy is that W505 may be involved in controlling the directionality of the SDI α -helix within the DPY-30 hydrophobic cleft as opposed to solely contributing to the interaction between both proteins. Intriguingly, this type of mechanism significantly differs from other X-bundle binding proteins. In fact, previous biochemical and structural studies showed that the 4-helix bundle proteins RI α and vinculin harbor specific binding pockets tailored to fit the shape of the side chains located on the hydrophobic ridge of the amphipathic α -helices of D-AKAP2 (bearing the motif LAxxIAxxIVxxVM) and RIAM (bearing the IxxMFxxLLxxMxxLTxxL) respectively (Sarma et al., 2009, Goult et al., 2013). In the D-

AKAP2/R1 α complex, two alanine residues on the N-terminal half of D-AKAP2 fit into shallow pockets on one side of the R1 α binding cleft, which could not accommodate larger side chains. In another example, a phenylalanine residue on the RIAM α -helix inserts into a deep pocket only found in one extremity of the vinculin four-helix bundle, and dictates the orientation of the helix within the four-helix bundle. Similarly, in a pilotin-secretin complex, three key hydrogen bonds, among which two involve residues located at the N- and C-termini of the amphipathic helix and one in the center, maintain the α -helix in a specific orientation within the pilotin four-helix bundle (Gu et al., 2012). Rotating this helix by 180° would place these hydrogen-bonding residues in a non-favorable conformation in the cleft. Our crystal structure suggests that W505 and D525 interact with both extremities of the longest α -helix in DPY-30. These unique asymmetric features, unlike the nearly symmetric hydrophobic residues forming the DPY-30 binding motif (VxxTLxxVLxxV), may confer binding of Ash2L^{SDI} to DPY-30 in a single orientation. Interestingly, interactions occurring on a single protomer is reminiscent of the RII α /D-AKAP2 complex wherein a loop on a single protomer of the RII α dimer forms unique interactions with D-AKAP2 amphipathic α -helix (Kinderman et al., 2006). Overall, the crystal structure of the Ash2L^{SPRY+SDI}-DPY-30 complex strongly suggest that W505 and D525 residues found at the N- and C-terminus of the SDI helix ensure that the nearly symmetrical amphipathic α -helix of Ash2L (VxxTLxxVLxxV) binds to DPY-30 in an asymmetric fashion.

While our new crystal structure of the DPY-30/Ash2L^{SPRY+SDI} complex unravels the role of the segment preceding the SDI, it raises interesting questions in regard to the

role of DPY-30 in controlling global histone H3K4 tri-methylation. In fact, several studies established that the deletion of DPY-30 impacts global histone H3K4 methylation and does so in several organismal models. Intriguingly, structural alignment of the DPY-30/Ash2L^{SPRY+SDI} heterotrimer onto the Ash2L SPRY domain co-crystallized in complex with MLL3 core complex places DPY-30 far from the SET domain of MLL3, and is likely unable to contact the SET domain of MLL3 to directly stimulate its enzymatic activity (**figure 16**). Moreover, our laboratory and several others could not recapitulate the importance of DPY-30 for H3K4 tri-methylation observed in knockdown studies in cells lines when performing methyltransferase assays with the SET domain of SET1 enzymes and the WRAD subcomplex (Patel et al., 2008, Patel et al., 2009). Considering that previous studies revealed that the first 44 residues on DPY-30 are dispensable for histone H3K4 tri-methylation, our analysis begs the question: how does DPY-30 allosterically regulate histone H3K4 tri-methylation? Several scenarios can be considered. Subtle differences exist in the structure of the SET domains of different SET1 enzymes, thus DPY-30 may have more affinity for certain SET1 paralogues and specifically stimulate their tri-methyltransferase activity. DPY-30 could directly contact the SET domain of SET1A/SET1B and promote H3K4 tri-methylation by stabilizing the SET domain to allow multiple rounds of methylation or by slightly restructuring the active site of the SET domain to accommodate a tri-methylated lysine on H3K4. DPY-30 could also accomplish this effect by directly interacting with tri-methylated H3K4 in holo-SET1A/B complexes and stabilizing the extra methyl group via hydrophobic interactions. Early studies in yeast and quantitative mass spectrometry in mammalian cells identified additional subunits in SET1 complexes other than WRAD, some of which

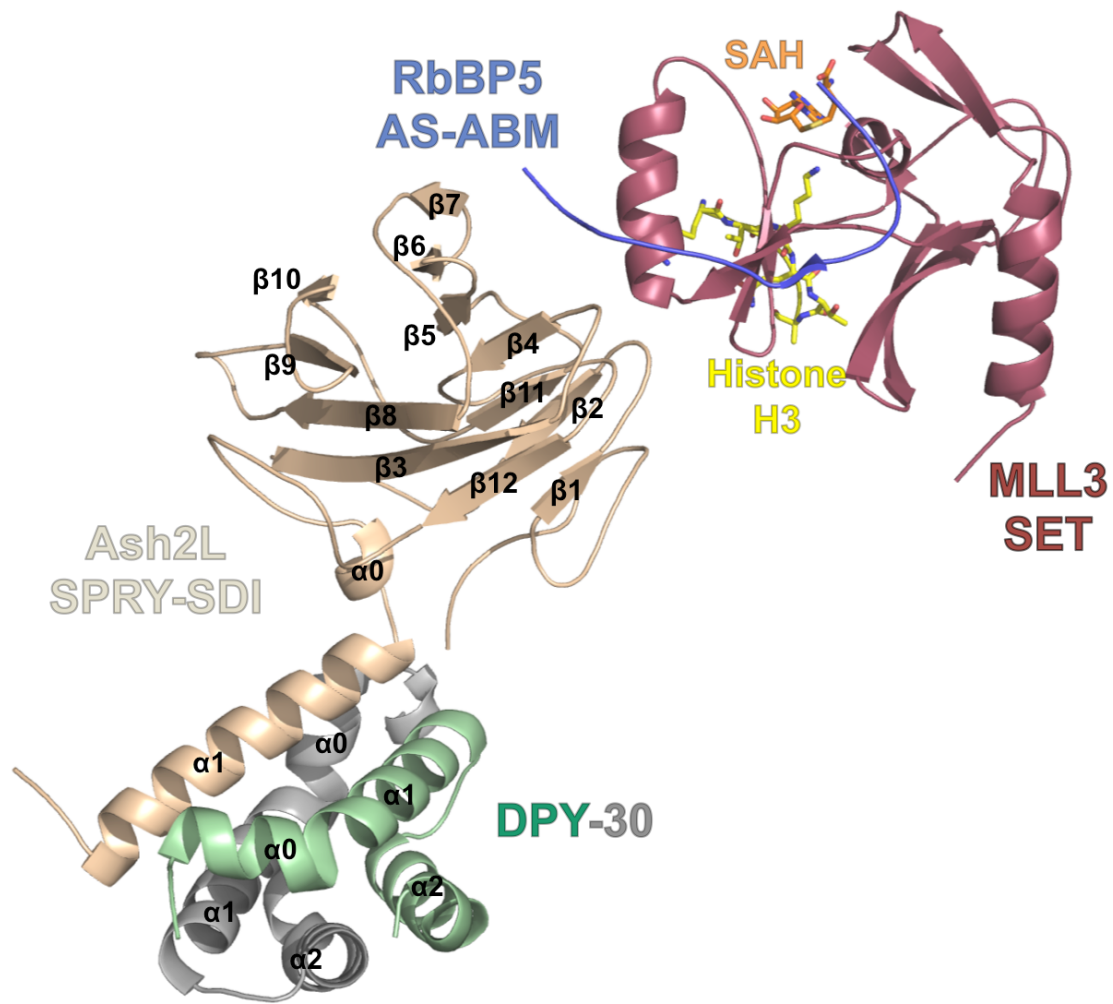


Figure 16. DPY-30 is physically distant from MLL3 SET domain. The crystal structure of MLL3 SET domain in complex with RbBP5 AS-AMB and Ash2L SPRY (PDB: 5f6k) and the DPY-30/Ash2L^{SPRY+SDI} were aligned using the SPRY domain from both structures as a template, and the product of the alignment is represented in cartoon diagram. All proteins are labeled and represented in different colors.

may cooperate with DPY-30 to enable its function (Dehe et al., 2006, Van Nuland et al., 2013). DPY-30 could function in SET1A/SET1B complexes by recruiting WDR82, a subunit that has been shown to specifically interact with SET1A/B and be crucial for their tri-methyltransferase activity on H3K4 (Wu et al., 2008).

This last suggestion is in agreement with the inability to reconstitute SET1/MLL complexes that tri-methylate H3K4 *in vitro*. WDR82 and its potential interaction with DPY-30 may be the missing link for the stimulation of the tri-methyltransferase activity of SET1A/B *in vitro*. However, additional structural and biochemical studies are needed to answer these questions.

2. Structural and biochemical characterization of the lipid-binding function of Ash2L and its impact on the assembly of the Ash2L/DPY-30 heterotrimer:

While solving the structure of the DPY-30/Ash2L heterotrimer, we observed a positive fourier at the Ash2L^{SDI}-DPY-30^{D/D} interface that did not belong to the protein and resembled the shape of a lipid head group. Initial binding studies using lipid arrays revealed that the Ash2L^{SPRY+SDI}-DPY-30 heterotrimer interacted with anionic lipids with a preference for the mitochondrial lipid cardiolipin. Liposome pelleting assays further confirmed the preferential binding of Ash2L to anionic lipids. We also showed that the pH could modulate the binding of the heterotrimer to lipids, as binding was only observed at a pH of 4.7. Pull-down experiments performed between the heterotrimer and CL indicated that Ash2L alone is sufficient for binding to cardiolipin. Moreover, a

construct of Ash2L encompassing the SPRY and SDI domains are essential for binding cardiolipin. Finally, binding experiments also suggested that Ash2L does not interact with lipids in complex with DPY-30, which was validated by an *in vitro* pull-down assay that showed a loss of interaction between Ash2L and DPY-30 upon addition of CL-containing liposomes.

Our lipid binding assays showed that Ash2L binds more efficiently to anionic lipid at an acidic pH of 4.7. In light of this result, we attempted to find residues with a pKa matching the pH profile of Ash2L binding to CL. Among selected residues, we found the side chain of glutamic acid E510 in close proximity to the positive peaks in the fourier transformation. However, substitution of this glutamic acid did not affect the ability of Ash2L to bind to CL-containing liposomes (data not shown). We also identified several residues that may coordinate for the binding of the lipids which include H329, H357, H511, and H519. Mutation of these residues to alanine followed by pull-down experiments using biotinylated CL would help us identify residues that confer CL binding activity to Ash2L. Another possibility is that our *in vitro* conditions may not be conducive to see the impact of the mutation. For instance, in the liposome binding assay, the concentration of lipids may be in saturating conditions for both the wild-type and the mutant and would thereby mask the impact of the mutations. Repeating the pull-down experiments using lower concentration of liposomes and Ash2L constructs harboring a single or several substitutions would enable us to identify the ionizable residues important for CL binding.

The impact of the pH on the lipid binding activity of a protein has been documented. Previous studies showed that a positively charged surface on cytochrome c constitutes a lipid-binding site that only appears at a pH lower than 7.5 (Kawai et al., 2005). Similarly, a large quantity of the Heat Shock Protein A1A (HSPA1A) (or HSP70) interacts with anionic lipid-containing liposomes (cardiolipin and BMP) when the binding reactions are performed at a pH of 4.5 (McCallister et al., 2016). Combined with our findings, these examples show that the pH can regulate the binding of lipids by proteins and potentially control its localization in a particular microenvironment or an organelle.

Despite having failed to identify the ionizable residues, the location of the positive fourier with respect to Ash2L and *in vitro* lipid binding assays with wild-type and an Ash2L construct lacking SDI provide some structural insights into the mode of binding of Ash2L to cardiolipin. In such model, the unidentified ionizable residues would cooperate with the first half of the SDI α -helix and W505 to stabilize CL. In membrane bilayers containing a conical-shaped cardiolipin on one side, the non-polar SDI face would insert and interact with the acyl chains while the capping tryptophan (W505) residue could anchor the helix into the bilayer and form a hydrogen bond with the polar head group of CL. However, additional structural studies will be required to fully elucidate the structural underpinnings underlying the binding of CL to Ash2L.

In our lipid-protein overlay assays, cardiolipin was first observed to bind the heterotrimer. Subsequent pull-down experiments with biotinylated-CL and CL-

containing liposomes indicated that only Ash2L binds to CL. In addition, competition assays revealed that increasing concentration of CL could disrupt the Ash2L/DPY-30 complex. While these results suggest two contradictory models, several explanations can be offered. First, before the crystallization of the heterotrimer, both Ash2L and DPY-30 were expressed in *E. coli* and purified separately by affinity chromatography. Binding of Ash2L to CL would have thus occurred within *E. coli* crude extract that contains CL from the plasma membrane, without requiring the presence of DPY-30. The concentration of DPY-30 and Ash2L being higher than that of co-purified CL, the integrity of the heterotrimer may not have been hampered by the low amounts of CL, which explains the partial co-crystallization of the heterotrimer with CL. Similarly, in the lipid overlay experiments, a lower quantity of CL was used compared to the biotin-CL and CL-containing liposome pull-downs, which may be too low to disrupt the interaction between Ash2L and DPY-30. Low concentrations of lipids are unable to form micelles, which could explain this discrepancy. This effect is observed when increasing amounts of CL-containing liposomes are added to a solution containing the DPY-30/Ash2L heterotrimer. In fact, with lower quantities of liposome, Ash2L remains bound to DPY-30, but as the concentration of liposomes increases, DPY-30 progressively dissociates from Ash2L.

The mutually exclusive binding of DPY-30 and CL to Ash2L and the absence of CL from nuclear membranes makes it unlikely that CL has a role in stabilizing the heterotrimer or stimulating the enzymatic activity of SET1 methyltransferases. CL marks the membranes of mitochondria, suggesting that CL-binding may confer Ash2L with mitochondrial functions. Interestingly, recent studies showed that the WRAD subunits

translocate from the nucleus to perform novel functions in other cellular organelles, such as the golgi apparatus, the midbody of cells, and mitochondria. Indeed, the Shu laboratory observed that WDR5 translocates from the nucleus to the mitochondria upon infection of HEK 293 cells with the sendai virus (Wang et al., 2009). The shuttling of WDR5 to mitochondria enabled the binding of the scaffolding protein to the VISA complex and assist in host cell defense mechanisms (Wang et al., 2009). A possible mechanism of recruitment during this process would be the initial targeting of Ash2L to mitochondria by binding to cardiolipin and the subsequent binding of WDR5 to Ash2L, via RbBP5, to the VISA complex. Consistent with a role for CL in the recruitment of proteins during cell stresses, the nlrp3 inflammasome also binds to CL found in the outer mitochondrial membrane. This complex then elicits a pro-inflammatory response by secreting specific cytokines and cleaving caspase 1. In another example, DPY-30, Ash2L, and RbBP5 have been linked to the trafficking of cargo in the trans-golgi network (TGN) (Xu et al., 2009). Initial screening with the lipid arrays revealed that the heterotrimer could bind to phosphatidylinositol 4-phosphate (PtdIns(4)P), a lipid mainly forming membranes in the Golgi (Graham and Burd, 2011). Ash2L can potentially bind to PtdIns(4)P in the TGN and recruit RbBP5 to fulfill their function in cargo packaging into the endosomes or in triggering the detachment of endosomes from the plasma membrane. Given that our binding experiments were performed on a limited set of phospholipids and phosphatidylinositols, we may have missed the genuine phospholipid binding to Ash2L. One such possibility is the Bis(Monoacylglycero)Phosphate (BMP), a lipid found in endosome membranes (Kobayashi et al., 1998). In this scenario, the binding of Ash2L to BMP would tether the protein to the TGN and regulate cargo trafficking. Ash2L may

bind to BMP found within endosomes and travel alongside cargo proteins to ensure that they attain their correct destination or may contribute to the progression of early endosomes into late endosomes. Additional studies will be necessary to address these questions.

As an attempt to measure the distribution of Ash2L, cellular fractionation could be performed on different cell lines (HEK293, HeLa, NIH 3T3, U2OS, etc), and the enrichment of Ash2L can be followed by western blotting. To differentiate between possible activities in the TGN and mitochondria, Ash2L could be knocked down in HeLa cells, and rescued with a dominant negative copy of Ash2L that can no longer bind to lipids or with wild-type Ash2L. In a first experiment, these cells could then be exposed to the sendai virus, and confocal microscopy could be used to verify if Ash2L and WDR5 are able to localize to mitochondria and participate in the host innate anti-viral response. Co-immunopurification experiments in mitochondrial extracts can assess whether WDR5 is still able to integrate into the VISA-associated complex in mitochondria, which activates anti-viral factors. The binding activity of Ash2L would implicate it in a role in cell defense mechanisms within mitochondria if WDR5 fails to localize to mitochondria and integrate into the VISA complex in cells containing the dominant negative copy of Ash2L compared to those containing the wild-type copy. Secondly, microscopy could be used to scan the cells for an enrichment of cargo proteins (like CIMPR or furin) at cellular protrusions, which is indicative of a failure to properly traffic cell cargo from the TGN to endosomes or between endosomes. The dominant negative copy of Ash2L would be expected to lose its ability to localize to the TGN or to circulating endosomes relative

to wild-type Ash2L if binding to lipids is essential for its recruitment to these organelles. Combining cellular localization studies and functional assays would allow us to establish the function that Ash2L is granted by binding to anionic lipids.

CONCLUSION:

In my thesis, I have shown that two residues flanking the SDI α -helix of Ash2L form non-symmetrical interactions with a single protomer of DPY-30 thereby highlighting a new mechanism controlling the asymmetric insertion of the nearly symmetric SDI amphipathic α -helix into the symmetrical DPY-30 D/D. I also discovered that Ash2L binds to anionic lipids *in vitro* with a preference for CL shedding new lights on a potentially novel non-nuclear functions of Ash2L.

REFERENCES:

- Adams, P.D., Afonine, P.V., Bunkóczi, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.-W., Kapral, G.J., Grosse-Kunstleve, R.W., et al. (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* *66*, 213–221.
- Akimoto, C., Kitagawa, H., Matsumoto, T., and Kato, S. (2008). Spermatogenesis-specific association of SMCY and MSH5. *Genes Cells* *13*, 623–633.
- Andreu-Vieyra, C.V., Chen, R., Agno, J.E., Glaser, S., Anastassiadis, K., Stewart, A.F., and Matzuk, M.M. (2010). MLL2 is required in oocytes for bulk histone 3 lysine 4 trimethylation and transcriptional silencing. *PLoS Biol.* *8*.
- Ang, Y.-S., Tsai, S.-Y., Lee, D.-F., Monk, J., Su, J., Ratnakumar, K., Ding, J., Ge, Y., Darr, H., Chang, B., et al. (2011). Wdr5 mediates self-renewal and reprogramming via the embryonic stem cell core transcriptional network. *Cell* *145*, 183–197.
- Arents, G., Burlingame, R.W., Wang, B.C., Love, W.E., and Moudrianakis, E.N. (1991). The nucleosomal core histone octamer at 3.1 Å resolution: a tripartite protein assembly and a left-handed superhelix. *Proc. Natl. Acad. Sci. U.S.A.* *88*, 10148–10152.
- Avdic, V., Zhang, P., Lanouette, S., Groulx, A., Tremblay, V., Brunzelle, J., and Couture, J.-F. (2011). Structural and biochemical insights into MLL1 core complex assembly. *Structure* *19*, 101–108.
- Ayton, P.M., and Cleary, M.L. (2001). Molecular mechanisms of leukemogenesis mediated by MLL fusion proteins. *Oncogene* *20*, 5695–5707.
- Bai, H., Li, Y., Gao, H., Dong, Y., Han, P., and Yu, H. (2016). Histone methyltransferase SMYD3 regulates the expression of transcriptional factors during bovine oocyte maturation and early embryonic development. *Cytotechnology* *68*, 849–859.
- Bailey, J.K., Fields, A.T., Cheng, K., Lee, A., Wagenaar, E., Lagrois, R., Schmidt, B., Xia, B., and Ma, D. (2015). WD repeat-containing protein 5 (WDR5) localizes to the midbody and regulates abscission. *J. Biol. Chem.* *290*, 8987–9001.
- Balakrishnan, A., Bleeker, F.E., Lamba, S., Rodolfo, M., Daniotti, M., Scarpa, A., van Tilborg, A.A., Leenstra, S., Zanon, C., and Bardelli, A. (2007). Novel somatic and germline mutations in cancer candidate genes in glioblastoma, melanoma, and pancreatic carcinoma. *Cancer Res.* *67*, 3545–3550.
- Banerjee, T., and Chakravarti, D. (2011). A peek into the complex realm of histone phosphorylation. *Mol. Cell. Biol.* *31*, 4858–4873.
- Barski, A., Cuddapah, S., Cui, K., Roh, T.-Y., Schones, D.E., Wang, Z., Wei, G., Chepelev, I., and Zhao, K. (2007). High-resolution profiling of histone methylations in the human genome. *Cell* *129*, 823–837.
- Beary, J.F. (1986). The polygraph: does it really detect lies? *Am Fam Physician* *33*, 98,

Bednar, J., Horowitz, R.A., Grigoryev, S.A., Carruthers, L.M., Hansen, J.C., Koster, A.J., and Woodcock, C.L. (1998). Nucleosomes, linker DNA, and linker histone form a unique structural motif that directs the higher-order folding and compaction of chromatin. *Proc. Natl. Acad. Sci. U.S.A.* *95*, 14173–14178.

Bernstein, E., and Hake, S.B. (2006). The nucleosome: a little variation goes a long way. *Biochem. Cell Biol.* *84*, 505–517.

Bernstein, B.E., Kamal, M., Lindblad-Toh, K., Bekiranov, S., Bailey, D.K., Huebert, D.J., McMahon, S., Karlsson, E.K., Kulbokas, E.J., Gingeras, T.R., et al. (2005). Genomic maps and comparative analysis of histone modifications in human and mouse. *Cell* *120*, 169–181.

Bernstein, B.E., Mikkelsen, T.S., Xie, X., Kamal, M., Huebert, D.J., Cuff, J., Fry, B., Meissner, A., Wernig, M., Plath, K., et al. (2006). A bivalent chromatin structure marks key developmental genes in embryonic stem cells. *Cell* *125*, 315–326.

Binda, O., LeRoy, G., Bua, D.J., Garcia, B.A., Gozani, O., and Richard, S. (2010). Trimethylation of histone H3 lysine 4 impairs methylation of histone H3 lysine 9: regulation of lysine methyltransferases by physical interaction with their substrates. *Epigenetics* *5*, 767–775.

Blower, M.D., and Karpen, G.H. (2001). The role of *Drosophila* CID in kinetochore formation, cell-cycle progression and heterochromatin interactions. *Nat. Cell Biol.* *3*, 730–739.

Boa, S., Coert, C., and Patterton, H.-G. (2003). *Saccharomyces cerevisiae* Set1p is a methyltransferase specific for lysine 4 of histone H3 and is required for efficient gene expression. *Yeast* *20*, 827–835.

Briggs, S.D., Bryk, M., Strahl, B.D., Cheung, W.L., Davie, J.K., Dent, S.Y., Winston, F., and Allis, C.D. (2001). Histone H3 lysine 4 methylation is mediated by Set1 and required for cell growth and rDNA silencing in *Saccharomyces cerevisiae*. *Genes Dev.* *15*, 3286–3295.

Byrd, K.N., and Shearn, A. (2003). ASH1, a *Drosophila* trithorax group protein, is required for methylation of lysine 4 residues on histone H3. *Proc. Natl. Acad. Sci. U.S.A.* *100*, 11535–11540.

Cao, R., and Zhang, Y. (2004). The functions of E(Z)/EZH2-mediated methylation of lysine 27 in histone H3. *Curr. Opin. Genet. Dev.* *14*, 155–164.

Cao, F., Chen, Y., Cierpicki, T., Liu, Y., Basrur, V., Lei, M., and Dou, Y. (2010). An Ash2L/RbBP5 heterodimer stimulates the MLL1 methyltransferase activity through coordinated substrate interactions with the MLL1 SET domain. *PLoS ONE* *5*, e14102.

- Cao, F., Townsend, E.C., Karatas, H., Xu, J., Li, L., Lee, S., Liu, L., Chen, Y., Ouillet, P., Zhu, J., et al. (2014). Targeting MLL1 H3K4 Methyltransferase Activity in Mixed-Lineage Leukemia. *Molecular Cell* 53, 247–261.
- Cellot, S., Hope, K.J., Chagraoui, J., Sauvageau, M., Deneault, É., MacRae, T., Mayotte, N., Wilhelm, B.T., Landry, J.R., Ting, S.B., et al. (2013). RNAi screen identifies Jarid1b as a major regulator of mouse HSC activity. *Blood* 122, 1545–1555.
- Chadwick, B.P., and Willard, H.F. (2002). Cell cycle-dependent localization of macroH2A in chromatin of the inactive X chromosome. *J. Cell Biol.* 157, 1113–1123.
- Chaturvedi, C.-P., Somasundaram, B., Singh, K., Carpenedo, R.L., Stanford, W.L., Dilworth, F.J., and Brand, M. (2012). Maintenance of gene silencing by the coordinate action of the H3K9 methyltransferase G9a/KMT1C and the H3K4 demethylase Jarid1a/KDM5A. *Proc. Natl. Acad. Sci. U.S.A.* 109, 18845–18850.
- Chen, J., Guo, Y., Zeng, W., Huang, L., Pang, Q., Nie, L., Mu, J., Yuan, F., and Feng, B. (2014). ER stress triggers MCP-1 expression through SET7/9-induced histone methylation in the kidneys of db/db mice. *Am. J. Physiol. Renal Physiol.* 306, F916–925.
- Chen, X., Xie, W., Gu, P., Cai, Q., Wang, B., Xie, Y., Dong, W., He, W., Zhong, G., Lin, T., et al. (2015). Upregulated WDR5 promotes proliferation, self-renewal and chemoresistance in bladder cancer via mediating H3K4 trimethylation. *Sci Rep* 5, 8293.
- Cho, Y.-W., Hong, T., Hong, S., Guo, H., Yu, H., Kim, D., Guszczynski, T., Dressler, G.R., Copeland, T.D., Kalkum, M., et al. (2007). PTIP associates with MLL3- and MLL4-containing histone H3 lysine 4 methyltransferase complex. *J. Biol. Chem.* 282, 20395–20406.
- Choi, J., Jang, H., Kim, H., Kim, S.-T., Cho, E.-J., and Youn, H.-D. (2010). Histone demethylase LSD1 is required to induce skeletal muscle differentiation by regulating myogenic factors. *Biochem. Biophys. Res. Commun.* 401, 327–332.
- Christensen, J., Agger, K., Cloos, P.A.C., Pasini, D., Rose, S., Sennels, L., Rappsilber, J., Hansen, K.H., Salcini, A.E., and Helin, K. (2007). RBP2 belongs to a family of demethylases, specific for tri- and dimethylated lysine 4 on histone 3. *Cell* 128, 1063–1076.
- Ciccone, D.N., Su, H., Hevi, S., Gay, F., Lei, H., Bajko, J., Xu, G., Li, E., and Chen, T. (2009). KDM1B is a histone H3K4 demethylase required to establish maternal genomic imprints. *Nature* 461, 415–418.
- Cleary, S.P., Jeck, W.R., Zhao, X., Chen, K., Selitsky, S.R., Savich, G.L., Tan, T.-X., Wu, M.C., Getz, G., Lawrence, M.S., et al. (2013). Identification of driver genes in hepatocellular carcinoma by exome sequencing. *Hepatology* 58, 1693–1702.
- Collins, R.E., Tachibana, M., Tamaru, H., Smith, K.M., Jia, D., Zhang, X., Selker, E.U., Shinkai, Y., and Cheng, X. (2005). In vitro and in vivo analyses of a Phe/Tyr switch

controlling product specificity of histone lysine methyltransferases. *J. Biol. Chem.* **280**, 5563–5570.

Corda, Y., Schramke, V., Longhese, M.P., Smokvina, T., Paciotti, V., Brevet, V., Gilson, E., and Géli, V. (1999). Interaction between Set1p and checkpoint protein Mec3p in DNA repair and telomere functions. *Nat. Genet.* **21**, 204–208.

Cosgrove, M.S., Boeke, J.D., and Wolberger, C. (2004). Regulated nucleosome mobility and the histone code. *Nat. Struct. Mol. Biol.* **11**, 1037–1043.

Couture, J.-F., and Trievel, R.C. (2006). Histone-modifying enzymes: encrypting an enigmatic epigenetic code. *Curr. Opin. Struct. Biol.* **16**, 753–760.

Couture, J.-F., Dirk, L.M.A., Brunzelle, J.S., Houtz, R.L., and Trievel, R.C. (2008). Structural origins for the product specificity of SET domain protein methyltransferases. *Proc Natl Acad Sci U S A* **105**, 20659–20664.

Dai, X., Guo, W., Zhan, C., Liu, X., Bai, Z., and Yang, Y. (2015). WDR5 Expression Is Prognostic of Breast Cancer Outcome. *PLoS ONE* **10**, e0124964.

Dehé, P.-M., Dichtl, B., Schaft, D., Roguev, A., Pamblanco, M., Lebrun, R., Rodríguez-Gil, A., Mkandawire, M., Landsberg, K., Shevchenko, A., et al. (2006). Protein interactions within the Set1 complex and their roles in the regulation of histone 3 lysine 4 methylation. *J. Biol. Chem.* **281**, 35404–35412.

Dey, B.K., Stalker, L., Schnerch, A., Bhatia, M., Taylor-Papadimitriou, J., and Wynder, C. (2008). The histone demethylase KDM5b/JARID1b plays a role in cell fate decisions by blocking terminal differentiation. *Mol. Cell. Biol.* **28**, 5312–5327.

Dhalluin, C., Carlson, J.E., Zeng, L., He, C., Aggarwal, A.K., and Zhou, M.M. (1999). Structure and ligand of a histone acetyltransferase bromodomain. *Nature* **399**, 491–496.

Dhar, S.S., Lee, S.-H., Kan, P.-Y., Voigt, P., Ma, L., Shi, X., Reinberg, D., and Lee, M.G. (2012a). Trans-tail regulation of MLL4-catalyzed H3K4 methylation by H4R3 symmetric dimethylation is mediated by a tandem PHD of MLL4. *Genes Dev.* **26**, 2749–2762.

Dhar, S.S., Lee, S.-H., Kan, P.-Y., Voigt, P., Ma, L., Shi, X., Reinberg, D., and Lee, M.G. (2012b). Trans-tail regulation of MLL4-catalyzed H3K4 methylation by H4R3 symmetric dimethylation is mediated by a tandem PHD of MLL4. *Genes Dev.* **26**, 2749–2762.

Dias, J., Van Nguyen, N., Georgiev, P., Gaub, A., Brettschneider, J., Cusack, S., Kadlec, J., and Akhtar, A. (2014). Structural analysis of the KANSL1/WDR5/KANSL2 complex reveals that WDR5 is required for efficient assembly and chromatin targeting of the NSL complex. *Genes Dev.* **28**, 929–942.

Dillon, S.C., Zhang, X., Trievel, R.C., and Cheng, X. (2005). The SET-domain protein

superfamily: protein lysine methyltransferases. *Genome Biol.* 6, 227.

Donaldson, R.P., and Beevers, H. (1977). Lipid composition of organelles from germinating castor bean endosperm. *Plant Physiol.* 59, 259–263.

Dou, Y., Milne, T.A., Tackett, A.J., Smith, E.R., Fukuda, A., Wysocka, J., Allis, C.D., Chait, B.T., Hess, J.L., and Roeder, R.G. (2005). Physical association and coordinate function of the H3 K4 methyltransferase MLL1 and the H4 K16 acetyltransferase MOF. *Cell* 121, 873–885.

Dou, Y., Milne, T.A., Ruthenburg, A.J., Lee, S., Lee, J.W., Verdine, G.L., Allis, C.D., and Roeder, R.G. (2006). Regulation of MLL1 H3K4 methyltransferase activity by its core components. *Nat. Struct. Mol. Biol.* 13, 713–719.

Drees, M., and Beyer, K. (1988). Interaction of phospholipids with the detergent-solubilized ADP/ATP carrier protein as studied by spin-label electron spin resonance. *Biochemistry* 27, 8584–8591.

Duan, Q., Chen, H., Costa, M., and Dai, W. (2008). Phosphorylation of H3S10 blocks the access of H3K9 by specific antibodies and histone methyltransferase. Implication in regulating chromatin dynamics and epigenetic inheritance during mitosis. *J. Biol. Chem.* 283, 33585–33590.

Eisen, J.A., Sweder, K.S., and Hanawalt, P.C. (1995). Evolution of the SNF2 family of proteins: subfamilies with distinct sequences and functions. *Nucleic Acids Res.* 23, 2715–2723.

Emsley, P., Lohkamp, B., Scott, W.G., and Cowtan, K. (2010). Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* 66, 486–501.

Ernst, P., Fisher, J.K., Avery, W., Wade, S., Foy, D., and Korsmeyer, S.J. (2004). Definitive hematopoiesis requires the mixed-lineage leukemia gene. *Dev. Cell* 6, 437–443.

Fang, R., Barbera, A.J., Xu, Y., Rutenberg, M., Leonor, T., Bi, Q., Lan, F., Mei, P., Yuan, G.-C., Lian, C., et al. (2010). Human LSD2/KDM1b/AOF1 regulates gene transcription by modulating intragenic H3K4me2 methylation. *Mol. Cell* 39, 222–233.

Ford, D.J., and Dingwall, A.K. (2015). The cancer COMPASS: navigating the functions of MLL complexes in cancer. *Cancer Genet* 208, 178–191.

Fossati, A., Dolfini, D., Donati, G., and Mantovani, R. (2011). NF-Y recruits Ash2L to impart H3K4 trimethylation on CCAAT promoters. *PLoS ONE* 6, e17220.

Fujii, T., Tsunesumi, S., Yamaguchi, K., Watanabe, S., and Furukawa, Y. (2011). Smyd3 is required for the development of cardiac and skeletal muscle in zebrafish. *PLoS ONE* 6, e23491.

Gan, Q., Thiébaud, P., Thézé, N., Jin, L., Xu, G., Grant, P., and Owens, G.K. (2011). WD repeat-containing protein 5, a ubiquitously expressed histone methyltransferase adaptor protein, regulates smooth muscle cell-selective gene activation through interaction with pituitary homeobox 2. *J. Biol. Chem.* 286, 21853–21864.

Ge, Z., Song, E.J., Kawasaki, Y.I., Li, J., Dovat, S., and Song, C. (2016). WDR5 high expression and its effect on tumorigenesis in leukemia. *Oncotarget* 7, 37740–37754.

Glaser, S., Schaft, J., Lubitz, S., Vintersten, K., van der Hoeven, F., Tufteland, K.R., Aasland, R., Anastassiadis, K., Ang, S.-L., and Stewart, A.F. (2006). Multiple epigenetic maintenance factors implicated by the loss of Mll2 in mouse development. *Development* 133, 1423–1432.

Glaser, S., Lubitz, S., Loveland, K.L., Ohbo, K., Robb, L., Schwenk, F., Seibler, J., Roellig, D., Kranz, A., Anastassiadis, K., et al. (2009). The histone 3 lysine 4 methyltransferase, Mll2, is only required briefly in development and spermatogenesis. *Epigenetics Chromatin* 2, 5.

Goo, Y.-H., Sohn, Y.C., Kim, D.-H., Kim, S.-W., Kang, M.-J., Jung, D.-J., Kwak, E., Barlev, N.A., Berger, S.L., Chow, V.T., et al. (2003). Activating Signal Cointegrator 2 Belongs to a Novel Steady-State Complex That Contains a Subset of Trithorax Group Proteins. *Mol. Cell. Biol.* 23, 140–149.

Goult, B.T., Zacharchenko, T., Bate, N., Tsang, R., Hey, F., Gingras, A.R., Elliott, P.R., Roberts, G.C.K., Ballestrem, C., Critchley, D.R., et al. (2013). RIAM and vinculin binding to talin are mutually exclusive and regulate adhesion assembly and turnover. *J. Biol. Chem.* 288, 8238–8249.

Graham, T.R., and Burd, C.G. (2011). Coordination of Golgi functions by phosphatidylinositol 4-kinases. *Trends Cell Biol.* 21, 113–121.

Grebien, F., Vedadi, M., Getlik, M., Giambruno, R., Grover, A., Avellino, R., Skucha, A., Vittori, S., Kuznetsova, E., Smil, D., et al. (2015). Pharmacological targeting of the Wdr5-MLL interaction in C/EBP α N-terminal leukemia. *Nat. Chem. Biol.* 11, 571–578.

Gregory, G.D., Vakoc, C.R., Rozovskaia, T., Zheng, X., Patel, S., Nakamura, T., Canaani, E., and Blobel, G.A. (2007). Mammalian ASH1L is a histone methyltransferase that occupies the transcribed region of active genes. *Mol. Cell. Biol.* 27, 8466–8479.

Gu, B., and Lee, M.G. (2013a). Histone H3 lysine 4 methyltransferases and demethylases in self-renewal and differentiation of stem cells. *Cell Biosci* 3, 39.

Gu, B., and Lee, M.G. (2013b). Histone H3 lysine 4 methyltransferases and demethylases in self-renewal and differentiation of stem cells. *Cell Biosci* 3, 39.

Gu, S., Rehman, S., Wang, X., Shevchik, V.E., and Pickersgill, R.W. (2012). Structural and functional insights into the pilotin-secretin complex of the type II secretion system. *PLoS Pathog.* 8, e1002531.

Guenther, M.G., Jenner, R.G., Chevalier, B., Nakamura, T., Croce, C.M., Canaani, E., and Young, R.A. (2005). Global and Hox-specific roles for the MLL1 methyltransferase. *Proc. Natl. Acad. Sci. U.S.A.* *102*, 8603–8608.

Guenther, M.G., Levine, S.S., Boyer, L.A., Jaenisch, R., and Young, R.A. (2007). A chromatin landmark and transcription initiation at most promoters in human cells. *Cell* *130*, 77–88.

Hahn, M., Dambacher, S., Dulev, S., Kuznetsova, A.Y., Eck, S., Wörz, S., Sadic, D., Schulte, M., Mallm, J.-P., Maiser, A., et al. (2013). Suv4-20h2 mediates chromatin compaction and is important for cohesin recruitment to heterochromatin. *Genes Dev.* *27*, 859–872.

Hake, S.B., and Allis, C.D. (2006). Histone H3 variants and their potential role in indexing mammalian genomes: the “H3 barcode hypothesis.” *Proc. Natl. Acad. Sci. U.S.A.* *103*, 6428–6435.

Hamamoto, R., Furukawa, Y., Morita, M., Iimura, Y., Silva, F.P., Li, M., Yagyu, R., and Nakamura, Y. (2004). SMYD3 encodes a histone methyltransferase involved in the proliferation of cancer cells. *Nat. Cell Biol.* *6*, 731–740.

Hanson, R.D., Hess, J.L., Yu, B.D., Ernst, P., van Lohuizen, M., Berns, A., van der Lugt, N.M., Shashikant, C.S., Ruddle, F.H., Seto, M., et al. (1999). Mammalian Trithorax and polycomb-group homologues are antagonistic regulators of homeotic development. *Proc. Natl. Acad. Sci. U.S.A.* *96*, 14372–14377.

He, S., Owen, D.R., Jelinsky, S.A., and Lin, L.-L. (2015). Lysine Methyltransferase SETD7 (SET7/9) Regulates ROS Signaling through mitochondria and NFE2L2/ARE pathway. *Sci Rep* *5*, 14368.

Heintzman, N.D., Stuart, R.K., Hon, G., Fu, Y., Ching, C.W., Hawkins, R.D., Barrera, L.O., Van Calcar, S., Qu, C., Ching, K.A., et al. (2007). Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. *Nat. Genet.* *39*, 311–318.

Herz, H.-M., Mohan, M., Garruss, A.S., Liang, K., Takahashi, Y., Mickey, K., Voets, O., Verrijzer, C.P., and Shilatifard, A. (2012). Enhancer-associated H3K4 monomethylation by Trithorax-related, the Drosophila homolog of mammalian Mll3/Mll4. *Genes Dev.*

Higa, L.A., Wu, M., Ye, T., Kobayashi, R., Sun, H., and Zhang, H. (2006). CUL4-DDB1 ubiquitin ligase interacts with multiple WD40-repeat proteins and regulates histone methylation. *Nat. Cell Biol.* *8*, 1277–1283.

Hohenauer, T., and Moore, A.W. (2012). The Prdm family: expanding roles in stem cells and development. *Development* *139*, 2267–2282.

Holmes, A., Roseaulin, L., Schurra, C., Waxin, H., Lambert, S., Zaratiegui, M., Martienssen, R.A., and Arcangioli, B. (2012). Lsd1 and lsd2 control programmed

replication fork pauses and imprinting in fission yeast. *Cell Rep* 2, 1513–1520.

Hong, L., Schroth, G.P., Matthews, H.R., Yau, P., and Bradbury, E.M. (1993). Studies of the DNA binding properties of histone H4 amino terminus. Thermal denaturation studies reveal that acetylation markedly reduces the binding constant of the H4 “tail” to DNA. *J. Biol. Chem.* 268, 305–314.

Hsu, D.R., and Meyer, B.J. (1994). The DPY-30 Gene Encodes an Essential Component of the *Caenorhabditis Elegans* Dosage Compensation Machinery. *Genetics* 137, 999–1018.

Hu, D., Gao, X., Morgan, M.A., Herz, H.-M., Smith, E.R., and Shilatifard, A. (2013). The MLL3/MLL4 Branches of the COMPASS Family Function as Major Histone H3K4 Monomethylases at Enhancers. *Mol Cell Biol* 33, 4745–4754.

Hu, J.-L., Zhou, B.O., Zhang, R.-R., Zhang, K.-L., Zhou, J.-Q., and Xu, G.-L. (2009). The N-terminus of histone H3 is required for de novo DNA methylation in chromatin. *Proc. Natl. Acad. Sci. U.S.A.* 106, 22187–22192.

Huang, J., Gurung, B., Wan, B., Matkar, S., Veniaminova, N.A., Wan, K., Merchant, J.L., Hua, X., and Lei, M. (2012). The same pocket in menin binds both MLL and JUND but has opposite effects on transcription. *Nature* 482, 542–546.

Hung, T., Binda, O., Champagne, K.S., Kuo, A.J., Johnson, K., Chang, H.Y., Simon, M.D., Kutateladze, T.G., and Gozani, O. (2009). ING4 mediates crosstalk between histone H3 K4 trimethylation and H3 acetylation to attenuate cellular transformation. *Mol. Cell* 33, 248–256.

Huntsman, D.G., Chin, S.F., Muleris, M., Batley, S.J., Collins, V.P., Wiedemann, L.M., Aparicio, S., and Caldas, C. (1999). MLL2, the second human homolog of the *Drosophila* trithorax gene, maps to 19q13.1 and is amplified in solid tumor cell lines. *Oncogene* 18, 7975–7984.

Iberg, A.N., Espejo, A., Cheng, D., Kim, D., Michaud-Levesque, J., Richard, S., and Bedford, M.T. (2008). Arginine methylation of the histone H3 tail impedes effector binding. *J. Biol. Chem.* 283, 3006–3010.

Iyer, S.S., He, Q., Janczy, J.R., Elliott, E.I., Zhong, Z., Olivier, A.K., Sadler, J.J., Knepper-Adrian, V., Han, R., Qiao, L., et al. (2013). Mitochondrial cardiolipin is required for Nlrp3 inflammasome activation. *Immunity* 39, 311.

Jacob, Y., Bergamin, E., Donoghue, M.T.A., Mongeon, V., LeBlanc, C., Voigt, P., Underwood, C.J., Brunzelle, J.S., Michaels, S.D., Reinberg, D., et al. (2014). Selective methylation of histone H3 variant H3.1 regulates heterochromatin replication. *Science* 343, 1249–1253.

Jenuwein, T. (2006). The epigenetic magic of histone lysine methylation. *FEBS J.* 273, 3121–3135.

- Jeyaprakash, A.A., Basquin, C., Jayachandran, U., and Conti, E. (2011). Structural basis for the recognition of phosphorylated histone h3 by the survivin subunit of the chromosomal passenger complex. *Structure* 19, 1625–1634.
- Jiang, H., Shukla, A., Wang, X., Chen, W., Bernstein, B.E., and Roeder, R.G. (2011). Role for DPY-30 in ES Cell-Fate Specification by Regulation of H3K4 Methylation within Bivalent Domains. *Cell* 144, 513–525.
- Jin, C., and Felsenfeld, G. (2006). Distribution of histone H3.3 in hematopoietic cell lineages. *Proc. Natl. Acad. Sci. U.S.A.* 103, 574–579.
- Jude, C.D., Climer, L., Xu, D., Artinger, E., Fisher, J., and Ernst, P. (2007). Unique and independent roles for MLL in adult hematopoietic stem cells and progenitors. *Cell Stem Cell* 1, 324–337.
- Kakimoto, Y., and Akazawa, S. (1970). Isolation and identification of N-G,N-G- and N-G,N'-G-dimethyl-arginine, N-epsilon-mono-, di-, and trimethyllysine, and glucosylgalactosyl- and galactosyl-delta-hydroxylysine from human urine. *J. Biol. Chem.* 245, 5751–5758.
- Karytinos, A., Forneris, F., Profumo, A., Ciossani, G., Battaglioli, E., Binda, C., and Mattevi, A. (2009). A novel mammalian flavin-dependent histone demethylase. *J. Biol. Chem.* 284, 17775–17782.
- Katz, D.J., Edwards, T.M., Reinke, V., and Kelly, W.G. (2009). A *C. elegans* LSD1 demethylase contributes to germline immortality by reprogramming epigenetic memory. *Cell* 137, 308–320.
- Kawai, C., Prado, F.M., Nunes, G.L.C., Mascio, P.D., Carmona-Ribeiro, A.M., and Nantes, I.L. (2005). pH-dependent Interaction of Cytochrome c with Mitochondrial Mimetic Membranes THE ROLE OF AN ARRAY OF POSITIVELY CHARGED AMINO ACIDS. *J. Biol. Chem.* 280, 34709–34717.
- Kerimoglu, C., Agis-Balboa, R.C., Kranz, A., Stilling, R., Bahari-Javan, S., Benito-Garagorri, E., Halder, R., Burkhardt, S., Stewart, A.F., and Fischer, A. (2013). Histone-methyltransferase MLL2 (KMT2B) is required for memory formation in mice. *J. Neurosci.* 33, 3452–3464.
- Kim, D.-H., Lee, J., Lee, B., and Lee, J.W. (2009). ASCOM controls farnesoid X receptor transactivation through its associated histone H3 lysine 4 methyltransferase activity. *Mol. Endocrinol.* 23, 1556–1562.
- Kim, J., Shin, S., Subramaniam, M., Bruinsma, E., Kim, T.-D., Hawse, J.R., Spelsberg, T.C., and Janknecht, R. (2010). Histone demethylase JARID1B/KDM5B is a corepressor of TIEG1/KLF10. *Biochem. Biophys. Res. Commun.* 401, 412–416.
- Kim, J.-H., Sharma, A., Dhar, S.S., Lee, S.-H., Gu, B., Chan, C.-H., Lin, H.-K., and Lee, M.G. (2014). UTX and MLL4 Coordinately Regulate Transcriptional Programs for Cell

- Proliferation and Invasiveness in Breast Cancer Cells. *Cancer Res* 74, 1705–1717.
- Kinderman, F.S., Kim, C., von Daake, S., Ma, Y., Pham, B.Q., Spraggon, G., Xuong, N.-H., Jennings, P.A., and Taylor, S.S. (2006). A dynamic mechanism for AKAP binding to RII isoforms of cAMP-dependent protein kinase. *Mol. Cell* 24, 397–408.
- Klein, B.J., Wang, X., Cui, G., Yuan, C., Botuyan, M.V., Lin, K., Lu, Y., Wang, X., Zhao, Y., Bruns, C.J., et al. (2016). PHF20 Readers Link Methylation of Histone H3K4 and p53 with H4K16 Acetylation. *Cell Rep* 17, 1158–1170.
- Klose, R.J., Gardner, K.E., Liang, G., Erdjument-Bromage, H., Tempst, P., and Zhang, Y. (2007). Demethylation of histone H3K36 and H3K9 by Rph1: a vestige of an H3K9 methylation system in *Saccharomyces cerevisiae*? *Mol. Cell. Biol.* 27, 3951–3961.
- Kobayashi, T., Stang, E., Fang, K.S., de Moerloose, P., Parton, R.G., and Gruenberg, J. (1998). A lipid associated with the antiphospholipid syndrome regulates endosome structure and function. *Nature* 392, 193–197.
- Krogan, N.J., Dover, J., Khorrami, S., Greenblatt, J.F., Schneider, J., Johnston, M., and Shilatifard, A. (2002). COMPASS, a histone H3 (Lysine 4) methyltransferase required for telomeric silencing of gene expression. *J. Biol. Chem.* 277, 10753–10755.
- Krogan, N.J., Dover, J., Wood, A., Schneider, J., Heidt, J., Boateng, M.A., Dean, K., Ryan, O.W., Golshani, A., Johnston, M., et al. (2003). The Paf1 complex is required for histone H3 methylation by COMPASS and Dot1p: linking transcriptional elongation to histone methylation. *Mol. Cell* 11, 721–729.
- Lachner, M., O’Carroll, D., Rea, S., Mechtler, K., and Jenuwein, T. (2001). Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins. *Nature* 410, 116–120.
- Längst, G., and Manelyte, L. (2015). Chromatin Remodelers: From Function to Dysfunction. *Genes (Basel)* 6, 299–324.
- Lanouette, S., Mongeon, V., Figeys, D., and Couture, J.-F. (2014). The functional diversity of protein lysine methylation. *Mol. Syst. Biol.* 10, 724.
- Lee, J.-H., and Skalnik, D.G. (2005). CpG-binding protein (CXXC finger protein 1) is a component of the mammalian Set1 histone H3-Lys4 methyltransferase complex, the analogue of the yeast Set1/COMPASS complex. *J. Biol. Chem.* 280, 41725–41731.
- Lee, J., Saha, P.K., Yang, Q.-H., Lee, S., Park, J.Y., Suh, Y., Lee, S.-K., Chan, L., Roeder, R.G., and Lee, J.W. (2008a). Targeted inactivation of MLL3 histone H3-Lys-4 methyltransferase activity in the mouse reveals vital roles for MLL3 in adipogenesis. *Proc. Natl. Acad. Sci. U.S.A.* 105, 19229–19234.
- Lee, J., Thompson, J.R., Botuyan, M.V., and Mer, G. (2008b). Distinct binding modes specify the recognition of methylated histones H3K4 and H4K20 by JMJD2A-tudor. *Nat. Struct. Mol. Biol.* 15, 109–111.

- Lee, J.-E., Wang, C., Xu, S., Cho, Y.-W., Wang, L., Feng, X., Baldrige, A., Sartorelli, V., Zhuang, L., Peng, W., et al. (2013). H3K4 mono- and di-methyltransferase MLL4 is required for enhancer activation during cell differentiation. *Elife* 2, e01503.
- Lee, J.-H., Tate, C.M., You, J.-S., and Skalnik, D.G. (2007a). Identification and characterization of the human Set1B histone H3-Lys4 methyltransferase complex. *J. Biol. Chem.* 282, 13419–13428.
- Lee, M.G., Villa, R., Trojer, P., Norman, J., Yan, K.-P., Reinberg, D., Di Croce, L., and Shiekhhattar, R. (2007b). Demethylation of H3K27 regulates polycomb recruitment and H2A ubiquitination. *Science* 318, 447–450.
- Lee, S., Lee, J., Lee, S.-K., and Lee, J.W. (2008c). Activating Signal Cointegrator-2 Is an Essential Adaptor to Recruit Histone H3 Lysine 4 Methyltransferases MLL3 and MLL4 to the Liver X Receptors. *Mol Endocrinol* 22, 1312–1319.
- Lee, Y.J., Han, M.-E., Baek, S.-J., Kim, S.-Y., and Oh, S.-O. (2015). Roles of DPY30 in the Proliferation and Motility of Gastric Cancer Cells. *PLOS ONE* 10, e0131863.
- van Leeuwen, F., Gafken, P.R., and Gottschling, D.E. (2002). Dot1p modulates silencing in yeast by methylation of the nucleosome core. *Cell* 109, 745–756.
- Lezina, L., Aksenova, V., Ivanova, T., Purmessur, N., Antonov, A.V., Tentler, D., Fedorova, O., Garabadiu, A.V., Talianidis, I., Melino, G., et al. (2014). KMTase Set7/9 is a critical regulator of E2F1 activity upon genotoxic stress. *Cell Death Differ.* 21, 1889–1899.
- Li, D., Sun, H., Sun, W.-J., Bao, H.-B., Si, S.-H., Fan, J.-L., Lin, P., Cui, R.-J., Pan, Y.-J., Wen, S.-M., et al. (2016a). Role of RbBP5 and H3K4me3 in the vicinity of Snail transcription start site during epithelial- mesenchymal- transition in prostate cancer cell. *Oncotarget* 7, 65553–65567.
- Li, J., Moazed, D., and Gygi, S.P. (2002). Association of the histone methyltransferase Set2 with RNA polymerase II plays a role in transcription elongation. *J. Biol. Chem.* 277, 49383–49388.
- Li, X.-X., Tsoi, B., Li, Y.-F., Kurihara, H., and He, R.-R. (2015). Cardiolipin and its different properties in mitophagy and apoptosis. *J. Histochem. Cytochem.* 63, 301–311.
- Li, Y., Han, J., Zhang, Y., Cao, F., Liu, Z., Li, S., Wu, J., Hu, C., Wang, Y., Shuai, J., et al. (2016b). Structural basis for activity regulation of MLL family methyltransferases. *Nature* 530, 447–452.
- Li, Y., Han, J., Zhang, Y., Cao, F., Liu, Z., Li, S., Wu, J., Hu, C., Wang, Y., Shuai, J., et al. (2016c). Structural basis for activity regulation of MLL family methyltransferases. *Nature* 530, 447–452.
- Lim, D.A., Huang, Y.-C., Swigut, T., Mirick, A.L., Garcia-Verdugo, J.M., Wysocka, J.,

Ernst, P., and Alvarez-Buylla, A. (2009). Chromatin remodelling factor Mll1 is essential for neurogenesis from postnatal neural stem cells. *Nature* 458, 529–533.

Logie, C., Tse, C., Hansen, J.C., and Peterson, C.L. (1999). The core histone N-terminal domains are required for multiple rounds of catalytic chromatin remodeling by the SWI/SNF and RSC complexes. *Biochemistry* 38, 2514–2522.

Lopez-Bigas, N., Kisiel, T.A., Dewaal, D.C., Holmes, K.B., Volkert, T.L., Gupta, S., Love, J., Murray, H.L., Young, R.A., and Benevolenskaya, E.V. (2008). Genome-wide analysis of the H3K4 histone demethylase RBP2 reveals a transcriptional program controlling differentiation. *Mol. Cell* 31, 520–530.

Lorenzo, A.D., and Bedford, M.T. (2011). Histone Arginine Methylation. *FEBS Letters* 585, 2024.

Lubitz, S., Glaser, S., Schaft, J., Stewart, A.F., and Anastassiadis, K. (2007). Increased apoptosis and skewed differentiation in mouse embryonic stem cells lacking the histone methyltransferase Mll2. *Mol. Biol. Cell* 18, 2356–2366.

Luger, K., Mäder, A.W., Richmond, R.K., Sargent, D.F., and Richmond, T.J. (1997). Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature* 389, 251–260.

Luger, K., Dechassa, M.L., and Tremethick, D.J. (2012). New insights into nucleosome and chromatin structure: an ordered state or a disordered affair? *Nat Rev Mol Cell Biol* 13, 436–447.

Lüscher-Firzlaff, J., Gawlista, I., Vervoorts, J., Kapelle, K., Braunschweig, T., Walsemann, G., Rodgarkia-Schamberger, C., Schuchlantz, H., Dreschers, S., Kremmer, E., et al. (2008). The human trithorax protein hASH2 functions as an oncoprotein. *Cancer Res.* 68, 749–758.

Macdonald, N., Welburn, J.P.I., Noble, M.E.M., Nguyen, A., Yaffe, M.B., Clynes, D., Moggs, J.G., Orphanides, G., Thomson, S., Edmunds, J.W., et al. (2005). Molecular basis for the recognition of phosphorylated and phosphoacetylated histone h3 by 14-3-3. *Mol. Cell* 20, 199–211.

Macfarlane, M.G. (1958). Structure of cardiolipin. *Nature* 182, 946.

Magerl, C., Ellinger, J., Braunschweig, T., Kremmer, E., Koch, L.K., Höller, T., Büttner, R., Lüscher, B., and Gütgemann, I. (2010). H3K4 dimethylation in hepatocellular carcinoma is rare compared with other hepatobiliary and gastrointestinal carcinomas and correlates with expression of the methylase Ash2 and the demethylase LSD1. *Hum. Pathol.* 41, 181–189.

McCallister, C., Kdeiss, B., and Nikolaidis, N. (2016). Biochemical characterization of the interaction between HspA1A and phospholipids. *Cell Stress & Chaperones* 21, 41.

- Miller, T., Krogan, N.J., Dover, J., Erdjument-Bromage, H., Tempst, P., Johnston, M., Greenblatt, J.F., and Shilatifard, A. (2001a). COMPASS: a complex of proteins associated with a trithorax-related SET domain protein. *Proc. Natl. Acad. Sci. U.S.A.* 98, 12902–12907.
- Miller, T., Krogan, N.J., Dover, J., Erdjument-Bromage, H., Tempst, P., Johnston, M., Greenblatt, J.F., and Shilatifard, A. (2001b). COMPASS: a complex of proteins associated with a trithorax-related SET domain protein. *Proc. Natl. Acad. Sci. U.S.A.* 98, 12902–12907.
- Milne, T.A., Kim, J., Wang, G.G., Stadler, S.C., Basrur, V., Whitcomb, S.J., Wang, Z., Ruthenburg, A.J., Elenitoba-Johnson, K.S.J., Roeder, R.G., et al. (2010a). Multiple interactions recruit MLL1 and MLL1 fusion proteins to the HOXA9 locus in leukemogenesis. *Mol. Cell* 38, 853–863.
- Milne, T.A., Kim, J., Wang, G.G., Stadler, S.C., Basrur, V., Whitcomb, S.J., Wang, Z., Ruthenburg, A.J., Elenitoba-Johnson, K.S.J., Roeder, R.G., et al. (2010b). Multiple interactions recruit MLL1 and MLL1 fusion proteins to the HOXA9 locus in leukemogenesis. *Mol. Cell* 38, 853–863.
- Mohan, M., Herz, H.-M., Smith, E.R., Zhang, Y., Jackson, J., Washburn, M.P., Florens, L., Eissenberg, J.C., and Shilatifard, A. (2011). The COMPASS family of H3K4 methylases in *Drosophila*. *Mol. Cell. Biol.* 31, 4310–4318.
- Morin, R.D., Mendez-Lago, M., Mungall, A.J., Goya, R., Mungall, K.L., Corbett, R.D., Johnson, N.A., Severson, T.M., Chiu, R., Field, M., et al. (2011). Frequent mutation of histone-modifying genes in non-Hodgkin lymphoma. *Nature* 476, 298–303.
- Murray, K. (1964). THE OCCURRENCE OF EPSILON-N-METHYL LYSINE IN HISTONES. *Biochemistry* 3, 10–15.
- Musri, M.M., Carmona, M.C., Hanzu, F.A., Kaliman, P., Gomis, R., and Párrizas, M. (2010). Histone demethylase LSD1 regulates adipogenesis. *J. Biol. Chem.* 285, 30034–30041.
- Natarajan, T.G., Kallakury, B.V., Sheehan, C.E., Bartlett, M.B., Ganesan, N., Preet, A., Ross, J.S., and Fitzgerald, K.T. (2010). Epigenetic regulator MLL2 shows altered expression in cancer cell lines and tumors from human breast and colon. *Cancer Cell Int.* 10, 13.
- Ng, H.H., Robert, F., Young, R.A., and Struhl, K. (2003). Targeted recruitment of Set1 histone methylase by elongating Pol II provides a localized mark and memory of recent transcriptional activity. *Mol. Cell* 11, 709–719.
- Nijwening, J.H., Geutjes, E.-J., Bernards, R., and Beijersbergen, R.L. (2011). The histone demethylase Jarid1b (Kdm5b) is a novel component of the Rb pathway and associates with E2f-target genes in MEFs during senescence. *PLoS ONE* 6, e25235.

- Nislow, C., Ray, E., and Pillus, L. (1997). SET1, a yeast member of the trithorax family, functions in transcriptional silencing and diverse cellular processes. *Mol. Biol. Cell* 8, 2421–2436.
- Noll, M. (1977). DNA folding in the nucleosome. *J. Mol. Biol.* 116, 49–71.
- van Nuland, R., Smits, A.H., Pallaki, P., Jansen, P.W.T.C., Vermeulen, M., and Timmers, H.T.M. (2013). Quantitative dissection and stoichiometry determination of the human SET1/MLL histone methyltransferase complexes. *Mol. Cell. Biol.* 33, 2067–2077.
- Ogihara, T., Vanderford, N.L., Maier, B., Stein, R.W., and Mirmira, R.G. (2009). Expression and function of Set7/9 in pancreatic islets. *Islets* 1, 269–272.
- Olins, A.L., and Olins, D.E. (1974). Spheroid chromatin units (v bodies). *Science* 183, 330–332.
- Olofsson, G., and Sparr, E. (2013). Ionization constants pKa of cardiolipin. *PLoS ONE* 8, e73040.
- O'Neill, L.A.J. (2013). Cardiolipin and the Nlrp3 inflammasome. *Cell Metab.* 18, 610–612.
- Ong, C.K., Subimerb, C., Pairojkul, C., Wongkham, S., Cutcutache, I., Yu, W., McPherson, J.R., Allen, G.E., Ng, C.C.Y., Wong, B.H., et al. (2012). Exome sequencing of liver fluke-associated cholangiocarcinoma. *Nat. Genet.* 44, 690–693.
- Oudet, P., Gross-Bellard, M., and Chambon, P. (1975). Electron microscopic and biochemical evidence that chromatin structure is a repeating unit. *Cell* 4, 281–300.
- Outchkourov, N.S., Muiño, J.M., Kaufmann, K., van Ijcken, W.F.J., Groot Koerkamp, M.J., van Leenen, D., de Graaf, P., Holstege, F.C.P., Grosveld, F.G., and Timmers, H.T.M. (2013). Balancing of histone H3K4 methylation states by the Kdm5c/SMCX histone demethylase modulates promoter and enhancer function. *Cell Rep* 3, 1071–1079.
- Ow, G.S., Ivshina, A.V., Fuentes, G., and Kuznetsov, V.A. (2014). Identification of two poorly prognosed ovarian carcinoma subtypes associated with CHEK2 germ-line mutation and non-CHEK2 somatic mutation gene signatures. *Cell Cycle* 13, 2262–2280.
- Paik, W.K., and Kim, S. (1967). E-N-dimethyllysine in histones. *Biochemical and Biophysical Research Communications* 27, 479–483.
- Paik, W.K., and Kim, S. (1968). Protein methylase I. Purification and properties of the enzyme. *J. Biol. Chem.* 243, 2108–2114.
- Paik, W.K., and Kim, S. (1970). Solubilization and partial purification of protein methylase 3 from calf thymus nuclei. *J. Biol. Chem.* 245, 6010–6015.
- Pasini, D., Hansen, K.H., Christensen, J., Agger, K., Cloos, P.A.C., and Helin, K. (2008).

Coordinated regulation of transcriptional repression by the RBP2 H3K4 demethylase and Polycomb-Repressive Complex 2. *Genes Dev.* 22, 1345–1355.

Pasqualucci, L., Trifonov, V., Fabbri, G., Ma, J., Rossi, D., Chiarenza, A., Wells, V.A., Grunn, A., Messina, M., Elliot, O., et al. (2011). Analysis of the coding genome of diffuse large B-cell lymphoma. *Nat. Genet.* 43, 830–837.

Patel, D.J., and Wang, Z. (2013). Readout of epigenetic modifications. *Annu. Rev. Biochem.* 82, 81–118.

Patel, A., Vought, V.E., Dharmarajan, V., and Cosgrove, M.S. (2008). A conserved arginine-containing motif crucial for the assembly and enzymatic activity of the mixed lineage leukemia protein-1 core complex. *J. Biol. Chem.* 283, 32162–32175.

Patel, A., Dharmarajan, V., Vought, V.E., and Cosgrove, M.S. (2009). On the mechanism of multiple lysine methylation by the human mixed lineage leukemia protein-1 (MLL1) core complex. *J. Biol. Chem.* 284, 24242–24256.

Patel, S.R., Kim, D., Levitan, I., and Dressler, G.R. (2007). The BRCT-domain containing protein PTIP links PAX2 to a histone H3, lysine 4 methyltransferase complex. *Dev. Cell* 13, 580–592.

Peña, P.V., Hom, R.A., Hung, T., Lin, H., Kuo, A.J., Wong, R.P.C., Subach, O.M., Champagne, K.S., Zhao, R., Verkhusha, V.V., et al. (2008). Histone H3K4me3 binding is required for the DNA repair and apoptotic activities of ING1 tumor suppressor. *J. Mol. Biol.* 380, 303–312.

Planas-Iglesias, J., Dwarakanath, H., Mohammadyani, D., Yanamala, N., Kagan, V.E., and Klein-Seetharaman, J. (2015). Cardiolipin Interactions with Proteins. *Biophys. J.* 109, 1282–1294.

Proserpio, V., Fittipaldi, R., Ryall, J.G., Sartorelli, V., and Caretti, G. (2013). The methyltransferase SMYD3 mediates the recruitment of transcriptional cofactors at the myostatin and c-Met genes and regulates skeletal muscle atrophy. *Genes Dev.* 27, 1299–1312.

Qi, J., Huo, L., Zhu, Y.T., and Zhu, Y.-J. (2014). Absent, small or homeotic 2-like protein (ASH2L) enhances the transcription of the estrogen receptor α gene through GATA-binding protein 3 (GATA3). *J. Biol. Chem.* 289, 31373–31381.

Rampalli, S., Li, L., Mak, E., Ge, K., Brand, M., Tapscott, S.J., and Dilworth, F.J. (2007). p38 MAPK signaling regulates recruitment of Ash2L-containing methyltransferase complexes to specific genes during differentiation. *Nat. Struct. Mol. Biol.* 14, 1150–1156.

Renner, L.D., and Weibel, D.B. (2011). Cardiolipin microdomains localize to negatively curved regions of Escherichia coli membranes. *Proc. Natl. Acad. Sci. U.S.A.* 108, 6264–6269.

- Robinson, N.C., Zborowski, J., and Talbert, L.H. (1990). Cardiolipin-depleted bovine heart cytochrome c oxidase: binding stoichiometry and affinity for cardiolipin derivatives. *Biochemistry* 29, 8962–8969.
- Rogakou, E.P., Pilch, D.R., Orr, A.H., Ivanova, V.S., and Bonner, W.M. (1998). DNA double-stranded breaks induce histone H2AX phosphorylation on serine 139. *J. Biol. Chem.* 273, 5858–5868.
- Rogakou, E.P., Boon, C., Redon, C., and Bonner, W.M. (1999). Megabase chromatin domains involved in DNA double-strand breaks in vivo. *J. Cell Biol.* 146, 905–916.
- Roguev, A., Schaft, D., Shevchenko, A., Pijnappel, W.W.M.P., Wilm, M., Aasland, R., and Stewart, A.F. (2001). The *Saccharomyces cerevisiae* Set1 complex includes an Ash2 homologue and methylates histone 3 lysine 4. *EMBO J* 20, 7137–7148.
- Rothbart, S.B., and Strahl, B.D. (2014). Interpreting the language of histone and DNA modifications. *Biochim. Biophys. Acta* 1839, 627–643.
- Santos-Rosa, H., Schneider, R., Bannister, A.J., Sherrieff, J., Bernstein, B.E., Emre, N.C.T., Schreiber, S.L., Mellor, J., and Kouzarides, T. (2002). Active genes are trimethylated at K4 of histone H3. *Nature* 419, 407–411.
- Sarma, G.N., Kinderman, F.S., Kim, C., von Daake, S., Chen, L., Wang, B.-C., and Taylor, S.S. (2010). Structure of D-AKAP2:PKA RI complex: insights into AKAP specificity and selectivity. *Structure* 18, 155–166.
- Sarvan, S., Avdic, V., Tremblay, V., Chaturvedi, C.-P., Zhang, P., Lanouette, S., Blais, A., Brunzelle, J.S., Brand, M., and Couture, J.-F. (2011). Crystal structure of the trithorax group protein ASH2L reveals a forkhead-like DNA binding domain. *Nat. Struct. Mol. Biol.* 18, 857–859.
- Schlame, M., and Greenberg, M.L. (2017). Biosynthesis, remodeling and turnover of mitochondrial cardiolipin. *Biochim. Biophys. Acta* 1862, 3–7.
- Schmitz, S.U., Albert, M., Malatesta, M., Morey, L., Johansen, J.V., Bak, M., Tommerup, N., Abarrategui, I., and Helin, K. (2011). Jarid1b targets genes regulating development and is involved in neural differentiation. *EMBO J.* 30, 4586–4600.
- Schneider, J., Wood, A., Lee, J.-S., Schuster, R., Dueker, J., Maguire, C., Swanson, S.K., Florens, L., Washburn, M.P., and Shilatifard, A. (2005). Molecular regulation of histone H3 trimethylation by COMPASS and the regulation of gene expression. *Mol. Cell* 19, 849–856.
- Schneider, R., Bannister, A.J., Myers, F.A., Thorne, A.W., Crane-Robinson, C., and Kouzarides, T. (2004). Histone H3 lysine 4 methylation patterns in higher eukaryotic genes. *Nat. Cell Biol.* 6, 73–77.
- Schübeler, D., MacAlpine, D.M., Scalzo, D., Wirbelauer, C., Kooperberg, C., van

- Leeuwen, F., Gottschling, D.E., O'Neill, L.P., Turner, B.M., Delrow, J., et al. (2004). The histone modification pattern of active genes revealed through genome-wide chromatin analysis of a higher eukaryote. *Genes Dev.* *18*, 1263–1271.
- Schuetz, A., Allali-Hassani, A., Martín, F., Loppnau, P., Vedadi, M., Bochkarev, A., Plotnikov, A.N., Arrowsmith, C.H., and Min, J. (2006). Structural basis for molecular recognition and presentation of histone H3 by WDR5. *EMBO J.* *25*, 4245–4252.
- Sheffield, P., Garrard, S., and Derewenda, Z. (1999). Overcoming expression and purification problems of RhoGDI using a family of “parallel” expression vectors. *Protein Expr. Purif.* *15*, 34–39.
- Shelby, R.D., Vafa, O., and Sullivan, K.F. (1997). Assembly of CENP-A into centromeric chromatin requires a cooperative array of nucleosomal DNA contact sites. *J. Cell Biol.* *136*, 501–513.
- Shen, Z., Ye, C., McCain, K., and Greenberg, M.L. (2015). The Role of Cardiolipin in Cardiovascular Health. *Biomed Res Int* *2015*, 891707.
- Shi, X., Hong, T., Walter, K.L., Ewalt, M., Michishita, E., Hung, T., Carney, D., Peña, P., Lan, F., Kaadige, M.R., et al. (2006). ING2 PHD domain links histone H3 lysine 4 methylation to active gene repression. *Nature* *442*, 96–99.
- Shi, Y., Lan, F., Matson, C., Mulligan, P., Whetstine, J.R., Cole, P.A., Casero, R.A., and Shi, Y. (2004). Histone demethylation mediated by the nuclear amine oxidase homolog LSD1. *Cell* *119*, 941–953.
- Shinsky, S.A., Monteith, K.E., Viggiano, S., and Cosgrove, M.S. (2015). Biochemical reconstitution and phylogenetic comparison of human SET1 family core complexes involved in histone methylation. *J. Biol. Chem.* *290*, 6361–6375.
- Sims, R.J., Chen, C.-F., Santos-Rosa, H., Kouzarides, T., Patel, S.S., and Reinberg, D. (2005). Human but not yeast CHD1 binds directly and selectively to histone H3 methylated at lysine 4 via its tandem chromodomains. *J. Biol. Chem.* *280*, 41789–41792.
- Sjöblom, T., Jones, S., Wood, L.D., Parsons, D.W., Lin, J., Barber, T.D., Mandelker, D., Leary, R.J., Ptak, J., Silliman, N., et al. (2006). The consensus coding sequences of human breast and colorectal cancers. *Science* *314*, 268–274.
- Son, M.J., Kim, W.K., Oh, K.-J., Park, A., Lee, D.S., Han, B.S., Lee, S.C., and Bae, K.-H. (2016). Methyltransferase and demethylase profiling studies during brown adipocyte differentiation. *BMB Rep* *49*, 388–393.
- South, P.F., Fingerman, I.M., Mersman, D.P., Du, H.-N., and Briggs, S.D. (2010). A conserved interaction between the SDI domain of Bre2 and the DPY-30 domain of Sdc1 is required for histone methylation and gene expression. *J. Biol. Chem.* *285*, 595–607.
- Southall, S.M., Wong, P.-S., Odho, Z., Roe, S.M., and Wilson, J.R. (2009). Structural

basis for the requirement of additional factors for MLL1 SET domain activity and recognition of epigenetic marks. *Mol. Cell* 33, 181–191.

Steward, M.M., Lee, J.-S., O'Donovan, A., Wyatt, M., Bernstein, B.E., and Shilatifard, A. (2006). Molecular regulation of H3K4 trimethylation by ASH2L, a shared subunit of MLL complexes. *Nat. Struct. Mol. Biol.* 13, 852–854.

Stoffel, W., and Schiefer, H.G. (1968). Biosynthesis and composition of phosphatides in outer and inner mitochondrial membranes. *Hoppe-Seyler's Z. Physiol. Chem.* 349, 1017–1026.

Strahl, B.D., Ohba, R., Cook, R.G., and Allis, C.D. (1999). Methylation of histone H3 at lysine 4 is highly conserved and correlates with transcriptionally active nuclei in *Tetrahymena*. *Proc. Natl. Acad. Sci. U.S.A.* 96, 14967–14972.

Stratmann, A., and Haendler, B. (2011). The histone demethylase JARID1A regulates progesterone receptor expression. *FEBS J.* 278, 1458–1469.

Stucki, M., Clapperton, J.A., Mohammad, D., Yaffe, M.B., Smerdon, S.J., and Jackson, S.P. (2005). MDC1 directly binds phosphorylated histone H2AX to regulate cellular responses to DNA double-strand breaks. *Cell* 123, 1213–1226.

Sun, G., Alzayady, K., Stewart, R., Ye, P., Yang, S., Li, W., and Shi, Y. (2010). Histone demethylase LSD1 regulates neural stem cell proliferation. *Mol. Cell. Biol.* 30, 1997–2005.

Sun, Y., Bell, J.L., Carter, D., Gherardi, S., Poulos, R.C., Milazzo, G., Wong, J.W.H., Al-Awar, R., Tee, A.E., Liu, P.Y., et al. (2015). WDR5 Supports an N-Myc Transcriptional Complex That Drives a Protumorigenic Gene Expression Signature in Neuroblastoma. *Cancer Res.* 75, 5143–5154.

Suto, R.K., Clarkson, M.J., Tremethick, D.J., and Luger, K. (2000). Crystal structure of a nucleosome core particle containing the variant histone H2A.Z. *Nat. Struct. Biol.* 7, 1121–1124.

Suzuki, S., Nozawa, Y., Tsukamoto, S., Kaneko, T., Imai, H., and Minami, N. (2015). Histone methyltransferase Smyd3 regulates early embryonic lineage commitment in mice. *Reproduction* 150, 21–30.

Tahiliani, M., Mei, P., Fang, R., Leonor, T., Rutenberg, M., Shimizu, F., Li, J., Rao, A., and Shi, Y. (2007). The histone H3K4 demethylase SMCX links REST target genes to X-linked mental retardation. *Nature* 447, 601–605.

Takahashi, Y., Westfield, G.H., Oleskie, A.N., Trievel, R.C., Shilatifard, A., and Skiniotis, G. (2011). Structural analysis of the core COMPASS family of histone H3K4 methylases from yeast to human. *Proc Natl Acad Sci U S A* 108, 20526–20531.

Takemoto, T., Daigo, K., and Takagi, N. (1964). [STUDIES ON THE HYPOTENSIVE

CONSTITUENTS OF MARINE ALGAE. I. A NEW BASIC AMINO ACID “LAMININE” AND THE OTHER BASIC CONSTITUENTS ISOLATED FROM LAMINARIA ANGUSTATA]. *Yakugaku Zasshi* 84, 1176–1179.

Tamori, A., Yamanishi, Y., Kawashima, S., Kanehisa, M., Enomoto, M., Tanaka, H., Kubo, S., Shiomi, S., and Nishiguchi, S. (2005). Alteration of gene expression in human hepatocellular carcinoma with integrated hepatitis B virus DNA. *Clin. Cancer Res.* 11, 5821–5826.

Tan, K., Shaw, A.L., Madsen, B., Jensen, K., Taylor-Papadimitriou, J., and Freemont, P.S. (2003). Human PLU-1 Has transcriptional repression properties and interacts with the developmental transcription factors BF-1 and PAX9. *J. Biol. Chem.* 278, 20507–20513.

Tang, L., Nogales, E., and Ciferri, C. (2010). Structure and function of SWI/SNF chromatin remodeling complexes and mechanistic implications for transcription. *Prog. Biophys. Mol. Biol.* 102, 122–128.

Tao, Y., Neppl, R.L., Huang, Z.-P., Chen, J., Tang, R.-H., Cao, R., Zhang, Y., Jin, S.-W., and Wang, D.-Z. (2011). The histone methyltransferase Set7/9 promotes myoblast differentiation and myofibril assembly. *J. Cell Biol.* 194, 551–565.

Taverna, S.D., Li, H., Ruthenburg, A.J., Allis, C.D., and Patel, D.J. (2007). How chromatin-binding modules interpret histone modifications: lessons from professional pocket pickers. *Nat. Struct. Mol. Biol.* 14, 1025–1040.

Thomas, L.R., Wang, Q., Grieb, B.C., Phan, J., Foshage, A.M., Sun, Q., Olejniczak, E.T., Clark, T., Dey, S., Lorey, S., et al. (2015). Interaction with WDR5 promotes target gene recognition and tumorigenesis by MYC. *Mol. Cell* 58, 440–452.

Tremblay, V., Zhang, P., Chaturvedi, C.-P., Thornton, J., Brunzelle, J.S., Skiniotis, G., Shilatifard, A., Brand, M., and Couture, J.-F. (2014). Molecular basis for DPY-30 association to COMPASS-like and NURF complexes. *Structure* 22, 1821–1830.

Trievel, R.C., Beach, B.M., Dirk, L.M.A., Houtz, R.L., and Hurley, J.H. (2002). Structure and catalytic mechanism of a SET domain protein methyltransferase. *Cell* 111, 91–103.

Tsukiyama, T., and Wu, C. (1995). Purification and properties of an ATP-dependent nucleosome remodeling factor. *Cell* 83, 1011–1020.

Ullius, A., Lüscher-Firzlauff, J., Costa, I.G., Walsemann, G., Forst, A.H., Gusmao, E.G., Kapelle, K., Kleine, H., Kremmer, E., Vervoorts, J., et al. (2014). The interaction of MYC with the trithorax protein ASH2L promotes gene transcription by regulating H3K27 modification. *Nucleic Acids Res.* 42, 6901–6920.

Varga-Weisz, P.D., Wilm, M., Bonte, E., Dumas, K., Mann, M., and Becker, P.B. (1997). Chromatin-remodelling factor CHRAC contains the ATPases ISWI and topoisomerase II. *Nature* 388, 598–602.

- Visel, A., Blow, M.J., Li, Z., Zhang, T., Akiyama, J.A., Holt, A., Plajzer-Frick, I., Shoukry, M., Wright, C., Chen, F., et al. (2009). ChIP-seq accurately predicts tissue-specific activity of enhancers. *Nature* *457*, 854–858.
- Wan, X., Liu, L., Ding, X., Zhou, P., Yuan, X., Zhou, Z., Hu, P., Zhou, H., Li, Q., Zhang, S., et al. (2014). Mll2 controls cardiac lineage differentiation of mouse embryonic stem cells by promoting H3K4me3 deposition at cardiac-specific genes. *Stem Cell Rev* *10*, 643–652.
- Wang, K.C., Yang, Y.W., Liu, B., Sanyal, A., Corces-Zimmerman, R., Chen, Y., Lajoie, B.R., Protacio, A., Flynn, R.A., Gupta, R.A., et al. (2011a). A long noncoding RNA maintains active chromatin to coordinate homeotic gene expression. *Nature* *472*, 120–124.
- Wang, P., Lin, C., Smith, E.R., Guo, H., Sanderson, B.W., Wu, M., Gogol, M., Alexander, T., Seidel, C., Wiedemann, L.M., et al. (2009a). Global analysis of H3K4 methylation defines MLL family member targets and points to a role for MLL1-mediated H3K4 methylation in the regulation of transcriptional initiation by RNA polymerase II. *Mol. Cell. Biol.* *29*, 6074–6085.
- Wang, W., Chen, Z., Mao, Z., Zhang, H., Ding, X., Chen, S., Zhang, X., Xu, R., and Zhu, B. (2011b). Nucleolar protein Spindlin1 recognizes H3K4 methylation and stimulates the expression of rRNA genes. *EMBO Rep.* *12*, 1160–1166.
- Wang, X., Lou, Z., Dong, X., Yang, W., Peng, Y., Yin, B., Gong, Y., Yuan, J., Zhou, W., Bartlam, M., et al. (2009b). Crystal structure of the C-terminal domain of human DPY-30-like protein: A component of the histone methyltransferase complex. *J. Mol. Biol.* *390*, 530–537.
- Wang, X.-X., Fu, L., Li, X., Wu, X., Zhu, Z., Fu, L., and Dong, J.-T. (2011c). Somatic mutations of the mixed-lineage leukemia 3 (MLL3) gene in primary breast cancers. *Pathol. Oncol. Res.* *17*, 429–433.
- Wang, Y.-Y., Liu, L.-J., Zhong, B., Liu, T.-T., Li, Y., Yang, Y., Ran, Y., Li, S., Tien, P., and Shu, H.-B. (2010). WDR5 is essential for assembly of the VISA-associated signaling complex and virus-triggered IRF3 and NF-kappaB activation. *Proc. Natl. Acad. Sci. U.S.A.* *107*, 815–820.
- Watanabe, Y., Castoro, R.J., Kim, H.S., North, B., Oikawa, R., Hiraishi, T., Ahmed, S.S., Chung, W., Cho, M.-Y., Toyota, M., et al. (2011). Frequent alteration of MLL3 frameshift mutations in microsatellite deficient colorectal cancer. *PLoS ONE* *6*, e23320.
- Wilson, J.R., Jing, C., Walker, P.A., Martin, S.R., Howell, S.A., Blackburn, G.M., Gamblin, S.J., and Xiao, B. (2002). Crystal structure and functional analysis of the histone methyltransferase SET7/9. *Cell* *111*, 105–115.
- Wirbelauer, C., Bell, O., and Schübeler, D. (2005). Variant histone H3.3 is deposited at sites of nucleosomal displacement throughout transcribed genes while active histone

modifications show a promoter-proximal bias. *Genes Dev.* *19*, 1761–1766.

Wittig, I., and Schagger, H. (2009). Supramolecular organization of ATP synthase and respiratory chain in mitochondrial membranes. *Biochim. Biophys. Acta* *1787*, 672–680.

Wright, K. (1989). *Antibodies a laboratory manual*: By E Harlow and D Lane. pp 726. Cold Spring Harbor Laboratory. 1988. \$50 ISBN 0-87969-314-2. *Biochem. Educ.* *17*, 220–220.

Wu, M., Wang, P.F., Lee, J.S., Martin-Brown, S., Florens, L., Washburn, M., and Shilatifard, A. (2008a). Molecular regulation of H3K4 trimethylation by Wdr82, a component of human Set1/COMPASS. *Mol. Cell. Biol.* *28*, 7337–7344.

Wu, M., Wang, P.F., Lee, J.S., Martin-Brown, S., Florens, L., Washburn, M., and Shilatifard, A. (2008b). Molecular regulation of H3K4 trimethylation by Wdr82, a component of human Set1/COMPASS. *Mol. Cell. Biol.* *28*, 7337–7344.

Wu, M.-Z., Tsai, Y.-P., Yang, M.-H., Huang, C.-H., Chang, S.-Y., Chang, C.-C., Teng, S.-C., and Wu, K.-J. (2011). Interplay between HDAC3 and WDR5 is essential for hypoxia-induced epithelial-mesenchymal transition. *Mol. Cell* *43*, 811–822.

Wysocka, J., Swigut, T., Xiao, H., Milne, T.A., Kwon, S.Y., Landry, J., Kauer, M., Tackett, A.J., Chait, B.T., Badenhorst, P., et al. (2006). A PHD finger of NURF couples histone H3 lysine 4 trimethylation with chromatin remodelling. *Nature* *442*, 86–90.

Xu, Z., Gong, Q., Xia, B., Groves, B., Zimmermann, M., Mugler, C., Mu, D., Matsumoto, B., Seaman, M., and Ma, D. (2009). A role of histone H3 lysine 4 methyltransferase components in endosomal trafficking. *J. Cell Biol.* *186*, 343–353.

Yang, Y.W., Flynn, R.A., Chen, Y., Qu, K., Wan, B., Wang, K.C., Lei, M., and Chang, H.Y. (2014a). Essential role of lncRNA binding for WDR5 maintenance of active chromatin and embryonic stem cell pluripotency. *Elife* *3*, e02046.

Yang, Z., Augustin, J., Chang, C., Hu, J., Shah, K., Chang, C.-W., Townes, T., and Jiang, H. (2014b). The DPY30 subunit in SET1/MLL complexes regulates the proliferation and differentiation of hematopoietic progenitor cells. *Blood* *124*, 2025–2033.

Yang, Z., Augustin, J., Hu, J., and Jiang, H. (2015). Physical Interactions and Functional Coordination between the Core Subunits of Set1/MLL Complexes and the Reprogramming Factors. *PLoS ONE* *10*, e0145336.

Yokoyama, A., and Cleary, M.L. (2008). Menin critically links MLL proteins with LEDGF on cancer-associated target genes. *Cancer Cell* *14*, 36–46.

Yokoyama, A., Kitabayashi, I., Ayton, P.M., Cleary, M.L., and Ohki, M. (2002). Leukemia proto-oncoprotein MLL is proteolytically processed into 2 fragments with opposite transcriptional properties. *Blood* *100*, 3710–3718.

- Yokoyama, A., Wang, Z., Wysocka, J., Sanyal, M., Aufiero, D.J., Kitabayashi, I., Herr, W., and Cleary, M.L. (2004). Leukemia proto-oncoprotein MLL forms a SET1-like histone methyltransferase complex with menin to regulate Hox gene expression. *Mol. Cell. Biol.* *24*, 5639–5649.
- Yu, C.A., and Yu, L. (1980). Structural role of phospholipids in ubiquinol-cytochrome c reductase. *Biochemistry* *19*, 5715–5720.
- Yu, B.D., Hess, J.L., Horning, S.E., Brown, G.A., and Korsmeyer, S.J. (1995). Altered Hox expression and segmental identity in Mll-mutant mice. *Nature* *378*, 505–508.
- Zentner, G.E., and Henikoff, S. (2013). Regulation of nucleosome dynamics by histone modifications. *Nat. Struct. Mol. Biol.* *20*, 259–266.
- Zhang, P., Lee, H., Brunzelle, J.S., and Couture, J.-F. (2012). The plasticity of WDR5 peptide-binding cleft enables the binding of the SET1 family of histone methyltransferases. *Nucleic Acids Res.* *40*, 4237–4246.
- Zhang, P., Chaturvedi, C.-P., Tremblay, V., Cramet, M., Brunzelle, J.S., Skiniotis, G., Brand, M., Shilatifard, A., and Couture, J.-F. (2015). A phosphorylation switch on RbBP5 regulates histone H3 Lys4 methylation. *Genes Dev.* *29*, 123–128.
- Zhou, W., Chen, H., and Zhang, L. (2009). The PcG protein hPc2 interacts with the N-terminus of histone demethylase JARID1B and acts as a transcriptional co-repressor. *BMB Rep* *42*, 154–159.
- Zhou, Y.B., Gerchman, S.E., Ramakrishnan, V., Travers, A., and Muyldermans, S. (1998). Position and orientation of the globular domain of linker histone H5 on the nucleosome. *Nature* *395*, 402–405.
- Zhu, P., and Li, G. (2016). Higher-order structure of the 30-nm chromatin fiber revealed by cryo-EM. *IUBMB Life* *68*, 873–878.
- Zwart, P.H., Afonine, P.V., Grosse-Kunstleve, R.W., Hung, L.-W., Ioerger, T.R., McCoy, A.J., McKee, E., Moriarty, N.W., Read, R.J., Sacchettini, J.C., et al. (2008). Automated structure solution with the PHENIX suite. *Methods Mol. Biol.* *426*, 419–435.

APPENDIX:

Table 3. List of p-Values for the binding of Ash2L constructs to various liposomes.
The p-Values for the comparative analysis of the binding of Ash2L FL, Ash2LSPRY+SDI, and Ash2LSPRY to PC:PE, PC:CL, PC:PA, and PC:PS are represented below.

Compared samples	p-Value
Ash2L ^{SPRY} binding to PC:CL vs Ash2L ^{SPRY} binding to PC:PE	0.00151865
Ash2L ^{SPRY} binding to PC:CL vs Ash2L ^{SPRY} binding to PC:PA	0.00195002
Ash2L ^{SPRY} binding to PC:CL vs Ash2L ^{SPRY} binding to PC:PS	0.004044343
Ash2L ^{SPRY+SDI} binding to PC:CL vs Ash2L ^{SPRY+SDI} binding to PC:PE	0.000011498
Ash2L ^{SPRY+SDI} binding to PC:CL vs Ash2L ^{SPRY+SDI} binding to PC:PA	0.0000237369
Ash2L ^{SPRY+SDI} binding to PC:CL vs Ash2L ^{SPRY+SDI} binding to PC:PS	0.0000748256
Ash2L FL binding to PC:CL vs Ash2L FL binding to PC:PE	0.000144531
Ash2L FL binding to PC:CL vs Ash2L FL binding to PC:PA	0.000151267
Ash2L FL binding to PC:CL vs Ash2L FL binding to PC:PS	0.000377104
Ash2L FL binding to PC:CL vs Ash2L ^{SPRY+SDI} binding to PC:CL	0.741232394
Ash2L FL binding to PC:CL vs Ash2L ^{SPRY} binding to PC:CL	0.031545152
Ash2L ^{SPRY+SDI} binding to PC:CL vs Ash2L ^{SPRY} binding to PC:CL	0.00111914

John Faissal Haddad

EDUCATION:

Master's in Biochemistry, University of Ottawa

BMI department (Faculty of Medicine), Ottawa Institute of Systems Biology (OISB)

May 2014-February 2017

Thesis topic: Characterizing the complete set of interactions between the Ash2L and DPY-30 subunits of SET1 methyltransferase complexes

- Learned to purify proteins belonging to SET1 methyltransferase complexes and reconstitute various subcomplexes to carry out crystallization trials
- Probing protein-protein interactions *in vitro* and in cells, and analyzing these interactions at the molecular level
- First in the Couture lab to study protein-lipid interactions and develop assays to characterize the lipid-binding properties of Ash2L
- Trained and supervised two undergraduate students (one COOP student, and one Honours student)
- Second author on a review chapter on methyltransferases, first author of a review currently in progress, and generated data to write two separate manuscripts.
- Attended 7 conferences in Canada
- Obtained the QEII-GSST award and excellence scholarship for the first year of my masters degree, and the OGS scholarship for the second year of my masters degree

Honours Bachelor of Science in Biochemistry

September 2010-April 2014

- Graduated Summa Cum Laude
- Faculty of Science Dean's Honor List (2011-2014)
- University of Ottawa Admission Scholarship for undergraduate studies of 3000\$ (2010-2014)
- Scholarship for studying in French, 1500\$ (2010)

PREVIOUS RESEARCH WORK:

- Institute of Research at Monfort Hospital, Unité de Recherche en Nutrition et Métabolisme (Ottawa, Ontario) *June-August, 2013*
 - Project : To assess the impact of the PCB pollutant on the expression of the complexes of the electron transport chain in soleus muscle in rats
 - Contribution : Helped establish the protocol for the homogenization of rat muscle tissue and immunoblotting of complexes I,II,III, and IV of the respiratory transport chain within the crude lysate.

PUBLICATIONS AND WORKS IN PROGRESS :

- S. Lanouette, **J. Haddad**, P. Zhang, J.-F. Couture. Chromatin Signaling and Diseases. Chapter 2 : Impacts of Histone Lysine Methylation on Chromatin. 2016.

- **J. Haddad**, S. Yeung and J.-F. Couture. Recognizing asymmetry in pseudo-symmetry; structural insights into the interaction between amphipathic α -helices and X-bundle proteins. Plan to submit to BBA journal February 2016.
- Working on a manuscript entitled “Structural insights into the insertion of the SDI amphipathic α -helix of Ash2L into the DPY-30 X-type four-helix bundle.”
- Working on a manuscript entitled “The Ash2L subunit of SET1 methyltransferases binds to anionic lipids and can function outside of its nuclear roles”

NATIONAL CONFERENCES :

Posters were presented at the following conferences/events :

- 14th PROTEO Annual Symposium, Université de Laval, Quebec City, Canada, May 2014
- 8th Annual Groupe de Recherche Axé sur la Structure des Protéines (GRASP) symposium, McGill University, Montreal, Canada, November 2015
- Annual BMI Poster day, University of Ottawa, Ottawa, Canada, May 2015
- 15th PROTEO Annual Symposium, Université de Laval, Quebec City, Canada, June 2015
- OISB Science Retreat, Montebello, Canada, June 2015 (Institutional conference)
- 9th Annual Groupe de Recherche Axé sur la Structure des Protéines (GRASP) symposium, McGill University, Montreal, Canada, November 2016
- 16th PROTEO Annual Symposium, Université de Laval, Quebec City, Canada, May 2016
- 2nd PEC conference, University of Ottawa, Ottawa, Canada, June 2016

A presentation was given at the following event :

- Annual BMI Seminar day, University of Ottawa, Ottawa, Canada, March 2016