

KU Leuven
Biomedical Sciences Group
Faculty of Medicine
Department of Pharmaceutical and Pharmacological Sciences
Laboratory of Molecular Virology and Gene Therapy



DEVELOPMENT OF A NEW THERAPEUTIC STRATEGY FOR THE TREATMENT OF MLL-REARRANGED ACUTE LEUKEMIA

Siska VAN BELLE

Jury

Supervisor: Prof. Zeger Debyser
Co-supervisor: Dr. Frauke Christ

Chair examining committee: Prof. John Creemers
Chair public defence: Prof. Pieter Annaert

Jury members: Prof. Kim De Keersmaecker
Dr. Arnaud Marchand
Prof. Tom Taghon
Dr. Ulrike Philippar
Prof. Vaclav Veverka

Dissertation presented in
partial fulfilment of the
requirements for the
degree of Doctor in
Biomedical Sciences

November 2022

Public PhD defense
Wednesday 16th of November 2022
Justus Lipsiuszaal (08.16)
Blijde Inkomststraat 21, 3000 Leuven

© 2022, Faculty of Medicine, KU Leuven
Uitgegeven in eigen beheer, Siska Van Belle
ON5, Herestraat 49 bus 1023, 3000 Leuven (Belgium)

Alle rechten voorbehouden. Niets uit deze uitgave mag worden vermenigvuldigd en/of openbaar gemaakt worden door middel van druk, fotokopie, microfilm, elektronisch of op welke andere wijze ook zonder voorafgaande schriftelijke toestemming van de uitgever.

All rights reserved. No part of the publication may be reproduced in any form by print, photoprint, microfilm, electronic or any other means without permission from the publisher.

Dankwoord

Twee verhuizen, veel tipjes, een hoop geduld, een snuifje frustraties en ruim 5 jaar rijping. Dat is mijn thesisrecept en het is klaar.

Ik heb de gewoonte om nogal snel iets te beschouwen als vanzelfsprekend, wat had ik anders moeten doen sinds ik hieraan begon? Maar neen, tijd om even stil te staan bij deze prestatie. Ik heb enorm veel geleerd en daar hou ik van. Een positief resultaat, een nieuw inzicht of een constructieve discussie spreken voor zich, maar het meeste leerde ik uit het oplossen van mijn fouten. Naast wetenschappelijk, was het ook op persoonlijk vlak een leerrijk traject. Vandaag overheerst vooral dankbaarheid. Dankbaarheid voor alle inzichten die ik verworven heb en connecties die ik heb kunnen maken. Want, niet te vergeten, dit recept was nooit tot stand gekomen zonder de volgende toevoegingen:

Zeger, je passie voor wetenschap is aanstekelijk. Bedankt dat ik de afgelopen jaren onderzoek heb mogen uitvoeren in je labo en dat ik kennis kon maken met zoveel verschillende technieken. De evenwichtige combinatie van vrijheid en begeleiding maakte het een fijne omgeving om in te werken. Ook wil ik je bedanken voor de extra verantwoordelijkheid en het vertrouwen dat ik de laatste maanden gekregen heb als labomanager, dit zet nieuwe deuren open om te groeien.

To Frauke, danke. Working with you was (and is) a pleasure. You were a source of motivation, stimulation and support. You were always available for input, from experimental setups to good restaurant suggestions. Your honest opinion and input have helped to shape the project to what it is now and I have the feeling there is much more to learn from you.

I would like to express my gratitude to all members of the jury. Thank you for the constructive feedback during my yearly evaluation moments and the thoroughly revision of my thesis. Your comments have improved the quality of this manuscript.

In een academische setting is er veel komen en gaan. Een aantal keer heb ik mijn dichtste collega's uitgezwaard, maar elk van hen heeft mee vorm gegeven aan dit resultaat. Jan and Sara, the first people I met in MolMed, thank you for such great supervisors in my first years as PhD student. I truly believe my PhD would have turned out differently without the two of you at the beginning. I am grateful for the journey we had together and the knowledge I obtained from both of you. Jolien en Tine, jullie waren een enorme steun en bron van informatie tijdens praktisch werk, meetings en bureau-momentjes! Bedankt voor alle hulp! Ik hoop dat ik jullie toch een beetje trots gemaakt heb met het gedeeltelijk overnemen van jullie projecten. Martine, dankzij jou waren de tripjes naar de muizen net wat aangener, kon ik net wat meer ontspannen op vakantie vertrekken en waren de incubatietijden

net wat gezelliger! Ik vond het enorm fijn om nauw met je samen te werken en geef je met plezier nog lab-updates over dinertjes.

Labbies, to be fair... without amazing MolMedians like you, this would not have been the same. You created a wonderful atmosphere to work in, I loved it. I want to THANK YOU for the yearly weekends, BBQ's, teambuilding days and all the laughter. Also, thanks to you I was able to manage a small sandwich-business anyway. Marjolein, een vaste en cruciale waarde in dit hele recept. Essentieel in wetenschappelijke termen. Je was er niet gewoon bij, maar je hebt dezelfde fases beleefd op ongeveer dezelfde momenten en de steun die ik bij jou heb kunnen vinden is immens. Ik bewonder je vermogen en je inzet en sta binnenkort bij jou op de eerste rij te supporteren. Je weet het normaal wel, maar nog eens: bedankt! To my male office-mates, Yannick and Rodrigo, they say women chit-chat a lot, but I am happy you know how to do it as well! Thank you for making dreary days more fun. Thatcher, you've only been around 20% of my PhD lifetime, but thank god, I'm happy you're in my team! You're a rock. Thank you for your positive vibes and enthusiasm and don't forget to use it on yourself.

Aan de behulpzame staf, merci! Ik kwam tot de vaststelling dat ik elk van jullie weleens benaderd heb met een vraag en dat jullie deze telkens met veel aandacht en vaak succesvol hebben beantwoord. Jullie zijn onmisbaar, en ik ben blij nu onderdeel uit te maken van jullie team! Paulien & Babs, ik vind jullie inzet voor het labo geweldig. Het is een geschenk om met zo'n ervaren en gedreven laboranten te werken. Bedankt voor de babbeltjes en de mentale ondersteuning! Rik, bedankt voor je wetenschappelijk enthousiasme en de wijze woorden waar ik van tijd tot tijd eens aan herinnerd moest worden. Chris en Joris, ook al zijn jullie Neuro's en dus nooit rechtstreeks betrokken geweest in mijn onderzoek verdienen jullie een extra dankjewel! Al was het maar voor de sport- en dierenverhalen ter afleiding of de redding toen ik opgesloten zat in de wc. Ik begrijp perfect waarom jullie de brandverantwoordelijke zijn in het labo. Ook Brigitte, Marc en Nadia hebben dit traject net wat dragelijker gemaakt. Door hen heb ik heel wat tijd kunnen uitsparen met administratie, bestellingen of het maken van liters LB en agar-plaatjes. Bedankt!

Een doctoraat is een hele uitdaging, eentje die ik *twenty-four seven* meedroeg en dus ook buiten de muren van MolMed. Vriendjes, en jullie weten best wie ik bedoel, bedankt. Misschien zonder dat jullie het weten, waren jullie een belangrijk onderdeel in de laatste jaren! Soms was ik vooral aan het klagen, maar bedankt voor al het plezier tijdens het lopen en voetballen, de lunch en diners, de (trouw)feesten, de babybezoekjes, de uitstappen en zoveel meer. Jullie waren één voor één een bron van positivisme, steun en welgekomen afleiding!

Aan mijn (schoon)familie, volgens mij dachten jullie vaak dat ik Chinees studeerde in plaats van onderzoek deed naar leukemie. Toch toonde jullie interesse en was het een leuke uitdaging om mijn

project verstaanbaar te maken naar jullie toe. Het was altijd fijn om eventjes zorgeloos omgeven te zijn door jullie warmte en te kunnen genieten van lekker eten. Bedankt Thomas, het was dan wel in voetbaltermen, hoe kan het ook anders, maar de ondersteuning die jij geboden hebt verdiend een speciaal plaatsje op deze pagina. Bedankt mama en bedankt papa. Zonder de vrijheid en de liefde die jullie aan mij gegeven hebben zou ik niet staan waar ik nu sta. Dankzij jullie kon ik mijn eigen pad kiezen en hierbij was de steun nooit ver zoek. De waarde die hiermee gepaard gaat, kent geen einde.

Tot slot, de persoon die alle geuren en kleuren bij dit recept vanop de eerste rij mee beleefd heeft, mijn liefste vriend (man, n.v.d.r.), Joren. Het was geen geheim wanneer ik het allemaal even beu was of de dag niet gelopen was zoals gepland. Vaak begreep je er inhoudelijk geen hol van, maar je hebt wel eindeloos naar mijn verhalen geluisterd en mij rust geboden wanneer ik deze het hardst nodig had. Je was altijd bereid mij te vergezellen (soms met pizza's) tijdens mijn avond-uitstappen naar het labo, zette de deur al open vlak voor ik laat thuiskwam en leende zonder probleem je truien uit – ook zonder je toestemming. Bedankt voor het geven van je liefde en steun, voor je geduld. Niet alleen tijdens dit doctoraat, maar tijdens de voorbije 13 jaar. Dit had ik met niemand anders gekund dan met jou ♡

Samengevat: bedankt allemaal!

Veel liefs,

Siska

Table of Contents

Table of Contents	I
Abstract	V
Samenvatting.....	VII
List of abbreviations	IX
List of figures	XV
List of tables	XVII
1. Introduction.....	1
1.1. Epigenetic regulation	2
1.1.1. The epigenetic writer KMT2A and KMT2 family.....	6
1.1.2. Structure and function of MLL1	7
1.2. MLL-rearranged leukemia: a distinct and aggressive leukemia.....	10
1.2.1. Pathogenetic mechanism of MLL-r.....	12
1.2.2. Treatment of MLL-r	14
1.2.3. Specific targets for MLL-r leukemia.....	15
1.2.3.1. Small molecules in clinical trials	15
1.2.3.2. Other therapeutic strategies to specifically target MLL-r leukemia	17
1.3. The epigenetic reader family of hepatoma-derived growth factor proteins.....	18
1.3.1. Lens Epithelium Derived Growth Factor	18
1.3.1.1. Structure and function of LEDGF.....	18
1.3.1.2. LEDGF/p75 in MLL-r leukemia	22
1.3.1.3. LEDGF/p75 in other diseases.....	25
1.3.2. HDGF-Related Protein 2	26
1.3.3. The other HDGF-family proteins	28
1.3.4. PWWP domain containing proteins	29
2. Objectives	33
2.1. Investigation of the role of HRP-2 in normal hematopoiesis and MLL-r leukemia	34
2.2. Development of small molecules that target the PWWP domain of LEDGF	34
3. Material and methods.....	37
3.1. Recombinant protein purification.....	38

3.2. Differential Scanning Fluorimetry (DSF).....	40
3.3. AlphaScreen and TR-FRET assay.....	40
3.4. Fragment design and screening	41
3.5. NMR spectroscopy	41
3.6. Cell culture	42
3.7. Co-immunoprecipitation.....	43
3.8. nanoBRET	43
3.9. Viral vector production and generation of stable cell lines	44
3.10. Western Blot	45
3.11. Quantitative PCR	45
3.12. Immunocytochemistry	46
3.13. Suspension and colony formation analysis.....	46
3.14. <i>HRP-2</i> knockout mouse model.....	47
3.15. Animal material processing.....	47
3.16. RNA-sequencing and bioinformatics.....	47
4. Results	51
4.1. <i>HRP-2</i> is dispensable for MLL-r leukemogenesis but important for leukemic cell survival	52
4.1.1. <i>HRP-2</i> interacts with MLL1 in the absence of menin	52
4.1.2. <i>HRP-2</i> and LEDGF/p75 interact with MLL1 through a conserved interface.....	53
4.1.3. Systemic <i>HRP-2</i> depletion in mice leads to increased postnatal mortality and decreased <i>in vitro</i> colony formation of hematopoietic stem cells.....	56
4.1.4. <i>HRP-2</i> depletion impairs the clonogenic growth of both human and mouse leukemic cell lines independently of MLL1 fusions.....	59
4.1.5. <i>HRP-2</i> overexpression rescues MLL1 clonogenic growth in LEDGF/p75-depleted cells	62
4.1.6. <i>HRP-2</i> is dispensable for the initiation of MLL-r leukemia.....	64
4.2. Development of small molecules targeting the PWWP domain of LEDGF/p75 in the context of mixed lineage leukemia.....	66
4.2.1. Targeting the PWWP domain of LEDGF inhibits clonogenic growth.....	66
4.2.2. Recombinant protein production for <i>in vitro</i> studies	67
4.2.3. AlphaScreen as a screening assay for the PWWP-nucleosome interaction <i>in vitro</i>	69
4.2.4. TR-FRET as <i>in vitro</i> screening assay for the PWWP-nucleosome interaction	75
4.2.5. Detection of nucleosome-LEDGF/p75 interaction in cells using nanoBRET.....	79
4.2.6. Structure based drug discovery of ligands of the <i>HRP-2</i> PWWP domain	81

4.2.7. Fragment hits bind to LEDGF _{PWWP} and HRP-2 _{PWWP} in a dose-dependent manner.....	84
4.2.8. Fragment hits bind to other PWWP-domains of PWWP domain containing proteins	85
4.2.9. Analysis of selected hits in AlphaScreen and TR-FRET	87
4.2.9.1. AlphaScreen.....	87
4.2.9.2. TR-FRET.....	89
5. Discussion	95
5.1. The function of HRP-2 in normal hematopoiesis	96
5.2. The function of HRP-2 in cell survival and MLL-r development.....	97
5.3. HRP-2 as target for MLL-r leukemia?	99
5.4. Validating PWWP as target for MLL-r leukemia.....	100
5.5. Assay development to detect the PWWP-nucleosome interaction.....	100
5.6. Discovery of small ligands targeting LEDGF _{PWWP} to treat MLL-r leukemia	102
5.7. The challenge to avoid toxicity and obtain LEDGF _{PWWP} specific activity	103
5.8. Conclusion and further perspectives	105
Acknowledgements, personal contribution and conflict of interest statements	109
References.....	113
Curriculum vitae	

Abstract

Mixed lineage leukemia-rearranged (MLL-r) leukemia is a genetically distinct subtype of leukemia characterized by chromosomal translocations in the MLL1 gene *KMT2A*. This rearrangement results in the expression of an oncogenic MLL-fusion protein (MLL-FP). To date, more than 100 fusion partners have been described and the majority of them enhance transcription elongation by stimulating RNA polymerase II. MLL-rearrangements are a poor prognostic factor associated with a low event-free survival. To obtain a better clinical outcome, multiple efforts are made to explore treatment options that specifically target key players in the MLL-r mechanism. One such target is the ternary complex of MLL1, menin and lens epithelium derived growth factor, isoform p75 (LEDGF/p75). Menin was described as molecular adaptor linking MLL1 to LEDGF/p75, whereas the latter serves as tether for the MLL1 complex to its target genes. At the amino terminus, LEDGF encompasses a LEDGF_{PWWP} domain that specifically reads methylated-lysine residue 36 of histone 3. We hypothesize that this pathological mechanism can be reversed by constraining the LEDGF/p75 chromatin binding site to reduce the MLL-FP driven expression.

In this thesis, we confirmed the LEDGF_{PWWP} domain as potential target to treat MLL-r leukemia. To investigate the interaction between nucleosomes and LEDGF/p75 we established two *in vitro* assays, AlphaScreen and TR-FRET and one interaction assay in a cellular context, the nanoBRET. The W21A, F44A double mutant of the LEDGF_{PWWP} domain that can no longer interact with nucleosomes was used to determine specificity. Interestingly, the LEDGF_{PWWP} protein domain was able to not only interact with the trimethylated mark of H3K36, but also interacted with non, mono and di-methylated nucleosomes but to a lesser extent in both *in vitro* assays. Following a structure-based drug discovery strategy, supported by X-ray crystallography and NMR, we aim to develop a small molecule that specifically targets the LEDGF_{PWWP} pocket. A small in-house screening and a medium-scale screening at the XChem facility in Oxford resulted in 106 hits. In this thesis, we characterized six confirmed fragments in differential scanning fluorimetry (DSF), AlphaScreen and TR-FRET. Over all experiments, fragments β and ζ showed the most promising results with temperature shifts of more than three degrees and moderate inhibition of the interaction.

In parallel, we considered the presence of hepatoma derived growth factor related protein 2 (HRP-2), the LEDGF/p75 paralog for which interchangeable functions in context of HIV-1 were reported. Before, no link between HRP-2 and MLL1 nor hematopoiesis or leukemia was described. HRP-2 co-immunoprecipitated with MLL1, but is less dependent on the presence of menin when compared to LEDGF/p75. Analysis of the solution structure of the HRP-2_{IBD} domain by NMR revealed a highly

conserved folding of the IBD domain in comparison to LEDGF_{IBD}. These structural similarities prompted us to investigate the role of HRP-2 in normal hematopoiesis using a systemic knockout mouse model. The *HRP-2* knockout mice presented with an increase in neutrophils in the peripheral blood. *Ex vivo* analysis of lineage depleted bone marrow cells by colony formation capacity and RNA sequencing in comparison to the wild type cells suggested a preferred stem-like state for HRP-2 expressing cells. The role of HRP-2 in MLL-r and other leukemic, human cell lines was evaluated by lentiviral-mediated miRNA depletion of HRP-2 followed by a growth monitoring in colony forming assay and liquid culture. Surprisingly, *HRP-2* knockdown affected growth of all leukemic cell types, irrespective of MLL-fusion proteins. However, HRP-2 overexpression in the absence of LEDGF/p75 suggests that both proteins can carry out similar functions, but are regulated differently. At last, to study whether HRP-2 influences the *MLL-r* leukemogenesis, we performed a bone marrow transplantation experiment to induce MLL-ENL driven leukemia. Interestingly, HRP-2 knockout cells transformed as efficiently as HRP-2 wild type cells in a colony forming assay, and resulted in a shorter life span for the transplanted mice compared to the wildtype.

In conclusion, this research provides novel information on the role of HRP-2 in mixed lineage leukemia. We believe that the HRP-2 interaction with MLL1 is not determinant for the development of MLL-r leukemia and that LEDGF/p75 is the main driver of MLL-r leukemia. Further development of the fragments with a focus on both PWWP domains will determine whether LEDGF_{PWWP} inhibitors are a feasible strategy for the treatment of MLL-r leukemia.

Samenvatting

Een specifieke, agressieve vorm van leukemie wordt gekenmerkt door een chromosomale translocatie in het *KMT2A* gen dat codeert voor het *mixed lineage leukemia 1* eiwit (MLL1). Bij leukemiepatiënten met deze chromosomale afwijking worden kankercellen met kenmerken van verschillende types bloedcellen gedetecteerd. Deze vorm van leukemie gaat gepaard met een ongunstige prognose en een lage overlevingskans, onder meer omdat er geen specifieke behandelingen beschikbaar zijn. De chromosomale herschikking resulteert in de expressie van een kanker-drijvend fusie-eiwit van MLL1 en een fusiepartner die zorgt voor een verhoogde expressie van onder andere *homeobox A* (*HoxA*) genen. Twee belangrijke eiwitten in dit ziektemechanisme zijn menin en *lens epithelium derived growth factor p75* (LEDGF/p75). Menin versterkt de binding tussen MLL1 en LEDGF/p75 waarbij deze laatste belangrijk is voor het richten van MLL1 en de MLL-fusie eiwitten naar het chromatine. Het PWWP-domein in het amino-terminaal uiteinde van LEDGF/p75 (LEDGF_{PWWP}) herkent specifiek lysine 36 op histon drie, gemerkt door twee of drie methyl groepen (H3K36me_{2/3}) en brengt zo de MLL1-eiwitcomplexen naar actief vertaalde genen. We veronderstellen dat het MLL-leukemie ziektemechanisme teruggedraaid kan worden door de binding tussen LEDGF_{PWWP} en histonen te verbreken met behulp van kleine moleculen.

Nadat verschillende publicaties LEDGF/p75 reeds valideerden als potentieel doelwit, tonen wij hier experimenteel aan dat het verstoren van het LEDGF_{PWWP}-domein voldoende is om de leukemiegroei in een kolonie-vormend experiment tegen te gaan, specifiek in MLL-herschikte leukemiecellen. Een dubbel gemuteerde vorm van het LEDGF_{PWWP}-domein (W21A, F44A) dat niet langer aan het chromatine kan binden is niet in staat de groei even sterk af te remmen. Twee *in vitro* experimenten (AlphaScreen en TR-FRET) en een cellulair experiment (nanoBRET) werden uitgewerkt om de directe interactie tussen LEDGF_{PWWP} en het nucleosoom te bestuderen. Het dubbel gemuteerde LEDGF_{PWWP} (W21A, F44A) eiwit werd gebruikt om de specificiteit te bepalen. Additioneel hebben we met deze technieken gezien dat LEDGF/p75 niet enkel aan dubbel of tripel gemethyleerd lysine 36 kan binden, maar in mindere mate ook aan niet- en mono-gemethyleerd lysine 36 bindt. Aan de hand van een op structuur-gebaseerde strategie, ondersteund door x-ray kristallografie en NMR, trachten we kleine moleculen te ontwikkelen die specifiek binden aan de LEDGF_{PWWP} pocket. Een tweedelige screening resulteerde in de detectie van 106 fragmenten die als basis gebruikt worden voor verdere ontwikkeling. In deze thesis werden 6 fragmenten gekarakteriseerd in *differential scanning fluorimetry* (DSF), AlphaScreen en TR-FRET. Samengevat vertonen fragmenten β en ζ het beste resultaat, met temperatuur verschuivingen van meer dan drie graden en een matige inhibitie van de interactie.

In parallel werd de rol van het zuster-eiwit *hepatoma derived growth factor 2* (HRP-2) onderzocht. Eerder onderzoek heeft aangetoond dat HRP-2 en LEDGF/p75 gelijkaardige functies kunnen uitvoeren, maar tot op heden werd geen link tussen HRP-2 en het MLL1 eiwit, noch hematopoëse of leukemie beschreven. Via immunoprecipitatie experimenten tonen we aan dat HRP-2 en MLL1 met elkaar in interactie treden en dat deze interactie minder afhankelijk is van menin dan het geval is voor de LEDGF/p75-MLL1 interactie. De structurele vergelijking tussen LEDGF_{IBD} en HRP-2_{IBD} laat zien dat de vouwing van het integrase binden domein (IBD) evolutionair behouden en sterk gelijkend is tussen beide eiwitten. Om de functionele rol van HRP-2 te bestuderen in hematopoëse werd een muismodel bestudeerd waarbij HRP-2 afwezig is. Een verhoogde hoeveelheid neutrofielen werd gedetecteerd in de HRP-2 knock-out muizen. *Ex vivo* analyse van de hematopoëtische stamcellen in een kolonie-vormend experiment en sequentie analyse van het RNA doet vermoeden dat cellen een voorkeur hebben voor een stamcel fenotype in aanwezigheid van HRP-2.

De rol van HRP-2 in MLL-r en andere humane, leukemie cellijnen werd bestudeerd aan de hand van miRNA-gedreven onderdrukking van HRP-2. In tegenstelling tot wat geobserveerd werd voor LEDGF/p75, tonen de resultaten aan dat minder HRP-2 resulteert in een tragere groei onafhankelijk van de aanwezigheid van MLL-fusie eiwitten in de cellen. Echter lijkt een overmaat van HRP-2 in de afwezigheid van LEDGF/75 een gelijkaardige functie te kunnen uitvoeren in een kolonie-vormend experiment. Tot slot, om te onderzoeken of HRP-2 een invloed heeft op het ontstaan van de MLL-r leukemie, werd een MLL-ENL fusie eiwit tot expressie gebracht en geëvalueerd in een kolonie-vormend experiment of getransplanteerd in muizen. Efficiënte transformatie werd zowel *in vitro* als *in vivo* vastgesteld aan de hand van respectievelijk de toename in kolonies en de verkorte levensduur van de muizen.

Samengevat draagt dit onderzoek bij aan de kennis van HRP-2 in hematopoëse en MLL-herschikte leukemie. We concluderen hieruit dat de HRP-2 interactie met MLL1 niet bepalend is voor de ontwikkeling van MLL-herschikte leukemie en dat LEDGF/p75 de voornaamste drijfkracht is in het mechanisme van deze leukemievorm. Verdere ontwikkeling van de fragmenten met aandacht voor beide PWWP-domeinen zal bepalen of LEDGF_{PWWP} inhibitoren gebruikt kunnen worden voor het bestrijden van MLL-herschikte leukemie.

List of abbreviations

aa	amino acid
AF10	ALL-fused gene on chromosome 10
AF4	ALL-fused gene on chromosome 4
AF6	ALL-fused gene on chromosome 6
AF9	ALL-fused gene on chromosome 9
ALL	acute lymphoid leukemia
ALL1	acute lymphoblastic leukemia gene 1
AlphaScreen	amplified luminescent proximity homogenous assay screen
AML	acute myeloid leukemia
Ash2L	absent, small or homeotic 2 like
ASK	activator of S-phase kinase
AT	AT-hook motif
BC	back complementation
BCR	breakpoint cluster region
BET	bromo-and extra-terminal domain
biot-N	biotinylated nucleosome, recombinant
BMC	bone marrow cell
BRD1-4	bromodomain containing proteins 1-4
BSA	bovine serum albumin
CBP	CREB-binding protein
cDNA	copy DNA
CFU	colony forming unit
CLP	common lymphoid progenitor
CML	chronic lymphoid leukemia
CMP	common myeloid progenitor
co-IP	co-immunoprecipitation
CR1-4	charged region 1-4
CtIP	C-terminal binding protein interacting protein
DAPI	diamidino-2-phenylindole
DM	double mutant (W21A, F44A)
DMEM	dulbecco modified eagle medium

DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
DOT1L	disruptor of telomeric silencing 1 like protein
DSF	differential scanning fluorimetry
DTT	dithiothreitol
EDTA	ethylenediaminetetraacetic acid
EFS	event-free survival
EG	ethylene glycol
ELL	eleven-nineteen lysine-rich leukemia
ENL	eleven-nineteen leukemia
FBS	fetal bovine serum
FP	fusion partner
FYRC	FY-rich C-terminal
FYRN	FY-rich N-terminal
GSEA	gene set enrichment analysis
H3K36	lysine 36 of histone H3
H3K4	lysine 4 of histone H3
H3K79	lysine 79 of histone H3
H4K20	lysine 20 of histone 4
HATH	homologous to amino-terminus of HDGF
HCC	hepatocellular carcinoma
HDGF	hepatoma derived growth factor
his ₆	6x histidine tag
HIV-1	human immunodeficiency virus 1
Hox	homeobox
HR3	homology region III
HRP1-4	HDGF-related proteins 1-4
HSC	hematopoietic stem cell
HTRX1	human trithorax 1
IBD	integrase binding domain
IBM	IBM binding motif
ICC	immunocytochemistry
IPTG	isopropyl- β -D-thiogalactopyranoside

IWS1	interacts-with-Spt6
JPO2	cell division cycle-associated 7-like protein
kb	kilobases
K_D	dissociation constant
KMT2(A-G)	lysine-methyltransferase 2 (A-G)
LB	lysogeny broth
LEDGF/p75	lens epithelium derived growth factor isoform p75
LEDGIN	LEDGF-integrase inhibitor
lin ⁻	lineage depleted hematopoietic stem cells
MBM	menin binding motif
MBP	maltose binding protein
mBU	milliBRET Units
MDS	myelodysplastic syndrome
me1, me2, me3	mono-, di-, and trimethylation
MeCP2	methyl CpG binding protein 2
MED1	mediator subunit 1
MEN1	multiple endocrine neoplasia type 1
(m)HSC	(murine) hematopoietic stem cell
MI	menin inhibitor
MLL Mut	mutant MLL (F9A, P10A and P13A)
MLL1	mixed lineage leukemia 1
MLL1 _C	320 kDa carboxy-terminal fragment
MLL1 _N	180 kDa amino-terminal fragment
MLL-r	mixed lineage leukemia-rearranged
MLL-FP	MLL-fusion partner
MPAL	mixed phenotype acute leukemia
MPP	multipotent progenitor
MSCV	murine stem cell virus
MSH6	mutS homolog 6
nanoBRET	nano-Bioluminescence resonance energy transfer
nanoLuc	nanoLuciferase
NLS	nuclear localization signal
NSD2/3	nuclear receptor binding SET domain protein 2/3

PAF1c	polymerase associated factor 1 complex
PBS	phosphate-buffered saline
PEI	polyethylenimine
PIN	percentage of inhibition
PogZ	pogo transposable element-derived protein with zinc finger
Pol II	polymerase 2
PPI	protein-protein interaction
PSIP1	PC4- and SFRS1-interacting protein 1
PTD	partial tandem duplication
PTM	post-translational modification
P-TEFb	positive transcription elongation factor b
PVDF	polyvinylidene fluoride
PWWP	proline - tryptophan - tryptophan - proline
(q)PCR	(quantitative) polymerase chain reaction
RbBP5	retinoblastoma protein 5
RFU	relative fluorescent units
RNA	ribonucleic acid
RPMI	Roswell Park memorial institute
SAH	S-adenosyl homocystein
SAM	S-adenosyl methionine
SET	su(var)3-9, enhancer-of-zeste and trithorax
SETD1A/B	SET domain containing protein 1A/1B
SLiM	single linear motifs
SO	SYPRO orange
SRD	supercoiled-recognition domain
ssDNA	single strand DNA
STE	sodium - tris - EDTA
TAD	transcriptional activation domain
TiMs	TFIIS-interaction motifs
T _m	melting temperature
TNDs	TFIIS N-terminal domains
TPG	translocation partner genes
TR-FRET	time-resolved fluorescence energy transfer

TSS	transcription start sites
TU	titer units
WBC	white blood cell
WDR5	WD-repeat protein 5
WHSC1	Wolf-Hirschhorn syndrome candidate 1
WHSC1L1	WHSC1-like protein 1
WRAD-complex	complex of WDR5, RbBP5, Ash2L and DPY30

List of figures

Introduction

Figure 1.1	Compaction of genetic DNA into chromosomes.....	2
Figure 1.2	Schematic representation of the different epigenetic players.....	4
Figure 1.3	Classification and function of the KMT2 family	6
Figure 1.4	Schematic representation of the MLL1 wild type complex	8
Figure 1.5	MLL1-menin-LEDGF/p75 ternary interaction model	9
Figure 1.6	The process of hematopoiesis	11
Figure 1.7	Schematic representation of MLL-rearranged fusion protein complex	13
Figure 1.8	Potential targets to treat MLL-r leukemia	16
Figure 1.9	Schematic structure of LEDGF/p75	19
Figure 1.10	Validation of LEDGF/p75 as potential target for MLL-r leukemia.....	23
Figure 1.11	Schematic structure of HRP-2	27

Results

Figure 4.1	Menin is dispensable for the interaction of MLL1 and HRP-2.	53
Figure 4.2	HRP-2 IBD and LEDGF/p75 IBD interact with MLL1 in a conserved manner	54
Figure 4.3	<i>Hrp-2</i> knockout mice show subtle hematopoietic defects	57
Figure 4.4	RNA sequencing analysis of <i>Hrp-2</i> wild type and knockout mice	58
Figure 4.5	<i>HRP-2</i> knockdown impairs growth in human leukemic cell lines	60
Figure 4.6	HRP-2 depletion variably affects growth of human cell lines in liquid culture....	61
Figure 4.7	<i>Hrp-2</i> knockdown impairs growth in murine leukemic cell lines.....	61
Figure 4.8	The LEDGF/p75-driven leukemic phenotype is rescued by HRP-2 overexpression.	62
Figure 4.9	Colony formation of Nalm6 cells is unaffected by LEDGF/p75 depletion and HRP-2 rescue	63
Figure 4.10	HRP-2 is not required for the initiation of leukemia <i>in vivo</i>	65
Figure 4.11	Validation of flag-LEDGF _{PWWP} overexpressing cell lines	67
Figure 4.12	Reduced colony formation upon LEDGF _{PWWP} WT overexpression in MLL-r cell lines	68
Figure 4.13	Validation and characterization of recombinant proteins used in this chapter..	70

Figure 4.14	AlphaScreen interaction between flag-LEDGF/p75 and his ₆ -LEDGF _{PWWP} and recombinant biotinylated H3K36 trimethylated nucleosome	72
Figure 4.15	Amino residues W21 and F44 of the PWWP domain are important for the nucleosome interaction	73
Figure 4.16	LEDGF _{PWWP} interacts with all methylation states of the H3K36 recombinant nucleosomes <i>in vitro</i>	74
Figure 4.17	HRP-2 _{PWWP} interacts with trimethylated recombinant nucleosomes <i>in vitro</i>	75
Figure 4.18	Amino residues W21 and F44 of the PWWP domain are important for the nucleosome interaction in TR-FRET	77
Figure 4.19	The binding of LEDGF/p75 _{DM} to nucleosomes is less dependent on the methylation state of H3K36	78
Figure 4.20	Optimization of the nanoBRET assay	80
Figure 4.21	Crystals of the HRP-2 _{PWWP} domain	82
Figure 4.22	Schematic representation of the followed screening strategy	83
Figure 4.23	DSF analysis of selected fragments with LEDGF _{PWWP} or HRP-2 _{PWWP}	85
Figure 4.24	DSF analysis of selected fragments with BRPF2 _{PWWP} or WHSC1 _{PWWP1}	86
Figure 4.25	Fragment testing in AlphaScreen interaction assay	88
Figure 4.26	Dose dependent inhibition of the LEDGF/p75-nucleosome interaction in TR-FRET	90

List of tables

Introduction

Table 1.1	Overview of most common epigenetic histone modifications and their effect	3
Table 1.2	Characteristics of common methylation sites linked to various cancers.	5
Table 1.3	KMT2 family characteristics	7
Table 1.4	Overview of PWWP-domain containing proteins.....	30

Materials and Methods

Table 3.1	Primers used for qPCR analysis.....	46
-----------	-------------------------------------	----

Results

Table 4.1	Genotype of heterozygous ($Hrp-2^{+/-}$) crossed off spring after birth or weaning (6 to 8 weeks).....	56
Table 4.2	Summary of fragment characteristics.....	91

1

Introduction

1.1. Epigenetic regulation

Epigenetics is known as the dynamic process that induces modifications of DNA or histones without changing the DNA nucleotide sequence. This can be obtained by post-translational modifications of DNA or histones [1] or non-coding RNA, including micro RNA [2]. These modifications influence cellular processes such as chromatin condensation, gene expression, DNA damage repair, cell cycle regulation, cell faith and many more. A balanced epigenome has proven to be crucial for homeostasis.

One important phenomenon, regulated by epigenetics, is chromatin condensation and the associated accessibility for gene transcription (Figure 1.1). Chromatin condensation is realized by the compact spatial organization of nucleosomes, which are histone octamers wrapped with DNA. A series of nucleosomes is called chromatin. Once the chromatin is packed together and highly condensed as heterochromatin, genes are not transcribed. In the euchromatin state, the DNA is less condensed and genes are accessible for transcription. Modifications of both DNA and histones influence the compactness of chromatin and thus gene expression and involve either the addition of chemical groups to one residue or structural conversion (Table 1.1).

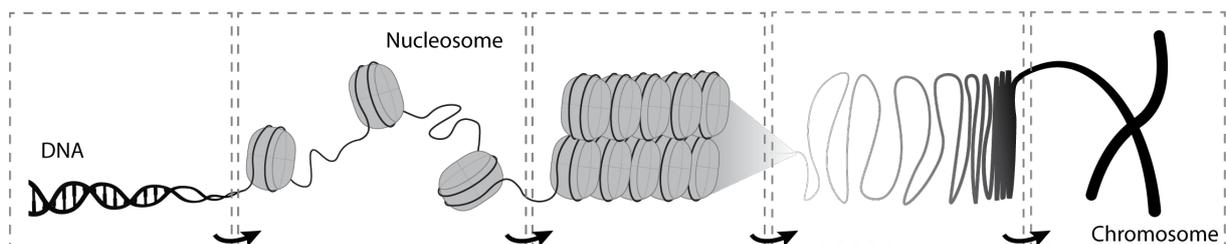


Figure 1.1: Compaction of genetic DNA into chromosomes

DNA is compacted at several levels, indicated by the black arrows. From left to right: double stranded DNA unbound to nucleosomes, DNA wrapped around nucleosomes in the euchromatin state, DNA wrapped around nucleosomes in the heterochromatin state, further compaction of chromatin into chromosomes. In the first two boxes, gene transcription is likely to occur. In the last three boxes, gene transcription mechanisms will be less active.

For DNA, only methylation has been described (Table 1.1) [3]. DNA methylation occurs mainly on cytosines belonging to CpG islands. Seventy percent of CpG islands are located in gene promoters and gene bodies and methylation is often related to gene silencing [4]. To date, more than ten different types of histone PTMs have been described (Table 1.1), of which methylation, acetylation, phosphorylation and ubiquitination are the most abundant modifications. Interestingly, some amino acids can be readily adjusted, such as lysine 4 on histone 3 (H3K4), which can be modified by addition of either an acetyl or methyl group. In addition, methylation can occur at four different levels: un-, mono-, di- or tri-methylated. One modification can induce opposite regulatory effects. Where

sumoylation was long described as inducing transcriptional repression, a more recent finding indicates sumoylation-driven transcriptional activation [5]. Strong evidence is given that histone modifications are context- and cell type-dependent. In addition, cross-regulation for different modifications has been described [6]–[10], underlining the complexity of gene regulation by post-translational modifications. Most modifications are involved in DNA damage repair (Table 1.1) [11]. Additionally, epigenetic modifications are contributing to the memory of cells and can be transferred from parental DNA to the daughter strands [12].

Table 1.1. Overview of most common epigenetic modifications and their effect

Epigenetic modification	Main target site(s)	Modification	Writer	Eraser	TXN reg. effect	dsDNA break repair	Ref.
DNA							
methylation	CpG island	methyl group	DNA methyltransferase	TET and TDG	↓	✓	[13], [14]
Histone							
acetylation	K	acetyl group	acetyltransferase	deacetylase	↑	✓	[15], [16]
methylation	R, K	methyl group	Histone methyltransferase	demethylase	↑, ↓	✓	[17], [18]
phosphorylation	S, T, Y	phosphate	kinase	phosphatase	↑, ↓	✓	[19]–[21]
ubiquitylation	K	ubiquitin	ubiquitin ligase	deubiquitinase	↑, ↓	✓	[22]–[24]
O-linked glycosylation	S, T	O-linked N-acetylglucosamine	O-GlcNAc transferase	O-GlcNAcase	↑, ↓	✓	[25], [26]
sumoylation	K	small ubiquitin-related modifier	sumo ligase	sumo-specific isopeptidases	(↑), ↓	✓	[5], [23]
crotonylation	K	crotonyl	crotonyl transferase	decrotonylase	↑, ↓	✓	[27], [28]
ADP-ribosylation	K	ADP-ribose	ADP-ribosyl transferase	ADP-ribose hydrolase	↑	✓	[29], [30]
isomerization	P	conformational change cis ↔ trans	isomerase		(↑), ↓		[6]
citrullination (deimination)	R	conversion to citrulline	arginine deiminase	--	(↑), ↓		[10], [31], [32]

TXN reg, transcriptional regulation; ↓, transcriptional repression; ↑, transcriptional activation; CpG island, cytosine and guanine repeats; K, lysine; R, arginine; S, serine; T, threonine; Y, tyrosine; P, proline; TET, ten-eleven translocation methyl-cytosine dioxygenase; TDG, thymine-DNA glycosylase; --, effect is irreversible.

Key players in the regulation of PTMs are proteins known as writers, readers and erasers of a histone mark (Figure 1.2). Addition of a chemical group to residues is carried out by epigenetic writers. A class of enzymes that introduces or 'writes' the modification on DNA and histones. Epigenetic readers are proteins that encompass structured domains that specifically read and thus interact with those DNA or histone modifications. Often, this interaction is crucial for protein activity or gene regulation. Finally, epigenetic modifications can be removed by epigenetic erasers, reversing the associated activity.

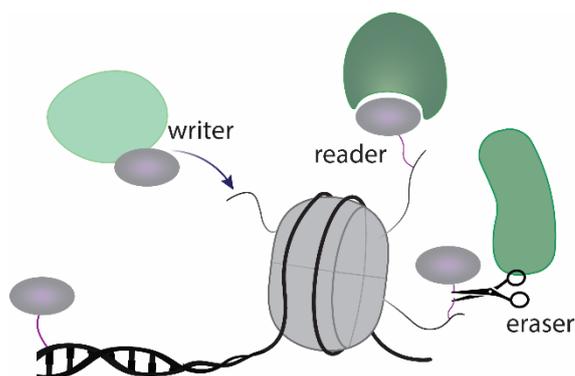


Figure 1.2: Schematic representation of the different epigenetic players

Epigenetic modifications (purple oval) leave the DNA sequence unchanged but can affect chromatin condensation and gene expression. This dynamic process, named epigenetics rely on three key players. On the one hand epigenetic writers place a modification on DNA or histones (light green). These modifications can be recognized by epigenetic readers (dark green), or removed by epigenetic erasers (right).

In light of this manuscript, some key methylations are listed in Table 1.2, foreseen of references to publications further elaborating on this methylation mark. Regulation of transcription by methylation is two sided and can be different between (i) different target residues, (ii) different methylation levels such as for H4K20 or (iii) a different environment, as is the case for H3K9 trimethylation. H3K9 methylation is generally considered to be associated with transcriptional repression, however Vakoc C. R. *et al.* suggest H3K9 methylation is associated with activation [33]. Overall, methylation is not only important for gene transcription regulation, but also plays an important role in the repair of DNA damage. Table 1.2 includes whether the methylation mark represses or stimulates gene transcription and which marks are involved in the DNA repair mechanism.

Growing evidence is given that an unbalanced epigenetic landscape impacts the pathogenesis of various cancers. Not solely genetic changes but also epigenetic changes contribute to the transformation and proliferation of malignant cells. Interestingly, alterations in the epigenome occur fast and are reversible, highlighting the importance of controlling these epigenetic changes in tumor cells. For all methylation marks listed in Table 1.2, de-regulation has been linked to leukemogenesis

and solid tumors. Often epigenetic writers, reader or erasers of the methylation mark are mutated or alternatively expressed. Furthermore, an aggressive form of high-grade glioma is characterized by a K27M mutation in histone 3 itself, reducing the possibility to methylate H3K27 [34], [35]. It follows that epigenetic writers, readers and erasers are being considered as emerging anti-cancer drug targets. Small molecule inhibitors against DOT1L [36]–[38], BET proteins [39]–[41] and histone deacetylases [42]–[44] serve as promising examples of targeting an epigenetic writer, reader or eraser, respectively, to treat associated diseases.

Table 1.2. Characteristics of common methylation sites linked to various cancers

Methylation mark	Associated region	Function	TXN reg. effect	Link to cancer?	Ref.
DNA	CpG islands (gene promotor)	Occupation and chromatin compaction	↓	AML, ALL, CML, CLL, MDS, solid tumors	[45]–[49]
H3K4me1-3	gene promotor, gene enhancer	Chromatin remodeling, DNA repair	↑	AML, ALL, lymphoma, solid tumors	[50]–[54]
H3K9me3	heterochromatin (repetitive non-coding sequences)	Chromatin compaction, maintenance of ES cells and organogenesis	↓,(↑)	AML, ALL, CLL, solid tumors	[55]–[59]
H3K27me3	gene promotor	Lineage commitment and differentiation	↓	AML, T-ALL, CLL, MDS, different lymphomas, high-grade glioma, solid tumors	[60]–[63]
H3K36me3	gene bodies	Gene expression stability, DNA damage repair	↑	AML, ALL, MLL, solid tumors	[64]–[68]
H3K79me3	transcribed gene regions	Promote transcription elongation, DNA repair, cell cycle regulation	↑	AML, MLL, MPAL, solid tumors	[69]–[73]
H4K20me	(me1, me2) end of gene bodies, damaged DNA (me3) promotor	(me1, me2) DNA repair and cell cycle (me3) formation and maintenance of heterochromatin	↑ ↓	Solid tumors	[74]–[76]

TXN reg, transcription regulation; ↑ transcriptional activation; ↓, transcriptional repression; MDS, myelodysplastic syndrome; AML, acute myeloid leukemia; (T-)ALL, (T-cell) acute lymphoblastic leukemia; CML, chronic myeloid leukemia; CLL, chronic lymphoblastic leukemia; BAL, bi-phenotypic leukemia; MLL, mixed lineage leukemia

1.1.1. The epigenetic writer KMT2A and KMT2 family

One protein family of interest, known to specifically write methyl groups on the lysine residue at position 4 of the histone three (H3K4), is the lysine-methyltransferase 2 (KMT2) family (Figure 1.3) [77], [78]. The KMT2 family encompasses all three methylation states of this mark. The seven protein family members (KMT2A to G) contain a Su(var)3-9, Enhancer-of-zeste and Trithorax (SET) domain, necessary to transfer a methyl group from S-adenosyl methionine (SAM) to H3K4 (Figure 1.3B) [79]. The KMT2 protein members are classified in four groups, KMT2A and KMT2B, KMT2C and KMT2D, KMT2E alone, and KMT2F and KMT2G (Figure 1.3A, Table 1.3) based on their domain-characteristics and slightly difference in substrate specificity and chromatin occupancy [80], which results in different physiological functions (reviewed in [81]). Unlike the other members, the KMT2E protein exhibits no activity to methylate nucleosome substrates [78]. For all KMT2 family members, disease-specific mutations have been identified [81], but the most extensively studied aberrations are the chromosomal rearrangements involving the *KMT2A* gene.

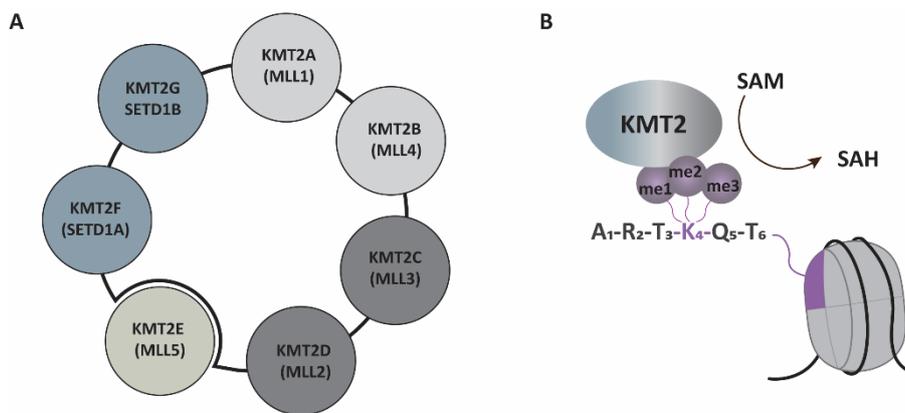


Figure 1.3: Classification and function of the KMT2 family

(A) The KMT2 protein family consist of seven members (KMT2A to KMT2G), classified into four groups, represented by different colors. Most common alternative names are mentioned between brackets. KMT2: lysine-methyl transferase 2; MLL, mixed lineage leukemia; SETD1, SET domain containing protein 1. (B) All family members are characterized by a Su(var)3-9, Enhancer-of-zeste and Trithorax (SET) domain that can mono-, di- or trimethylate (me) lysine 4 on histone three (H3K4, blue) in addition of SAM. One exception is the KMT2E protein, for which no methyl-transferase activity was described. KMT2: lysine methyltransferase 2, SAM: S-Adenosyl Methionine, SAH: S-Adenosylhomocysteine

Table 1.3. KMT2 family characteristics

	size (aa)	DNA (or RNA) binding domains	functional SET domain	interaction to WRAD complex
KMT2A	3696	AT, Bromo, CXXC, PHD	✓	✓
KMT2B	2715	AT, Bromo, CXXC, PHD	✓	✓
KMT2C	4911	PHD, HMG	✓	✓
KMT2D	5537	PHD, HMG	✓	✓
KMT2E	1858	PHD	-	-
KMT2F	1707	<u>RRM</u>	✓	✓
KMT2G	1966	<u>RRM</u>	✓	✓

KMT2, lysine methyltransferase 2; *aa*, amino acid; *SET*, *Su(var)3-9*, *Enhancer-of-zeste* and *Trithorax*; *WRAD*, complex of *WDR5*, *RbBP5*, *Ash2L* and *DPY30*; *AT*, AT hook motif; *Bromo*, bromo-domain; *HMG*, high mobility group; *RRM*, RNA recognition motif

1.1.2. Structure and function of MLL1

The earliest and most extensively studied member of the KMT2 family is the KMT2A protein [82], encoded by the *KMT2A* gene. Synonyms for this protein are mixed lineage leukemia 1 (MLL1) [82], human trithorax 1 (HTRX1) and acute lymphoblastic leukemia gene 1 (ALL1), but throughout this manuscript the preferred name will be MLL1. The *KMT2A* gene was first identified in 1991 in the context of acute leukemia [82], [83] and it displays high sequence homology to the trithorax gene present in *Drosophila* (hence alternative name HTRX1) [84]. The human *KMT2A* gene spans a sequence of 87.43 kilobases (kb) on the 11q23 locus. This sequence is divided in 37 exons, coding a 3969 amino acid-long protein with a molecular weight of 431 kDa (Figure 1.4A) [85]. MLL1 is widely expressed in different tissues [86].

After translation, MLL1 is cleaved by *taspase-1* twice [87], [88], resulting in a 180 kDa amino-terminal (MLL1_N) and a 320 kDa carboxy-terminal fragment (MLL1_C) (Figure 1.4). Both fragments form a heterodimer through the interaction of the FY-rich N-terminal (FYRN) and FY-rich C-terminal (FYRC) domains on the N- and C-terminus, respectively [87]. MLL1 comprises many other interaction motives for protein-protein and protein-DNA interactions (Figure 1.4B). The multi-protein complex formation is necessary for the physiological functions of the MLL1 protein. MLL1-driven methylation of H3K4 through the MLL1 SET domain is 600-fold enhanced when MLL1_C interacts with WD-repeat protein 5 (WDR5) in complex with retinoblastoma protein 5 (RbBP5) and absent, small or homeotic 2 like (Ash2L) and additionally, binding of DPY30 to the latter [89]. This complex of WDR5, RbBP5, Ash2L and DPY30 is called the WRAD-complex [90]. Next to the SET domain, a transcriptional activation domain (TAD) is

present in MLL1_c [91]. As the name suggests, TAD increases the gene expression and this is mediated by interaction with the CREB-binding protein (CBP), a known histone acetyltransferase protein associated with active transcription [91], [92].

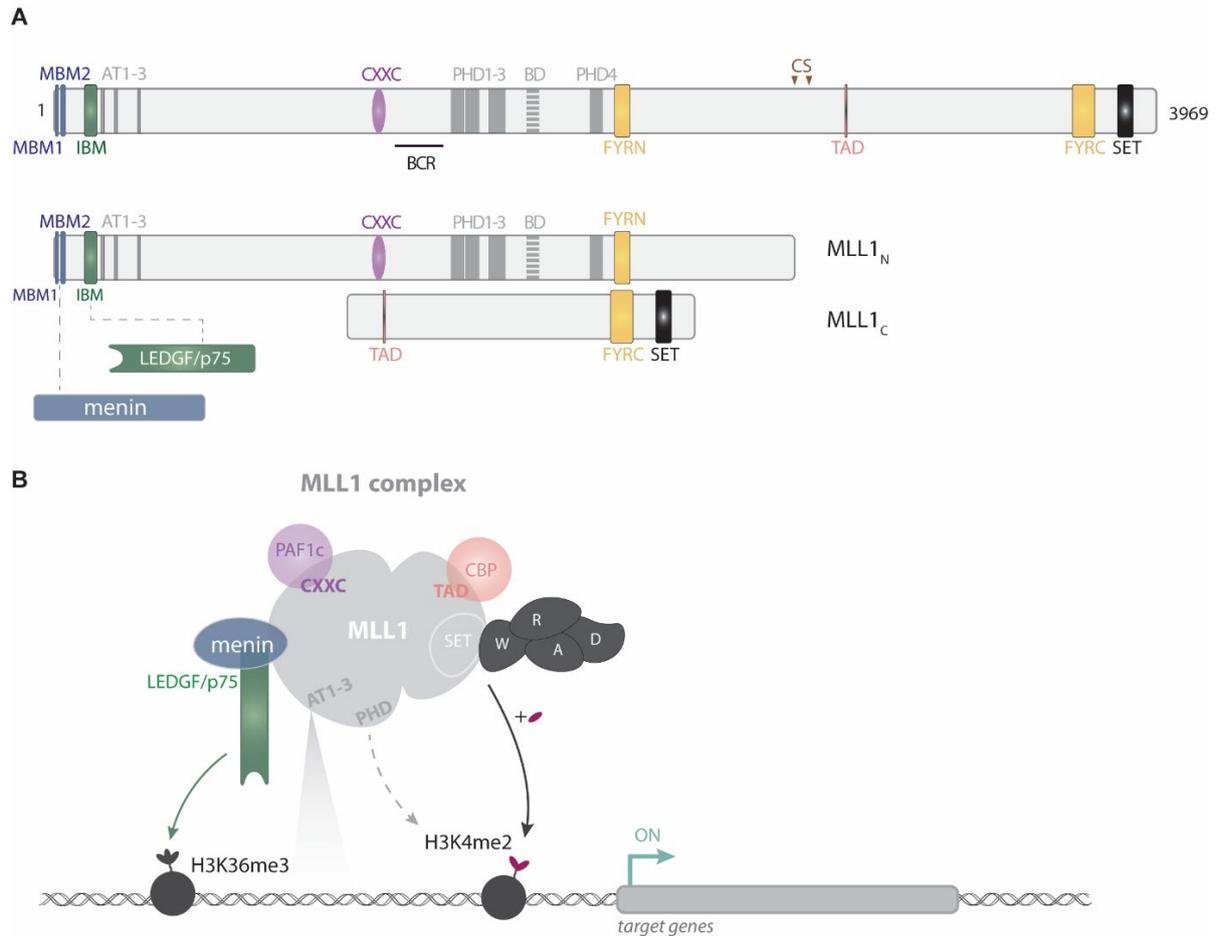


Figure 1.4: Schematic representation of the MLL1 wild type complex

(A) domain structure of MLL1 before (upper) and after (lower) taspase-1 cleavage at the two cleavage sites (CS, arrowheads). MBM, menin binding domain; LBD, LEDGF binding domain; AT, AT-hook motifs; CXXC, zinc finger domain CXXC; BCR, breakpoint cluster region; PHD, PHD fingers; BD, bromodomain; FYRN, FY-rich N-terminal; FYRC, FY-rich C-terminal; TAD, transcriptional activation domain; SET, Su(var)3-9, Enhancer-of-zeste and Trithorax. (B) representation of the functional MLL1 wild type multiprotein complex. N-terminally, MLL1 interacts with menin, LEDGF/p75 and polymerase associated factor 1 complex (PAF1c). At the C-terminus, the WRAD-complex and CREB-binding protein (CBP) associate with MLL1. AT-hook motifs and PHD fingers recognize DNA and methylation marks on lysine 4 of histone 3 (H3K4me2), respectively. LEDGF/p75 binds methylated lysine 36 at histone three (H3K36me3). PHD finger four recognizes H3K4 methylation marks and guides the SET domain to its target to methylate H3K4. Epigenetic regulation by this protein complex is associated with gene regulation of target genes. W, WD-repeat protein 5; R, retinoblastoma protein 5; A, absent, small or homeotic 2 like protein; D, DPY30.

More protein interaction domains, as well as DNA-interaction domains are described for MLL1_N. At the N-terminus, MLL1 harbors two menin binding motifs (MBMs) [93] and a LEDGF/p75 binding motif called IBD-binding motif (IBM) [94]. Interaction with both menin and lens epithelium derived growth factor isoform p75 (LEDGF/p75) results in a well-defined ternary complex. Although *in vitro* experiments indicate that MLL1 interacts with both proteins [95], binding of menin to MLL1 creates a binding pocket for the LEDGF/p75–MLL1 interaction [96]. By formation of this complex, LEDGF/p75 tethers MLL1 to actively transcribed genes by the methyl-lysine reading function in its PWWP domain. The MLL-menin-LEDGF/p75 ternary complex has been studied extensively and modeled by the combination of X-ray crystallography and NMR (Figure 1.5) [96], [97]. Initial interaction studies and crystallography suggested that LEDGF/p75 could not directly interact with a MLL1 peptide (aa 4-153) [96]. Binding of menin to MLL1 induced a V-shaped groove in which the LEDGF_{IBD} domain binds. However, by NMR, Cermakova K. *et al.* showed later that an extended MLL1 fragment up to R160 provides a secondary interaction interface between MLL1 and LEDGF_{IBD} [97]. In line with earlier reported experiments [95], this longer MLL1 fragment interacts with LEDGF/p75 and revealing a second, direct interaction surface (Figure 1.5).

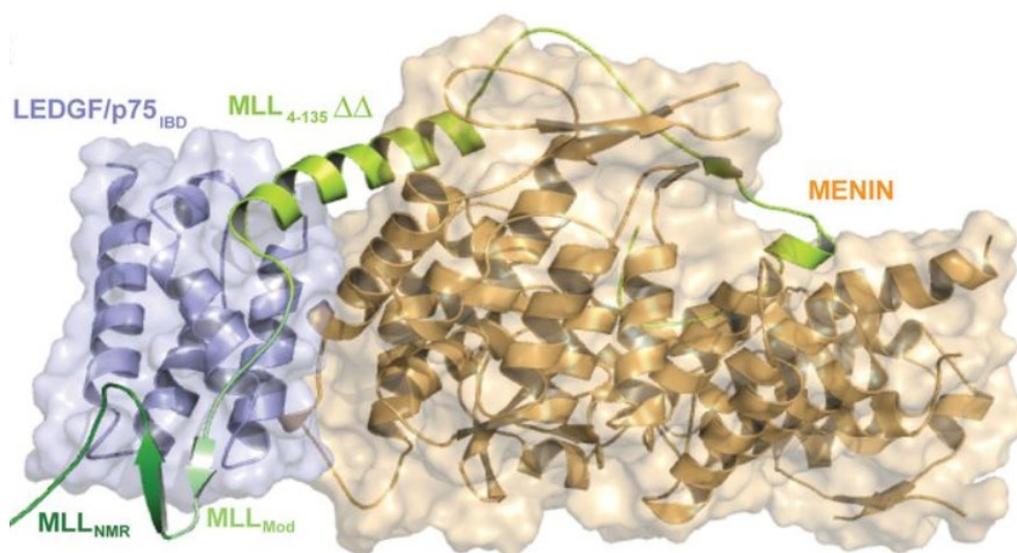


Figure 1.5: MLL1-menin-LEDGF/p75 ternary interaction model

This structural representation of the MLL1-menin-LEDGF/p75 ternary complex was obtained by combining the structural information of the menin, LEDGF_{IBD} and part of MLL1 (aa 4 to 135 with two deletion regions aa 16–22 and 36–102; MLL₄₋₁₃₅ΔΔ) obtained by X-ray crystallography (PDB ID 3U8) with a slightly longer MLL1 fragment (MLL₁₄₀₋₁₆₀) which was studied by NMR (PDB ID 2MSR). This latter revealed an extra interaction interface (MLL_{NMR}, dark green). The connection between both interaction sites was modeled (MLL_{Mod}). Figure copied with permission from [97].

In addition to this epigenetic-driven tethering by LEDGF/p75, MLL1_N also interacts directly with the DNA or modified histones via three AT-hook motifs and different types of Zinc finger domains [98]–[100]. One of these zinc-finger domains, PHD3, is important for the recognition of lysine 4 on histone three, and thus important to properly target the SET domain toward this residue [101]. The Zinc finger CxxC domain interacts with polymerase associated factor 1 complex (PAF1c) [102], [103]. This complex influences a wide range of biological processes [104] and is essential for MLL1 activity [102].

Methylation of H3K4 is associated with active transcription and similar to its trithorax ortholog, MLL1 targeted genes are associated with early development and hematopoiesis as revealed by mouse studies. While complete depletion of *Kmt2a* was lethal [105], [106], *Kmt2a* heterozygous mice presented with retarded growth and hematopoietic abnormalities such as anemia, decreased blood platelets and B-cell populations [105]. In addition, these mice showed a change in homeobox (Hox) gene expression over the anterior-posterior axis [105]. Knocking out *Kmt2a* during developmental stages resulted in defective postnatal hematopoiesis [107]. Interestingly, a conditional *Kmt2a*^{-/-} mouse model in the hematopoietic system presented with normal hematopoiesis. However, these cells failed to reconstitute hematopoiesis after bone marrow transplantation in irradiated recipient mice [106]. Milne T. *et al.* demonstrated that MLL1 regulates H3K4 methylation at Hox gene promoters [108], which are the drivers of body plan formation and hematopoietic stem cell differentiation. The positive Hox-gene regulation by MLL1 is the best studied, but is not the sole group of MLL1 target genes. Comparison of wild type murine stem cells to induced and conditional *Kmt2a* knockout mice in the hematopoietic system revealed an affected set of genes by *Kmt2a* deletion that control self-renewal and proliferation, including *Prdm16*, *Mecom*, *Pbx1* and *Eya1* [109].

Overall, several studies confirmed the importance of the MLL1 protein in regulation of development and normal hematopoiesis. Consequently, alterations in the MLL1 function have been associated with several diseases. Many neurological disorders such as autism spectrum disorders and schizophrenia have shown an abnormal H3K4 methylation pattern [110] and mutations in MLL1 have also been linked to the Wiedemann–Steiner syndrome [111], [112] and cancer [113]. However, the best-known pathological condition associated with MLL1 is MLL-rearranged (MLL-r) leukemia, a distinct subset of bi-phenotypic leukemias [114].

1.2. MLL-rearranged leukemia: a distinct and aggressive leukemia

Leukemia is an umbrella name for many different blood cancers in which the hematologic hierarchy of white blood cells is disturbed, mostly resulting in enormous amounts of immature and nonfunctional blood cells called blasts. The classical description of hematopoiesis (Figure 1.6), the process in which

new functional blood cells are generated, has recently been challenged. This theory is based on determinism and starts from the multipotent hematopoietic stem cell (HSC) in the bone marrow. HSCs have the ability to renew themselves or to differentiate into either the common myeloid progenitor (CMP) or common lymphoid progenitor (CLP) lineage. Through external stimuli and expression of transcription factors, the cells further differentiate into more mature blood cells. During this process, the maturing cells lose the potency to convert in any type of blood cell. However, these initial steps that imply unidirectional and irreversible differentiation had to make way for a stochastic hematopoietic theory [115]–[117]. Here, evidence has been given that hematopoietic stem cells are multipotent but lineage-biased [118]–[120] and that multipotent progenitors (MPPs) feature self-renewal capacity [121]. However, these early progenitors show plasticity and lineage-specific fate can be remodeled as is reported for lymphoid progenitors which commit to the myeloid lineage powered by specific transcription factors [122], [123].

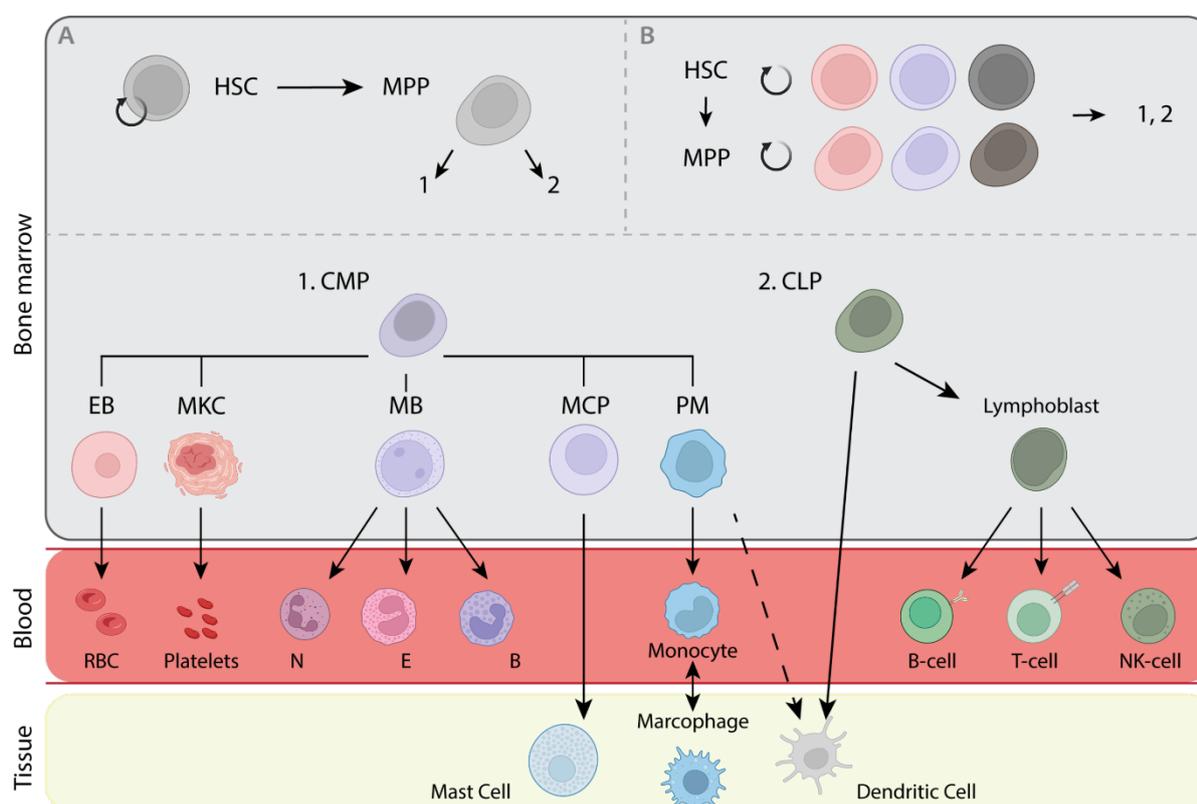


Figure 1.6: The process of hematopoiesis

(A) classical hematopoiesis. Hematopoietic stem cells (HSC) in the bone marrow can renew themselves or differentiate into common progenitor cells, committed to one of the two main lineages: (i) the myeloid (left) or (ii) lymphoid (right) lineage. Common progenitor cells further differentiate into various functionally mature cells that navigate in the blood vessels. (B) stochastic hematopoiesis. Due to plasticity, the whole pool of HSCs and multipotent progenitors with self-renewal capacity can differentiate into the different lineages. CMP, common myeloid progenitor; EB, erythroblast; MKC, megakaryocyte; MB, myeloblast; MCP, mast cell precursor; PM, promonocyte; RBC, red blood cell; N, neutrophil; E, eosinophil; B, basophil; CLP, common lymphoid progenitor; NK, natural killer.

After maturation, most blood cells migrate into the blood circulation where they perform their function or are transported to infected tissues. Red blood cells (RBC) transport oxygen and platelets are important for blood clotting, whereas the other cell types are important for immune responses.

The conventional classification of leukemia recognizes distinct types of leukemia. The first hallmark is determined by which lineage is affected. Secondly, the rate at which symptoms occur determines the type of leukemia at diagnosis. The leukemia is called acute in case of a rapid increase of symptoms or chronic in case the disease progression develops slowly. The combination of these parameters results in the following main types of leukemia: (i) acute myeloid leukemia (AML), (ii) acute lymphoid leukemia (ALL), (iii) chronic myeloid leukemia (CML) and (iv) chronic lymphoid leukemia (CLL). Overall the prognosis for acute leukemias is worse [124], [125]. Less frequently, both lineages of the hematopoietic hierarchy are affected, resulting in so called bi-phenotypic or mixed phenotype acute leukemia (MPAL) [126]. In these patients, oncogenic changes occur in early progenitors or blasts and present with markers of both lineages [127], [128], or patients who are originally diagnosed with myeloid leukemia acquire a leukemia with lymphoblastic character [129], [130]. The most common abnormalities in the MPAL class are t(9;22) translocations involving the ABL and BCR gene and 11q23 rearrangements causing mixed lineage leukemia-rearranged (MLL-r) leukemia [131], of which the latter will be described in more detail below.

1.2.1. Pathogenetic mechanism of MLL-r

The underlying cause leading to frequent chromosomal translocations in the *MLL1* locus remains elusive but today up to 135 translocation partner genes (TPG) have been described to engage in chromosomal *MLL1* rearrangements causing the MLL-r leukemia [132]. Remarkable, a major breakpoint cluster region (BCR) in *MLL1* was identified between exon 8 and 14 [133] (Figure 1.4A), which is affected in 93.5% of all MLL-r characterized leukemias [132]. More recently, next generation sequencing lead to the identification of a second, smaller BCR more upstream, between intron 19 and exon 24 [134]. To map the different TPGs, more than 2000 patient samples obtained over 13 years were analyzed for their chromosomal rearrangement and characterized into subgroups based on the patients age and leukemic subtype (reviewed in [132]). When the chromosomal translocation results in an in-frame fusion between the open reading frame of *MLL1* and the TPG, a fusion protein is transcribed consisting of *MLL1* and a fusion partner (FP, Figure 1.7A). Remarkably, the six most recurrent TPGs/FPs account for 84.4% of all translocation partner genes found in the study (Figure 1.7B). Five of these six genes, namely ALL-fused gene on chromosome 4 (*AF4*, 36%), *AF9* (19%), eleven-nineteen leukemia (*ENL*, 13%), *AF10* (8%) and eleven-nineteen lysine-rich leukemia (*ELL*, 4%) [132] express proteins with a role in transcription regulation. In more detail, the positive transcription

elongation factor b (P-TEFb) is recruited to the MLL-FP complex through the interaction of P-TEFb bound AF4 with MLL-fused AF4, AF9, ENL or ELL (Figure 1.7C) (reviewed in [135]). P-TEFb stimulates gene transcription by phosphorylation of RNA polymerase II (pol II) to re-activate paused pol II and to stimulate transcription [136]. P-TEFb associated with AF4 and ENL is called an AEP complex [137]. With addition of ELL, the AEP complex forms a super elongation complex (SEC) [138]. Furthermore, AF9, ENL and AF10 accelerate transcription by a direct interaction with the disruptor of telomeric silencing 1 like protein (DOT1L) [135]. DOT1L is a methyltransferase that specifically methylates lysine 79 on histone three (H3K79, Figure 1.7C) [139], [140], a marker for actively transcribed chromatin. This methylation plays a role in several biological processes such as DNA damage, cell cycle regulation and transcription elongation, as well as tumor development [140]. Recruitment of P-TEFb or DOT1L to the MLL-FP target genes is the best described mechanism, however evidence is given for other mechanistic pathways that involve recruitment of the PAF1c or the polycomb repressive complex [136].

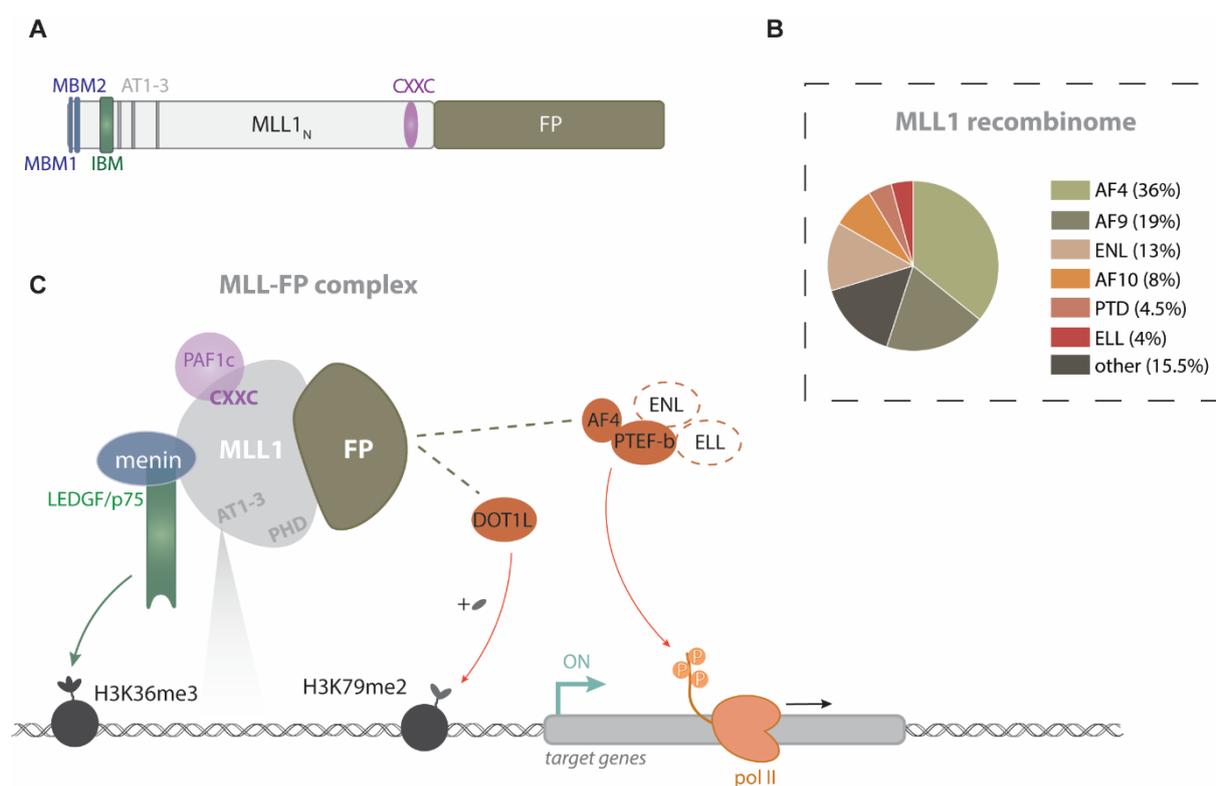


Figure 1.7: Schematic representation of MLL-rearranged fusion protein complex

(A) Schematic representation of rearranged MLL1_N to the fusion protein partner (FP) after chromosomal translocation. (B) distribution of the six most common MLL1 translocation partner genes or fusion proteins. AF4, ALL-fused gene on chromosome 4; AF9, ALL-fused gene on chromosome 9; ENL, eleven-nineteen leukemia; AF10, ALL-fused gene on chromosome 10; ELL, eleven-nineteen lysine-rich leukemia; PTD, partial tandem duplication. (C) mechanism by which MLL-FP upregulates gene expression of target genes is upregulated. Fusion partners often associate with DOT1L or the AF4/PTEF-b/ENL complex. DOT1L methylates lysine 79 on histone 3 (H3K79) and is associated with active gene transcription. The PTEF-b complex stimulates transcription by releasing polymerase II (pol II) from pausing after phosphorylation of the c-terminal tail of pol II. At the N-terminus, MLL1 interaction with LEDGF/p75 and menin is retained.

Thus, in case of MLL-r leukemia, the C-terminal part of MLL1 is often replaced through chromosomal translocation with (part of) one of the previously mentioned TPGs. This implies a loss of SET-domain and KMT function in the MLL-FP, but a gain of function due to the fusion partner. Meanwhile, the direct (AT-hook motifs, Zinc fingers) and indirect (MLL1-menin-LEDGF/p75 ternary complex) DNA binding sites remain in the N-terminal part of MLL1, meaning the MLL-FP complex is tethered towards identical target genes as wild type MLL1, accompanied with RNA transcription elongation stimulating factors (Figure 1.7C). This gain of function results in the oncogenic MLL-FP, driver of the MLL-r leukemia. Due to the prolonged elongation of the MLL1 target genes, such as the Hox cluster genes, the protein levels are upregulated and disturb normal cell function [141]. Increased Hox expression results in a blockage in the differentiation of the hematopoietic stem cells, impairing maturation and resulting in an increase in non-functional cells [142].

Besides the relevance of the fusion partner, a crucial role for the MLL1-menin-LEDGF/p75 ternary complex (Figure 1.5) at the N-terminus of MLL1 is apparent. Previous studies using deletion constructs of MLL1 have indicated that loss of the MBM resulted in a reversal of the MLL-r leukemogenic phenotype [143], [144]. Moreover, similar effects were observed upon knockdown of LEDGF/p75 in MLL-r transformed cells [94], [95]. This indicated that it is possible to target the MLL-FP complex by impairing chromatin association through indirect interactions instead of directly targeting the oncogenic protein itself.

Of note, not only rearrangements, but also duplication and sometimes deletions are seen within the *KMT2A* gene. Up to 4.5% of patients in the screening study of 2017 presented with partial tandem duplication (PTD) [132], where most often exons 3 to 9 are duplicated [145]. Furthermore amplifications of the *KMT2A* gene in the form of intra-chromosomal homogeneously staining regions and ring-chromosomes are detected in several AML and myelodysplastic syndrome (MDS) patient populations [146]–[148]. The exact mechanism by which these MLL-PTD and amplifications lead to leukemia is not yet understood.

1.2.2. Treatment of MLL-r

Unlike most cancers, where multiple events sum up to the perturbation of the malignancy of the cell, MLL-r cells have a very short latency time and aggressive behavior. Whole-genome analysis of MLL-r patients have indicated a low presence of additional genomic changes, underlining the high potency of *KMT2A* gene rearrangements to induce uncontrolled growth of the affected cells [149]–[151]. This may explain why the incidence rate of infants is very high with percentage up to 80% in ALL [152]. The incidence of *MLL1* rearrangements in adults is lower, however not absent. Around 10 percent of adults

with AML or ALL present with a chromosomal translocation involving the *MLL1* gene [152], [153]. In many adult cases, the MLL-r phenotype arises as a secondary event after treatment with topoisomerase II inhibitors such as etoposide, a commonly used anti-leukemic chemotherapeutic [154]. Remarkably, MLL1 rearrangement is associated with a poor prognosis and the young leukemic patients suffer from relapses despite the treatment [124], [155]. Current treatment options are based on standardized AML and ALL chemotherapeutics, but no specific treatment targeting the MLL-r pathogenetic mechanism are available [152]. Multiple centers follow an international protocol, interfant 06 [156], to better monitor the outcome of the treatment. Recent study results revealed an event-free survival (EFS) of 73.9% for patients without MLL1 rearrangements, whereas this EFS percentage decreased to 44.5% and 20.9% for patients with a rearranged MLL1 that were classified as medium risk or high risk, respectively [157]. High risk patients presented with MLL1 rearrangements at an age below 6 months and a highly elevated white blood cell (WBC) count of more than 300×10^9 WBC per liter, making these parameters poor prognostic factors. Why treatment and cure of MLL-r leukemic patients is inefficient is not entirely clear.

1.2.3. Specific targets for MLL-r leukemia

The high unmet medical need has triggered numerous efforts to identify potential targets and the development of inhibitors or small molecules that specifically target the MLL-r driven mechanism to improve the response to chemotherapy and expectancy of life.

1.2.3.1. Small molecules in clinical trials

One decade ago, in 2012, the first menin inhibitors were described [158]. Menin, encoded by the multiple endocrine neoplasia type 1 (*MEN1*) gene, is classified as a tumor suppressor protein because loss of function mutations in the *MEN1* gene often lead to the eponymous syndrome, made up of tumor growth in one or more endocrine glands [159]. Although targeting menin to treat MLL-r leukemia thus appears contraindicative, several publications indicated that the interaction of menin and MLL1 is crucial for the pathogenesis of MLL-r and thus an interesting target to treat MLL-r leukemia (Figure 1.8A). As mentioned before, menin is the stabilizing factor for the interaction between LEDGF/p75 and MLL-FP at the MLL1 amino-terminus. Despite the fact that menin inhibitors (MI) affect the wild type MLL1-menin-LEDGF/p75 ternary complex, they target all chromosomal rearrangements as the N-terminal part of MLL1 is retained in all MLL-FPs.

Over the past 10 years, a gradual improved affinity was obtained for small molecules binding menin [158], [160]–[162]. For the most recent MI, sub-nanomolar potency *in vitro* against the MLL1-menin interaction was reported [160]. The MI-3454 inhibitor specifically targets cell lines with MLL-rearrangements and obtained a complete remission of the leukemia in a patient-derived xenograft mouse model [160]. Additionally, many other groups have started to design diverse menin inhibitors reaching similar affinity for menin and inhibition activity on proliferation of MLL-r cell lines [161], [162]. Moreover, menin inhibitors are investigated in other cancers [163], [164].

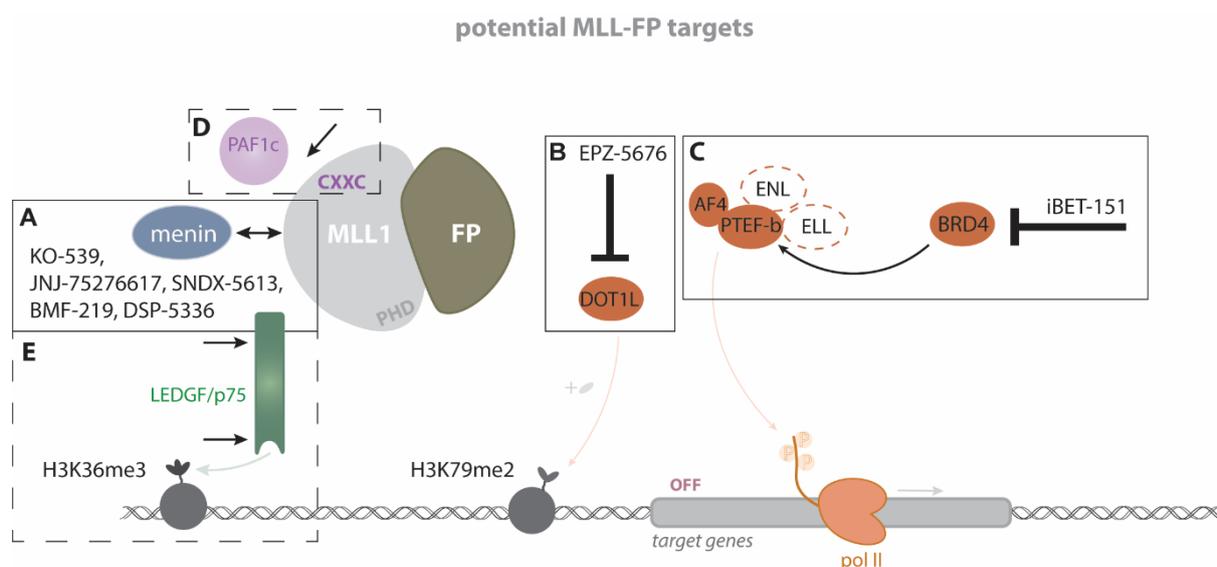


Figure 1.8: Potential targets to treat MLL-r leukemia

Schematic representation of potential strategies to hamper gene expression of the MLL-FP target genes to prevent leukemic transformation. Developed small molecules that inhibit the interaction between menin and MLL1 (box A) dissociate the MLL1-menin-LEDGF/p75 ternary complex and replace MLL-FP from the chromatin. DOT1L inhibitors (box B) aim to inhibit the DOT1L methylation activity and BRD4 inhibitors (box C) inhibit the PTEF-b phosphorylation of pol II, both to reduce target gene expression. Novel potential strategies are indicated by dotted squares. Inhibition of the CXXC-PAF1c interaction (box D) could help to replace the complex from the DNA. Box E represents the dissociation of LEDGF/p75 from either the ternary complex or the binding to histones to prevent the MLL-FP from association with actively transcribed chromatin.

As a result, five phase 1/2 clinical trials are currently listed at the ClinicalTrial.gov online database, investigating menin inhibitors in MLL-r leukemic patients [165]–[169], as well as patients with NPM1 mutations [165], [167], [169], diffuse large B-cell lymphoma or multiple myeloma [168].

Apart from the menin inhibitors directly targeting one of the MLL-FPs, promising results were obtained by targeting specific players of the transcription elongation stimulatory machinery, which is brought to the MLL1 target genes by various fusions. One such player is DOT1L (Figure 1.8B), a methyltransferase that methylates lysine 79 on histone 3 and is involved in several regulatory functions

(reviewed in [140]). Bernt K. *et al.* [170] reported that the H3K79 methylation pattern is altered in MLL-AF9 target loci and shortly after it was confirmed that the potent DOT1L inhibitor Pinometostat (EPZ-5676) selectively reduces H3K79 methylation and specifically inhibits MLL-r mediated leukemogenesis *in vitro* and *in vivo* [37], [171]. EPZ-5676 was also investigated in a phase 1 clinical trial including patients with *MLL1*-rearrangements [172]. Through complete remissions in two out of 51 patients, it was proven that targeting DOT1L has potential to treat MLL-r leukemia, however with low efficacy when given as single agent [172]. The effect of combination therapies are now explored preclinically, for instance the combination of DOT1L with menin inhibitors [163].

Similarly, the iBET-151 inhibitor against bromodomain containing proteins 2 to 4 (BRD2-4) of the bromo- and extra-terminal domain (BET)-protein family (Figure 1.8C) showed potent inhibition of MLL-r leukemic cells [39]. BRD4 is a positive regulator of P-TEFb, enhancing the Pol II elongation during transcription [173], [174]. Unfortunately, resistance against BET inhibitors has been reported by various mechanisms such as the suppression of BRD4 degradation pathways [175], [176].

1.2.3.2. Other therapeutic strategies to specifically target MLL-r leukemia

Although promising steps are taken towards a more specific treatment in MLL-r leukemia, novel strategies are being explored.

One theory suggested by Liang K. *et al.* is to tackle MLL-r leukemogenesis by stabilizing the wild type MLL1 protein such that the MLL-FP is replaced from the target sites [177]. This could be achieved by adding IRAK inhibitors to prevent wild type MLL1 degradation [177] or inhibiting casein kinase II to hamper the caspase1 cleavage, which would result in wild type MLL1 proteins with higher stability [178]. Next, the CXXC domain was labeled as a key player in MLL-r leukemogenesis. As the CXXC domain precedes the breakpoint cluster region, it remains present in all MLL-FPs. Through interaction with PAF1c, the CXXC domain is important for MLL1 association with non-methylated CpG DNA sites and necessary for MLL1 activity [102], serving as potential target for MLL-r leukemia [100] (Figure 1.8D). Another interesting target for MLL-r leukemogenesis is LEDGF/p75 (Figure 1.8E). As demonstrated by Yokoyama A. *et al.*, replacing the IBM on MLL-FP with the PWWP domain of LEDGF is sufficient to induce MLL-r transformation in cells, bypassing menin and the MLL1-menin-LEDGF/p75 triple complex formation [94]. Depletion of LEDGF/p75 significantly reduced the growth of MLL-r leukemic cells [94], [95]. Specific deletion of LEDGF in the hematopoietic system of mice prevented transformation of MLL-rearranged hematopoietic stem cells and subsequently any leukemic development in the animals, indicating that LEDGF is required for the development of MLL-ENL driven leukemia [179]. In none of these experimental settings, knocking down or knocking out of LEDGF/p75

seemed to affect viability of MLL1 wild type cells. Despite the promising target validation, in detail discussed in section 1.3.1.2. (page 22), development of small molecules targeting LEDGF are not reported in literature yet. A more detailed discussion on this strategy and considerations when targeting LEDGF/p75 are discussed in the next sections.

1.3. The epigenetic reader family of hepatoma-derived growth factor proteins

Together with the hepatoma-derived growth factor protein (HDGF) and HDGF-related proteins 1 to 4 (HRP1-4, or also named HDGF1-4), Lens Epithelium Derived Growth Factor (LEDGF) is part of the HDGF family [180]. Structurally, the protein family is characterized by the conserved sequence of the N-terminal Homologue to Amino Terminus of HDGF (HATH) region, which encompasses the proline-tryptophan-tryptophan-proline or PWWP domain. The cellular functions of the HDGF-family members have not been unraveled completely. All family members will be discussed in more detail in the sections below, with a focus on LEDGF (Figure 1.9).

1.3.1. Lens Epithelium Derived Growth Factor

1.3.1.1. Structure and function of LEDGF

LEDGF is a transcriptional co-activator involved in many different physiological processes which are not yet fully unraveled. The protein is encoded by the PC4- and SFRS1-interacting protein 1 (*PSIP1*) gene, located on chromosome 9 at the p22.2 locus. The gene contains 15 exons and expresses a long (p75) and shorter (p52) splice variant (Figure 1.9A). The LEDGF/p52 protein consists of 333 amino acids and harbors one structured domain at the amino-terminus, the PWWP-domain. For the splice variant LEDGF/p75, the first 325 amino acids are identical to p52, but p75 has a longer C-terminal domain which is largely folded as the integrase binding domain (IBD). Expression of LEDGF/p75 and p52 is regulated by different transcription start sites (TSS), of which the p75 TSS is 12-fold more active than the TSS of p52 [181].

The Proline-Tryptophan-Tryptophan-Proline domain

At the amino-terminus, the LEDGF protein contains a PWWP-domain (LEDGF_{PWWP}), named after the conserved amino acid sequence of consecutively a Proline, Tryptophan, Tryptophan and Proline residue. However, the complete PWWP domain spans a 93 amino acid region, which is also known as the homologous to amino-terminus of HDGF or HATH domain and is well conserved between all HDGF family members. Paradoxically, LEDGF/p75 expresses a PHWP sequence.

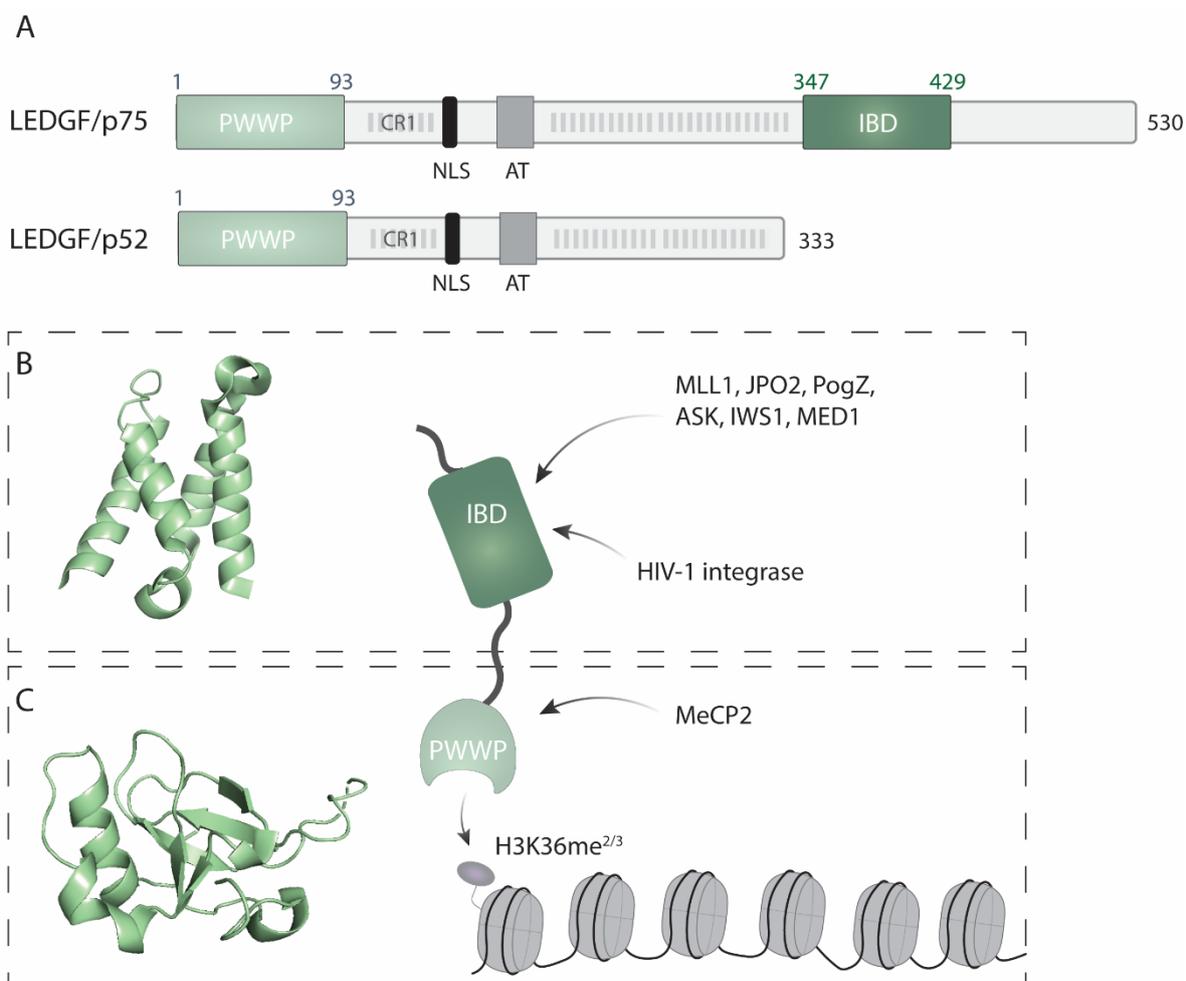


Figure 1.9: Schematic structure of LEDGF/p75

(A) Schematic domain structures of LEDGF/p75 and LEDGF/p52. Both proteins harbor a structured proline-tryptophan-tryptophan-proline (PWWP) domain at the N-terminus and present with a different C-terminus. For LEDGF/p75 the C-term comprises a second structured domain, the integrase binding domain (IBD). (B) Solution structure of the LEDGF_{IBD} domain (PDB 1Z9E) and indication of IBD interacting proteins. MLL1, mixed lineage leukemia 1; JPO2, cell division cycle-associated 7-like protein; PogZ, pogo transposable element-derived protein with zinc finger; ASK, activator of S-phase kinase; IWS1, interacts-with-Spt6; MED1, mediator subunit 1; HIV-1, human immunodeficiency virus 1. (C) Solution structure of LEDGF_{PWWP} domain (PDB 4FU6). A small pocket is present in the LEDGF_{PWWP} domain that specifically recognizes di- and trimethylated lysine 36 of histone three (H3K36me^{2/3}). Methyl CpG binding protein 2 (MeCP2) interacts with the N-terminal domain of LEDGF.

Both solution and crystal structures of the LEDGF_{PWWP} domain revealed an anti-parallel β -barrel of five strands with one α -helix between beta strands four and five and two α -helices at the c-terminal end, a structure conformation shared with other PWWP-domains [182] (Figure 1.9C). The domain forms a horseshoe-shaped pocket, with a hydrophobic surface and an opening that perfectly fits a post-translational modified lysine. Moreover, several studies demonstrated a preference for the H3K36me³ mark, giving LEDGF the characteristic of a chromatin reader [183]. More recently it was found that the PWWP domain of LEDGF not solely recognizes trimethylated H3K36 marks, but can additionally read di-methylated H3K36 [184]. The H3K36me^{2/3} mark is often associated with active

chromatin [185]. Knowing that the methyl-pocket interaction is a rather weak interaction with low affinity, it appears that a charged-driven interaction between the outside of the PWWP pocket and chromosomal DNA enhances the affinity of the LEDGF_{PWWP} domain to nucleosomes 10,000-fold [183]. This finding of a bivalent interaction was recently confirmed and visualized by a cryo-EM structure between a recombinant nucleosome and PWWP domain of LEDGF [186]. The cryo-EM structure confirmed that the PWWP domain binds both gyres of the DNA [183], [186], but does not interact with specific DNA bases. Instead, the interaction occurs with the phosphodiester backbone of the nucleotide sequence and lacks specificity [186]. Analysis of the structurally formed pocket revealed that the side chains of tryptophan at position 21 (W21) and the phenylalanine at position 44 (F44) are positioned at the base of the cavity [182] and an electron mobility shift assay with alanine-mutated pocket residues indicated that both W21 and F44 are crucial for the interaction of the LEDGF_{PWWP} domain with H3K36me3 of the nucleosome [187].

NLS and DNA-binding domains

Downstream of the PWWP domain, both LEDGF/p75 and p52 contain a nuclear localization signal (NLS) and additional DNA binding motifs. Nuclear import of LEDGF is mediated by the classical-type NLS¹⁴⁸GRKRKA EKQ¹⁵⁶ binding to the importin α /importin β import receptors [188]. Closely following the NLS, LEDGF harbors two AT-hook motifs. As the name suggests, this domain can interact with the minor DNA groove between adenine and thymine rich DNA and is composed of mainly positively charged amino acid residues. Mutagenesis studies revealed that these domains bind DNA non-specifically [189], [190]. Furthermore, LEDGF/p75 harbors four charged regions (CR1-4), of which the final two form a supercoiled-recognition domain (SRD) [191]. Of note, two publications indicate that interaction of the AT-hook motifs, the preceding CR1 and the SRD are necessary but insufficient for chromatin interaction of LEDGF, referring to a LEDGF-chromatin interaction domain beyond the PWWP fragment [189], [192]. Strikingly, two independent groups have shown that the LEDGF_{PWWP} and the CR1 domain engage in a protein-protein interaction with Methyl CpG binding protein 2 (MeCP2, Figure 1.9C) [193], [194], but whether this results in loss of H3K36me3 binding is not investigated.

The Integrase Binding Domain

For LEDGF/p75 but not p52, a second structured domain is present at the carboxyl-terminal, named the integrase binding domain (IBD, Figure 1.9B). The IBD is formed by a compact formation of two hairpins consisting of two alpha helices, connected by a smaller and differently oriented alpha helix

[195]. Two small pockets and a positively charged surface area were reported as binding site for protein-protein interactions [196]. The IBD was first reported to interact with integrase of the human immunodeficiency virus 1 (HIV-1), hence its name. Nonetheless, not only HIV-1 integrase, but multiple other cellular binding partners are described to interact with the IBD of LEDGF/p75 (Figure 1.9B). So far, the following cellular proteins have been reported to have an IBD-binding motif (IBM), necessary for interaction with LEDGF/p75: cell division cycle-associated 7-like protein (JPO2) [197], [198], pogo transposable element-derived protein with zinc finger (PogZ) [199], activator of S-phase kinase (ASK) [200], interacts-with-Spt6 (IWS1) [196], mediator subunit 1 (MED1) [201] and MLL1 [94], [202]. All IBMs are mainly unstructured, but share an acidic linker followed by a FxGF sequence [201] and display single linear motif (SLiM) characteristics. Upon IBD binding, all IBM motifs fold into a conserved and structured interaction interface, suggesting a conformation change of the protein necessary to interact with the LEDGF_{IBD} [201]. In addition, phosphorylation of the IBM increased affinity of the cellular interaction partners to the IBD revealing a control mechanism for context-regulated partner binding [201]. Recently, it was discovered that the same interaction mode between IBD and IBM was seen in other protein-protein interactions and more general names of respectively TFIIIS N-terminal domains (TNDs) and TFIIIS-interaction motifs (TIMs) were given [203]. The LEDGF/p75 interaction partners JPO2 and ASK harbor two IBMs. Two LEDGF_{IBD} domains were detected per JPO2_{IBM} by isothermal titration calorimetry [196]. Furthermore, it was published that LEDGF/p75 dimerization affects DNA binding [204] and later NMR-based evidence was given that dimerization occurs through two LEDGF_{IBD} domains [205]. The biological relevance of this dimerization is not yet fully understood.

LEDGF in health

Unlike the name suggests, LEDGF is not limited to the lens epithelium, but is ubiquitously expressed through the human body. A systemic mouse model where *Psip1* is knocked out presented with high perinatal lethality and multiple phenotypic abnormalities such as low fertility, motor defects and skeletal abnormalities for the surviving mice [206], denoting that LEDGF is essential for development.

As the structural organization of LEDGF/p75 enables binding to both the chromatin at actively transcribed regions marked with H3K36me3 on the one hand and cellular proteins on the other hand, LEDGF/p75 can be considered as an ideal transcriptional co-activator. However, the exact physiological mechanism by which both LEDGF/p75 and p52 regulate gene transcription is not completely understood [207] and further research revealed that LEDGF/p75 activity extends in multiple cellular processes such as stress response, apoptosis, growth and auto-immunity. Shortly after the discovery of the LEDGF protein and its transcriptional co-activator function in 1998 [207], it was reported that

LEDGF/p75 plays a role in stress response in chicken cells [208]. In the meantime, additional evidence was given that LEDGF/p75 is upregulated by induced environmental heat- or oxidative stress and occupies mainly promotor elements of stress-related genes [209], resulting in the upregulation of genes such as alphaB-crystallin in astrocytes and lens epithelial cells [210], [211] and protection against stress-induced cell death. Interestingly, alphaB-crystallin binds to an intermediate cleaved, non-active product of caspase 3, inhibiting caspase 3-induced apoptosis [212]. Not only has LEDGF/p75 a pro-survival or anti-apoptotic effect via alphaB-crystallin, but LEDGF/p75 promotes several other stress-related genes such as heat shock protein 27 [211], anti-oxidant protein 2A [213], involucrin [214] and alcohol dehydrogenase 1 and 4 [215]. Interestingly, upon apoptosis LEDGF/p75 is cleaved by caspase 3 and 7, resulting in smaller LEDGF fragments of 65 and 58 kDa, lacking the N-terminal PWWP domain and the anti-apoptotic effect [216]. In addition, LEDGF/p75 promotes cell survival through binding to the heterodimer of ASK and Cdc7, influencing the S-phase of the cell cycle [200]. Not much later, a general response for LEDGF/p75 in DNA damage and repair was suggested [217]. Daugaard M. *et al.* found that after DNA damage, LEDGF/p75 is important for the stabilization of single strand DNA (ssDNA), which is essential for the recruitment of the ssDNA-binding replication protein A complex to the restorative DNA of the homologous recombination-mediated repair mechanism. Without LEDGF/p75, C-terminal binding protein interacting protein (CtIP) is no longer recruited to the double strand DNA break and thus cannot be repaired. Defective repair of DNA by the lack of LEDGF/p75 resulted in decreased migration [218] and impaired colony formation [217]. More recently, Le Roy G. *et al.* suggested that LEDGF/p75 is also engaged in RNA transcription by lifting RNA polymerase II to overcome the nucleosome-induced barrier at the initiation of transcription [219].

1.3.1.2. LEDGF/p75 in MLL-r leukemia

As described earlier, LEDGF/p75 is involved in the normal and oncogenic functional activities of MLL1 and MLL-FP, respectively, by engaging in a ternary MLL1-menin-LEDGF/p75 complex. Inhibition of menin has proven to impair the leukemogenesis of MLL-r *in vivo* and menin inhibitors are tested in a clinical trial (summarized in section 1.2.3., page 15). Evidence is given for an alternative target in this ternary complex as a strategy to treat MLL-r leukemia, namely by targeting LEDGF/p75. Several publications indicate the mechanistic importance of the ternary complex formation of MLL1, menin and LEDGF/p75 in MLL-r leukemia (Figure 1.10A) [94], [95], [97], [179]. In all cases, abolishing the association of the oncogenic fusion protein with the active chromatin reduced or even eliminated the clonogenic growth and transformation potential of cells. The functional and structural involvement of LEDGF/p75 in MLL-r leukemia is bipartite and thus two strategies are possible to prevent tethering of the MLL-FP complex to the target genes.

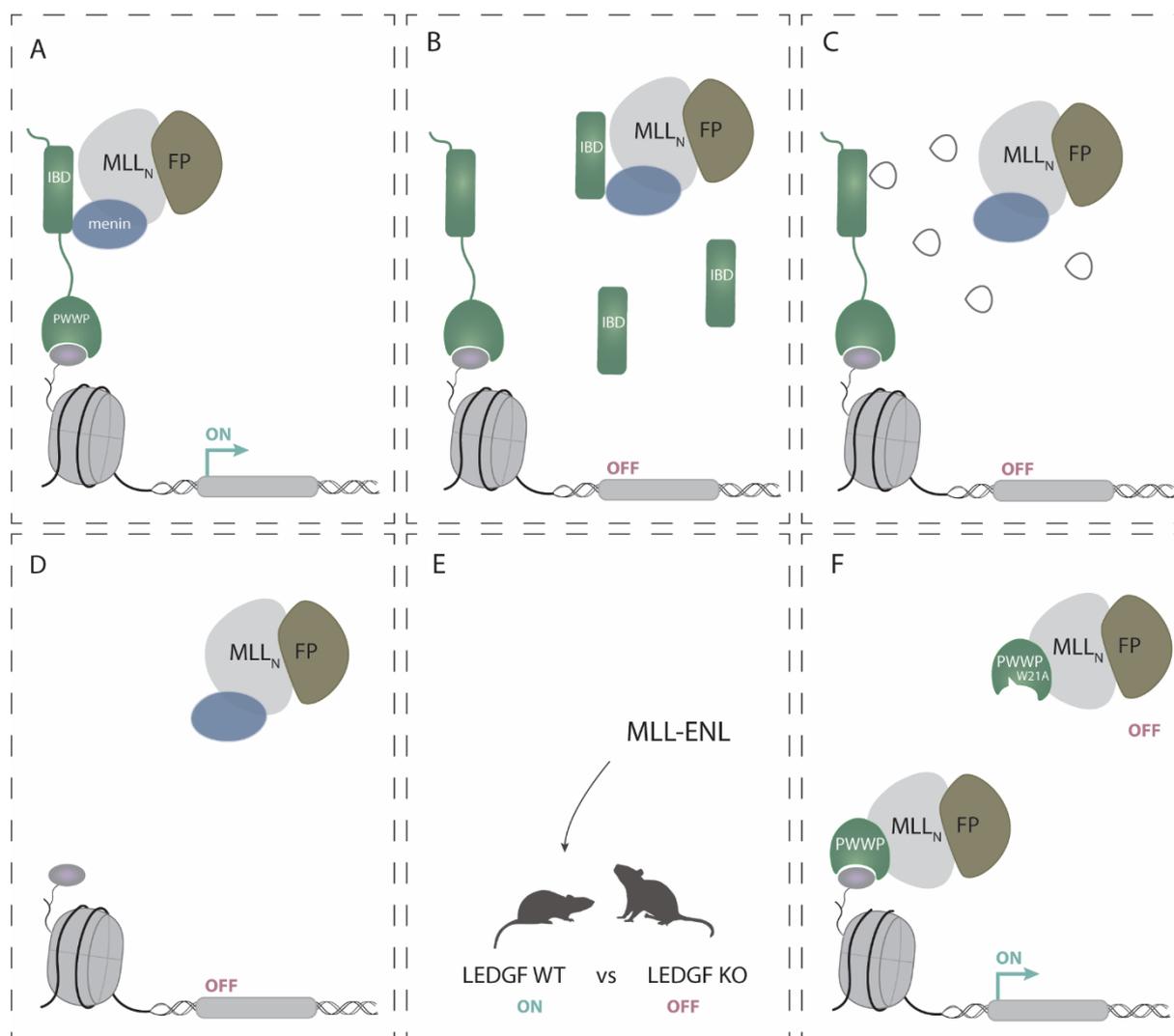


Figure 1.10: Validation of LEDGF/p75 as potential target for MLL-r leukemia

(A) the MLL-FP is tethered to the chromatin by formation of a ternary complex between MLL, LEDGF and menin. Association of the MLL-FP with active transcription results in leukemic transformation. (B) Overexpression of the IBD domain competes with endogenous LEDGF/p75 for binding to the MLL-FP. When MLL-FP is bound to the IBD alone, a decrease in leukemic transformation is detected. (C) Expression of cyclic peptides binding to the IBD domain of LEDGF/p75 replace the MLL-FP from the IBD interaction surface and reduce leukemic transformation. (D) depletion of LEDGF/p75 prevents the association of MLL-FP to the chromatin. (E) Mice that received a bone marrow transplantation with LEDGF knockout (KO) cells transduced with an MLL-ENL fusion did not develop leukemia. (F) Fusing the PWWP domain of LEDGF to MLL-FP resulted in MLL-r leukemia, whereas the same construct including a W21A point mutation in the PWWP domain did not result in cellular transformation.

On the one end, one would opt to interfere with the MLL1 interaction with the IBD domain of LEDGF/p75 (Figure 1.10B and C). This strategy is supported by the research of Méreau H. *et al.* demonstrating that the overexpression of the eGFP-fused IBD domain was sufficient to reduce colony formation *in vitro* and development of leukemia *in vivo* (Figure 1.10B) [95]. Interestingly, this observation was specific for MLL-r cell lines and did not affect the proliferation or colony formation of primary bone marrow cells (BMCs) and the HL-60 cell line [95]. In addition, Desimmie B. *et al.* identified

small cyclic peptides that specifically bind to the LEDGF IBD domain (Figure 1.10C) [220], which later resulted in a specific inhibition of the clonogenic growth of MLL-AF9 primary BMCs, but not wild type BMCs [97].

These two studies suggest that targeting the IBD domain of LEDGF/p75 is a potential strategy to treat MLL-r leukemia. Although perinatal depletion of LEDGF in mice is lethal [206], depletion of LEDGF/p75 significantly reduced the growth of MLL-r leukemic cells *in vitro* (Figure 1.10D) [21,22]. Although the deviating differential blood count in a conditional, hematopoietic LEDGF knockout mouse model suggests that LEDGF affects hematopoietic progenitor cells. No other phenotypic abnormalities were observed, suggesting that LEDGF is dispensable for normal hematopoiesis [179]. In addition, bone marrow transplantation experiments showed that LEDGF is necessary to initiate and maintain the MLL-r leukemia *in vivo* (Figure 1.10E) [179]. Secondly, supported by the same evidence that LEDGF/p75 depletion impairs MLL-r leukemia, we hypothesize that MLL-FPs can be replaced from the chromatin by disconnecting the protein complex from the chromatin. Although the LEDGF/p75-chromatin interaction domain spans a rather large amino acid sequence encompassing the PWWP domain, CR1 and AT-hook motifs, it was published that the affinity of the individual domains for chromatin or DNA is rather weak or insufficient [189], [192]. This suggests that targeting one domain might be sufficient to displace LEDGF/p75 from the chromatin. The most prominent druggable candidate of this interaction interface is the PWWP domain as this domain is the only structured region in the amino-terminal sequence of LEDGF/p75 which forms a small pocket with a specific interaction preference for the H3K36 methylation mark, associated with actively transcribed genes [183], [186]. Expression of LEDGF_{PWWP} directly fused to an MLL-ENL oncogenic protein resulted in tethering of the MLL-ENL fusion to the chromatin and leukemic transformation (Figure 1.10F) [94]. Moreover, mutating the two amino acids (W21 and F44) in the PWWP pocket that are important for chromatin binding [183] to alanine residues completely abolished this leukemic transformation and observed *Hoxa9* upregulation [94]. These findings support the feasibility to target the PWWP domain to treat MLL-r leukemia. However, no experiment directly targeting or competing with the PWWP domain of LEDGF has been reported as target validation of the PWWP pocket.

Many challenges arise during the process of drug discovery. Technical challenges regarding small molecule design, compound synthesis and solubility as well as challenges in light of toxicity and ADME pharmacokinetic tests can occur, however it is of interest to properly understand the cellular functions and relations of the target protein in health and disease, beyond the pathological mechanism of interest. In the context of this PhD manuscript, attention should be given to cellular proteins that harbor one or both similarly structured domains as of LEDGF/p75. Notwithstanding, it is also of interest to reflect on other diseases in which LEDGF/p75 is involved.

1.3.1.3. LEDGF/p75 in other diseases

Targeting the transcriptional co-activator LEDGF/p75, which has many cellular interaction partners, might come with a risk of toxicity. Nevertheless, published results indicate that replacing or depleting LEDGF/p75 in a cellular context or adult hematopoietic system does not affect overall viability of the cells or mice, respectively [95], [179]. Real toxicity of small molecules can only be assessed when available.

Meanwhile, it is of interest to take in consideration other diseases in which LEDGF/p75 contributes to the disease phenotype. Although the LEDGF/p75 effect on MLL-r cell lines seems to be specific for the MLL1 rearrangement, LEDGF/p75 has been identified as important player in several other cancers such as prostate cancer [221]–[223], breast cancer [224], cervical cancer [225], colon cancer and thyroid cancer [226]. Both LEDGF/p75 transcripts and protein levels were elevated in most tumor types as screened by Basu A. *et al.* [226], but no unifying mechanism has been described. In addition, LEDGF/p75 expression levels were reported to influence chemosensitivity of cancer cells to treatment [222], [227], [228]. Potentially, LEDGF/p75 is also involved in the diseases related to the cellular binding partners of LEDGF/p75, such as medulloblastoma (related to JPO2) [229], intellectual disability and autism spectrum disorder (related to PogZ) [230], but also MeCP2-related Rett Syndrome [231]. Higher LEDGF/p75 levels have also been linked to human papillomavirus E6 and E7-induced malignant cell transformation, where LEDGF/p75 supports resistance against stress [232].

Much less understood is the presence of anti-LEDGF/p75 antibodies, in literature referred to as auto-dense fine speckled 70 (DSF70) antibodies [233]. These antibodies were initially associated with atopic diseases and other inflammatory conditions, however more sensitive antibody detection revealed that these DSF antibodies are present in healthy individuals as well [234]. The biological function and clinical relevance of these autoantibodies is an unsolved matter. Twenty years of research on anti-LEDGF/p75 antibodies is reviewed in [234].

The disease in which LEDGF/p75 was investigated the most intensively is HIV-1. More specifically, the viral protein HIV-1 integrase was the first interaction partner described of the IBD domain (Figure 1.9B), hence its name [235]. After binding to LEDGF_{IBD}, HIV-1 integrase hijacks the tethering function of LEDGF/p75 towards actively transcribed genes at the chromatin, where the viral DNA integrates into the host genome [236], [237]. Both depletion of LEDGF/p75 and overexpression of the IBD domain resulted in a reduction of the HIV-1 replication [238], [239]. Additionally, LEDGF/p75 depletion negatively impacted the reactivation of the HIV-1 latent reservoir [240], which is the major barrier in finding a cure for HIV-1 patients. Drug development efforts have resulted in the discovery of LEDGINS, small molecules that specifically target HIV-1 integrase and hamper binding to the LEDGF_{IBD} domain

[240], [241]. Use of LEDGINs in cellular assays has shown effects similar as LEDGF/p75 knockdown, targeting HIV-1 integration away from active sites, albeit not in a completely random manner [240]. Furthermore, next to these so-called early effects, LEDGINs hamper reactivation of latently infected cells and produce defective viral particles by oligomerization of HIV integrase, affecting the HIV-1 replication efficiency [242].

1.3.2. HDGF-Related Protein 2

The most interesting HDGF-family member related to this manuscript is HDGF-related protein 2 (HRP-2), also known as HDGF2. HRP-2 is the only member for which the shared structural feature is not limited to the PWWP-domain, but the protein additionally harbors an IBD at the C-terminus (Figure 1.11A) [243]. Therefore, it is of interest to investigate the role of HRP-2 in normal hematopoiesis, as well as MLL-r leukemia to predict the potential effect generated by HRP-2 when targeting the highly similar LEDGF/p75 structured domains. Although HRP-2 was reported one year before LEDGF/p75 [244] and both sequences show high similarity, HRP-2 is less extensively investigated.

The HRP-2 protein is encoded by the *HDGFL2* gene at the 19p13.3 locus. With 671 amino acids, the HRP-2 protein is slightly longer than LEDGF/p75 and has a calculated weight of 74.3 kDa. Of note, HRP-2 appears as a larger protein of around 140 kDa by gel electrophoresis [243], [245]. Next to the PWWP and IBD domain, HRP-2 harbors an AT hook motif and homology region III (HR3) domain [243]. No NLS has been reported in the HRP-2 sequence. Two splice variants of the HRP-2 gene exist [246], one with nine additional amino acids (isoform b) and one with a 53 amino acid deletion at the N-terminus (isoform c), partly affecting the PWWP domain [246].

N-terminally, the HRP-2_{PWWP} domain sequence is highly similar to the LEDGF_{PWWP} domain and crystal structures indicate the typical PWWP-fold of this domain (PDB 3EAE, Figure 1.11B) [247]. In accordance, the HRP-2_{PWWP} domain interacts with the di- and trimethylated H3K36 modification mark [219]. Zhu X. *et al.* report that HRP-2 prefers binding to H3K36me2 [248]. However, the HRP-2_{PWWP} domain is described to interact with several other epigenetic marks. Wu H. *et al.* showed that HRP-2 weakly binds to H4K20me3 and H3K79me3 peptides [247]. The structure of the complex of HRP-2 with H3K79me3 was resolved by crystallization (PDB 3QJ6). Interestingly, H4K20me3 is associated with transcription repression and Baude A. *et al.* have shown that HRP-2 binds stronger to peptides carrying the heterochromatin marks H3K9me3 and H3K27me3, which are marks associated with no transcriptional activity [249], hinting toward a functional difference in chromatin binding between LEDGF and HRP-2.

Although no NLS region was reported in HRP-2, the protein is nuclear and plays a role in DNA double strand break repair in a similar way as LEDGF [249]. Since the effect on DNA repair after knocking down one or the other is larger than 50%, it was suggested that both proteins do not perform identical functions [249]. In relation to the heterochromatin binding, it was suggested that HRP-2 repairs DNA double strand breaks in heterochromatin, whereas LEDGF is active in euchromatin. Additionally, differential binding of LEDGF and HRP-2 to chromatin has been observed by Vanega M. *et al.* [245]. While LEDGF tightly binds condensed DNA, HRP-2 localizes in the nucleus independently from chromosomes [245].

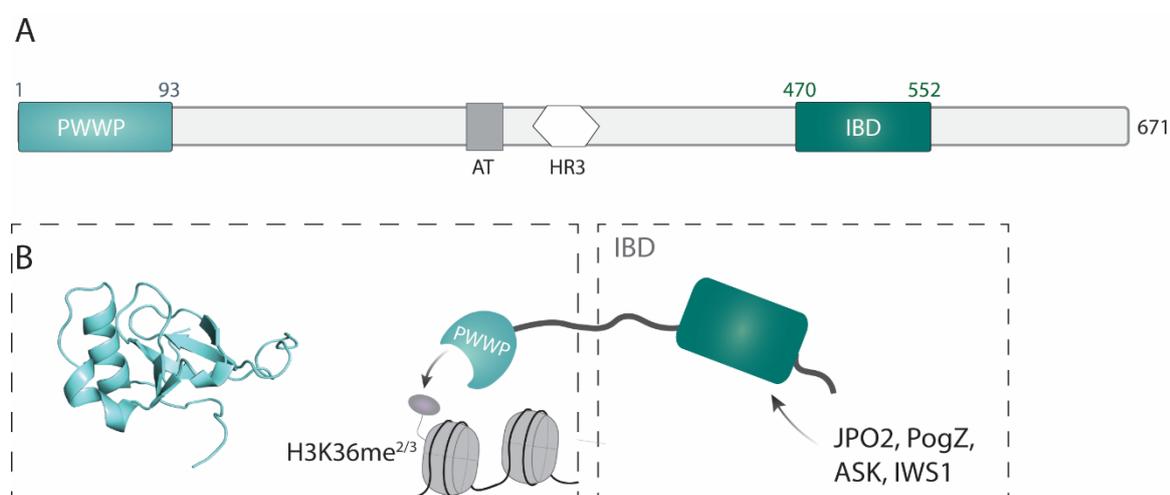


Figure 1.11: Schematic structure of HRP-2

Schematic domain structures HRP-2. Alike LEDGF, HRP-2 harbors a structured proline, tryptophan, tryptophan, proline (PWWP) domain at the N-terminus that specifically interacts with di- and trimethylated lysine 36 of histone three (H3K36me^{2/3}). The solution structure of the PWWP domain is depicted in light blue (PDB 3EAE). At the C-terminus, HRP-2 contains an integrase binding domain (IBD), known to interact with JPO2 (cell division cycle-associated 7-like protein), PogZ (pogo transposable element-derived protein with zinc finger) and IWS1 (interacts-with-Spt6).

No crystal structure of the IBD domain of HRP-2 is published, but several cellular proteins have been described to interact with this domain (Figure 1.11B). By co-immunoprecipitation, PogZ and JPO2 were identified as binders [199], whereas the HRP-2 IBD was described to also functionally activate the ASK in complex with Cdc7 and thereby stimulate DNA replication [200]. IWS1, an RNA processing regulator, co-precipitates and co-localizes with HRP-2 [250]. Alike LEDGF, HRP-2 was reported to relieve the nucleosome-induced barrier for transcription elongation [219] and here, Le Roy *et al.* suggest a disparity between HRP-2 and LEDGF activity in sequential cell types.

As several cellular interaction partners and their function are redundant with the LEDGF/p75 interactome, it is plausible that both proteins also show redundancy in cellular processes. Also, in relation to diseases, it is alluded that both LEDGF/p75 and HRP-2 have similar but not identical functions. Furthermore, the role of HRP-2 in diseases has not been studied intensively. Like for other HDGF-family members, one publication reported that both the mRNA and protein expression levels of HRP-2 are upregulated in 40% of hepatocellular carcinoma (HCC) patient tissues [250]. *In vitro* studies indicate that silencing of the HRP-2 expression reduces cell growth and *vice versa*, exogenous HRP-2 expression promotes HCC cell growth and these findings were corroborated *in vivo* [250]. HRP-2 is best studied in parallel to the role of LEDGF/p75 in HIV infection. LEDGF/p75 is responsible for the integration of the viral DNA into the host genome and LEDGF/p75 knockdown efficiently impairs integration and viral replication. Studies by Schrijvers R *et al.* have indicated that HRP-2 can rescue the impeded viral infection and integration in the absence of LEDGF/p75 when overexpressed [251], [252]. Concurrent results were obtained by Wang H. *et al.* [253].

Following this putative role in HIV-1 and cancer and the highly related sequence and activities of HRP-2 and LEDGF/p75, it is of interest to investigate the role of HRP-2 in MLL-r leukemia; more specifically to explore whether HRP-2 serves as potential target for the treatment of MLL-r leukemia.

1.3.3. The other HDGF-family proteins

The other members of the HDGF family: HDGF, HRP-1 and HRP-3 are less important in light of the MLL-r leukemia mechanism. As the name suggests, HDGF was described to stimulate growth of hepatoma cells [254]. In addition, a dual role in angiogenesis and both a pro- and anti-apoptotic function were reported for HDGF [255]. The HDGF engagement in these vital pathways could explain why the HDGF protein is reported to play a role in various cancers, where HDGF often appears upregulated and has been associated with metastasis and recurrence, radio sensitivity, chemoresistance, and/or microvascular density [256].

The HRP-1 protein is specifically expressed in the testis of mice [244], [257] and aside from three genome-wide screening studies [258]–[260], no other publications report about this protein, hence the function and importance of HRP-1 in humans remains unknown. Similarly, the last registered HDGF family member to date, which was discovered early 2002, was bovine HRP-4. Identical to HRP-1, HRP-4 was detected only in the testis. Since then, no additional functional reports have been provided and it is not known whether HRP-4 is expressed in humans.

The HRP-3 protein is expressed in, but not limited to the testis. In particular, the highest expression of HRP-3 was found in the brain [261], [262] where it was reported to have a neurotrophic [262] as well

as angiogenic effect in retinal endothelium [263] similar to HDGF. However, HRP-3 expression is highly upregulated in hepatocellular carcinoma (HCC) cells. Reducing this HRP-3 expression in malignant cells increased the level of apoptosis, sensitized cells to treatment and reduced tumor growth both *in vitro* and *in vivo* [264].

1.3.4. PWWP domain containing proteins

Besides the six HDGF-family members, 18 additional proteins are known to harbor a PWWP domain (Table 1.4). Similar to the PHWP sequence of the LEDGF_{PWWP} domain, several proteins of the PWWP-protein family have a slightly different PWWP sequence after which the domain was named (Table 1.4, PWWP variation). For many PWWP domains a 3D structure has already been acquired (PDB entries in Table 1.4) and comparison of these structures revealed a conserved beta-barrel consisting of 5 β -strands, oriented antiparallel as described for the PWWP domain of LEDGF/p75. A difference in the amino acid sequence and length can be observed between beta-strand two and three. In addition, the most striking variation is observed at the end of the PWWP-domain, where the amino acids are structured in two to six α -helices. However, all PWWP domains encompass a nicely shaped pocket for interaction with the methylated lysine. In addition to this common PWWP-fold, the positively charged surface at the outside of the pocket is comparable for all proteins and facilitates the binding to DNA. Overall, based on this sequence and structural homology, it is likely that all PWWP domain containing proteins interact with their PWWP domain to the nucleosome in a similar way as presented by the cryo-EM [186]. Having multiple PWWP domains circulating in the cell, obtaining specificity can be challenging. The PWWP-domain containing proteins can be subdivided into several families, although some proteins are reported in different families by different publications [186], [247], [265].

The HDGF family was already discussed earlier. The largest subfamily of proteins harboring a PWWP domain are the MutS homolog 6 (MSH6)-like proteins (Table 1.4). Interestingly, three proteins of this family harbor two PWWP domains. The first PWWP domain, more located towards the amino-terminal end (PWWP₁) belongs to the MSH6-like protein subgroup whereas the second PWWP domain (PWWP₂) of these proteins is classified in a separate Wolf-Hirschhorn syndrome candidate 1 (WHSC1) related subgroup. This latter subgroup consists of methyltransferases of lysine 36 at histone 3 of which in several cancers (reviewed in [266]) and efforts in finding specific inhibitors encountered difficulties in obtaining specificity and low nanomolar activity [267]. Interestingly, very recent, specific WHSC1 inhibitors of the N-terminal PWWP domain (WHSC1_{PWWP1}) were reported to disrupt the interaction with nucleosomes *in vitro*, with a dissociation constant of 3.4 μ M for the WHSC1_{PWWP1} domain [268], [269]. Specificity was shown for the WHSC1_{PWWP1} domain over the closest related protein WHSC1-like

protein 1 (WHSC1L1), or NSD3. *Vice versa*, inhibitors of the N-terminal PWWP domain of WHSC1L1 subgroup. This latter subgroup consists of methyltransferases of lysine 36 at histone 3 of which WHSC1, often referred to as the nuclear receptor binding SET domain protein 2 (NSD2), was the first protein for which a PWWP domain was described [270]. WHSC1 is often mutated [149] or upregulated (WHSC1L1_{PWWP1}) are reported to specifically bind the WHSC1L1_{PWWP1} given the specific temperature shift in a denaturation assay [271]. Furthermore, these inhibitors show cellular activity in a selection of cell lines by downregulating the expression of myc [271]. These findings illustrate the feasibility to target PWWP domains in a specific manner.

Table 1.4. Overview of PWWP-domain containing proteins.

Related family	Protein name	PWWP variation	PDB entry
HDGF	HDGF	PHWP	1RI0
	HRP-1	AHWP	-
	HRP-2	PHWP	3EAE
	HRP-3	PHWP	6IIP
	LEDGF	PHWP	2M16
BRPF	BRPF1	PSYP	2X35
	BRPF2	PSYP	3LYI
	BRPF3	PSYP	3PFS
DNMT3	DNMT3A	SWWP	3LLR
	DNMT3B	SWWP	3FLG
	MBD5	TSWP	-
	PWWP2A	PWWP	-
	PWWP2B	PWWP	4LD6
MSH6	NSD1 _{PWWP1}	PWWP	-
	WHSC1 _{PWWP1}	PWWP	5VC8
	WHSC1L1 _{PWWP1}	PWWP	6G3T
	ZCWPW1	PWWP	-
	ZCWPW2	PSWP	-
	MSH6	PWWP	6OQM
WHSC1	NSD1 _{PWWP2}	RWWP	-
	WHSC1 _{PWWP2}	RWWP	7CRO
	WHSC1L1 _{PWWP2}	RWWP	4RXJ
BS69	ZMYND8	PFWP	4COS
	ZMYND11	GFWP	4NS5
	MUM1	PFWP	3PMI
	NP60	PPWP	-

2

Objectives

Mixed lineage leukemia-rearranged (MLL-r) leukemia is a genetically distinct subset of leukemia with poor prognostic outcome. Although the mechanism by which wild type MLL1 and most of the formed oncogenic MLL-fusion proteins function is extensively investigated, no specific treatment option is available for MLL-r leukemia patients that efficiently eradicates the leukemic cell population, demonstrating that this disease faces an unmet medical need. The MLL1-menin-LEDGF/p75 ternary complex plays an important role in the mechanism driving MLL-r leukemia and previous findings have validated LEDGF/p75 as a potential target to treat MLL-r leukemia. During my PhD, the overall aim is to understand whether the LEDGF/p75 paralog HRP-2 was involved in this mechanism and to develop first-in-class small molecules that specifically target LEDGF/p75 to treat MLL-r leukemia. The specific objectives in this manuscript are:

2.1. Investigation of the role of HRP-2 in normal hematopoiesis and MLL-r leukemia

Due to high sequence homology and similar functional roles in health and disease, we want to investigate whether the LEDGF/p75 paralog, HRP-2, plays a role in normal hematopoiesis and in MLL-r leukemia. First, I will examine whether HRP-2 interacts with MLL1 or the MLL1-menin complex through immunoprecipitation experiments. In addition, we will aim to obtain the NMR solution structures of the HRP-2_{IBD} to compare with the previously published LEDGF_{IBD} domain. To study the role of HRP-2 in normal hematopoiesis a systemic HRP-2 knockout mouse model will be used. We will analyze peripheral blood counts and stem cell proliferation in a colony formation assay. Next the impact of HRP-2 depletion on leukemic cell growth *in vitro* will be examined by following the growth of the cells in liquid culture or by performing a colony formation assay. To examine the impact of HRP-2 depletion on the survival *in vivo*, a bone marrow transplantation to induce leukemia in the presence or absence of HRP-2 will be performed. These findings will help to determine whether HRP-2 is a target for MLL-r leukemia and will potentially clarify the associated risks of small molecules with activity towards HRP-2.

2.2. Development of small molecules that target the PWWP domain of LEDGF

The role of LEDGF/p75 in MLL-r leukemia is well described. Inhibiting either the LEDGF_{PWWP} binding to nucleosomes or the LEDGF_{IBD} binding to MLL-FPs could prevent tethering of the oncogenic fusion protein to its target genes. In this project, we aim to develop small molecules that specifically target the LEDGF_{PWWP} domain. Although LEDGF/p75 knockdown or competition experiments in a cellular context show a reversal of the leukemic phenotype and LEDGF is essential for leukemogenesis *in vivo*, additional validation of the LEDGF_{PWWP} domain by itself as a potential target for MLL-r leukemia is required. As proof of concept, the LEDGF_{PWWP} domain will be overexpressed in an MLL-r leukemic cell

line to induce competition with endogenous LEDGF for H3K36me3 binding. The effect of this transduction on growth will be validated by a colony formation assay. As control, a similar experimental setup will be performed in an MLL1 wild type cell line.

We aim to identify small molecules that specifically interact with LEDGF_{PWWP} by a structure-based drug discovery strategy using the crystal structure of HRP-2_{PWWP}. Subsequently, hits will be tested for binding to both the HRP-2_{PWWP} and LEDGF_{PWWP} domain in DSF. During my PhD, I will optimize the AlphaScreen and TR-FRET assays to detect the interaction between recombinant nucleosomes and LEDGF protein. These assays can be used to study the interaction inhibition potential of fragments during hit to lead optimization. To study the interaction in a cellular context, I will optimize the LEDGF-nucleosome interaction in a nanoBRET assay. When targeting a methyl-lysine binding pocket, it can be challenging to obtain specificity and therefore several proteins of the PWWP-protein family will be assayed in parallel in the early stages of the screening.

3

Materials and methods

Part of the materials and methods has previously been published in S. Van Belle et al., "Unlike Its Paralog LEDGF/p75, HRP-2 Is Dispensable for MLL-r Leukemogenesis but Important for Leukemic Cell Survival," Cells, 2021 [272].

3.1. Recombinant protein purification

Flag-LEDGF/p75 WT and the W21A, F44A double mutant were expressed from pCPnat-3xFlag-LEDGF/p75 after transformation in BL21 Star (DE3) competent bacteria. The bacteria were grown in lysogeny broth (LB, Sigma) medium supplemented with 100 µg/mL ampicillin (Sigma-Aldrich). Before induction, the bacteria were grown at 37°C until an OD₆₀₀ of 0.6. Protein expression was induced by adding 0.5 mM isopropyl-β-D-thiogalactopyranoside (IPTG, Sigma-Aldrich). After the IPTG was added, the cultures were kept at 29°C for three to four hours before harvesting. The bacteria were spun down at 5,000 rpm for 15 minutes and resuspended in STE buffer (100 mM NaCl (Sigma-Aldrich), 10 mM Tris-HCl (Sigma-Aldrich) pH 7.4 and 0.1 mM ethylenediaminetetraacetic acid (EDTA, Sigma-Aldrich)) and stored as pellet at -20°C. To purify the Flag-LEDGF/p75 protein from the bacterial pellet, the pellet was resuspended in 500 mM NaCl, 30 mM Tris-HCl pH 7.4, 1 mM dithiotreitol (DTT, Merck) and cOmplete protease inhibitor cocktail, EDTA free (Merck) and followed by 1 minute of sonication with 2s on/4s off pulses at 50% amplitude using the SFX250 sonifier (Branson). After sonication, 0.1 µg/mL DNase (Roche) was added and incubated on ice for 15 to 20 minutes. The lysate was cleared by a 30-minute centrifugation at 15,000 rpm. The supernatant was filtered through a 0.22 µm Millex-GS Syringe Filter Unit (Merck) before purification over a 5 mL HiTrap Heparin HP Column (Cytiva), equilibrated in 150 mM NaCl, 30 mM Tris-HCl pH 7.0 and 1 mM DTT. The protein was eluted by increasing the salt concentration from 150 mM up to 2 M using the AKTA purifier (GE Healthcare) and Unicorn v5 software. Peak fractions were analyzed on an in house SDS-PAGE. Fractions containing flag-LEDGF/p75 were pooled and loaded on a superpose™ 6 10/300 GL size exclusion column (GE Healthcare) to further purify. The size exclusion column was equilibrated in 150 mM NaCl, 30 mM Tris-HCl, pH 7.4 and 1 mM DTT. Peak fractions were analyzed on an in house SDS-PAGE followed by a Coomassie stain (Coomassie Brilliant Blue G250, Merck). The fractions containing flag-LEDGF/p75 were supplemented with 10% (v/v) glycerol (VWR chemicals) and stored at -80°C.

The His₆-LEDGF_{PWWP} wild type, double mutant (W21A, F44A) and His-BRPF2_{PWWP} proteins were expressed from the pET16b plasmid using BL21(DE3) cells. The LEDGF_{PWWP} protein is encoded by amino acids (aa) 1 to 110 of full length LEDGF/p75. For the BRPF2_{PWWP} protein amino acid 925 to 1057 were expressed. The cultures were grown in LB medium supplemented with 100 µg/mL ampicillin. For each, the culture was grown to an OD₆₀₀ of 0.8 at 30°C. The culture was supplemented with 1 mM IPTG to induce protein expression and incubated overnight at 15°C. The cells were washed in STE buffer and stored as pellet at -20°C. For purification, the bacterial pellet was resuspended in 50 mM HEPES (Merck), pH 7.0, 500 mM NaCl, 1 mM DTT, 0.1% CHAPS (Thermo Scientific) and cOmplete protease inhibitor cocktail. The bacterial lysate was sonicated at identical conditions as flag-LEDGF/p75 described before. After sonication, 0.1 µg/mL DNase was added and incubated for 15 minutes on ice.

The cell debris was pelleted by centrifugation for 30 minutes at 15,000 rpm and the supernatant was loaded on a disposable column provided with a resin of His-Select HC Nickel Affinity Gel (Sigma-Aldrich) equilibrated with wash buffer (50 mM HEPES, pH 7.0, 500 mM NaCl, 1 mM DTT and 12.5 mM imidazole (Acros Organics, Thermo Fisher Scientific)). After loading the supernatant, the column was washed and the protein was eluted using wash buffer with an increased imidazole concentration of 250 mM. During elution, the eluate was fractionated and spotted on a Whatman paper, stained with Coomassie Brilliant Blue G 250 (Merck) dissolved in 50% (v/v) methanol (VWR chemicals), 10% (v/v) acetic acid (Sigma-Aldrich) and distilled water to analyze the protein content. The fractions with highest protein expression were pooled and dialyzed overnight in 50 mM HEPES, 150 mM NaCl and 1 mM DTT. The next day, the protein solution was spun down at 13,000 rpm for 10 minutes at 4°C and the supernatant was loaded onto the AKTA pure (Cytiva) connected to an HiTrap SP HP cation exchange column (GE Healthcare) equilibrated in 50 mM HEPES, 150 mM NaCl and 1 mM DTT. The protein was eluted by increasing salt concentrations and peak fractions were analyzed by an in-house SDS-PAGE. Fractions containing His-BRPF2_{PWWP} were pooled and supplemented with 10% glycerol and stored at -80°C. For His-LEDGF_{PWWP}, the protein containing fractions were further purified on a size exclusion column equilibrated in 50 mM HEPES, 150 mM NaCl and 1 mM DTT. Finally, the peak fractions were analyzed by an in-house SDS-PAGE and Coomassie stain and supplemented with 10% glycerol before storage at -80°C.

The His-WHSC1_{PWWP1} domain (aa 211 to 362) was expressed from pET28 plasmid and cultured in 50 µg/mL kanamycin (Thermo Fisher Scientific). The same purification protocol was followed as described for His-BRPF2_{PWWP}.

The recombinant, purified his-SUMO-HRP-2_{PWWP} (used for interaction experiments) and untagged HRP2_{PWWP} (used for Differential Scanning Fluorimetry or DSF) were a gift from the Laboratory of Biocrystallography, KU Leuven, Belgium.

The expression and purification of LEDGF/p75₃₄₅₋₄₂₆ was described earlier [202]. Identical conditions were used for expression and purification of HRP-2₄₆₉₋₅₄₉. Maltose binding protein (MBP) alone or MBP-tagged constructs MBP-LEDGF₃₂₅₋₅₃₀ and MBP-HRP-2₄₄₇₋₅₅₂ were purified similarly as described for the latter in [239]. The MLL₁₋₁₆₀-GST construct was purified as described in [95]. These proteins were expressed in *E.coli* Rosetta2 (DE3), grown in Lysogeny Broth (LB) medium (Sigma-Aldrich) supplemented with 10 µg/mL ampicillin (Sigma-Aldrich).

3.2. Differential Scanning Fluorimetry (DSF)

The melting temperature (T_m) of the various PWWP protein domains was assessed in the CFX Opus 96 Real-Time PCR Instrument (Bio-Rad, Belgium) using pre-treated 8-tube polymerase chain reaction (PCR) strips and optical flat caps (Bio-Rad, Belgium). These PCR tubes were incubated at 95°C for 10 minutes to reduce stickiness. In a reaction volume of 20 μ L, 10 to 30 μ M of protein was incubated with 10X to 20X of SYPRO™ Orange Protein Gel Stain (ThermoFisher Scientific) for 20 minutes at room temperature in DSF buffer (50 mM HEPES and 150 or 600 mM NaCl). Optimized running conditions for each protein domain were determined and reported in Figure 4.13. After incubation, the sample was heated gradually (0.2°C steps) from 20°C to 95°C. A melting curve was generated by measuring fluorescence from the SYPRO™ Orange (SO) dye in each step. The T_m of each protein was analyzed using the CFX Maestro software of Bio-Rad.

3.3. AlphaScreen and TR-FRET assay

The Amplified Luminescent Proximity Homogenous (AlphaScreen®) assay was performed in a 384-well OptiPlate (PerkinElmer®, Massachusetts, U.S.) and final volume of 25 μ L. Proteins and reagents were diluted in AlphaScreen buffer containing 150 mM NaCl, 25 mM Tris-HCl pH 7.4, 1 mM, 0.1% Tween®-20 (Thermo Fisher Scientific, Massachusetts, U.S.) and 0.1% Bovine Serum Albumin (BSA, Sigma-Aldrich, Missouri, U.S.). First, both partners of the protein-protein or protein-nucleosome interaction were incubated for one hour at 4°C. Optimal protein concentrations were determined by cross-titration and final protein concentrations are indicated per experiment in the results. Depending on the protein tags, corresponding 10 μ g/mL of donor and 10 μ g/mL of acceptor beads (PerkinElmer, Massachusetts, U.S.) were mixed together and added to each well. In addition of the beads, the plate was incubated another hour at room temperature and read using the Envision Xcite Multilabel Reader (PerkinElmer®, Belgium). Counts were plotted and graphs were generated using GraphPad Prism version 9.2.0.

In out-competition assays, we study the possibility of a related protein or fragment (= the out-competitor) to interfere with a protein-protein or protein-nucleosome interaction. The out-competitor is expected to bind one of the interaction partners and is either untagged or not recognized by donor nor acceptor bead. The out-competitor is titrated to a protein-protein or protein-nucleosome interaction at fixed concentrations. For this assay, a pre-incubation step of 30 minutes was performed with out-competitor and its interaction partner. After the second interaction partner was added, the protocol continued as described above.

For assays where related proteins were compared, the following requirements were met: proteins were expressed from identical plasmid backbones, parallel performance of purification and

concentration determination. In addition, the experiment was performed using identical buffer compositions, experimental set-up and interaction partner dilutions. For that reason, proteins with different tags cannot be compared.

For the time-resolved fluorescence energy transfer (TR-FRET[®]) assay, identical buffer composition and protocol were followed. The TR-FRET assay differs from AlphaScreen through the addition of 5 nM LANCE Eu-W8044 labelled streptavidin donor (PerkinElmer[®], Zaventem, Belgium) and 25 nM anti-tag specific ULIGHT[™] acceptor fluorophores (PerkinElmer[®], Zaventem, Belgium) instead of AlphaScreen beads. In our assays, the ULIGHT acceptor was attached to anti-6xhis or anti-flag antibodies (both Perkin Elmer[®]), dependent on the protein.

3.4. Fragment design and screening

Fragments were modelled and generated in the Laboratory of Biocrystallography (KU Leuven, Belgium) and the Laboratory of Medicinal Chemistry (KU Leuven, Belgium) and labelled by the Greek alphabet. For confidentiality reasons, structures of the fragments will not be made available. Fragments were dissolved in ethylene glycol (EG, ThermoFisher Scientific) at a stock concentration of 50 mM and stored at -20°C. The fragments were added during the incubation step in the DSF or added as described for the 'out-competitor' in AlphaScreen and TR-FRET protocol. In these experiments, the percentage of EG for all fragment concentrations and controls was kept at 1%.

3.5. NMR spectroscopy

NMR data were generated at the Laboratory of structural biology at the Institute of Organic Chemistry and Biochemistry (IOCB, Prague).

Nuclear Magnetic Resonance (NMR) spectra were acquired at 25°C on the 850 MHz Bruker Avance spectrometer equipped with a triple-resonance (¹⁵N/¹³C/¹H) cryoprobe. For structure determination, the sample volume was 0.35 mL, with a concentration of 500 μM HRP-2₄₆₉₋₅₄₉ in the NMR buffer (25 mM Tris pH 7.0, 150 mM NaCl, 1 mM TCEP), 5% D₂O/95% H₂O. The sequence-specific backbone and side-chain resonance assignment were obtained using a series of standard triple-resonance spectra (HNCO, HN(CA)CO, HNCACB, CBCA(CO)NH, HBHA(CO)NH, CCC(CO)HN and HCCH-TOCSY [273], [274]). ¹H-¹H distance constraints for structural determination were obtained from intensities of NOE cross peaks in the 3D ¹⁵N/¹H NOESY-HSQC and ¹³C/¹H NOESY-HMQC spectra that were acquired using a NOE mixing time of 100ms.

The families of converged structures were initially calculated in Cyana 3.98 using the combined automated NOE assignment and structure determination protocol [275]. In addition, backbone torsion angle constraints, generated from assigned chemical shifts using the program TALOS+ [276] were included in the calculations. Subsequently, five cycles of simulated annealing combined with redundant dihedral angle constraints were used to produce sets of converged structures with no significant restraint violations (distance and van der Waals violations $<0.2 \text{ \AA}$ and dihedral angle constraint violation $< 5^\circ$), which were further refined in explicit solvent using the YASARA software with the YASARA forcefield [277]. The 30 HRP-2 IBD structures with the lowest total energy were selected, analyzed and validated using the Protein Structure Validation Software suite (http://psvs-1_5-dev.nesg.org). The constraints and structural quality statistics for the final set of water-refined HRP-2 IBD structures is summarized in Supplementary Table S1. The structure, NMR constraints, and resonance assignments were deposited in the Protein Data Bank (PDB, accession number 6T3I) and Biological Magnetic Resonance Bank (BMRB, accession number 34442).

In titration experiments $20 \mu\text{M}$ of ^{15}N -labeled HRP-2_{469–549} or LEDGF/p75_{345–426} were mixed with various concentrations of unlabeled synthetic MLL1 peptide (123–160) or DMSO as a control. For each titration point, the chemical shift perturbations (CSP) in $^{15}\text{N}/^1\text{H}$ HSQC spectra measured in the SOFAST fashion were calculated and the dissociation constant (K_D) was determined by a non-linear least squares analysis using GraphPad Prism and the equation

$$\text{CSP}_{\text{obs}} = \text{CSP}_{\text{max}} \times \frac{[\text{L}] + [\text{P}] + K_D - \sqrt{([\text{L}] + [\text{P}] + K_D)^2 - 4 \times [\text{L}] \times [\text{P}]}}{2 \times [\text{P}]}$$

where CSP_{obs} is the observed CSP at the given total ligand concentration $[\text{L}]$, CSP_{max} is the CSP at saturation, and $[\text{P}]$ is the total concentration of protein [278].

The peptides used in section 4.1.2. were synthesized by solid-phase synthesis in the Laboratory of Medicinal chemistry, IOCB, ASCR (Prague, Czech Republic).

3.6. Cell culture

The human cell lines THP1, SEM, K562, Kasumi-1 and the murine cell line MLL-AF9 were a kind gift from Dr. Jürg Schwaller (Laboratory of childhood leukemia, Basel University, Switzerland). The human Nalm6 cell line was a kind gift from Prof. Noritaka Adachi (Yokohama City University, Yokohama, Japan). The HEK293T cells were bought from LGC Standards LTd (code CRL-11268 293T/17). The E2A-HLF cells were generated by harvesting lineage depleted cells from 8-weeks-old C57Black6 mice (animalium KU Leuven) and transduced with pMSCVneo-E2A-HLF vector. All generated cell lines tested negative for mycoplasma contamination.

The human suspension cells (THP1, SEM, K562, Kasumi-1 and Nalm6) were cultured in Roswell Park Memorial Institute (RPMI) medium (GIBCO, Thermo Fisher Scientific, Massachusetts, U.S.) supplemented with 0.01% (v/v) gentamicin (GIBCO, Thermo Fisher Scientific, Massachusetts, U.S.) and 10% (v/v) fetal bovine serum (FBS, GIBCO, Thermo Fisher Scientific, Massachusetts, U.S.), except for the Kasumi-1 cell line that was cultured in 20% FBS. The adherent cell line, HEK293T, was cultured in Dulbecco Modified Eagle Medium (DMEM, GIBCO, Massachusetts, U.S.) supplemented with 0.01% (v/v) gentamicin and 5% (v/v) FBS. The cultured murine cells (MLL-AF9 and E2A-HLF) were held in RPMI medium supplemented with 0.01% (v/v) gentamicin and 20% FBS, with the addition of the murine cytokine interleukin 3 (6 ng/mL), interleukin 6 (10 ng/mL) and murine stem cell factor (100 ng/mL, all from Peprotech, London, UK) to maintain stemness of the cells.

3.7. Co-immunoprecipitation

The MLL-ELL fusion was expressed from pCL-neo-Flag-MLL-ELL, a kind gift from Dr. Jürg Schwaller (Laboratory of childhood leukemia, Basel University, Switzerland). The MLL1 fragment 1-330 was expressed from p3xFlag-MLL₁₋₃₃₀ and point mutations blocking the MLL1-menin interaction (F9A, P10A and P13A) were introduced in this construct by oligonucleotide annealing (IDT). The sense strand (5'-TCGAGGAGTTTATGGCACATAGCTGTCGTTGGCGTGCCGCCGCACGTGCGGGTACAA-3') annealed to the antisense strand 5'-CCGTTGTACCCGCACGTGCGGCGGCACGCCAACGACAGCTATGTGCCATAAACTCC-3' when the temperature decreased after incubation of both single strand oligonucleotides at 95°C, creating compatible overhangs to the backbone when restricted with XhoI and AgeI (Thermo Fisher Scientific). To perform the co-IP, 6 million HEK293T cells were plated in three 8.5 cm petri dishes/condition and transfected with 20 µg of each indicated plasmid. The cells were lysed one day later as described in [201].

To study the effect of menin inhibitors (MI) in co-immunoprecipitation (co-IP) experiments, an excess of menin inhibitor (100 µM MI-2 and 50 µM MI-538; both MedChem Express) or dimethyl sulfoxide (DMSO, Sigma-Aldrich) as control, were added during both lysis and overnight incubation. Immunoprecipitated protein was eluted with SDS-PAGE loading buffer and visualized by western blotting.

3.8. nanoBRET

The nano-bioluminescence resonance energy transfer (nanoBRET) was set up using the NanoBRET™ Flexi® starter vectors kit of Promega (Wisconsin, U.S.). The coding sequence of full length LEDGF/p75 and LEDGF_{PWWP} (amino acid 1 to 110) were amplified from their protein expression constructs and

ligated in the NanoBRET Flexi constructs after enzyme digestion. The Histone H3.3-HaloTag® Fusion Vector-plasmid for histone H3 expression was commercially available (Promega). To start, HEK293T cells were seeded in a 24-well plate at a concentration of 400 000 cells/mL. After 6h of incubation, the cells were co-transfected with plasmids encoding a nanoLuc-tagged and HaloTag-tagged protein using branched polyethylenimine (PEI) (10 mg/mL, polyscience) at a ratio of 5 µL PEI per µg DNA. After 20 hours, the cells were trypsinized (Gibco™ Trypsin-EDTA (0.25%), Thermo Fisher Scientific) and plated in a 96-well plate at a concentration of 200 000 cells/mL in Opti-MEM™ I Reduced Serum Medium without phenol red (Fisher Scientific, cat n 11058021) supplemented with 4% FBS, in addition of either the HaloTag® NanoBRET™ 618 ligand (Promega) or DMSO (Sigma-Aldrich) control. The next morning, the luciferase NanoBRET™ Nano-Glo® Substrate (Promega, Wisconsin, U.S.) was added to the wells and fluorescent signal was detected after two minutes using the Envision Xcite Multilabel Reader (PerkinElmer®, Zaventem, Belgium). The fluorescent signal was converted to milliBRET Units (mBU) using the following formula:

$$\left(\text{Ligand} \left(\frac{\text{abs 618 nm}}{\text{abs 460 nm}} \right) - \text{DMSO control} \left(\frac{\text{abs 618 nm}}{\text{abs 460 nm}} \right) \right) * 1000 = \text{mBU}$$

where abs stand for the measured absorbance at the corresponding wavelength

3.9. Viral vector production and generation of stable cell lines

The Murine Stem Cell Virus (MSCV) retroviral expression vectors encoding MLL-AF9 fusion or eGFP (pMSCV IRES-eGFP-PGK-Puro) and the pMSCV-E2A-HLF-Neo were provided by Prof. Jürg Schwaller (Laboratory of childhood leukemia, Basel University, Switzerland). The pMSCV retroviral expression vector pMSCV-MLL-ENL-Neo and pMSCV-Neo were provided by Prof. Akihiko Yokoyama (Laboratory for Malignancy Control Research, Kyoto University, Japan). The lentiviral vector expressing mi30-based RNA targeting HRP-2 or eGFP control and retroviral vectors expressing Lcdgf/p75 miRNA were described earlier [95], [251]. For experiments where miRNA was used and no back complementation was performed, stable cell lines were generated at least twice.

For human cell lines, viral vector productions were performed as previously described [179], [279]. Titer units (TU) were determined by a p24 ELISA test (Fujirebio, Belgium). Human cell lines were transduced in a 1:1 volume ratio with concentrated lentiviral supernatants. Murine cells were transduced with pMSCV-based vectors as previously described [179]. Forty-eight hours post-transduction, cells were selected with the corresponding selection antibiotic at the following concentration: 1 µg/mL puromycin, 10 µg/mL blasticidin or 320 µg/mL hygromycin (all purchased from Invivogen).

3.10. Western Blot

To visualize proteins from whole cell lysates, cell pellets (10^6 to 2.10^6 cells) were washed twice in phosphate-buffered saline (PBS, Thermo Fisher Scientific, Massachusetts, U.S.) and lysed using RIPA Lysis and Extraction buffer (Thermo Fisher Scientific) during a 30 minutes incubation on ice. Cell debris was removed from the solution after centrifugation at 13 000 rpm, 4°C. The total protein content was assessed using the bicinchoninic acid assay (BCA) protein assay kit (Thermo Fisher Scientific) to equally load the samples on a 4–15% Criterion™ TGX™ Precast Midi Protein Gel (Bio-Rad, California, U.S.). The gel was run in Tris-Glycine SDS running buffer (Novex, Life technologies) at 80 V during the stacking phase and the voltage was increased to 140 V for further separation. The proteins were transferred from the SDS-PAGE to a polyvinylidene fluoride (PVDF, Bio-Rad) or nitrocellulose (GE Healthcare, Illinois, U.S.) membrane using the Trans-Blot SD Semi-Dry Transfer Cell (Bio-Rad). PVDF membranes were activated in methanol before transfer. Nitrocellulose membranes were stained after the transfer with ponceau S (Sigma) to evaluate the equal loading. After blocking the membranes in 5% (w/v) milk in PBS for 30 to 60 minutes the blots were incubated with primary antibody overnight at 4°C. To detect multiple proteins with variable molecular weight, the blot was cut accordingly. The molecular marker (PageRuler™ Prestained Protein Ladder, Thermo Scientific) above and below the expected protein height were present on the blot piece. The following primary antibodies were used: LEDGF/p75 (A300-848A, Bethyl Laboratories), HRP-2 (15134-1-AP, Proteintech Europe), Menin (A300-105A, Bethyl Laboratories), HA (ab215069, Abcam), Flag (F7425, Sigma) or GAPDH (ab9485, Abcam). After washing the membrane in PBS-T (PBS + 0.1% (v/v) triton X-100 (Acros Organics, Thermo Fisher Scientific)) for five minutes, three times, the blot was incubated with a dilution of secondary antibody (Dako, Aligent Technologies, California, U.S.) in 5% (w/v) milk for one hour at room temperature on a shaking device. Protein bands were visualized using ECL Prime Western Blot Detection Reagent (GE Healthcare, ref RPN2236) or ECL clarity Max (Bio-Rad, ref 1705062) in case of weak detection and ImageQuant 800 (GE Healthcare, Belgium).

3.11. Quantitative PCR

To obtain copy DNA (cDNA), RNA was extracted from cell pellets according to the Aurum Total RNA mini kit (Bio-Rad, California, U.S.). The RNA yield was measured using a spectrophotometer (IMPLEN, München, Germany) and 5 µg of RNA was reverse transcribed to cDNA using the High Capacity cDNA Reverse Transcription Kit (Thermo Fisher Scientific). Specific primers were used (Table 3.1) to detect and amplify the sequence of interest. Expression levels are determined relative to a standard curve and normalized to a housekeeping gene (*β-Actin* for human samples, *Gapdh* for mouse samples).

Table 3.1. Primers used for qPCR analysis

Gene, species		Primer sequence (5' → 3')
<i>Hrp-2</i> , mouse	FW	TGAGTCGGAGAAGACCAGTGACC
	RV	AATCCGAGGCTGATGGCACCTTC
<i>Psip1</i> , mouse	FW	CCTCAAACATGACTCGCGATTC
	RV	GCTCCATCAGGAACTTCATCTAC
<i>Gapdh</i> , mouse	FW	TGTGTCCGTCGTGGATCTGA
	RV	CCTGCTTCACCACCTTCTTGA
<i>HRP-2</i> , human	FW	AGGATGGAGAGCGACTCAGA
	RV	CGAGTTCTCCTCTTCGGATG
<i>PSIP1</i> , human	FW	GAACTTGCTTCACTTCAGGTCACA
	RV	TCGCCGTATTTTTTTCAGTGTAGT
<i>PWWP</i> , human	FW	CCAGCTCGAGTAGACGAAGT
	RV	TTGGAACTCATGAGACTGCT
<i>HoxA9</i> , human	FW	ATGGCATTAAACCTGAACCG
	RV	GTCTCCGCCGCTCTCATTC
β -Actin, human	FW	CACTGCGCGAGGCTACAGC
	RV	TTGATGTCGCGCACGATTT

3.12. Immunocytochemistry

8-well chambers were coated with poly-D-lysine (Sigma-Aldrich, Missouri, U.S.) before plating 50 000 cells/well. After attachment, the cells were fixated in 4% (v/v) paraformaldehyde (Sigma-Aldrich, Missouri, U.S.) for 15 minutes at room temperature and washed twice with PBS (ThermoFisher Scientific). Cells were permeabilized using 0.1% (v/v) tritonX-100 (Acros Organics, Thermo Fisher Scientific) in PBS to permeabilize the membrane and blocked for 30 minutes using PBS-based blocking buffer containing 1% BSA and 0.1% (v/v) Tween[®]-20 (Thermo Fisher Scientific, Massachusetts, U.S.) at room temperature. Primary antibody (1/400 anti-mouse Flag M2 F1804, Sigma) was incubated overnight in blocking buffer at 4°C before washing the cells with PBS three times. After a 30 minutes incubation in the dark with the secondary antibody (1/500 goat-anti-mouse-Alexa 555) and 4',6-diamidino-2-phenylindole (DAPI, 1/1000, Invitrogen, Massachusetts, U.S.) in blocking buffer and at room temperature, three additional PBS wash steps were performed. The sample was stored in the dark at 4°C until microscopic analysis. Immunocytochemistry (ICC) pictures were analyzed and colored using ImageJ software.

3.13. Suspension and colony formation analysis

To monitor the growth of all suspension cells, 50 000 cells/mL were plated six-fold in a 24-well plate. Every 24 hours, cells from one well were counted using trypan blue (Thermo Fisher Scientific) and the

TC20™ Automated Cell Counter from Bio-Rad. Ln-transformed cell counts were analyzed using GraphPad Prism 8 to determine the doubling time.

For the colony forming unit assay (CFU), the human cell lines were cultured in MethoCult™ H4230 (STEMCELL Technologies, Vancouver, Canada) and scored after 10 days in culture by manual counting through the microscope. Primary murine lineage depleted (lin^{-}) cells were cultured and scored in the CFU assay as described before [179].

3.14. *HRP-2* knockout mouse model

C57BL/6N-Hdgfrp2<tm1b (KOMP)Wtsi>/Tcp (*HRP-2*^{tm1b}) mice were ordered at Toronto Centre for Phenogenomics after they were generated as part of the NorCOMM2 phenotyping project [280]. All animal experiments were approved by the KU Leuven ethical committee (P201/2014).

3.15. Animal material processing

For the blood sampling of the mice, a minimum of 30 μ L blood was harvested from the submandibular vein in the cheek and collected in Microtubes for Automated Process with 1 mg K2EDTA (Becton Dickinson, New Jersey, U.S.). The blood was diluted in 1/10 in sterile PBS and counted using the Siemens Advia 2120 (Siemens, Munich, Germany) hematology analyzer.

Lineage depleted (lin^{-}) progenitor cells were isolated using the EasySep Mouse Hematopoietic Progenitor Cell Isolation Kit (STEMCELL Technologies, Vancouver, Canada) on red blood cell lysed bone marrow cells harvested from femur and tibia from 6 to 8-week-old mice. Cells were cultured or used for RNA extraction (Aurum™ Total RNA Mini Kit, Biorad).

3.16. RNA-sequencing and bioinformatics

Total RNA samples (500 ng) were cleaned using the DNase I kit (Thermo Fisher Scientific) according to the Rapid out removal DNA kit instruction and converted into cDNA by using the QuantSeq 3' mRNA-seq reverse 4 Library Prep Kit (Lexogen, Wien, Austria) according to manufacturer's instructions [281] to generate a compatible library for Illumina sequencing. Briefly, library generation was initiated by oligodT priming for first strand cDNA which generated one fragment per transcript. The second strand cDNA was subsequently synthesized using random primers. Illumina-specific linker sequences were introduced by the primer with barcoding indices for different samples. The quality of cDNA libraries was determined using a High Sensitivity DNA Assay 2100 Bioanalyzer (Agilent, California, U.S.) for quality control analysis. Sequencing of the cDNA library with 75bp single end reads was performed

using an Illumina NextSeq 500 system. Reads were aligned to the reference genome GRCm38 using STAR-2.4.2a with default settings [282]. STAR was also used for gene expression quantification on the Ensembl GTF file version 84. Differential expression analysis was performed using DESeq2 in R [283]. The RNA-sequencing data are available in NCBI's Gene Expression Omnibus (GSE154202)

4

Results

4.1. HRP-2 is dispensable for MLL-r leukemogenesis but important for leukemic cell survival

Initial findings in this project were published in the doctoral thesis of Sara El Ashkar [284]. This research was further expanded with immunoprecipitations, NMR data, additional genotyping, RNA sequencing, additional cellular data and in vivo experiments. Results reported in this section has previously been published in S. Van Belle et al., "Unlike Its Paralog LEDGF/p75, HRP-2 Is Dispensable for MLL-r Leukemogenesis but Important for Leukemic Cell Survival," Cells, 2021 [272].

4.1.1. HRP-2 interacts with MLL1 in the absence of menin

Given the interchangeable role of HRP-2 and LEDGF/p75 in HIV infection [243], [251] and given as well that LEDGF/p75 is important for MLL-r [94], [179], we investigated the involvement of HRP-2 in hematopoiesis and MLL-mediated transformation. Menin stabilizes the interaction between MLL1 and LEDGF/p75 and is thus essential for the formation of the triple transcription-regulatory complex [94]. Here, we investigated the potential interaction of HRP-2 with the MLL-menin complex by *in vitro* co-immunoprecipitation (IP) using HEK293T cells transfected with flag-tagged MLL-ELL and/or menin-HA expression constructs (Figure 4.1A-C). Binding of endogenous HRP-2 and LEDGF/p75 was detected using anti-HRP-2 or anti-LEDGF antibodies. Despite the poor detection of flag-MLL-ELL in the precipitate in the absence of ectopic menin expression, it is clear that upon overexpression of menin and MLL, both HRP-2 and LEDGF/p75 were precipitated. To further clarify whether the binding of HRP-2 to MLL-ELL is dependent on menin, as is the case for LEDGF/p75, we treated cellular lysates with 100 μ M of MI-2 [285], a MLL1-menin interaction inhibitor. Although we used double the concentration of MI-2 as published before [285], MI-2 treatment resulted in a partial loss of menin binding upon MLL1 precipitation (Figure 4.1B, top). HRP-2 still co-precipitated with MLL1 upon MI-2 treatment. Similar results were obtained using 50 μ M of the 20 times more potent MI-538 inhibitor (Figure 4.1B, bottom) [286]. Since both menin inhibitors did not fully abrogate the MLL1-menin interaction, we introduced three point mutations (F9A, P10A and P13A) into a flag-MLL-ELL₁₋₃₃₀ (MLL Mut) construct, known to completely abolish binding to menin [96]. Whereas the MLL1 mutations did not interfere with the precipitation of HRP-2 (Figure 4.1C), the MLL1 mutant failed to co-IP LEDGF/p75 despite the presence of overexpressed menin, suggesting HRP-2 is less dependent on menin for binding to MLL1. Whereas ectopic overexpression of menin resulted in increased co-precipitation of LEDGF/p75, binding to HRP-2 was reduced, supporting a competition between HRP-2 and LEDGF/p75 for binding to MLL1 under control of menin (Figure 4.1A and C).

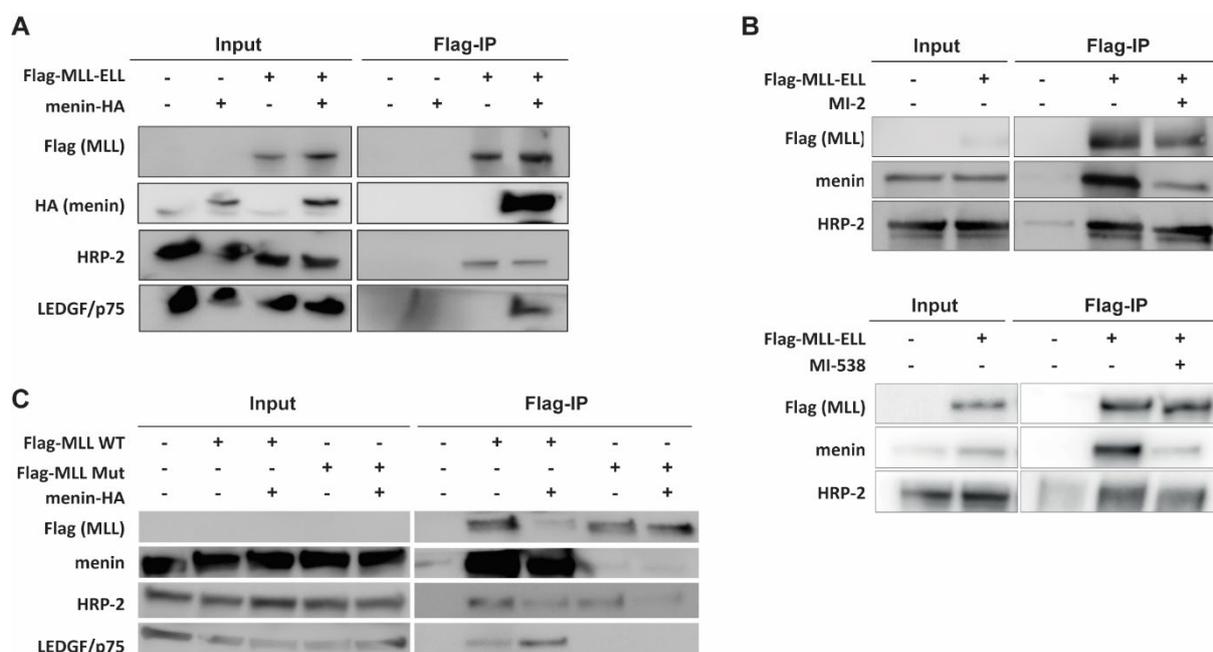


Figure 4.1: Menin is dispensable for the interaction of MLL1 and HRP-2.

(A) HEK293T cells were transfected with flag-tagged MLL-ELL and/or menin-HA expression constructs as indicated. Flag-MLL-ELL was immunoprecipitated using anti-flag beads and analyzed using antibodies against flag. Menin, endogenous HRP-2 and LEDGF/p75 were detected using specific antibodies; (B) HEK293T cells were transfected with flag-tagged MLL-ELL as indicated. MLL-ELL was immunoprecipitated using anti-flag beads in the presence of a previously described menin-MLL1 interaction inhibitor (top) MI-2 [285] or (bottom) more potent MI-538 [286], at concentrations of 100 and 50 μ M respectively or DMSO control. Precipitated proteins were analyzed by western blot. Flag antibodies were used for the detection of MLL-ELL. Endogenous levels of menin, HRP-2 and LEDGF/p75 were detected using specific antibodies; (C) HEK293T cells were transfected with flag-tagged wild-type MLL-ELL₁₋₃₃₀ (MLL WT) or a menin interaction-deficient construct 'MLL Mut' with point mutations F9A, P10A and P13A and/or menin-HA as indicated. Due to low expression levels, MLL₁₋₃₃₀ is not detected in the input by flag antibody. Endogenous levels of menin, LEDGF/p75 and HRP-2 are detected using specific antibodies.

4.1.2. HRP-2 and LEDGF/p75 interact with MLL1 through a conserved interface

To obtain a detailed insight into the mechanism of the MLL1 and HRP-2 interaction, the solution structure of the HRP-2-IBD domain (amino acids 469-549, Figure 4.2A) was determined. The solution structure of the HRP-2-IBD revealed a compact right-handed oriented bundle composed of five α -helices, comparable to other members of the TFIIIS N-terminal domain family and demonstrated a high degree of structural conservation between LEDGF/p75 and HRP-2 IBDs (Figure 4.2C).

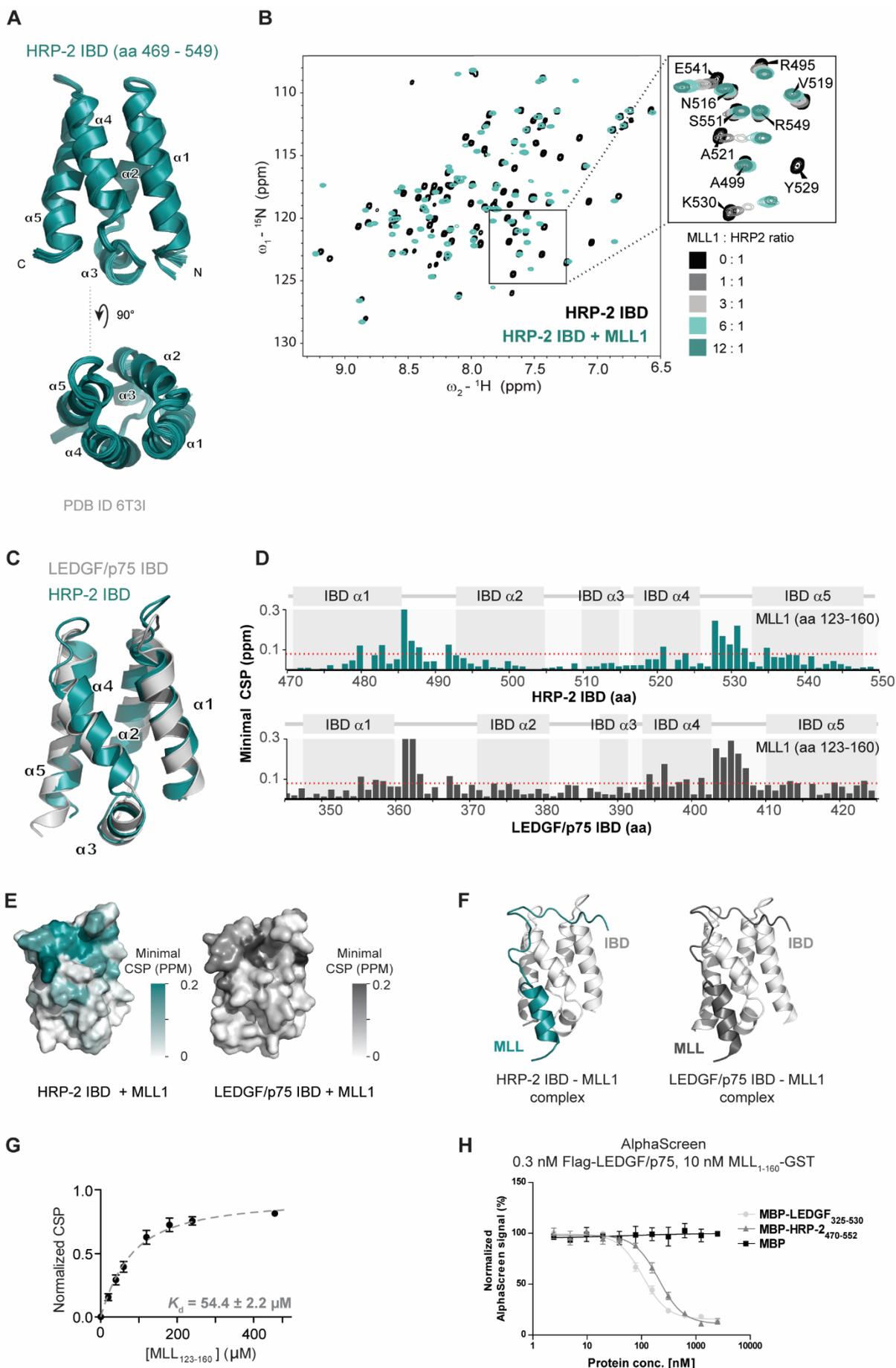
To validate and characterize the direct interaction between HRP-2-IBD and MLL1 in NMR, we followed the changes in positions of backbone NMR signals of ¹⁵N-labeled HRP-2-IBD either in absence or presence of different concentrations of a synthetic MLL-derived peptide (amino acids 123-160). MLL₁₂₃₋₁₆₀ induced significant chemical shift perturbations of the IBD backbone signals (Figure 4.2B). Moreover, this experiment revealed that MLL1 recognizes HRP-2-IBD through the same interface and

binds with similar affinity as LEDGF/p75 (Figure 4.2D-H). In particular, the chemical shift perturbations in the IBD backbone induced by binding of MLL₁₂₃₋₁₆₀ were found in two regions (amino acids 479-492 and 520-535) (Figure 4.2D). As for the LEDGF/p75 IBD-MLL₁₂₃₋₁₆₀ interaction [202], these regions form two interhelical loops connecting IBD α helices α 1- α 2 and α 4- α 5, respectively (Figure 4.2E,F). Additionally, analysis of the chemical shift perturbations in the HRP-2 IBD backbone induced by binding of JPO2 (amino acids 1-130) and POGZ (amino acids 1117-1410) revealed a pattern remarkably similar to that induced upon addition of MLL1 confirming that these protein fragments bind to HRP-2 in the same conserved structural mode as LEDGF/p75 IBD [272].

Importantly, the dissociation constants for the MLL₁₂₃₋₁₆₀ interaction with HRP-2-IBD ($54.4 \pm 2.2 \mu\text{M}$, Figure 4.2G) obtained by NMR titration experiments are comparable with those of LEDGF/p75 ($64.0 \pm 6.5 \mu\text{M}$) obtained in earlier studies [201]. In addition, affinity of HRP-2 and LEDGF/p75 to MLL1 were compared in an AlphaScreen assay with MBP-fused recombinant C-terminal fragments of LEDGF/p75 (LEDGF₃₂₅₋₅₃₀) and HRP-2 (HRP-2₄₇₀₋₅₅₂), purified from *E. coli* and used to outcompete the interaction between recombinant flag-tagged LEDGF/p75 and GST-tagged N-terminal fragment of MLL1 (MLL₁₋₁₆₀-GST). Unlike MBP alone, both IBD domains of HRP-2 and LEDGF/p75 efficiently outcompeted the interaction between MLL₁₋₁₆₀-GST and full-length flag-LEDGF/p75 (Figure 4.2H). Altogether, our data revealed that the overall binding mechanism used by HRP-2 and LEDGF/p75 IBDs is highly conserved.

Figure 4.2: HRP-2 IBD and LEDGF/p75 IBD interact with MLL1 in a conserved manner

(A) Solution structure of HRP-2 IBD (PDB ID 6T3I); (B) HRP-2-IBD directly interacts with the identical consensus motif of MLL1 (amino acids 123-160) alike LEDGF/p75-IBD as determined by NMR spectroscopy. Comparison of the $^{15}\text{N}/^1\text{H}$ HSQC spectra of the 20 μM HRP-2 IBD in the absence (black) and presence (green) of 120 μM MLL₁₂₃₋₁₆₀. On the right, detail HRP-2 IBD titration with MLL₁₂₃₋₁₆₀. HSQC spectra are colored based on MLL₁₂₃₋₁₆₀ concentration as indicated in the figure. The spectra were obtained from the ^{15}N -labeled recombinant IBD and the unlabeled synthetic MLL-derived peptide; (C) Superposition of HRP-2-IBD (green) and LEDGF/p75-IBD (light grey) solution structures; (D, E) Comparison of the HRP-2-IBD – MLL₁₂₃₋₁₆₀ and LEDGF/p75-IBD – MLL₁₂₃₋₁₆₀ interaction surfaces. Representation of the minimal chemical shift perturbation (CSP) in backbone amide signals of the IBDs upon addition of MLL₁₂₃₋₁₆₀ peptide in panel D. Amino acid residues that are significantly perturbed upon addition of MLL₁₂₃₋₁₆₀ to HRP-2-IBD or LEDGF/p75-IBD (as determined by NMR spectroscopy) are highlighted in green or gray on the surface of the IBD structures in panel E; (F) Comparison of homology model of MLL-HRP-2-IBD and solution structure of MLL-LEDGF/p75-IBD solution structure (PDB ID 6emq); (G) K_d fit from NMR titrations of HRP-2-IBD with MLL₁₂₃₋₁₆₀. Dissociation constant was determined by following the chemical shift perturbations of the HRP-2-IBD backbone amide signals induced upon titration with MLL₁₂₃₋₁₆₀. Error bars represent the error of the fit for most perturbed residues (n=10); (H) Alpha Screen. Full length flag-tagged LEDGF/p75 (0.3 nM) was incubated with GST-tagged MLL₁₋₁₆₀ (10 nM) in an out-competition AlphaScreen assay. C-terminal fragments of LEDGF/p75 (LEDGF₃₂₅₋₅₃₀) and HRP-2 (HRP-2₄₇₀₋₅₅₂) were titrated to outcompete the interaction.



4.1.3. Systemic HRP-2 depletion in mice leads to increased postnatal mortality and decreased *in vitro* colony formation of hematopoietic stem cells

To address the role of HRP-2 in postnatal hematopoiesis, we investigated a systemic knockout mouse model from the Toronto Centre Phenogenomics, Canada. Depletion of HRP-2 mRNA and protein levels in these knockout (*Hrp-2^{-/-}*) mice were confirmed by quantitative PCR (qPCR) and Western blot in bone marrow cells (Figure 4.3A). Observation of inbred crossings revealed a distribution of 61% and 29.5% of heterozygous (*Hrp-2^{+/-}*) and wild type (*Hrp-2^{+/+}*) mice, respectively, at age of 6 to 8 weeks (Table 4.1). These percentages deviate from the expected Mendelian inheritance pattern. Moreover, *Hrp-2^{-/-}* mice were present at a lower percentage (9.5%) than expected. At one day after birth, the ratio of *Hrp-2^{-/-}* mice corresponded more closely to the expected Mendelian ratio by representing 20% of newborn pups.

Table 4.1. Genotype of heterozygous (*Hrp-2^{+/-}*) crossed off spring after birth or weaning (6 to 8 weeks)

Age	Number (%) of mice			
	<i>Hrp-2^{+/+}</i>	<i>Hrp-2^{+/-}</i>	<i>Hrp-2^{-/-}</i>	
1 day	36 (23.2)	88 (56.8)	31 (20.0)	*
6-8 weeks	31 (29.5)	64 (61.0)	10 (9.5)	
Expected distribution	25%	50%	25%	

*, $p = 0.012$ in Chi square test compared to expected mendelian distribution

In contrast to an earlier report by Wang et al. [287], we observed that *Hrp-2^{-/-}* pups presented with increased mortality before 6-8 weeks of age, suggesting that HRP-2 is important for postnatal survival early after birth. Read-through of the gene trap could explain discrepancies between both models, since Wang et al. showed a 5-20% residual *Hrp-2* expression in their model, while *Hrp-2* mRNA levels in hematopoietic stem cells (HSC) of our few surviving mice were undetectable by qPCR (Figure 4.3A).

Differences in steady-state hematopoiesis of weaned mice, was evaluated by a differential blood count for each genotype. No significant differences were observed in total white and red blood cell count (Figure 4.3B), whereas the differential blood count revealed a significant increase in neutrophils in the *Hrp-2^{-/-}* mice compared to the wild type ($p=0.042$). Other cell types did not differ between genotypes (Figure 4.3B). To explore the functionality of the cells, we sought to compare the colony-forming capacity of *Hrp-2* wild type, heterozygous and knockout lineage depleted (*lin⁻*) cells using myeloid CFU assays. The *lin⁻* cells were harvested from *Hrp-2^{+/+}*, *Hrp-2^{+/-}* and *Hrp-2^{-/-}* mice and serially plated. After

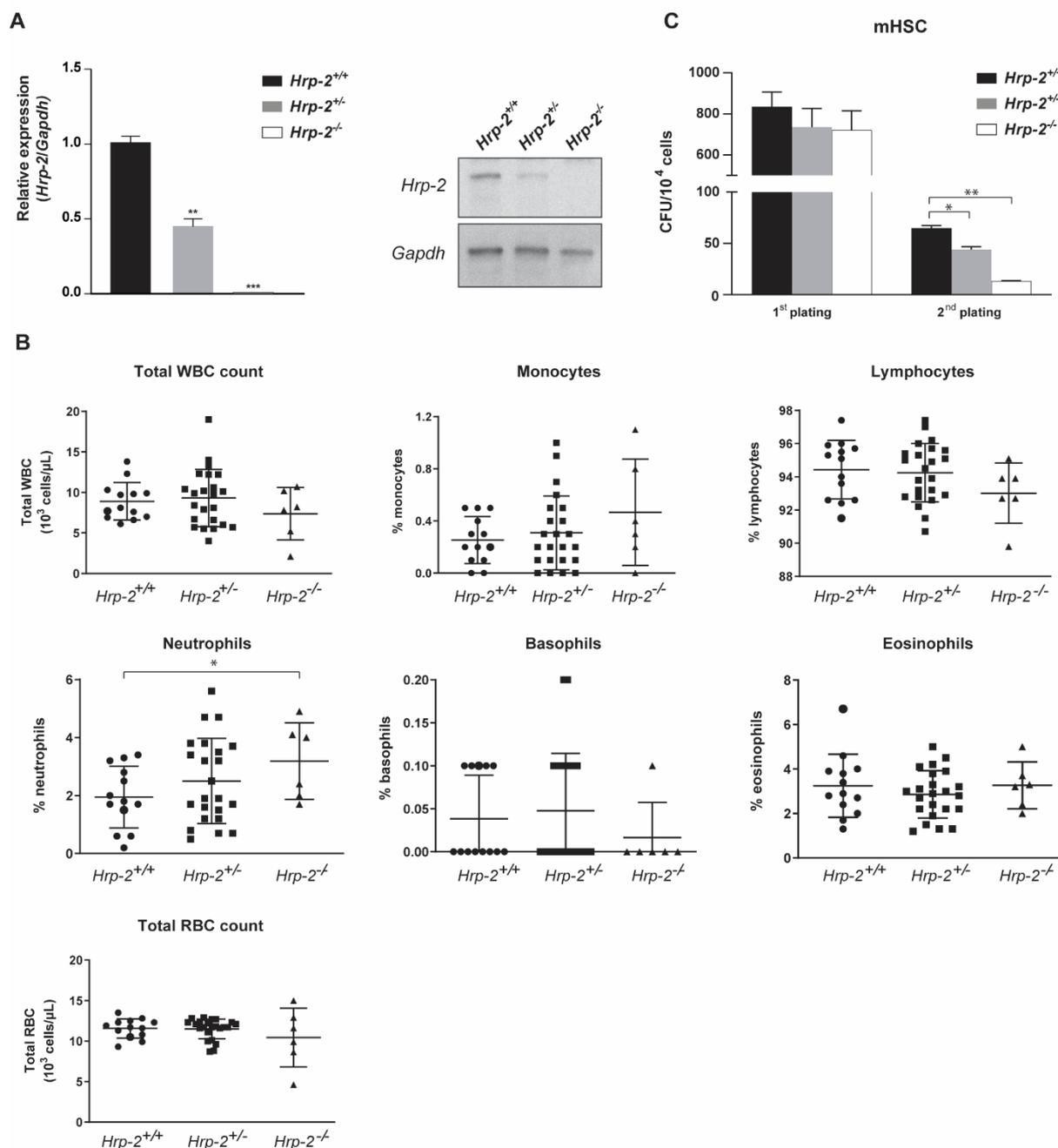


Figure 4.3: *Hrp-2* knockout mice show subtle hematopoietic defects

(A, left) qRT-PCR of lineage depleted bone marrow cells to validate *Hrp-2* mRNA expression in knockout mice (*Hrp-2*^{-/-}), compared to wild type (*Hrp-2*^{+/+}) and heterozygous (*Hrp-2*^{+/-}) animals. Expression levels were normalized to *Gapdh*. Error bars indicate standard deviations of triplicate measurements. Differences were determined using Student's t-test; ** $p < 0.01$, *** $p < 0.001$; (A, right) Western Blot of whole bone marrow cells to validate HRP-2 protein expression levels in HRP-2 wild type (*Hrp-2*^{+/+}), heterozygous (*Hrp-2*^{+/-}) and knockout (*Hrp-2*^{-/-}) animals. GAPDH was used as loading control. (B) Peripheral blood counts in wild type (*Hrp-2*^{+/+}), heterozygous (*Hrp-2*^{+/-}) and knockout (*Hrp-2*^{-/-}) mice. Average and standard deviation are indicated. Significance level was determined using two-sided student's t-test (* $p = 0.042$); (C) Number of colonies for 10⁴ lineage depleted cells harvested from *Hrp-2*^{+/+}, *Hrp-2*^{+/-} and *Hrp-2*^{-/-} mice in two consecutive platings in a myeloid CFU assay. Error bars indicate standard deviations of duplicate measurements.

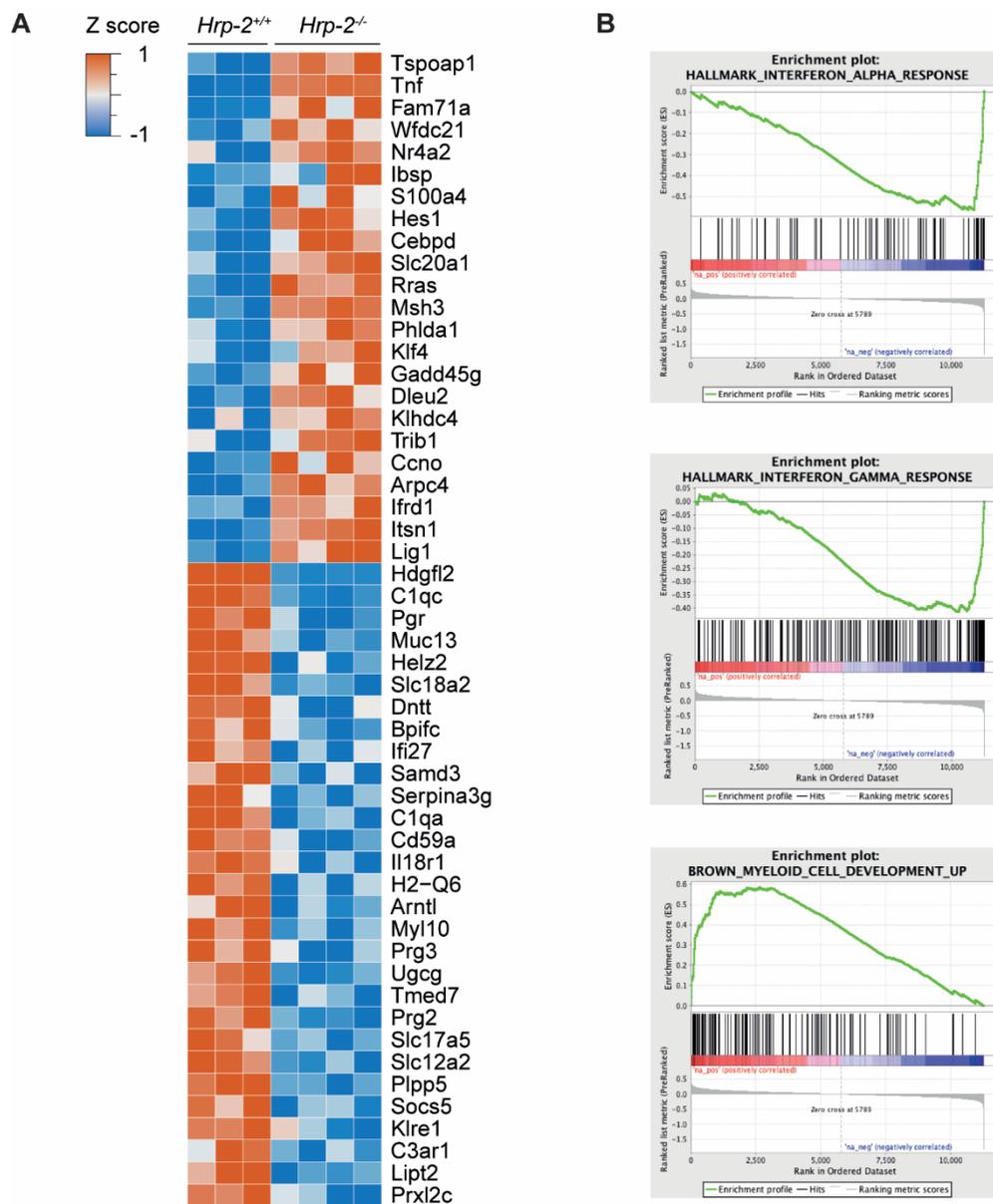


Figure 4.4: RNA sequencing analysis of *Hrp-2* wild type and knockout mice

(A) Heat map of differentially expressed genes detected by RNA-seq for *Hrp-2*^{-/-} lineage depleted bone marrow (lin⁻) cells compared to *Hrp-2*^{+/+} lin⁻ cells. Color scale indicates the upregulated (red) and downregulation (blue) genes. Each row represents one animal. (B) Gene set enrichment analysis (GSEA) showing a correlation between the RNA profile of *Hrp-2*^{-/-} lineage depleted (lin⁻) bone marrow cells compared to the gene signature of myeloid differentiation and down-regulation of interferon pathways alpha (middle) and gamma (right).

two rounds of plating, the number of colonies derived from *Hrp-2*^{-/-} and *Hrp-2*^{+/+} bone marrow cells were respectively 80% and 33% lower compared to the wild-type control (Figure 4.3C). To gain more detailed knowledge on the *Hrp-2* knockouts, we performed a gene expression profile analysis,

comparing RNA of lin^- bone marrow cells of *Hrp-2^{+/+}* and *Hrp-2^{-/-}* mice. We found a total of 52 differentially expressed genes (FDR 0.25, Figure 4.4A), of which 23 were upregulated and 29 were downregulated in *Hrp-2^{-/-}* cells. Gene set enrichment analysis (GSEA) uncovered that *Hrp-2^{-/-}* cells display a trend towards a gene signature of myeloid differentiation ($p < 2.2e-16$, Figure 4.4B). Moreover, down-regulation of interferon alpha and gamma pathways was observed in multiple gene sets ($p < 2.2e-16$, Figure 4.4B). Taken together, these results hint that HRP-2 depletion reduces the colony formation capacity and induces myeloid differentiation, suggesting that HRP-2 is involved in maintaining the stem-like state of bone marrow cells.

4.1.4. HRP-2 depletion impairs the clonogenic growth of both human and mouse leukemic cell lines independently of MLL1 fusions

Next, the role of HRP-2 in leukemic transformation induced by oncogenic MLL1 fusions was investigated in human leukemic cell lines harbouring the MLL-fusions MLL-AF9 (THP1 cells) or MLL-AF4 (SEM cells), as well as the MLL1 wild-type leukemic cell lines Kasumi1, K562 and Nalm6. All cell lines were transduced with a lentiviral vector to deplete HRP-2 or with a control vector and mRNA and protein expression levels were validated by qPCR and western blot respectively (Figure 4.5B-C). Of note, LEDGF/p75 mRNA and protein levels remained unaffected upon HRP-2 depletion. Stably transduced cells were plated in methylcellulose-based medium to monitor colony growth. After 12 days, a decrease in the number of colonies of THP1 ($32.8 \pm 2.1\%$) and SEM ($51.5 \pm 7.0\%$) cells was observed (Figure 4.5A). Interestingly, the number of colonies in MLL1 wild type cell lines also decreased upon HRP-2 knockdown (Figure 4.5A) with $44.4 \pm 27.4\%$ (K562), $38.3 \pm 14.2\%$ (Kasumi1) and $19.3 \pm 8.8\%$ (Nalm6). Experiments with higher vector titers resulted in an even more pronounced drop in the number of colonies (Figure 4.5, inserted table), suggesting a concentration-dependent effect. In liquid culture, an impaired cellular growth was observed for SEM and Kasumi1 cells but not for THP1, K562 and Nalm6 cells, excluding that impaired cell growth by HRP-2 depletion is affected in an MLL-r dependent way (Figure 4.6). As well as for cultured murine lin^- bone marrow cells expressing an MLL-AF9 fusion or a control fusion E2A-HLF, respectively 70% and 56% less colonies were observed after a lentiviral-induced *Hrp-2* knockdown (Figure 4.7). Taken together, these observations suggest that loss of HRP-2, in contrast to LEDGF/p75, generally impairs growth of human and murine leukemic cells even in the absence of MLL1 fusions. Since LEDGF/p75 depletion results in decreased expression of HoxA9 in MLL-FP cells [94], [95], we wanted to further emphasize whether the observed phenotype was independent of MLL. HoxA9 levels assessed by qPCR (Figure 4.5D) were not significantly affected by HRP-2 depletion.

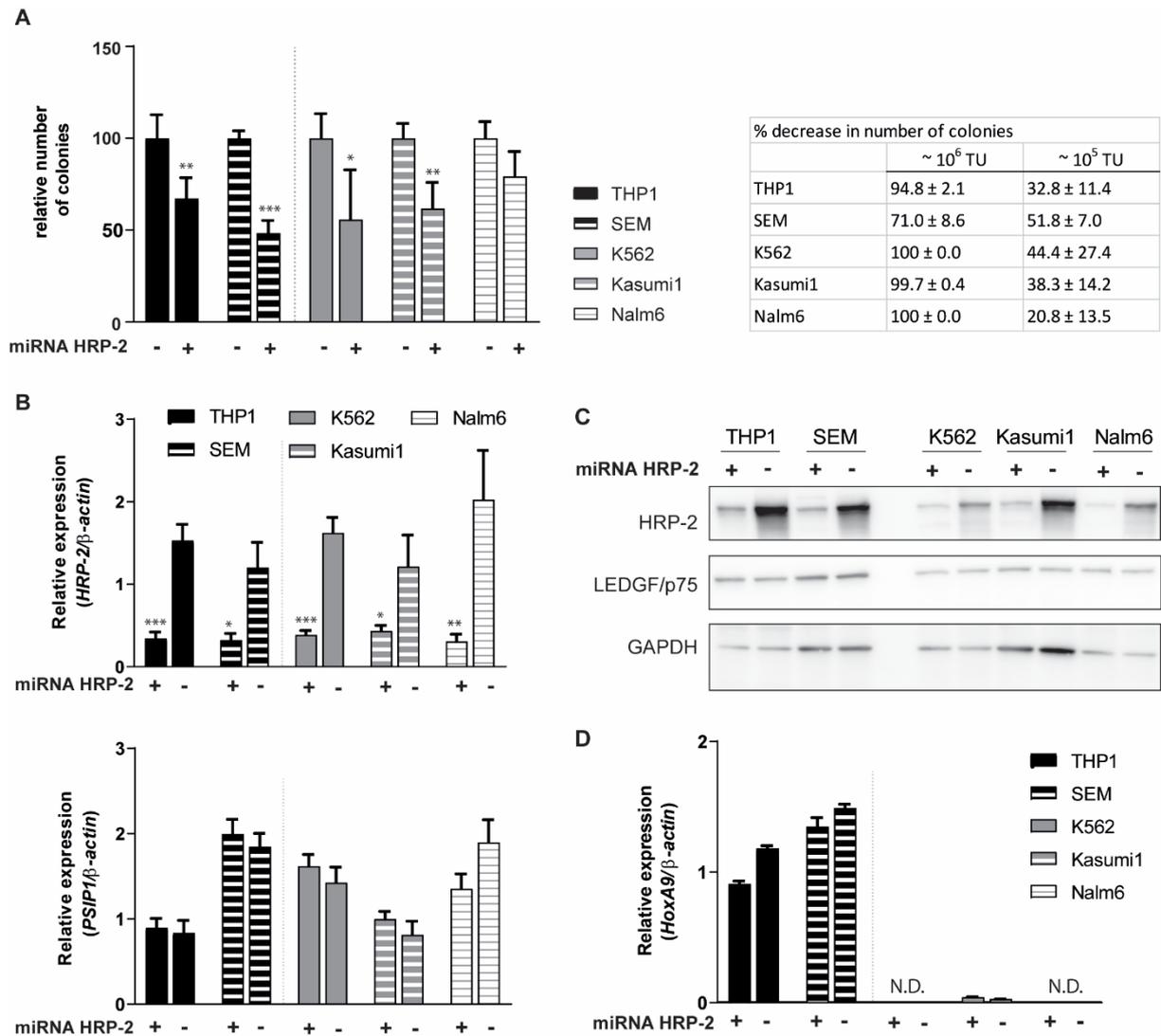


Figure 4.5: *HRP-2* knockdown impairs growth in human leukemic cell lines

(A) Human MLL-transformed cell lines THP1 (MLL-AF9) and SEM (MLL-AF4), as well as wild type MLL1 cells (Nalm6, Kasumi1 and K562), separated by the dotted line, were transduced with a lentiviral vector encoding a miRNA to knockdown *HRP-2* (+) or a control vector (-). After 12 days in methylcellulose the number of colonies was scored. Counts were normalized to their associated control. Error bars indicate standard deviations of four replicates. Differences were determined using Mann-Whitney U test; * $p < 0.05$. Inserted table describes average percentage decrease in number of colonies for indicated vector titers. TU = titer units (p24 $\mu\text{g/mL}$) \pm S.D.; (B) mRNA levels of *HRP-2* (top) and *PSIP1* (LEDGF/p75, bottom) in whole cell lysates. Expression levels were verified by qPCR and normalized to β -actin. One representative experiment shown ($n=3$). Error bars indicate the standard deviation of three replicates; (C) Western blot analysis of *HRP-2* and LEDGF/p75 protein expression levels after transduction with miRNA to knockdown *HRP-2* (+) or control (-). GAPDH was used as loading control; (D) qPCR analysis of *HoxA9* expression levels in human cell lines, normalized to β -actin. One representative experiment shown ($n=3$). Error bars indicate the standard deviation of duplicates. N.D. = not detectable

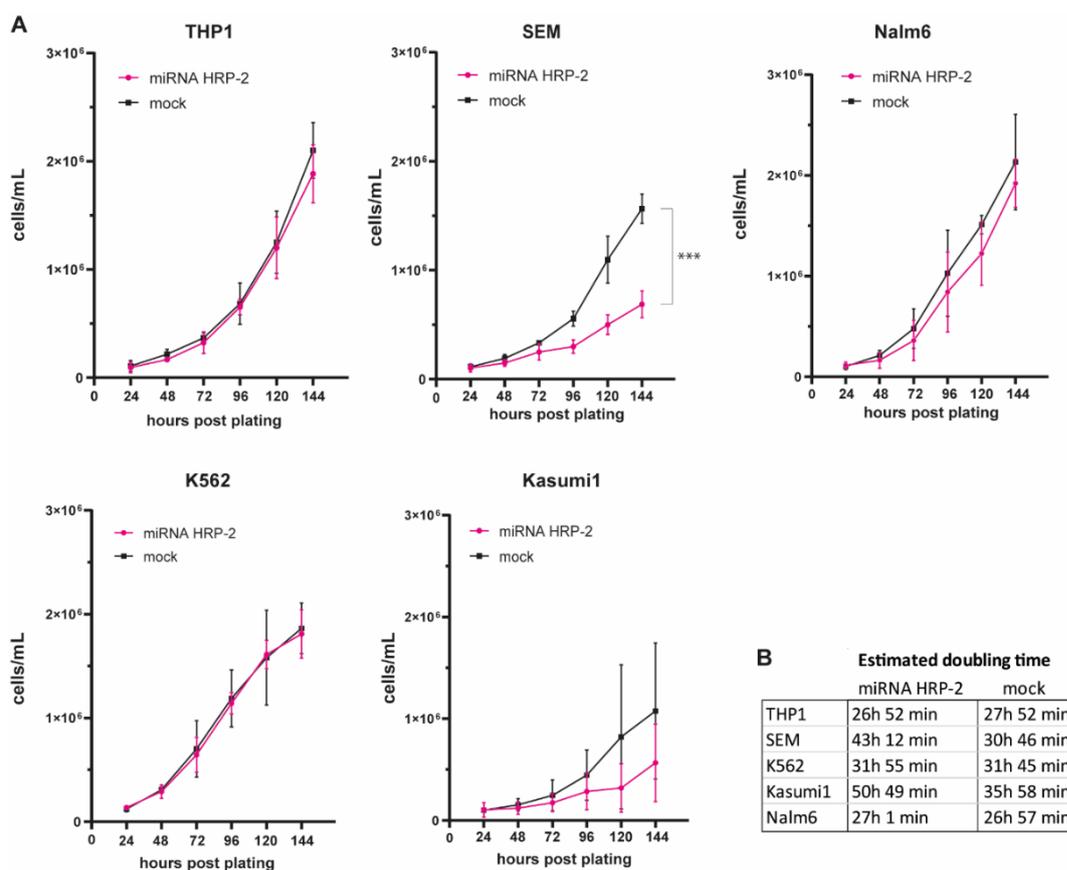


Figure 4.6: HRP-2 depletion variably affects growth of human cell lines in liquid culture

(A) Growth of cultured cells was monitored for six consecutive days by counting the number of cells per mL culture. Error bars indicate the standard deviation of three independent replicates. Differences at the end point were determined using Mann-Whitney U test; *** $p < 0.003$; (B) Inserted table comparing doubling time between HRP-2 depleted (miRNA HRP-2) and control (mock) cells lines from the represented graphs.

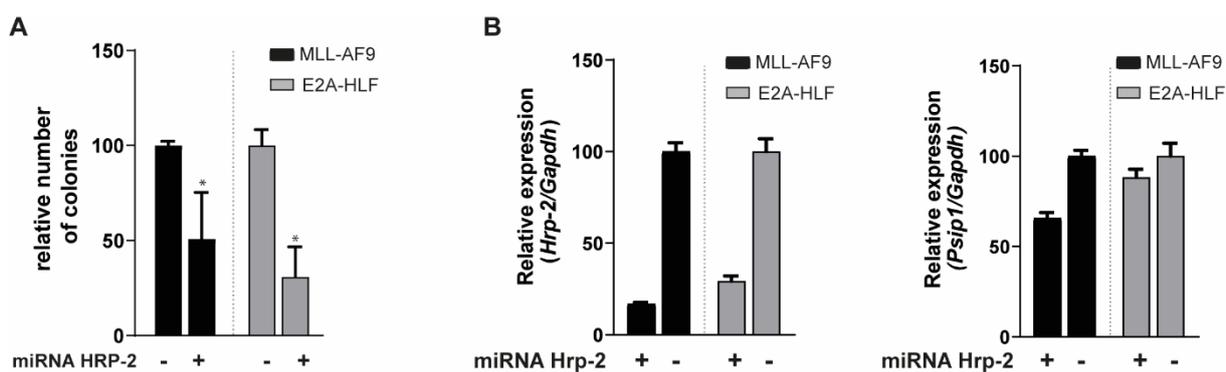


Figure 4.7: *HRP-2* knockdown impairs growth in murine leukemic cell lines

(A) Colony forming assay (CFU) after *HRP-2* knockdown (+) or control (-) of primary bone marrow cells harvested from leukemic mice transplanted with MLL-AF9 or E2A-HLF transduced cells. Counts were normalized to their associated control. Error bars indicate standard deviations of four replicates. Differences were determined using Mann-Whitney U test; * $p < 0.05$; (B) mRNA levels of *HRP-2* (left) and *PSIP1* (LEDGF/p75, right) in whole cell lysates. Expression levels were verified by qPCR and normalized to *Gapdh*. One representative experiment shown ($n=3$).

4.1.5. HRP-2 overexpression rescues MLL1 clonogenic growth in LEDGF/p75-depleted cells

Since HRP-2 depletion affects clonogenic growth of all leukemic cell lines tested, we investigated whether HRP-2 overexpression could rescue the colony-forming capacity of an MLL-AF9 leukemic cell line (THP1) in the absence of LEDGF/p75. First, we stably expressed miRNA-resistant LEDGF/p75, HRP-2 or a mock control. Subsequently, cell lines were transduced with a miRNA-expressing lentiviral vector to specifically knockdown the LEDGF/p75 encoding gene *Psip1* or control (Figure 4.8A) and expression levels were verified by qPCR and western blot (Figure 4.8B-C). In line with previous reports, LEDGF/p75 depletion caused a ~65% decrease in the number of colonies formed compared to the control [94], [95], [179] (Figure 4.8A).

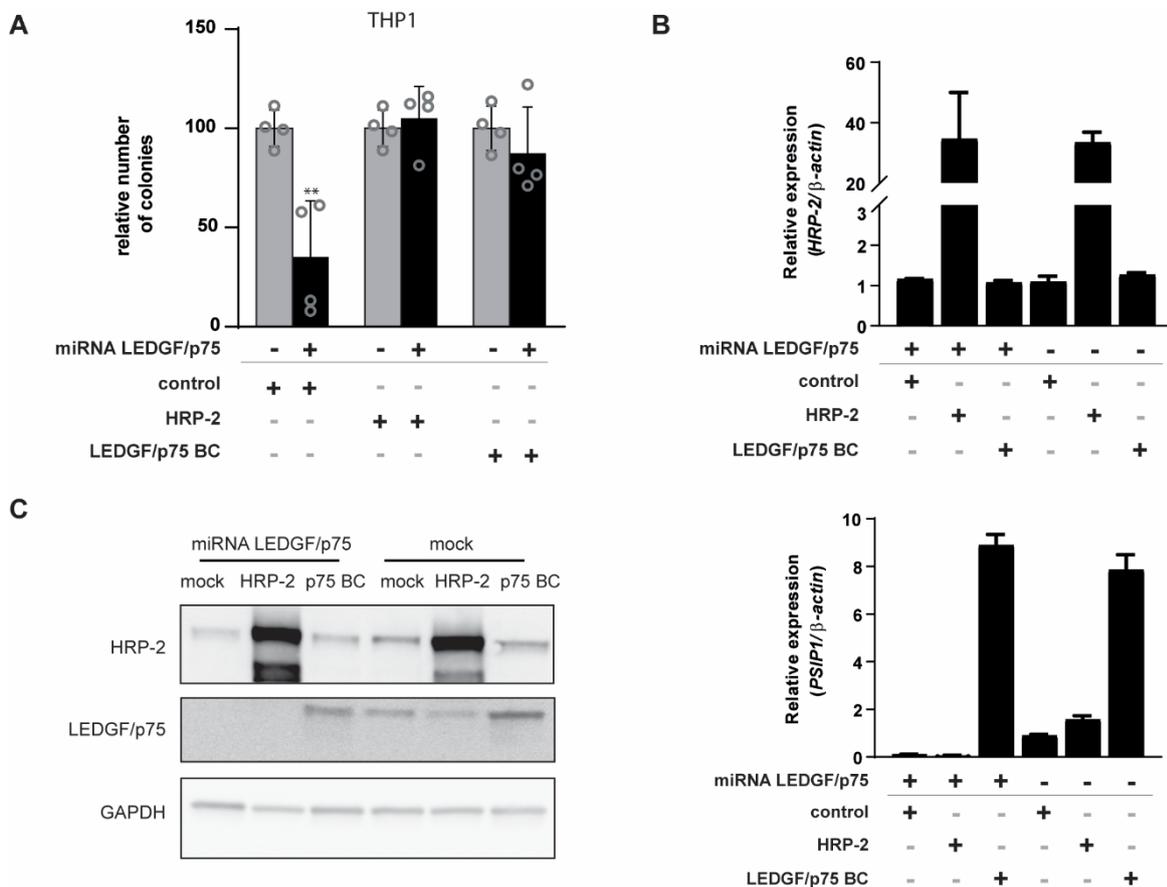


Figure 4.8: The LEDGF/p75-driven leukemic phenotype is rescued by HRP-2 overexpression

(A) Relative number of colonies per 500 plated cells for the THP1 cell line overexpressing miRNA resistant LEDGF/p75, HRP-2 or empty vector (control) after transduction with a lentiviral vector expressing a LEDGF/p75-miRNA to knockdown LEDGF/p75 or a control (black -). Error bars indicate standard deviations of triplicate measurements. Differences were determined using Mann-Whitney U test; * $p < 0.05$. (B) qPCR analysis for *HRP-2* (top) and *PSIP1* (LEDGF/p75, bottom) for all generated cell lines. One representative experiment is shown ($n=3$). Error bars represent standard deviation of duplicates. Expression levels were normalized to β -actin; (C) Whole cell lysates of all generated cell lines in THP1 were validated using specific antibodies in a western blot. One representative experiment is shown ($n=3$).

The defect was rescued by LEDGF/p75 back-complementation (BC). Of interest, overexpression of HRP-2 fully restored the CFU activity to wild-type levels, indicating that overexpression of HRP-2 can functionally compensate for the absence of LEDGF/p75 in MLL-transformed cells. The experiment was repeated in the MLL1 wild-type Nalm6 cell line (Figure 4.9). In line with published data [95], [179], depletion of LEDGF/p75 did not affect the colony forming capacity of the Nalm6 cell line, as the LEDGF/p75-dependent drop in colonies is MLL-specific. Although both HRP-2 and LEDGF/p75 overexpression increased the colony formation of Nalm6 cells, the effect was not significant.

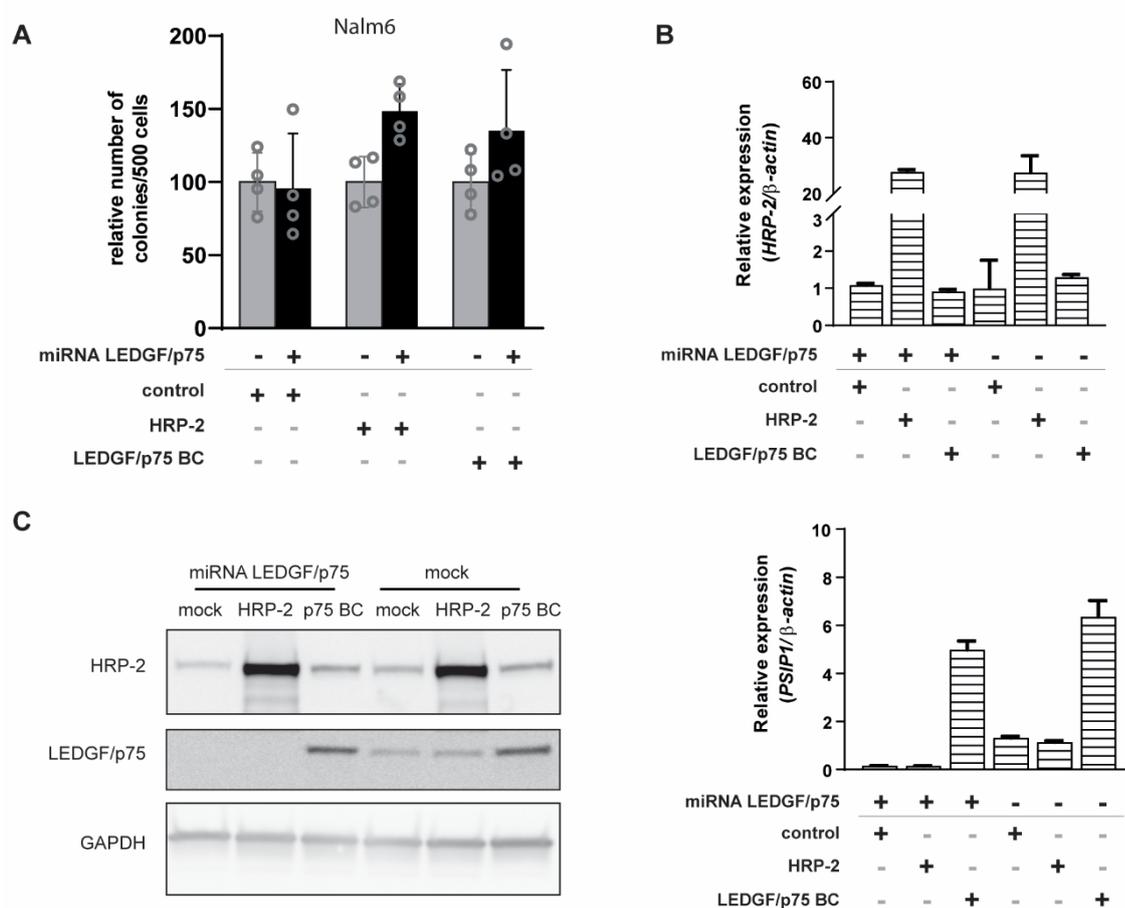


Figure 4.9: Colony formation of Nalm6 cells is unaffected by LEDGF/p75 depletion and HRP-2 rescue

(A) Relative number of colonies per 500 plated cells for the Nalm6 cell line overexpressing miRNA resistant LEDGF/p75, HRP-2 or empty vector (control) after transduction with a lentiviral vector expressing a LEDGF/p75-miRNA to knockdown LEDGF/p75 or a control (black -). Error bars indicate standard deviations of triplicate measurements. Differences were determined using Mann-Whitney U test; * $p < 0.05$. (B) qPCR analysis for *HRP-2* (top) and *PSIP1* (LEDGF/p75, bottom) for all generated cell lines. One representative experiment is shown ($n=3$). Error bars represent standard deviation of duplicates. Expression levels were normalized to β -actin; (C) Whole cell lysates of all generated cell lines in Nalm6 were validated using specific antibodies in a western blot. One representative experiment is shown ($n=3$).

4.1.6. HRP-2 is dispensable for the initiation of MLL-r leukemia

These experiments imply that HRP-2 and LEDGF/p75 display physiological different roles in cells. Therefore, we examined the involvement of HRP-2 in the initiation of *MLL*-r leukemia *in vivo*. HSC harvested from *Hrp-2^{+/+}*, *Hrp-2^{+/-}* and *Hrp-2^{-/-}* mice were transduced with lentiviral vectors encoding one of the most common MLL-fusion proteins (MLL-ENL) or a control fusion inducing acute lymphoblastic leukemia (E2A-HLF) and their transformation potentials were compared in serial plating assays (Figure 4.10A). Remarkably, *Hrp-2* wild type, heterozygous and knockout cells were efficiently transformed by MLL-ENL and E2A-HLF as revealed by the increased number of colonies after three rounds in the CFU assay (Figure 4.10B). The drop of colonies in the second round in MLL-ENL is likely due to a low transduction efficiency and represents a selection step prior to cellular transformation. Finally, we performed an *in vivo* bone marrow transplantation, where MLL-ENL-expressing *Hrp-2^{+/+}*, *Hrp-2^{+/-}* and *Hrp-2^{-/-}* cells were transplanted into lethally irradiated recipient mice and monitored for leukemogenesis. Kaplan-Meier survival plots reveal that mice transplanted with MLL-ENL transformed *Hrp-2^{+/-}* and *Hrp-2^{-/-}* cells died significantly faster than the transformed wild-type cells ($p = 0.02$ and $p = 0.006$, respectively) (Figure 4.10C), excluding the requirement of HRP-2 in leukemogenesis.

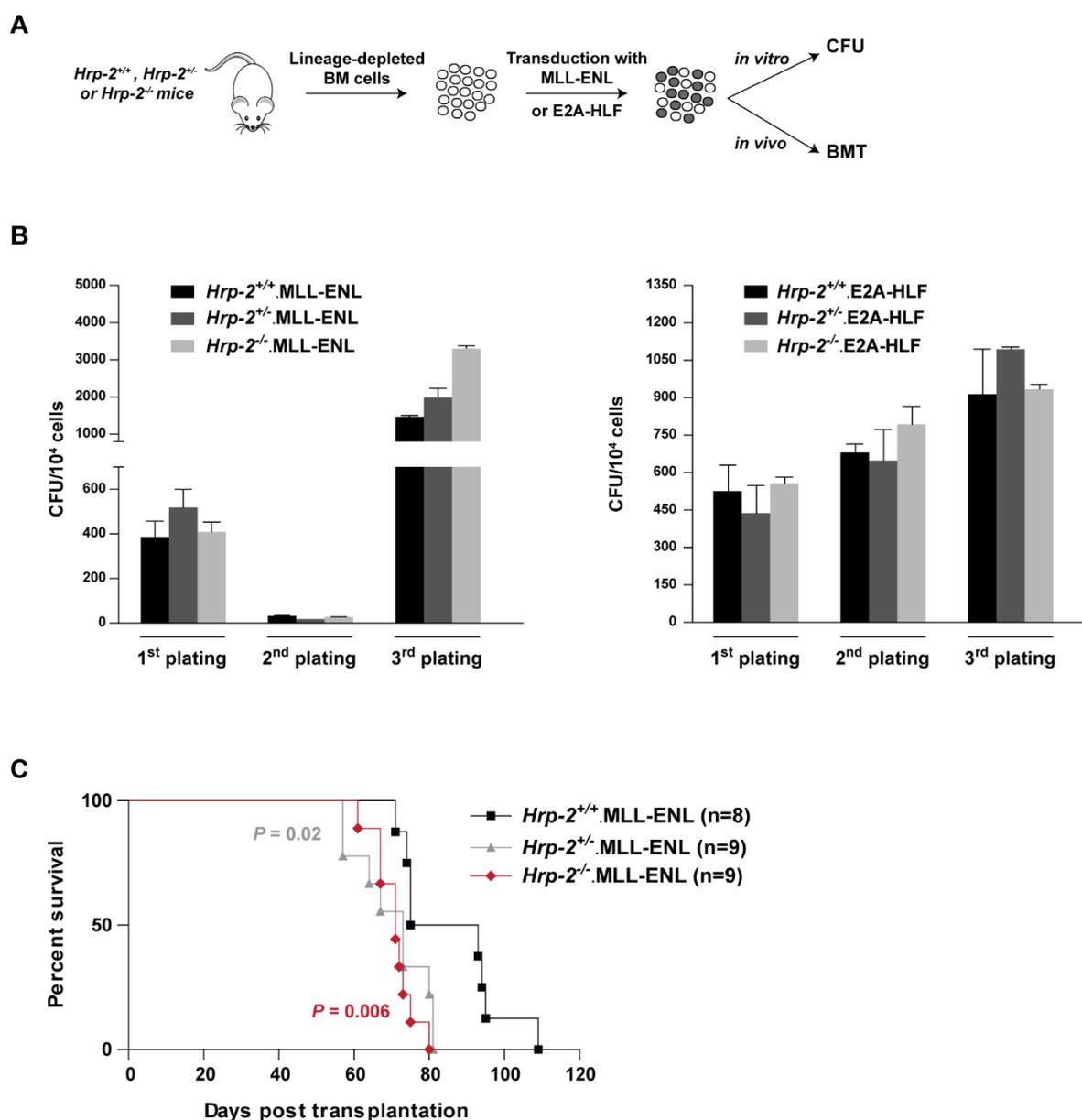


Figure 4.10: HRP-2 is not required for the initiation of leukemia *in vivo*

(A) Schematic representation of the experimental set up. BM = bone marrow; CFU = colony forming assay; BMT, bone marrow transplantation; (B) Colony-forming assay (CFU) for 10⁴ *Hrp-2*^{+/+}, *Hrp-2*^{+/-} and *Hrp-2*^{-/-} cells transduced with a retroviral vector encoding the MLL-ENL (B, left) or E2A-HLF (B, right) fusion. Error bars represent standard deviation of triplicate measurements; (C) Kaplan-Meier survival curve from bone marrow transplantation experiments of irradiated mice who received *Hrp-2* wild type (^{+/+}), heterozygous (^{+/-}) or knockout (^{-/-}) cells, transduced with mouse stem cell virus (MSCV) MLL-ENL expression vector. Number of transplanted animals (n) per group is indicated. Statistical differences were determined using GraphPad prism.

4.2. Development of small molecules targeting the PWWP domain of LEDGF/p75 in the context of mixed lineage leukemia.

We have shown previously that the structurally and functionally related proteins LEDGF and HRP-2 play different physiological roles in the development and maintenance of MLL-r leukemia [95], [179], [272]. Clear evidence has been given for the importance of LEDGF/p75 in the pathogenesis of MLL-r leukemia, which pinpoints LEDGF/p75 as an interesting drug target for the treatment of MLL-r leukemia [94], [95], [179]. When looking at the structure of LEDGF/p75 (Figure 1.9), two structured domains can be recognized, both containing druggable pockets.

4.2.1. Targeting the PWWP domain of LEDGF inhibits clonogenic growth

Earlier studies by Méreau H. et al. [95] indicated that overexpression of the C-terminal IBD reduces the clonogenic growth of MLL-r derived cell lines *in vitro*. To investigate whether the LEDGF_{PWWP} domain induces a similar phenotype, an expression construct was generated to overexpress a flag-tagged PWWP fragment (amino acids (aa) 1 to 110) of LEDGF (LEDGF_{PWWP}) in cells. Parallel to the wild type LEDGF_{PWWP} domain, a mutant PWWP domain with either the W21A (LEDGF_{PWWP W21A}) or the F44A (LEDGF_{PWWP F44A}) mutation or both point mutations (W21A, F44A; LEDGF_{PWWP DM}), were generated. These mutations are located in the methyl lysine binding pocket of LEDGF and are known to be important for the interaction with methylated H3K36 [183]. Via viral vector transduction, all flag-LEDGF_{PWWP} constructs were stably expressed in MLL-r cell lines THP1 (MLL-AF9) and SEM (MLL-AF4), as well as the MLL wild type cell line Nalm6. Expression in all cell lines was validated using qPCR to verify the PWWP mRNA levels and western blot detection with antibodies directed against the Flag-tag to assess Flag-PWWP expression (Figure 4.11). Unfortunately, protein expression of the single mutants was much lower compared to the wild type. For both single mutants, low or no protein was detected on western blot despite the high mRNA levels, implying that this mutant affects the translation or stability of the protein. For that reason, only the W21A, F44A double mutant (DM) was used to study the importance of the PWWP domain in further experiments (LEDGF_{DM} or LEDGF_{PWWP DM}). After validation of expression, these cells were plated in semi-solid medium for a colony-forming unit (CFU) assay to assess proliferation and differentiation (Figure 4.12A). Seven days after plating these cells, the number of colonies was manually counted using a microscope. For the Nalm6, no significant difference was observed between the control condition and cells overexpressing Flag-LEDGF_{PWWP} or Flag-LEDGF_{PWWP DM}. Interestingly, in the MLL-r cells THP1 and SEM a respective reduction of 57.5% and 56.9% of colonies was detected upon Flag-LEDGF_{PWWP} expression compared to the control vector.

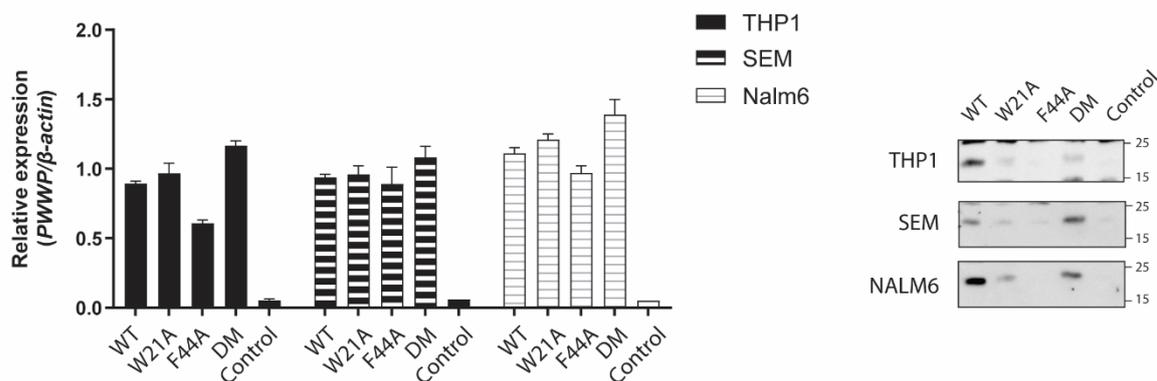


Figure 4.11: Validation of flag-LEDGF_{PWWP} overexpressing cell lines

Validation of stably transduced THP1 (MLL-AF9), SEM (MLL-AF4) and Nalm6 (MLL1 wildtype) cells with flag-LEDGF_{PWWP} wild type (WT), single (W21A or F44A) and double mutant (DM, W21A and F44A), or control. Left: relative mRNA levels of the PWWP domain in whole cell lysates, normalized to the housekeeping gene β -actin. One representative experiment is shown. Error bars indicate standard deviation (SD) between triplicates. Right: western blot detection of Flag-PWWP expression with the flag antibody. One representative experiment is shown. Molecular weight marker indicated at the right (kDa);

This result mimics the phenotype in MLL-r obtained after LEDGF/p75 knockdown [95]. The reduction in clonogenic growth was less pronounced when the Flag-LEDGF_{PWWP DM} domain was overexpressed (31.3% in THP1 and 40.6% in SEM), suggesting that the inhibition relates to the LEDGF_{PWWP} binding to methylated H3K36. To further elucidate this hypothesis, the cells were stained with a flag-specific antibody to investigate the localization of the over-expressed flag-LEDGF_{PWWP} domain (Figure 4.12B). Staining results indicate that the wild-type protein fragment is more abundant within the nucleus, stained with DAPI (Figure 4.12B, blue). On the contrary, the LEDGF_{PWWP DM} protein was localized to both nucleus and cytoplasm. This staining supports the CFU data and the theory that the Flag-LEDGF_{PWWP} protein is bound to methylated nucleosomes, occupying the binding site for endogenous LEDGF/p75.

4.2.2. Recombinant protein production for *in vitro* studies

To further study the interaction of LEDGF with nucleosomes and to screen potential interaction inhibitors, we purified recombinant proteins from *E. coli* bacterial cultures. Proteins were purified from the bacterial lysate by affinity chromatography, ion-exchange chromatography and/or size exclusion chromatography, explained in the materials and methods, chapter 3. Purity of the protein was assessed by separating the purified protein sample on an SDS-PAGE gel, followed by Coomassie staining (Figure 4.13A). Identification of the proteins was obtained by western blot with protein-specific or tag-specific antibodies (data not shown).

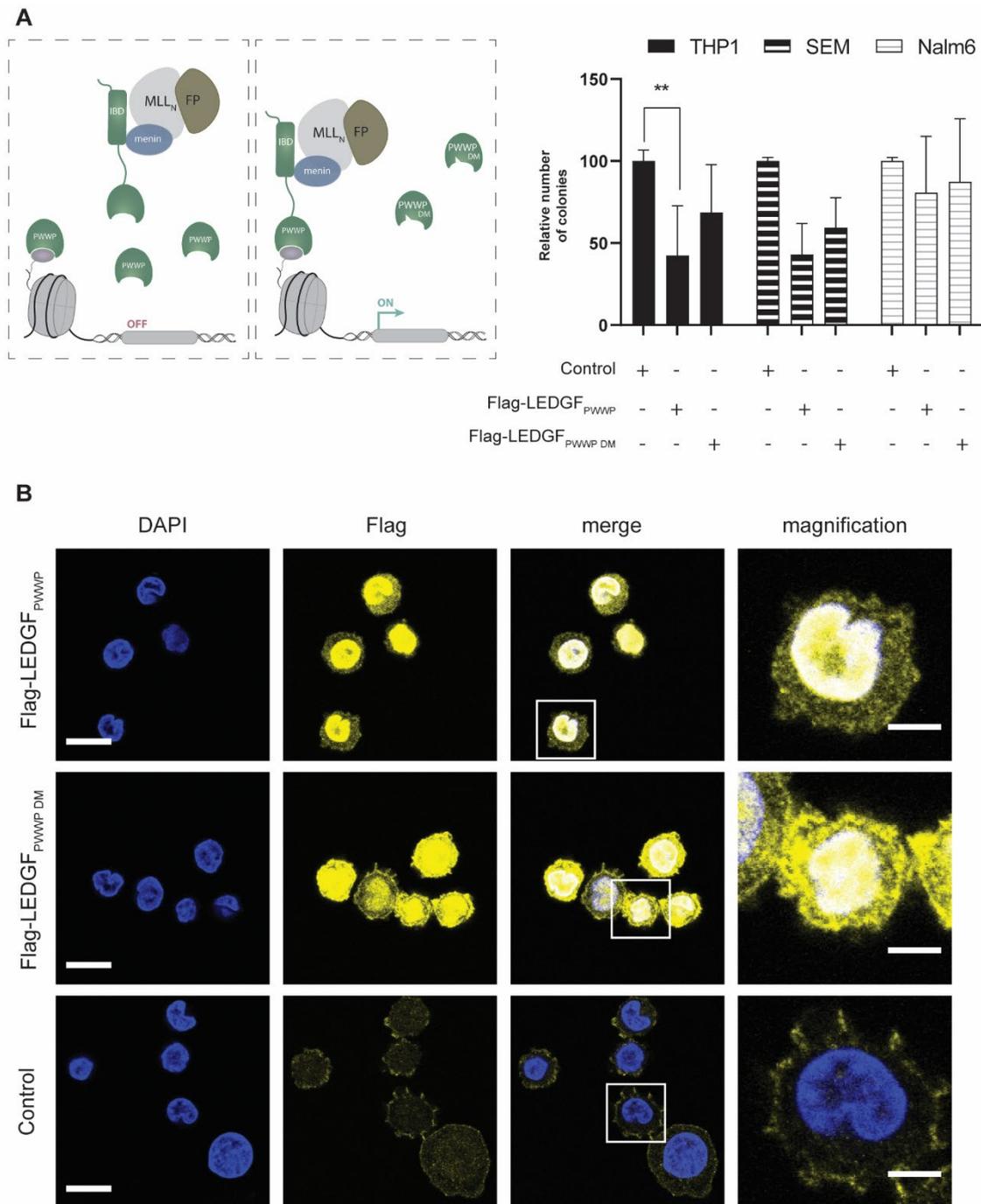


Figure 4.12: Reduced colony formation upon LEDGF_{PWWP} WT overexpression in MLL-r cell lines

(A, left) Schematic representation of experimental set-up and hypothesis. (A, right) Number of colonies relative to the control condition in colony-formation unit (CFU) assay. THP, n=4; SEM and Nalm6, n=2. Error bars indicate the SD of the independent experiments. A student's t-test was performed between every control and experimental condition, ** p<0.001. (B) Immunocytochemistry (ICC) staining of THP1 cells, stably expressing Flag-LEDGF_{PWWP}, flag-LEDGF_{PWWP DM} or control vector expressing eGFP (control). Fluorescence imaging of DNA (DAPI, blue) or flag-PWWP (primary flag antibody, secondary Alexa 555, yellow). Representative images of two independent experiments are shown. The white square in the merged images (third column) indicates the region of the magnification image (fourth column). Scale bar in individual pictures, 15 μ m. Scale bar in magnification, 5 μ m.

For structured domains, such as the PWWP domains of LEDGF, HRP-2, WHSC1_{PWWP1} and BRPF2, the thermal stability of the domains was assessed via a differential scanning fluorimetry (DSF) assay. In DSF, the proteins are gradually heated in the presence of Sypro Orange (SO) to a temperature of 95°C. The SO is quenched in the aqueous buffer conditions but when the protein undergoes denaturation, hydrophobic patches will be exposed and occupied by the dye, resulting in a fluorescent emission which is measured every 0.2°C. A fluorescent maximum is obtained when all proteins in the sample are unfolded and at higher temperatures, the fluorescent signal decreases because of protein aggregation. The melting curve of one pure, structured domain will present as a smooth line that forms one peak. From the double derivative of these data, the melting temperature (T_m) of the protein, defined as the temperature in which half of the sample is folded and half of the sample unfolded, can be determined.

For each PWWP domain, the optimal combination of protein concentration and SO was determined by a small cross-titration (Figure 4.13B) and defined by straightness of the melting curve, signal to background ratio and protein consumption. Overall, the combination of the highest protein concentration and SO results in the highest relative fluorescent units (RFU) and for both LEDGF_{PWWP} and HRP-2_{PWWP} this resulted in the best melting curves. Because of a small difference between background and peak RFU (< 500) for the HRP-2 domain at usual salt concentrations of 200 mM, we investigated the effect of different salt concentration in the buffer. Results indicate that the higher salt concentration of 600 mM (Figure 4.13B, top right, green and red lines) resulted in the best behavior of the HRP-2_{PWWP} domain during denaturation. On the contrary, unfolding of the PWWP domains of bromodomain and PHD finger-containing protein 2 (BRPF2_{PWWP}) and wolf-Hirschhorn syndrome candidate 1 protein (WHSC1_{PWWP1}) domain resulted in high RFU values and 10 μ M of protein in combination with the 20X concentration of SO was enough to obtain curves with good characteristics. The obtained melting temperatures were higher for HRP-2_{PWWP} (56.6°C) and BRPF2_{PWWP} (56.4°C) than for LEDGF_{PWWP} (43.0°C) and WHSC1_{PWWP1} (42.0°C), indicating that the different PWWP domains have distinctive characteristics.

4.2.3. AlphaScreen as a screening assay for the PWWP-nucleosome interaction *in vitro*

We used two related assay formats to study the interaction of recombinant LEDGF with nucleosomes, namely Amplified Luminescent Proximity Homogenous Assay (AlphaScreen, described below) and time resolved fluorescent resonance energy transfer (TR-FRET, section 4.2.4.).

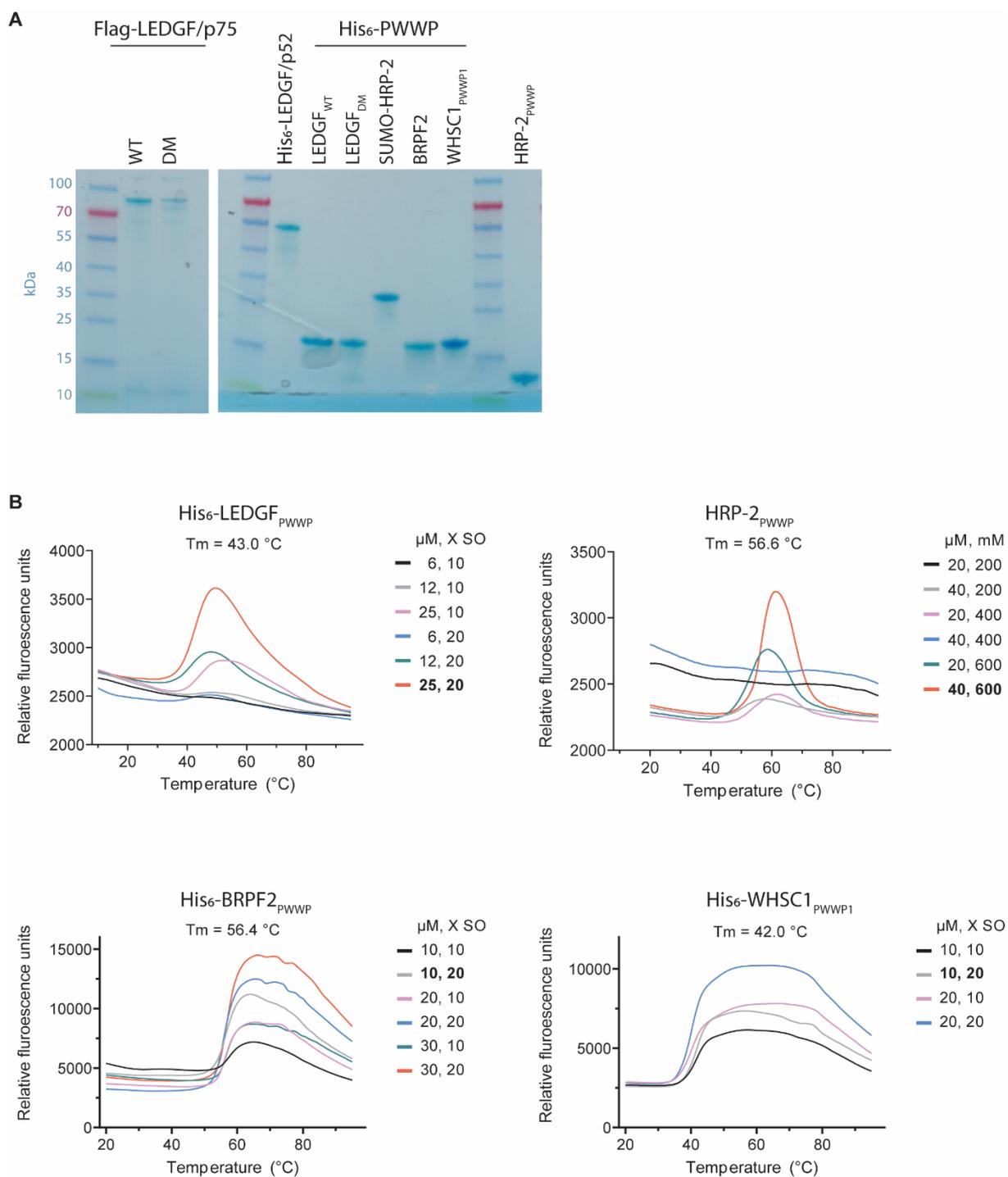


Figure 4.13: Validation and characterization of recombinant proteins used in this chapter

(A) Coomassie stained SDS-PAGE gel of recombinant proteins. Each set of proteins (Flag-tagged, His₆-tagged and untagged) is preceded by a protein marker for which the corresponding molecular weights are indicated at the left. Equal loading of 2 μg. (B) Melting curves of various PWWP domains tested in differential scanning fluorimetry. The selected combination of concentration Sypro Orange (X SO) or salt concentration (mM) and protein (μM) is indicated in bold and the corresponding melting temperature (T_m) is given on top of the graph.

In AlphaScreen, the interaction between two recombinant, tagged proteins is detected using tag-specific and light sensitive donor and acceptor beads (Figure 4.14A). After light stimulation at 680 nm of the photosensitizer in the donor bead, a singlet oxygen is generated and released. This reactive oxygen will initiate a chemical cascade in the acceptor bead, resulting in the emission by fluorophore rubrene at 520-620 nm. Proximity of donor and acceptor beads depends on the protein-protein interaction between the tagged proteins and is limited to a distance of 200 nm for energy transfer. This emission by the acceptor bead is detected and referred to as the AlphaScreen signal. AlphaScreen beads have a multivalent character, referred to as avidity (Figure 4.14A, right), which means that several target recognition sites occur on one bead

First, the optimal interaction conditions were determined by performing a cross-titration with both interaction partners. Different concentrations of biotinylated nucleosome (biot-N) with a trimethylation mark on lysine 36 of histone 3 (H3K36me3) were tested against a concentration range of either flag-tagged LEDGF/p75 (Flag-LEDGF/p75, Figure 4.14B, left) or His₆-tagged LEDGF_{PWWP} (aa 1-110, His₆-LEDGF_{PWWP}, Figure 4.14B, right). Results indicate that both LEDGF/p75 and LEDGF_{PWWP} interact in a concentration-dependent manner with the nucleosome. Interestingly, a good signal to background ratio for the Flag-LEDGF/p75 and biot-N(H3K36me3) interaction starts at lower protein concentrations and reaches a higher maximum compared to the interaction of His₆-LEDGF_{PWWP} with the nucleosomes.

The recombinant nucleosome consists of histone proteins H2A, H2B, H3.1 and H4 and is wrapped by the 147-base pair long Widom 601 core sequence which has high affinity for histone octamers [288], biotinylated at the 5' end. To exclude that the binding of flag-LEDGF/p75 and His₆-LEDGF_{PWWP} occurs solely with (free) DNA in the nucleosome batch, rather than with the trimethylated tail of the nucleosome, a Widom 601 DNA fragment was generated by PCR and tested for its interaction with both proteins (Figure 4.14C). An interaction between protein and Widom 601 DNA was observed for both flag-LEDGF/p75 and His₆-LEDGF_{PWWP}, although at much lower AlphaScreen counts compared with the complete nucleosomes. Based on these results, the interaction optimum between 5 nM of biot-N(H3K36me3) and 10 nM of protein (either Flag-LEDGF/p75 or His₆-LEDGF_{PWWP}) was selected for fragment screening (section 4.2.9.). To investigate whether both Flag-LEDGF/p75 and His₆-LEDGF_{PWWP} interact with the recombinant nucleosome in a similar way, we performed reciprocal out-competition assays. In this assay, the concentration of both the nucleosome (4 nM) and His₆-LEDGF_{PWWP} (30 nM) were fixed and incubated with different concentrations of Flag-LEDGF/p75 (Figure 4.15A). Detection of this interaction between nucleosome and His₆-LEDGF_{PWWP} decreased with increasing concentration of Flag-LEDGF/p75, indicating that an excess of Flag-LEDGF/p75 is able to displace the His₆-LEDGF_{PWWP} protein from the interaction with the nucleosome and that the interaction sites likely overlap.

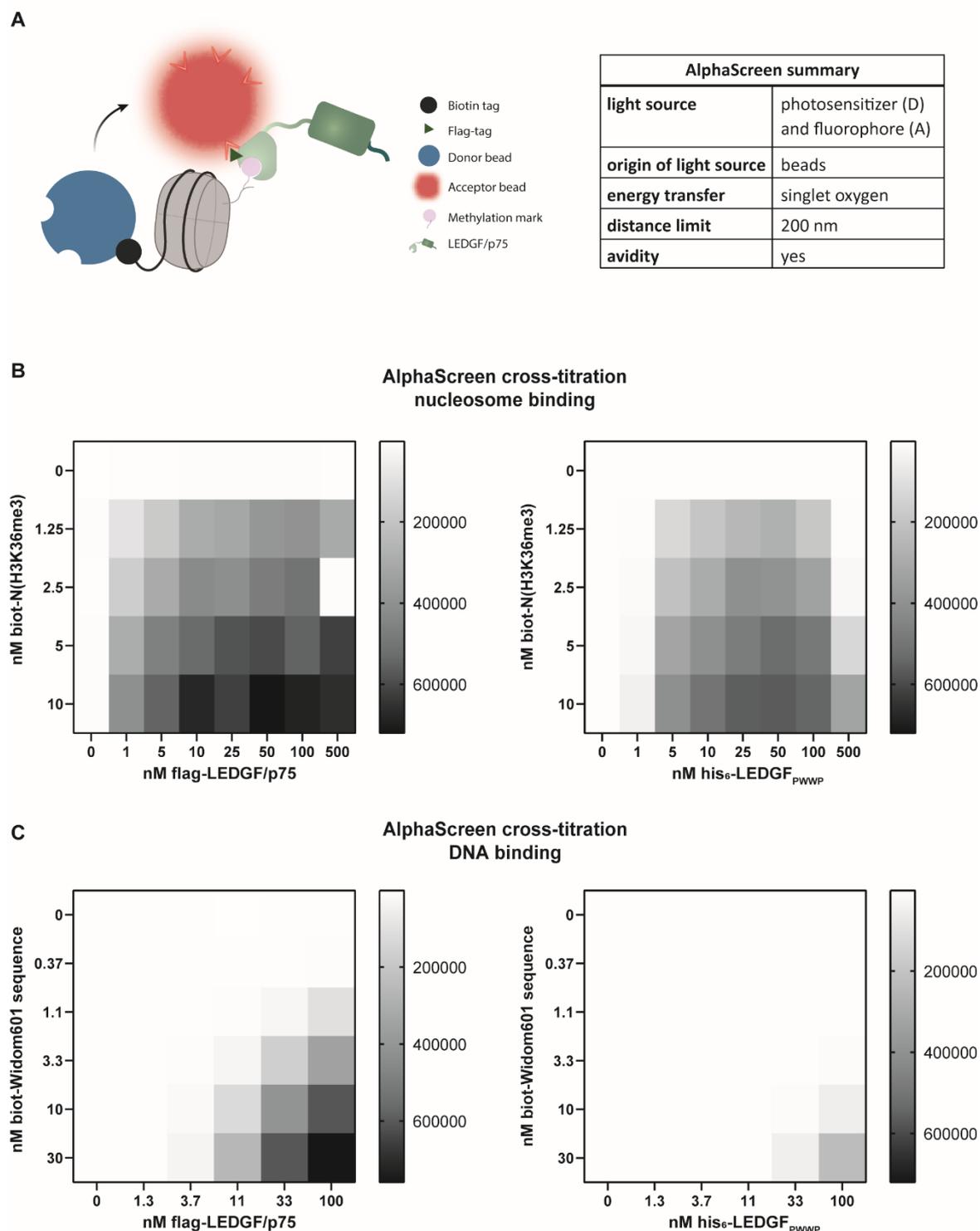


Figure 4.14: AlphaScreen interaction between flag-LEDGF/p75 and his₆-LEDGF_{PWWP} and recombinant biotinylated H3K36 trimethylated nucleosome

(A) Schematic representation of the AlphaScreen interaction between trimethylated nucleosome and Flag-LEDGF/p75 (left). Summary of assay characteristics (right) with D = donor and A = acceptor. (B) Heat map of the cross-titration between recombinant, biotinylated nucleosomes trimethylated on lysine 36 of histone 3 (biot-N(H3K36me₃)) and the flag-tagged LEDGF/p75 (left) or the His₆-tagged LEDGF_{PWWP} domain (right). (C) Heat map of the cross-titration between the biotinylated Widom 601 sequence and flag-tagged full length LEDGF/p75 (left) or the His₆-tagged LEDGF_{PWWP} domain (right).

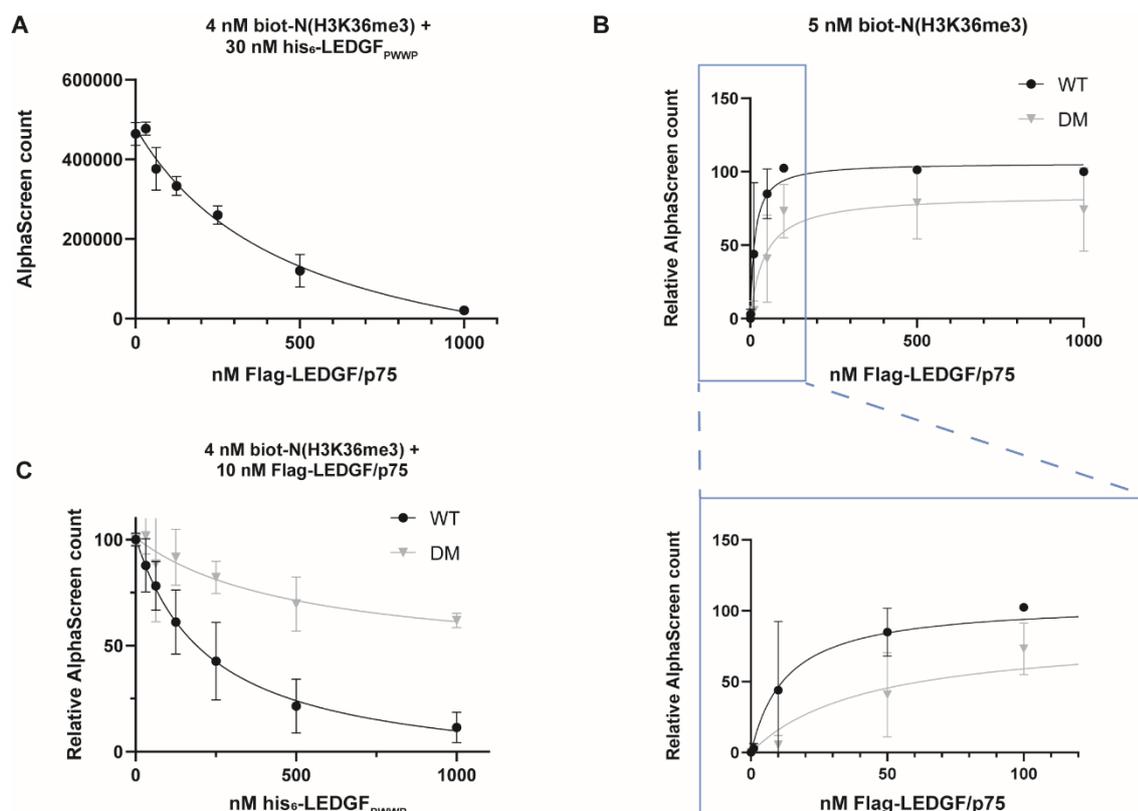


Figure 4.15: Amino residues W21 and F44 of the PWWP domain are important for the nucleosome interaction

(A) Comparison of the direct interaction of wild type or mutant His₆-LEDGF_{PWWP} with biotinylated H3K36me3 nucleosomes. Error bars represent SD of duplicate measurements; (B) Interaction between His₆-LEDGF_{PWWP} or His₆-LEDGF_{PWWP}^{DM} (W21A, F44a) and biotinylated nucleosomes with H3K36me3 mark. The blue box is a magnification of the lower concentrations. (C) Out competition with His₆-LEDGF_{PWWP} wild type (WT) and double mutant (DM) of the Flag-LEDGF/p75-nucleosome interaction (H3K36me3).

To further investigate whether the nucleosome interaction is based on the LEDGF_{PWWP} domain, the PWWP double mutant (DM) with the W21A and F44A point mutations that binds with a lower affinity to nucleosomes was expressed and purified as recombinant protein in both the full length (flag-LEDGF/p75_{DM}) and the PWWP domain (His₆-LEDGF_{PWWP}^{DM}) constructs. Titration of flag-LEDGF/p75 and flag-LEDGF/p75_{DM} protein against a fixed concentration of biot-N(H3K36me3) indicates that the flag-LEDGF/p75_{DM} is able to bind to the nucleosome to a reduced extent (Figure 4.15B). This suggests that the nucleosome-LEDGF/p75 interaction is not solely dependent on the PWWP domain. However, the His₆-LEDGF_{PWWP}^{DM} protein was not able to outcompete the nucleosome-LEDGF/p75 interaction to a similar extent as the wild type protein (Figure 4.15C). These results indicate that the LEDGF_{PWWP} domain is sufficient for the interaction with the recombinant nucleosome.

Next, we assessed the specificity of the trimethylation mark for the nucleosome-LEDGF interaction *in vitro*. Binding of His₆-LEDGF_{PWWP} was tested in parallel to biotinylated nucleosomes with either a mono-methylation (me1), di-methylation (me2) or trimethylation (me3) or without methylation mark on lysine 36 of histone three (H3K36) (Figure 4.16). In line with previous findings [289], His₆-LEDGF_{PWWP} prefers the trimethylated nucleosomes (H3K36me3, Figure 4.16, blue), but also shows a clear binding to the di-methylated nucleosomes (H3K36me2, dark gray). For both non-methylated and monomethylated nucleosomes, the counts were weaker, but still detectable *in vitro*.

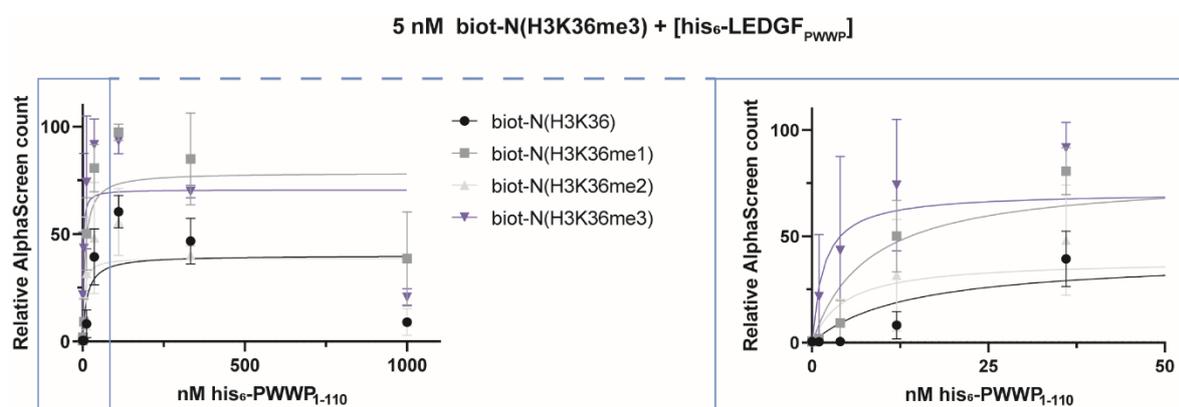


Figure 4.16: LEDGF_{PWWP} interacts with all methylation states of the H3K36 recombinant nucleosomes *in vitro*

Titration of His₆-LEDGF_{PWWP} against a fixed (5 nM) concentration of biotinylated nucleosome with different methylation states. All methylation states are positioned at lysine 36 on histone 3 (H3K36): not methylated, black; me1, monomethylated, dark grey; me2, demethylated, light gray; me3, trimethylated, blue. Error bars indicate the standard deviation of two independent experiments (n=2). The blue box is a magnification of the His₆-LEDGF_{PWWP} concentrations below 50 nM.

In light of the strong similarity between the LEDGF_{PWWP} and HRP-2_{PWWP} domains, we investigated whether HRP-2_{PWWP} also interacts with the recombinant nucleosomes *in vitro*. To study this, the His₆-SUMO-HRP-2 protein fragment (aa 1 to 93) was purified and tested in an AlphaScreen (Figure 4.17). The HRP-2_{PWWP} domain was sumoylated to facilitate the purification process. Different concentrations of biot-N(H3K36me3) and His₆-SUMO-HRP-2_{PWWP} were titrated against each other resulting in an increasing signal with increasing concentration (Figure 4.17, left). Interestingly, His₆-LEDGF_{PWWP} was replaced in the interaction with nucleosomes by increasing concentration of the HRP-2_{PWWP} domain, which suggests that both PWWP domains bind at the same site on the nucleosome and binding of both PWWP domains is mutually exclusive.

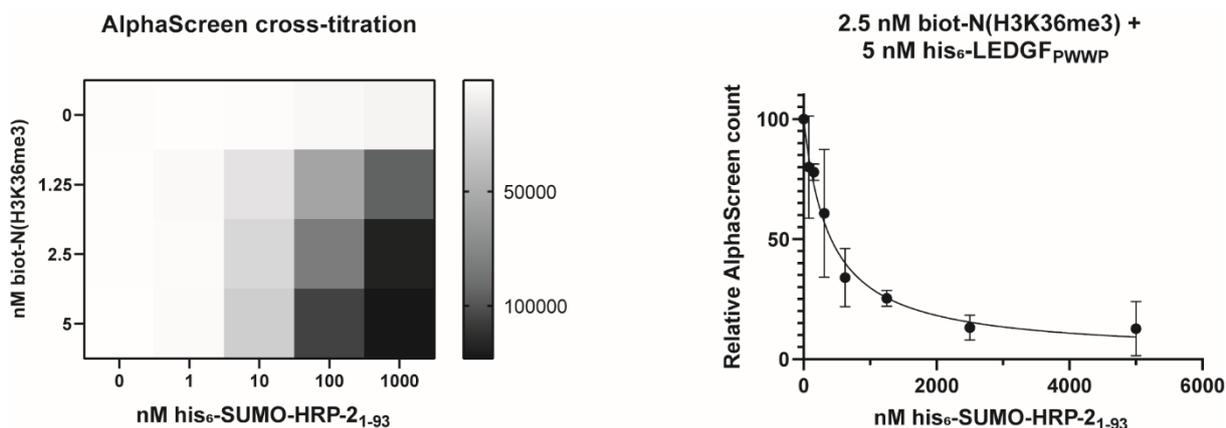


Figure 4.17: HRP-2_{PWWP} interacts with trimethylated recombinant nucleosomes *in vitro*

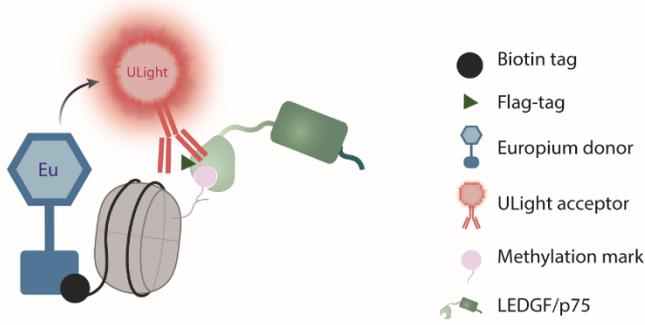
(A) AlphaScreen cross titration between biotinylated, recombinant nucleosomes with the trimethylation mark on lysine 36 of histone 3 (biot-N(H3K36me3)) and the HRP-2_{PWWP} domain (aa 1-93). (B) Out competition of the interaction between the biot-N(H3K36me3) nucleosome and the LEDGF_{PWWP} domain by the HRP-2_{PWWP} domain. N=2, error bars indicate SD of two independent experiments.

4.2.4. TR-FRET as *in vitro* screening assay for the PWWP-nucleosome interaction

Next to AlphaScreen, a TR-FRET assay was optimized. In TR-FRET, tags are recognized by antibodies or conjugates bound to fluorophores instead of beads (Figure 4.18A). While multiple tags can interact with one AlphaScreen bead, in TR-FRET one fluorophore binds to one single protein. The maximum distance for transfer of energy between donor and acceptor in TR-FRET is 9 nm in contrast to a distance of 200 nm in AlphaScreen. This makes TR-FRET less sensitive compared to AlphaScreen. According to previously published data, assembled nucleosomes have a spherical form with a diameter of ~10 nm. To identify whether the TR-FRET is useful to study the interaction between nucleosomes and LEDGF, different cross titration experiments were performed (Figure 4.18B).

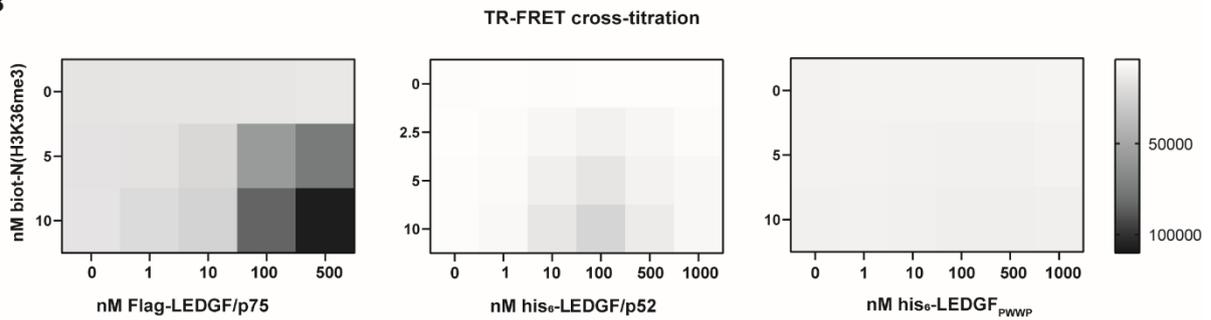
The assay was validated by titrating increasing concentrations of H3K36 trimethylated nucleosomes with Flag-LEDGF/p75 (Figure 4.18B, left), which resulted in a concentration dependent interaction. Similarly, the shorter isoform of LEDGF, p52, was able to interact with the nucleosomes in a concentration-dependent manner (Figure 4.18B, middle). Unfortunately, no interaction was detected between nucleosomes and the His₆-LEDGF_{PWWP} domain (Figure 4.18B, right). The importance of the PWWP domain in the interaction was evaluated by using the mutant Flag-LEDGF/p75_{DM} protein (Figure 4.18C). For the wild type LEDGF_{PWWP} protein domain the TR-FRET signal was saturated from the concentration of 100 nM on and higher. For the Flag-LEDGF/p75_{DM} mutant a similar interaction curve was observed, with saturation from the same concentration on. Yet, for the corresponding concentrations, the TR-FRET signal went up to only 50% of the wild type protein signal.

A

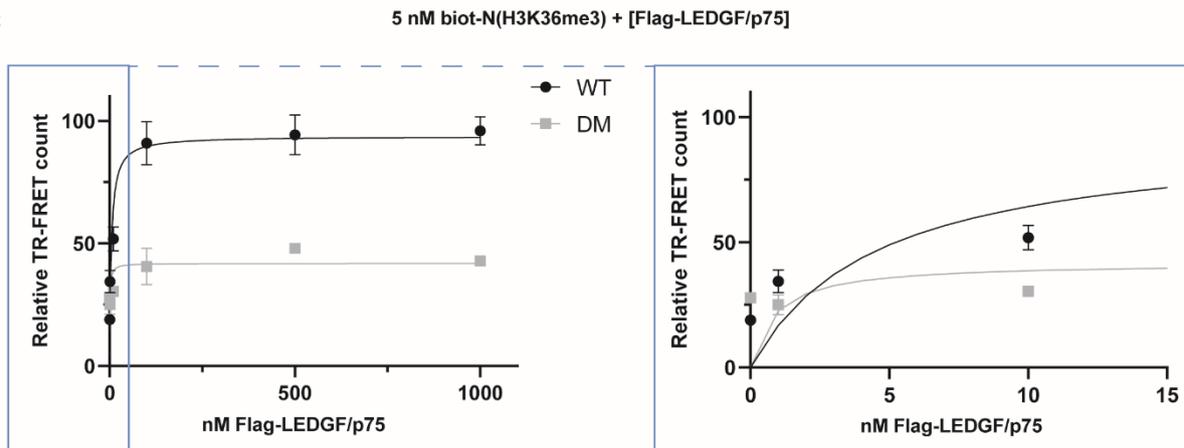


TR-FRET summary	
light source	fluorophores
origin of light source	anti-tag antibody or other conjugate
energy transfer	fluorescence
distance limit	9 nm
avidity	only streptavidin-Eu

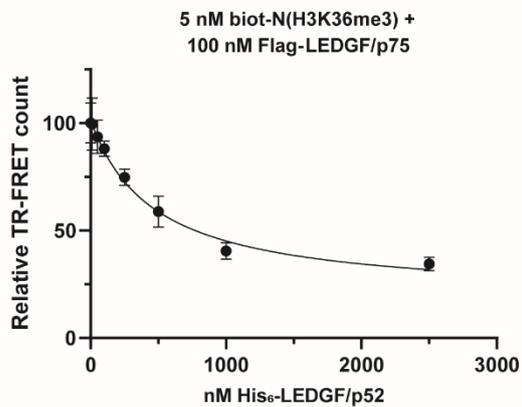
B



C



D



E

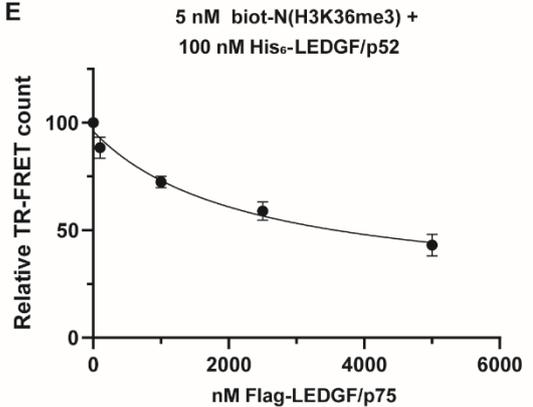


Figure 4.18: Amino residues W21 and F44 of the PWWP domain are important for the nucleosome interaction in TR-FRET

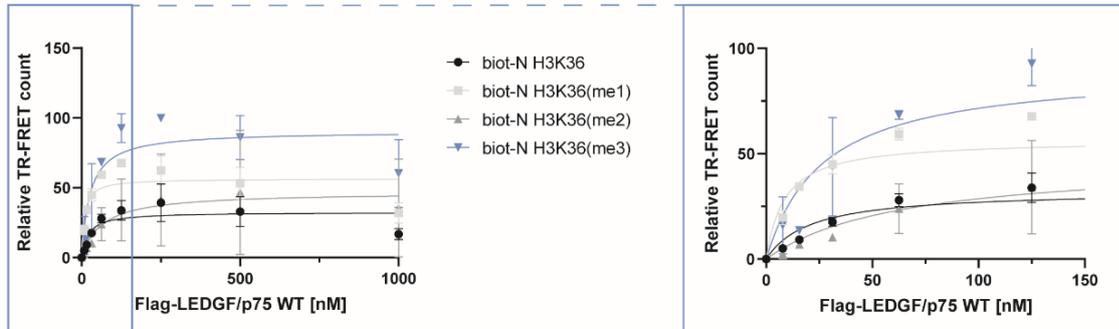
(A) Cartoon of the TR-FRET interaction between trimethylated nucleosomes with biotin tag attached to the DNA, recognized by Europium (Eu) labelled streptavidin (pink square) and recombinant Flag-LEDGF/p75, recognized by an anti-flag antibody, labelled with an ULight acceptor (left). Summary of the assay characteristics on the right. D = donor; A = acceptor (B) Different cross titrations between biotinylated nucleosomes with trimethylated lysine 36 on histone three and full length LEDGF/p75 (left), p52 (middle) or LEDGF_{PWWP} domain (right) were performed. One representative experiment is shown for each condition; (C) Comparison of the direct interaction of wild type or double mutant Flag-LEDGF/p75 to biotinylated H3K36me3 nucleosomes. Error bars represent the SD of independent experiment. The blue box is a magnification of the lower concentrations; (D) Out competition of the interaction between 5 nM of biotinylated nucleosome, tri-methylated on H3K36 and 100 nM of Flag-LEDGF/p75 by titration of His₆-LEDGF/p52. Error bars represent the SD of independent experiment. (E) Out competition of the interaction between 5 nM of biotinylated nucleosome, tri-methylated on H3K36 and 100 nM of His₆-LEDGF/p52 by flag-LEDGF/p75.

This supports the previous findings that point mutations W21A and F44A in the PWWP domain reduce binding to the nucleosomes [183]. To validate this assay for the detection of inhibitors, we performed out-competition assays with LEDGF/p52. His₆-LEDGF/p52 was titrated to a fixed concentration of 5 nM biot-N(H3K36me3) and 100 nM Flag-LEDGF/p75 (Figure 4.18D). A decrease in TR-FRET signal was observed upon increasing concentrations of His₆-LEDGF/p52. *Vice versa*, when titrating flag-LEDGF/p75 at a fixed concentration of 5 nM biot-N(H3K36me3) and 100 nM His₆-LEDGF/p52 (Figure 4.18E), the interaction was reduced again with increasing concentration of LEDGF/p75, supporting the feasibility to identify inhibitors of the interaction between nucleosome and LEDGF. In both out-competition experiments, the signal did not drop below 34.4%, which corresponds to a TR-FRET signal of around 6000 counts. Of note, similar TR-FRET counts were obtained in the cross-titration background of the Flag-LEDGF/p75 and biot-N(H3K36me3) interaction in Figure 4.18B.

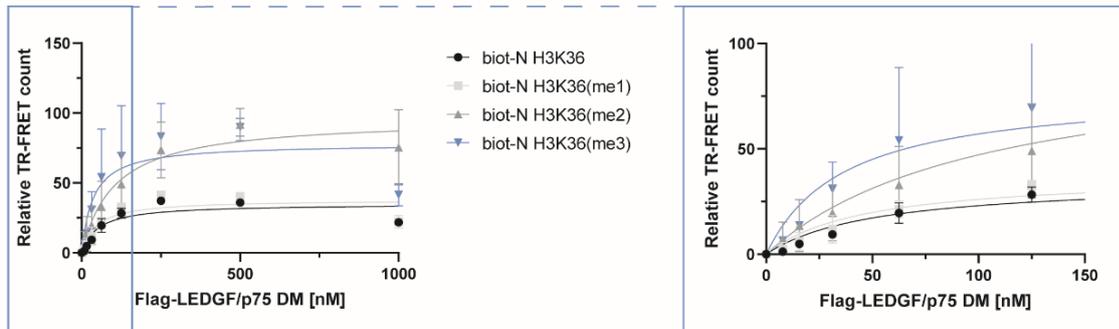
As our previous experiments indicate that the LEDGF/p75-nucleosome interaction can be measured by TR-FRET, we next assessed the binding of both flag-LEDGF/p75 WT and DM to the different methylation states of H3K36 (Figure 4.19). Results suggest LEDGF/p75 prefers binding to H3K36 di- and trimethylated nucleosomes (Figure 4.19A), correlating with the results obtained for the His₆-LEDGF_{PWWP} domain in AlphaScreen (Figure 4.16) and previously published data [183]. An interaction to non- and monomethylated nucleosomes was detected at much lower counts. Interestingly, the flag-LEDGF/p75_{DM} showed a similar interaction pattern, demonstrating the best binding to di- and trimethylated nucleosomes (Figure 4.19B). Since the experiments with wild type and double mutant were run in parallel, binding of both LEDGF/p75 proteins to each methylation state could be compared (Figure 4.19C). Results indicate that the difference between WT and DM decreases when the methylation state is reduced. For the di-methylated nucleosomes, a high variability between two assays was obtained. A third experiment confirmed higher AlphaScreen counts for the LEDGF/p75 WT interaction to H3K36me2 than the LEDGF/p75 DM, however due to a different experimental setup

this result could not be integrated. Overall, the data suggest that the level of methylation has more impact on the interaction between nucleosomes and wild type LEDGF/p75 than for LEDGF/p75_{DM}.

A LEDGF/p75 WT



B LEDGF/p75 DM



C

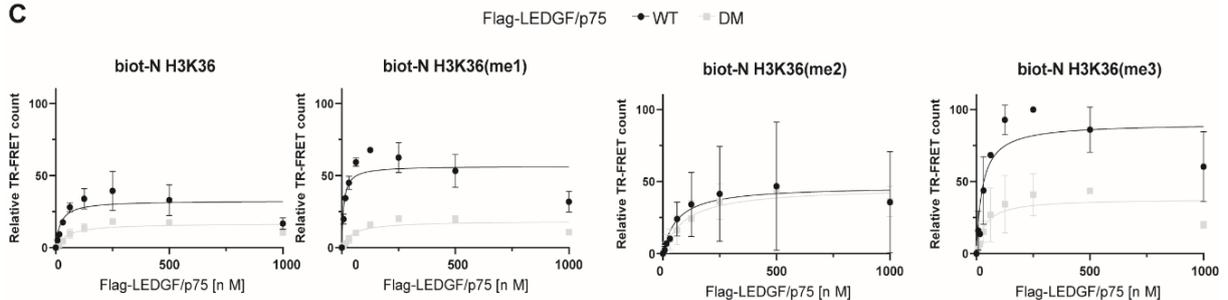


Figure 4.19: The binding of LEDGF/p75_{DM} to nucleosomes is less dependent on the methylation state of H3K36

(A,B) Titration of WT (A) or DM (W21A, F44A) LEDGF/p75 (B) at a fixed concentration of 5 nM biotinylated nucleosomes with variable levels of methylation; non-methylated, black; monomethylated (me1), light gray; di-methylated (me2), dark gray; trimethylated (me3) blue. Counts are relative to the highest TR-FRET count detected in the corresponding experiment. (C) Other representation of the data in (B). Comparison of the binding of WT or LEDGF/p75 to different methylation states of biot-N H3K36 nucleosomes. In each experiment, the WT and W21A, F44A flag-LEDGF titration was performed in parallel. Counts are relative to the highest TR-FRET count detected for the WT in the corresponding experiment. (B,C) Error bars indicate standard deviation (n=2).

4.2.5. Detection of nucleosome-LEDGF/p75 interaction in cells using nanoBRET

Although these *in vitro* data clearly show the interaction between LEDGF and H3K36me3 nucleosomes, a nano-bioluminescent resonant energy transfer assay (nanoBRET) was established to study the interaction in a cellular context [290]. In this assay, protein interactions are detected using a nanoLuciferase (nanoLuc) as bioluminescent donor and a fluorophore labelled HaloTag as energy acceptor. In brief, HEK293T cells were transfected with two plasmids, one encoding histone 3 (H3) tagged with a HaloTag, while the second plasmid encodes LEDGF/p75 or LEDGF_{PWWP} fused to nanoLuc. After protein expression, the HaloTag fluorophore (product Promega) is added to the cellular medium and covalently bound to the HaloTag. To determine whether an interaction occurs, the luciferase substrate, furimazine, is added to the well and activated by nanoLuc. The converted furimamide serves as a light donor to the HaloTag fluorophore within a range of 9 nm [291]. To analyze the data, emission by nanoLuc (input) is divided by emission of the HaloTag fluorophore (output) and corrected for background signal (see chapter material and methods). Results are expressed in milliBret Units (mBU).

For our purpose we used the expression construct in which the Halo-Tag was fused to the C-terminus of histone H3 (Figure 4.20A, design by Promega). To optimize the detection of the histone-LEDGF interaction, the optimal position of the nanoLuc on LEDGF/p75 had to be assessed (Figure 4.20B). Our results indicate that the position of nanoLuc at either the N- or C- terminus does not significantly affect the detection of the interaction, as both LEDGF/p75 constructs resulted in 7 and 5 mBU values, respectively (Figure 4.20B, left). However, to be consistent with the *in vitro* assays, the N-terminally tagged construct was selected for further use. Next, the optimal ratio of acceptor to donor was determined by titrating the nanoLuc donor construct (Figure 4.20B, middle). The optimal ratio, resulting in the highest nanoBRET units, was obtained at a 100-fold excess of donor (NanoLuc-LEDGF/p75) to acceptor (H3-HaloTag). The interaction was abolished when full length flag-LEDGF/p75 or the PWWP domain alone (flag-PWWP WT) was overexpressed to compete for the interaction with HaloTag-H3 (Figure 4.20B, right). Remarkably, overexpression of the LEDGF_{PWWP DM} domain also interfered with the nanoBRET signal to a similar extent. This result is consistent with the AlphaScreen data where the LEDGF_{DM} still binds nucleosomes and the previous finding of a bivalent interaction surface [186].

In addition, similar optimization steps were performed for the interaction between HaloTag-H3 and nanoLuc-LEDGF_{PWWP} (Figure 4.20C). Although the C-terminal position of nanoLuc seemed favorable in the initial experiment (Figure 4.20C, left), mBU values of 15 were obtained after optimizing the donor to acceptor ratio with nanoLuc at the N-terminal position (Figure 4.20C, middle). In accordance to the set-up with full length LEDGF/p75 and recombinant proteins for *in vitro* assays, the N-terminal

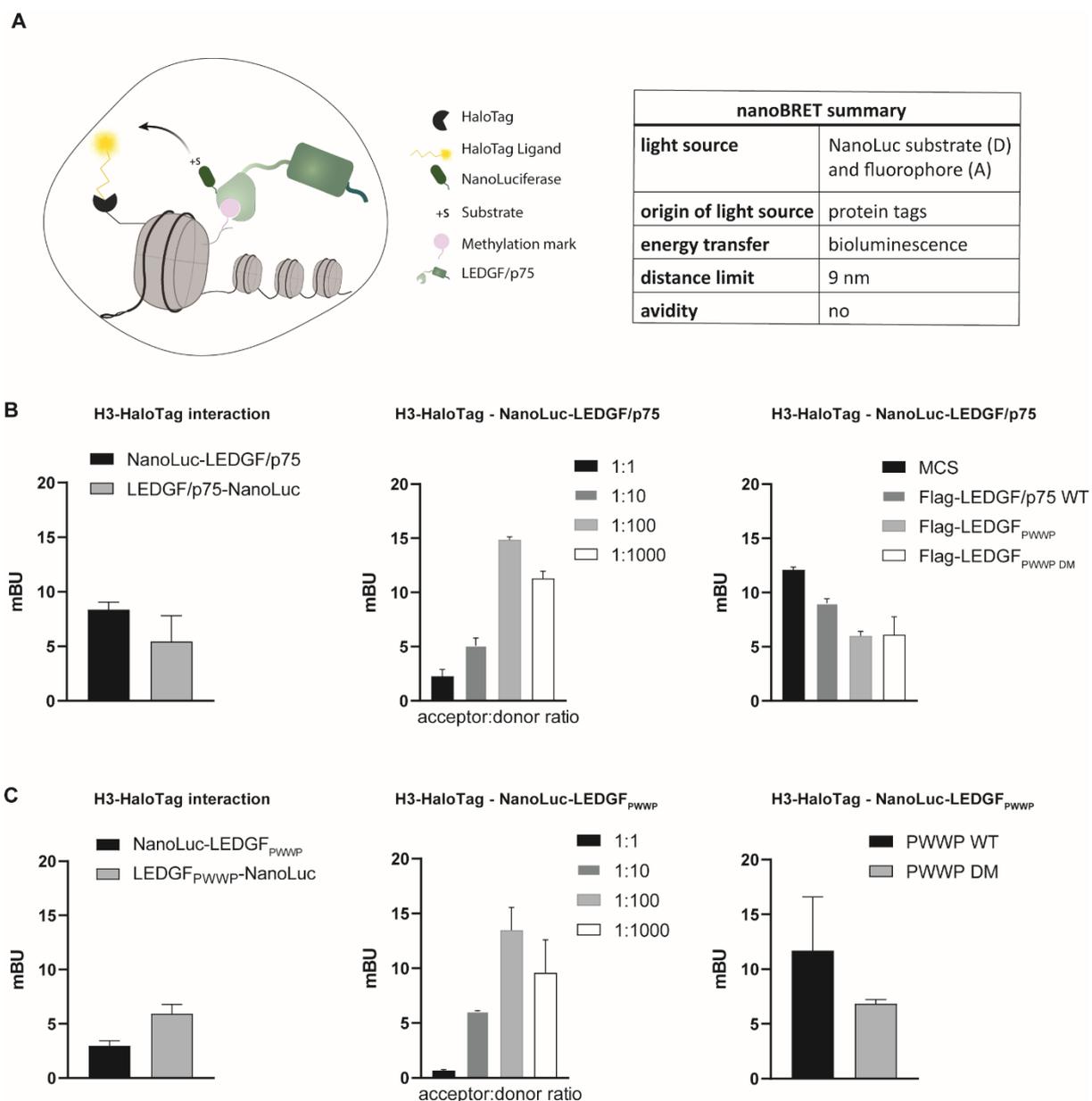


Figure 4.20: Optimization of the nanoBRET assay

(A) Representation of a nanoBRET interaction between a histone 3 (H3) HaloTag-containing nucleosome and NanoLuciferase (nanoLuc)-LEDGF/p75 in addition of the luciferase substrate (S). Summary of the assay characteristics on the right; (B) NanoBRET assays to compare tag position of nanoLuciferase (NanoLuc) to LEDGF/p75 (left) and to determine optimal acceptor to donor ratio (middle). Error bars indicate the standard deviation of three replicates ($n=1$). Right; inhibition of the nanoBRET signal by overexpression of flag-tagged full length LEDGF/p75, LEDGF_{PWWP} WT or LEDGF_{PWWP} DM. Error bars indicate the standard deviation of three replicates ($n=1$); (C) NanoBRET assay to compare tag position of nanoLuciferase to LEDGF_{PWWP} (left) and to determine optimal acceptor to donor ratio (middle). Error bars indicate the standard deviation of three replicates ($n=1$). Right; comparison of the interaction between WT or DM NanoLuc-LEDGF_{PWWP} with HaloTag-Histone H3 in nanoBRET. Error bars indicate the standard deviation of two independent experiments ($n=2$).

construct was used for further experiments. The LEDGF_{PWWP DM} bound less to nucleosomes when compared to wild type LEDGF_{PWWP} in *in vitro* assays. To verify reduced interaction between the mutant and nucleosomes in a cellular context, the LEDGF_{PWWP DM} domain was cloned into the nanoLuc expression construct for nanoBRET. As could be suspected, a decrease in mBU signal was observed for the nanoLuc-LEDGF_{PWWP DM} (Figure 4.20C, right). These data indicate that the LEDGF interaction with histone H3 and presumably the nucleosome can be detected in nanoBRET and that this assay can be used for further studies of the interaction in a cellular context.

4.2.6. Structure based drug discovery of ligands of the HRP-2 PWWP domain

There is a growing interest in epigenetic drugs, small molecule drugs that interfere with the epigenetic regulation of gene expression. BRD4 antagonists have paved the way for this research field [39]. LEDGF/p75 and HRP-2 are epigenetic readers recognizing H3K36me2/3. LEDGF/p75 has been associated with HIV replication and our previous results suggest that the interaction between LEDGF/p75 and nucleosomes is essential for the development of mixed lineage leukemia [94], [95], [179]. In collaboration with the Laboratory of Biocrystallography at KU Leuven (Prof. S. Strelkov) and the NMR expertise from the Laboratory of Structural Biology at the IOCB in Prague (Prof. V. Veverka), we aim to develop small molecules interfering with the interaction between the LEDGF_{PWWP} domain and the nucleosome.

Earlier in this chapter, we manifested the interaction between LEDGF/p75 and H3K36me3 modified nucleosomes in three independent assays. The out-competition experiments provide proof of concept that inhibitors of the interaction can be identified, pending sufficient affinity to overcome the binding energy of the protein-protein interaction (PPI). PPIs that are defined and limited in interaction surface, such as the interaction between HIV-1 integrase and the IBD of LEDGF/p75, are best suited for drug discovery. The development of small molecule inhibitors of the PWWP domain NSD3 and NSD2 [269], [271] corroborates the feasibility of our approach to identify small molecule inhibitors of LEDGF_{PWWP} and/or HRP-2_{PWWP}.

A fragment and structure-based drug design approach was initiated using a selected set of fragments based on the fragLites library described in [292]. Fragments are distinguished from compounds by a molecular weight below 300 Da, a logP value below three and a maximum of three hydrogen bond donors and acceptors. The fragLites are drug-like chemical fragments used for bottom-up optimization to develop a lead compound. In a first step, the fragments were soaked to the HRP-2_{PWWP} domain at the Laboratory of Biocrystallography (KU Leuven). Of note, our collaborators successfully crystallized the LEDGF_{PWWP} domain, however, due to occupancy of the aromatic cage of LEDGF_{PWWP} by a second,

adjacent molecule in the crystal, it was impossible to use the LEDGF_{PWWP} domain for soaking fragments or compounds. Interestingly, this issue did not arise when crystallizing the HRP-2_{PWWP} domain (Figure 4.21A). Yet both HRP-2_{PWWP} and LEDGF_{PWWP} share a similarity percentage of 87.1% in protein sequence between the first 93 amino acids (Figure 4.21B) and show highly identical pocket formation. Therefore, the HRP-2_{PWWP} domain was used to soak fragments and identify binders.

After soaking, binding fragments presented with a clear structure density and positioned in the HRP-2_{PWWP} pocket. Confirmed hits were used to virtually fine tune the occupancy of the fragment in the crystal structure and to guide medicinal chemistry to synthesize improved and extended ligands (Figure 4.22). To link structure with biological activity, the six fragments of the initial FragLite screen with best pocket occupancy were selected for further characterization in *in vitro* assays (sections 4.2.7 – 4.2.9.). These six fragments belong to four different structural scaffolds, that will not be shown for confidentiality reasons (Figure 4.22, bottom).

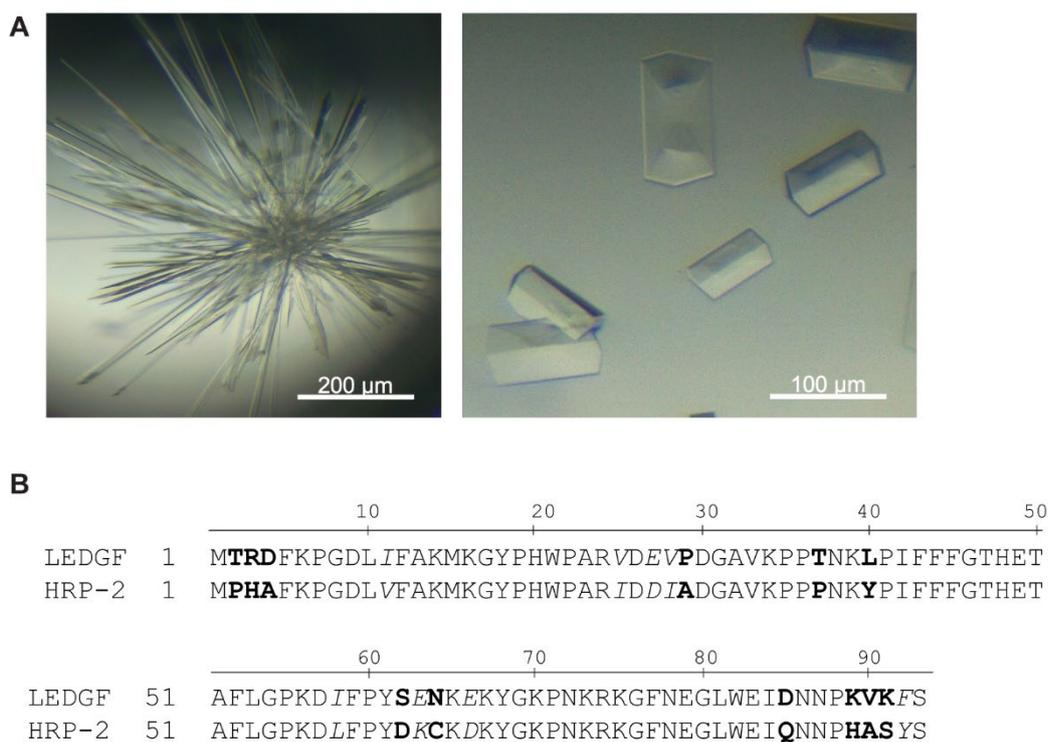


Figure 4.21: Crystals of the HRP-2_{PWWP} domain

(A) Two crystal forms obtained under different crystallization conditions. Left: spike crystals resulting in an HRP-2_{PWWP} structure with a resolution of 1.7 Å, used for in-house screening of fragments. Right: prism crystals resulting in an HRP-2_{PWWP} structure with a resolution of 1.9 Å. Due to a higher stability, these crystal forms and conditions were used at the XChem facility at Diamond Light Source, Oxford. Images courtesy of the Laboratory of Biocrystallography at the KU Leuven (Dr. S. Strelkov). (B) Protein amino acid sequence alignment of amino acids 1 to 93 of LEDGF and HRP-2. Bold: different, non-similar amino acids; Italic = different but similar amino acid, analyzed by SnapGene[®] v5.2.

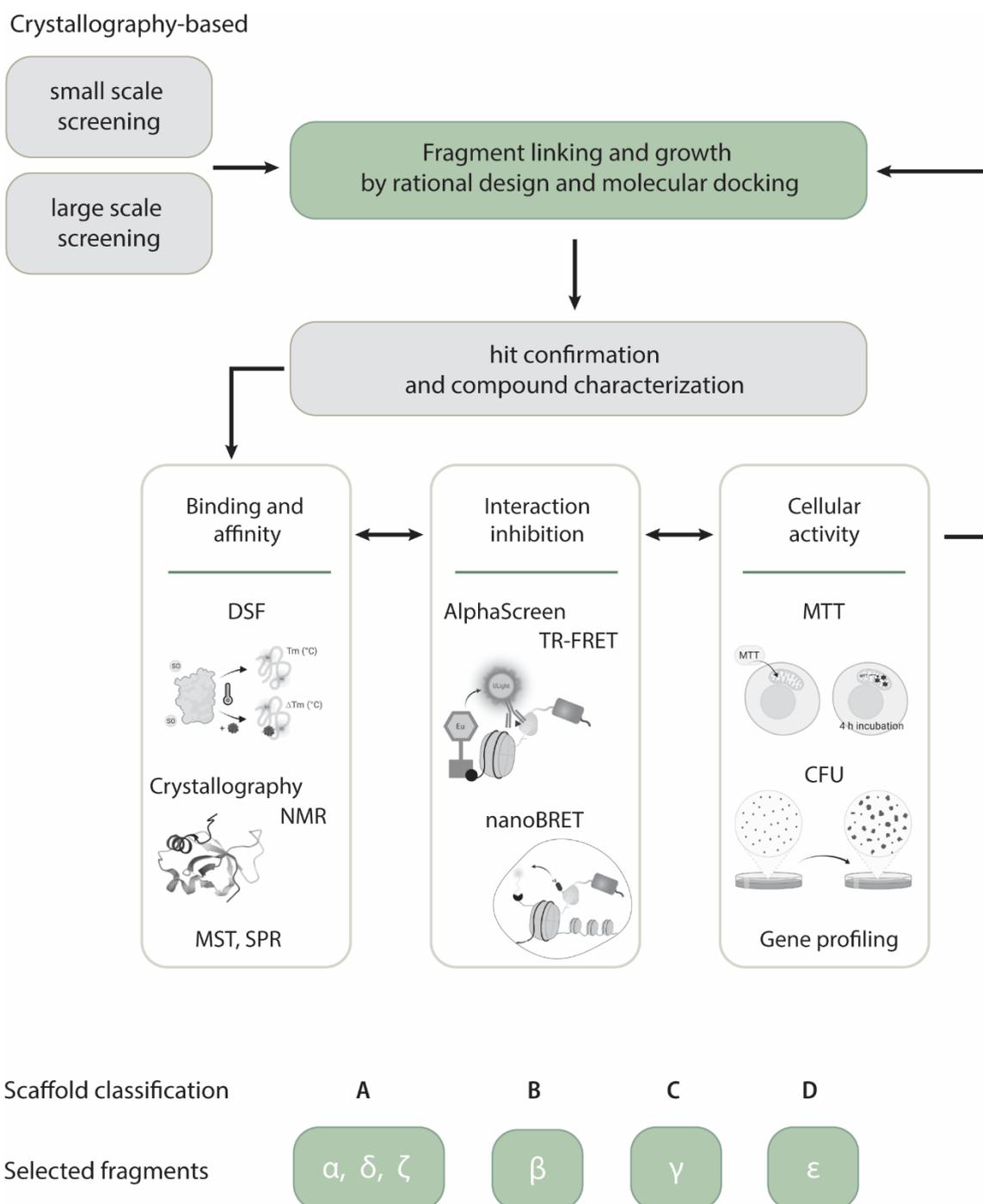


Figure 4.22: Schematic representation of the followed screening strategy

Initially, a crystallography-based screening was performed at the Laboratory of Biocrystallography (KU Leuven) and the XChem facility (Diamond, Oxford). Crystal structures provided information on fragment composition and was used to further grow the fragments. Hits were validated and further analyzed in an additional set of experiments. Binding and/or affinity of the fragments can be assessed in crystallography, differential scanning fluorimetry (DSF), Nuclear Magnetic Resonance (NMR), microscale thermophoresis (MST) and Surface Plasmon Resonance (SPR). Interaction inhibition potency will be tested in AlphaScreen and TR-FRET using recombinant proteins or nanoBRET in cellular context. The effect on cellular viability will be tested for the best fragments in an MTT experiment. Leukemic specific effect can be assessed using colony formation assay (CFU) and gene profiling upon treatment with a fragment. Bottom: classification by scaffold (1 to 4) of selected fragments tested in this manuscript.

These promising results using the initial fragment hits gave the impetus to screen a fragment library of 558 fragments at the X-Chem facility at Diamond Light Source, Oxford [293]. Overall, around 70 fragments bound to the HRP-2_{PWWP} domain. However, these fragments were not available within the timeframe of this study and will not be reported here. Our strategy to further profile and optimize the fragments into a lead compound is shown in Figure 4.22. Primary hits will be validated and used for design of new, growing fragments by rational design and molecular docking. Newly synthesized analogues will be soaked to the HRP-2_{PWWP} domain to obtain structural information, affinity will be evaluated in NMR and the analogues will be ranked for temperature shift in DSF. Results obtained from these assays will be used to further optimize fragments. This iterative process will sustain the development of more potent inhibitors. In second instance, promising fragments will further be characterized in AlphaScreen and TR-FRET.

4.2.7. Fragment hits bind to LEDGF_{PWWP} and HRP-2_{PWWP} in a dose-dependent manner

First, the six hits from the screen were tested for binding to the PWWP domain of LEDGF in the optimal conditions of the DSF assay, determined in section 4.2.2. (page 66). Binding of a fragment to the LEDGF_{PWWP} domain is indicated by a shift in the melting temperature of the protein domain compared to the protein domain on its own.

To better identify binders, different concentrations of fragments were incubated with the protein before heating and melting temperatures were assessed (Figure 4.23A). Interestingly, at a concentration of 500 μM , three fragments showed a positive temperature shift of at least 0.6°C. Of note, **fragment γ** showed the highest temperature shift of $1.23 \pm 0.36^\circ\text{C}$ at 150 μM and a clear dose-dependent increase in the shift up to 500 μM ($2.03 \pm 0.08^\circ\text{C}$) (Figure 4.23A, table). **Fragment ζ** displayed strongest binding activity at 500 μM with a temperature shift of $3.23 \pm 0.36^\circ\text{C}$.

Since the fragments were designed in the pocket of HRP-2_{PWWP} (and not LEDGF_{PWWP}), we analyzed the fragments as well in a DSF with HRP-2_{PWWP} (Figure 4.23B). Of note, a similar DSF pattern as for LEDGF_{PWWP} was observed. Hit-**fragments γ and ζ** proved to be the strongest binders to the HRP-2_{PWWP} domain with respectively a temperature shift of $2.62 \pm 0.91^\circ\text{C}$ and $3.32 \pm 0.20^\circ\text{C}$ (Figure 4.23B, table). Remarkably, **fragment β** , with a clear dose-dependent effect for binding to LEDGF_{PWWP}, did not show clear binding to the HRP-2_{PWWP} domain (Figure 4.23A and B).

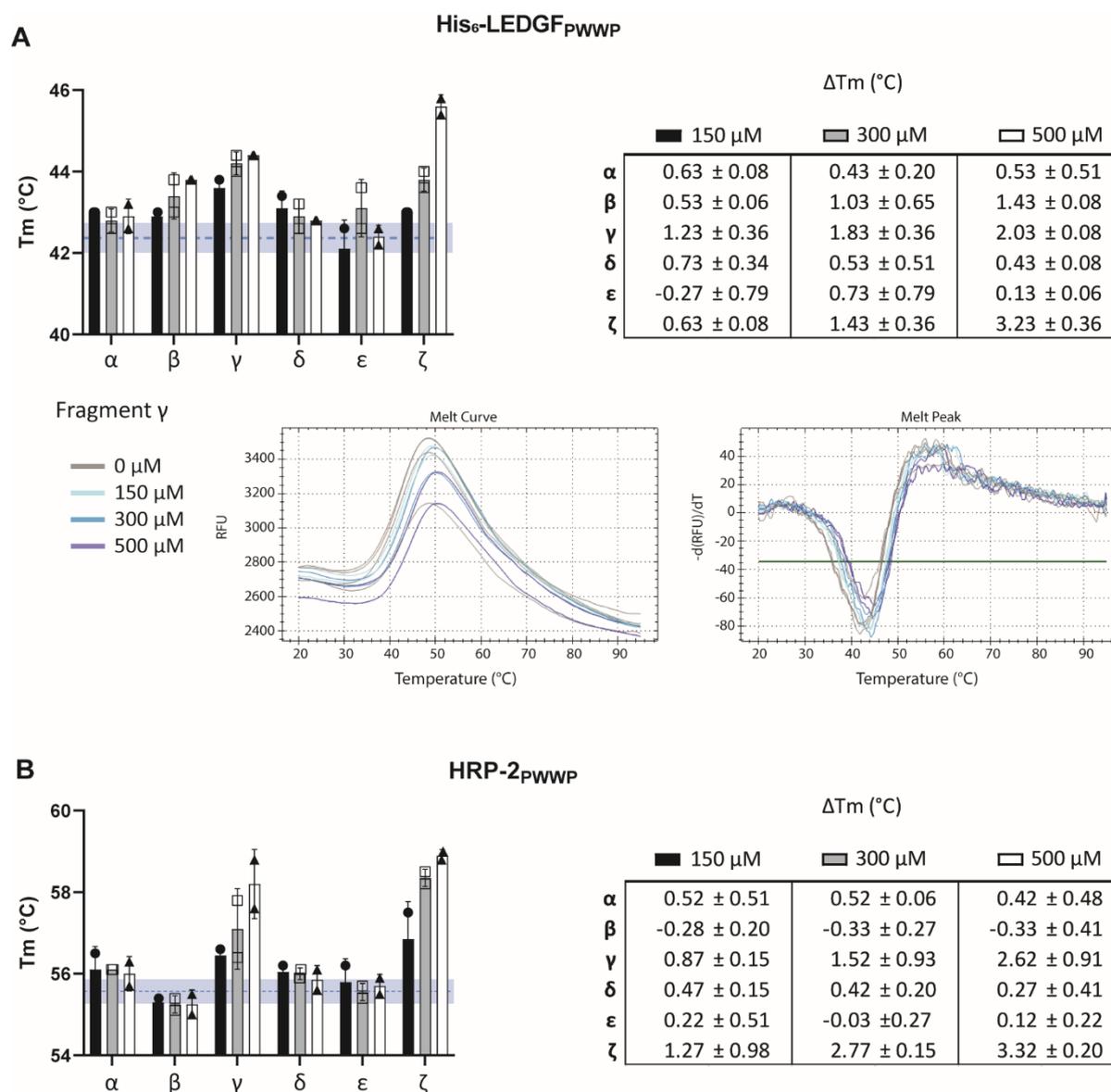


Figure 4.23: DSF analysis of selected fragments with LEDGF_{PWWP} or HRP-2_{PWWP}

Increasing concentrations of fragments (150 μ M, black; 300 μ M, grey; 500 μ M, white) or control (1% ethylene glycol (EG), blue) were added to recombinant and purified preparations of PWWP domains of LEDGF/p75 (A) and HRP-2 (B). The melting temperature for each domain was determined by DSF as well as the shift induced by the addition of ligands. An example of a DSF experiment for fragment γ and the LEDGF_{PWWP} domain is given in A (bottom). For each experiment, the melting temperatures are indicated by a symbol in the graph. Each DSF was run twice ($n=2$) and averages of both assays are given. The average of the control condition is given by the dotted line and the standard deviation is indicated with a blue area. Calculated difference in melting temperature (°C) for each condition is given in the tables on the right ($\Delta T_m \pm \text{SD}$).

4.2.8. Fragment hits bind to other PWWP-domains of PWWP domain containing proteins

Next to determining binding affinity, we investigated the specificity of these fragments for different proteins harboring a PWWP domain. Therefore, all six fragments were tested for binding to two PWWP domains of unrelated PWWP-domain harboring proteins: BRPF2 (Figure 4.24A) and WHSC1

(Figure 4.24B). Due to a higher variability in the melting temperature of BRPF2_{PWWP} (Figure 4.24A, blue bar), a temperature shift transcending the variation was only induced by **fragment β** at 500 μM ($-0.46 \pm 0.16^\circ\text{C}$). Although this decrease is smaller than 0.6°C and within the variability, a dose dependent effect was observed at the lower concentrations 300 μM and 150 μM (Figure 4.24A, table). An identical observation was made for **fragment α and δ**, where respectively marginal positive and negative temperature shifts occur in a dose-dependent manner.

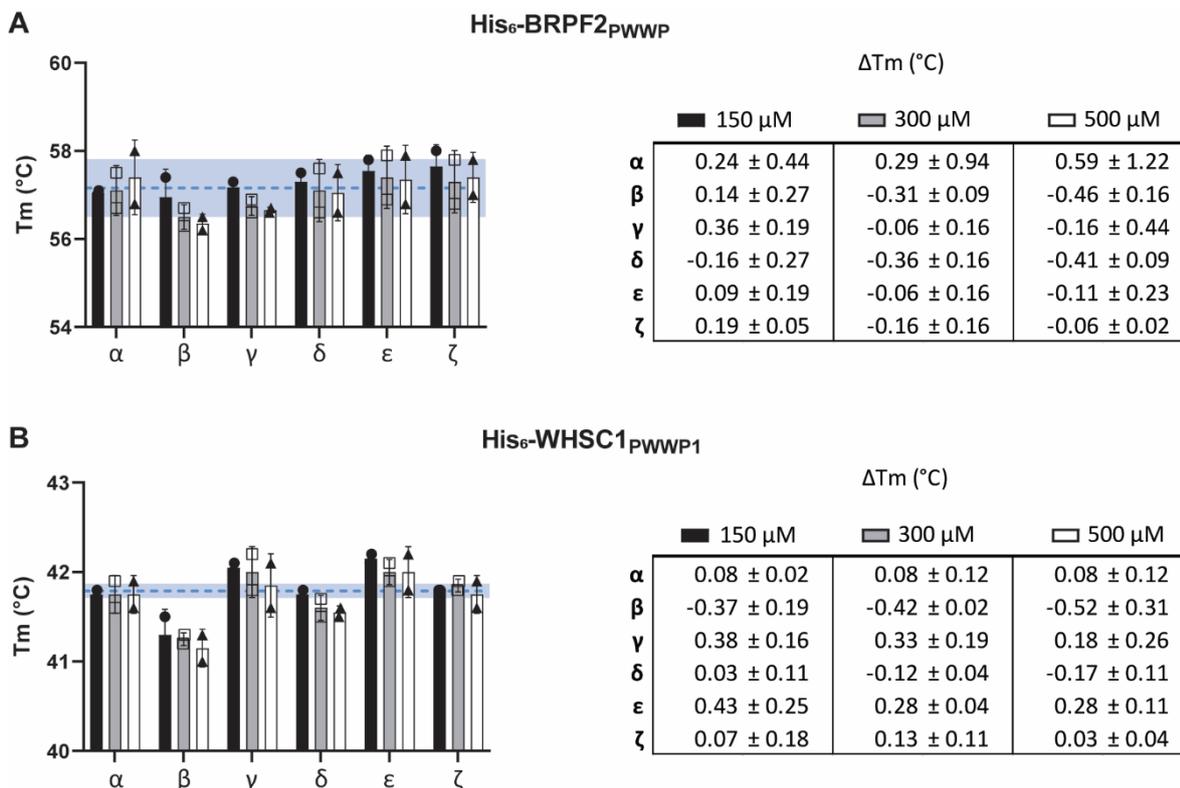


Figure 4.24: DSF analysis of selected fragments with BRPF2_{PWWP} or WHSC1_{PWWP1}

Increasing concentrations of fragments (150 μM, black; 300 μM, grey; 500 μM, white) or control (1% ethylene glycol (EG), blue) were added to recombinant and purified preparations of PWWP domains of BRPF2 (A) and WHSC1 (B). The melting temperature for each domain was determined by DSF as well as the shift induced by the addition of ligands. An example of a DSF experiment for fragment γ and the LEDGF_{PWWP} domain is given in A (bottom). For each experiment, the melting temperatures are indicated by a symbol in the graph. Each DSF was run twice (n=2) and averages of both assays are given. The average of the control condition is given by the dotted line and the standard deviation is indicated with a blue area. Calculated difference in melting temperature (°C) for each condition is given in the tables on the right ($\Delta T_m \pm \text{SD}$).

From the two PWWP domains present in the WHSC1 protein, we investigated fragment binding to the first PWWP domain (WHSC1_{PWWP1}, Figure 4.24B). Similar to the BRPF2_{PWWP} domain, **fragment β** displayed the largest effect, inducing a negative temperature shift of $-0.52 \pm 0.31^\circ\text{C}$ with the WHSC1_{PWWP1} domain at 500 μM. Here, no negative shift in temperature was induced by fragment γ.

Surprisingly, **fragment ζ**, with the largest temperature shift in DSF assays with LEDGF_{PWWP} and HRP-2_{PWWP}, had no effect on either of the unrelated PWWP domains. For **fragment ε**, no clear shift in any experiment was observed. On the contrary, temperature shift observations at 500 μM of **fragment δ** were very minimal but visible in each experiment (Figure 4.23 and Figure 4.24), suggesting this fragment may constitute a general binder to the PWWP domains.

4.2.9. Analysis of selected hits in AlphaScreen and TR-FRET

A disadvantage of DSF is that no additional information is obtained in regard to the location where the fragment binds and its potential to inhibit the interaction with the nucleosome. Although hits were selected by crystallography, in theory a ligand binding to PWWP outside the *bonafide* binding pocket may affect the melting temperature as well. Moreover, neither crystallography or DSF tell us whether the ligands function as inhibitors of the interaction with H3K36me3. Hence, the potential of fragments to inhibit the interaction of the PWWP domain with nucleosomes was assessed *in vitro* using the previously described interaction assays AlphaScreen and TR-FRET.

4.2.9.1. AlphaScreen

The fragments were first tested against the nucleosome-LEDGF/p75 interaction in AlphaScreen. AlphaScreen is a useful technique to detect weak protein-protein interactions because of an enzymatic cascade to amplify the detection signal. However, this mechanism in combination with the multivalent character of the beads hampers the detection of weak inhibitors as all protein-protein interactions bound to the bead should be interrupted for a clear decrease in the AlphaScreen signal. However, the fragments were tested in the presence of a fixed concentrations of 5 nM biotinylated nucleosome with the H3K36 trimethylation and 10 nM of Flag-LEDGF/p75. The six fragments were tested for their potency to inhibit the nucleosome-LEDGF/p75 interaction at the same concentration as before (150 μM, 300 μM and 500 μM; Figure 4.25A) and the outcome is reported as percentage of inhibition (PIN), ± the standard deviation. This PIN represents the percentage of AlphaScreen signal that was lost after incubation with the related fragment compared to the interaction without addition of fragment. Of note, this reference interaction and all fragment conditions were performed in the presence of 1% ethylene glycol (EG).

Results indicate a high variability between assay repeats, resulting in a spread of the counts over different experiments (n=4 or 5, duplicates are represented). Notwithstanding the assay variability

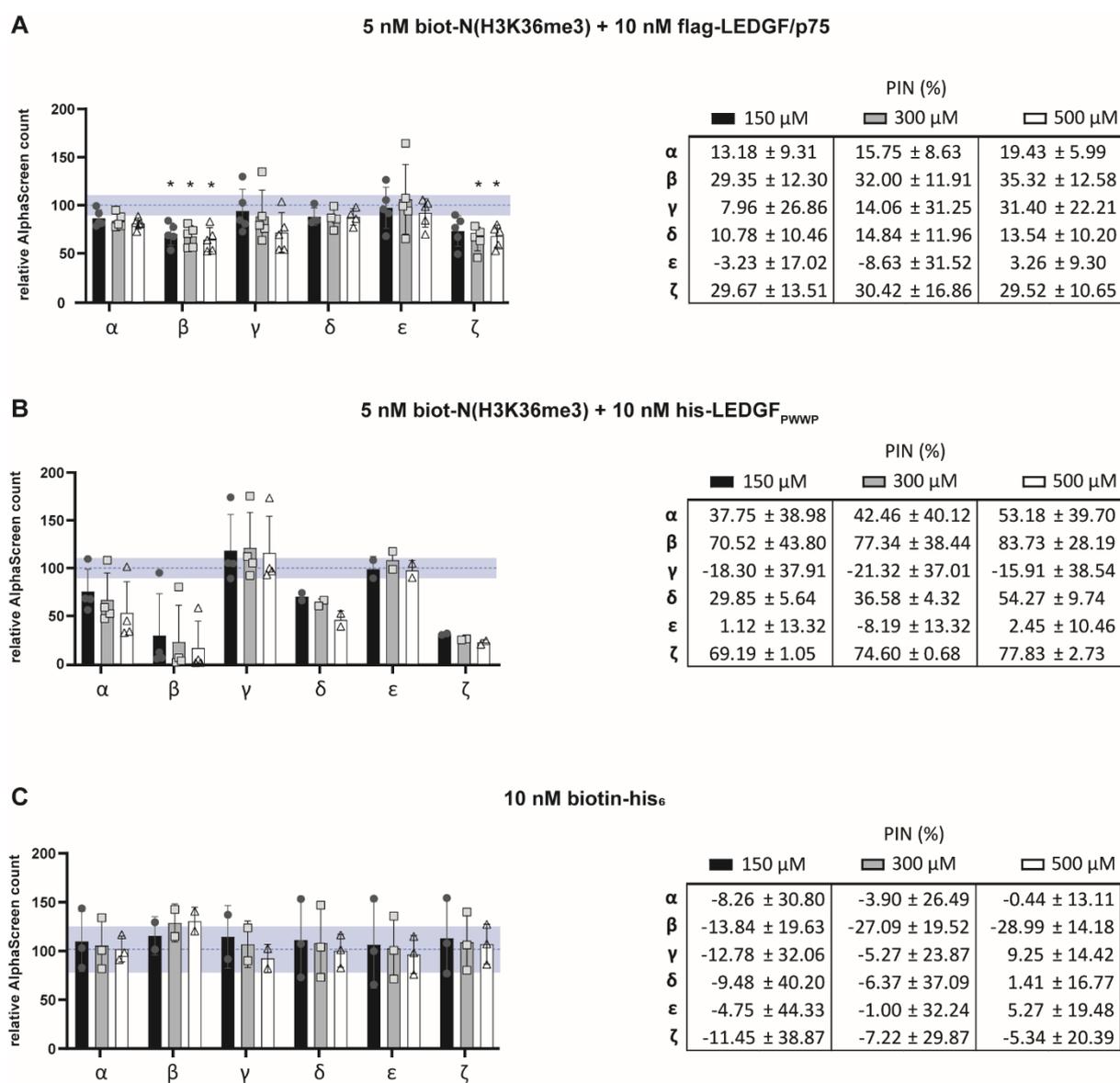


Figure 4.25: Fragment testing in AlphaScreen interaction assay

The percentage of inhibition (PIN) for each fragment at different concentrations (150 μ M, black; 300 μ M, grey; 500 μ M, white) or proper control (1% EG, blue) in (A) the biotinylated and trimethylated nucleosome (biot-N(H3K36me3), 5 nM) and Flag-LEDGF/p75 (100 nM) interaction ($n=5$ for all fragments except for fragment α $n=3$), (B) the biot-N(H3K36me3) – LEDGF_{PWWP} interaction ($n=3$ for α , β and γ , $n=2$ for δ , ϵ and ζ) or (C) quenching assay with biotinylated-His₆ (biotin-his₆, 10 nM) ($n=3$ for α , δ , ϵ and ζ ; $n=2$ for β and γ). Average of the control condition is given by the dotted line and standard deviation is indicated with a blue area. Calculated PIN (%) of the detected TR-FRET signal is given in the tables on the right (PIN \pm SD). The Kruskal-Wallis test with Dunn's correction was used to measure statistical differences between each fragment concentration and 0 μ M where $n \geq 3$; * $p < 0.05$.

with an average of $100 \pm 10.55\%$ for the untreated interaction (Figure 4.25A, blue bar), a clear reduction of the nucleosome-LEDGF/p75 interaction was observed for **fragment β** ($35.32 \pm 12.58\%$), **fragment γ** ($31.40 \pm 22.21\%$) and **fragment ζ** ($29.52 \pm 10.65\%$) at the concentration of $500 \mu\text{M}$. Fragment γ showed a dose-dependent effect with a decrease of $14.03 \pm 31.25\%$ and $7.96 \pm 26.86\%$ at $300 \mu\text{M}$ and $150 \mu\text{M}$ respectively (Figure 4.25A).

The selectivity and specificity of the fragment for the LEDGF_{PWWP} domain is important for the drug design program. This interaction can be detected in AlphaScreen and hence all fragments were tested in the presence of 5 nM biotinylated nucleosomes (H3K36me3) and 10 nM of the respective His₆-LEDGF_{PWWP} domain. The variability of the interaction signal without fragment ($100 \pm 25.88\%$, blue bar Figure 4.25B), as well as the variable results obtained after incubation with the fragments make it very hard to interpret the data. For **fragment β** , three out of four experiments resulted in an almost complete loss of signal, whereas an inhibition of 58% was reached in the fourth experiment. For **fragment γ** , three out of four experiments resulted in an AlphaScreen signal around 100% of the untreated interaction. For one experiment with fragment γ , the AlphaScreen counts increased together with the fragment concentration up to 194% at $300 \mu\text{M}$. For **fragment δ** and **fragment ζ** , a dose-dependent inhibition of the AlphaScreen signal was obtained with PINs of $54.27 \pm 9.74\%$ and $77.83 \pm 2.73\%$ at $500 \mu\text{M}$, respectively.

To exclude false-positive results from assay interference, a quenching assay was performed in which the fragments were incubated with 10 nM of biotinylated-6x histidine (biotin-his₆, Figure 4.25C). For this quenching assay, the acceptor and donor bead conjugates are anchored together and thus in close proximity, resulting in an AlphaScreen signal resistant to any fragment inhibition. In the event of a decreased AlphaScreen signal of this biotin-his₆ fusion, it is likely that the fragment interferes with the assay itself rather than the assay interaction. The quenching experiment characterized by a high variability for assay repeats of both the treated and untreated interaction ($100 \pm 23.6\%$). Based on the averages of three experiments, none of the six fragments interferes with the assay (Figure 4.25C).

4.2.9.2. TR-FRET

In addition to the AlphaScreen, all fragments were tested in the TR-FRET assay.

In a first set-up, all fragments were tested for inhibition of the nucleosome-LEDGF/p75 interaction in the TR-FRET assay. Increasing concentrations of fragments ($150 \mu\text{M}$, $300 \mu\text{M}$ and $500 \mu\text{M}$) were added to a fixed concentration of 5 nM biotinylated nucleosome and 100 nM Flag-LEDGF/p75, for which the TR-FRET signal was established at 100% (Figure 4.26A, dotted line). Interestingly, **fragments β , γ and ζ**

showing the most pronounced effect in the LEDGF_{PWWP} DSF (Figure 4.23A), demonstrate a clear dose-dependent effect in the TR-FRET assay, suggesting that the fragment prevents the interaction between the full length LEDGF/p75 protein and the H3K36 me3 nucleosome. **Fragment γ** displayed the highest percentage of dose dependent inhibition (PIN) at all concentrations. The nucleosome-LEDGF/p75 interaction was inhibited for $23.01 \pm 1.6\%$ at $500 \mu\text{M}$ (Figure 4.26A, table). Additionally, **fragment ζ** , with the highest shift for LEDGF_{PWWP} in DSF at $500 \mu\text{M}$ ($3.23 \pm 0.36^\circ\text{C}$, Figure 4.23A table), reached a PIN of $10.33 \pm 4.7\%$ at $500 \mu\text{M}$ in TR-FRET. Next is **fragment β** , that showed a dose-dependent effect at all concentrations with a $7.52 \pm 4.4\%$ inhibition at $150 \mu\text{M}$. Remarkably, **fragment δ** appeared as a dose-dependent inhibitor of the nucleosome-LEDGF interaction, with a PIN of $13.09 \pm 0.0\%$ at $500 \mu\text{M}$, although it induced minimal temperature shifts in DSF. Unfortunately, as mentioned before, the interaction between the nucleosome and the LEDGF_{PWWP} domain was not detectable in TR-FRET, so no fragments could be tested for the interaction here.

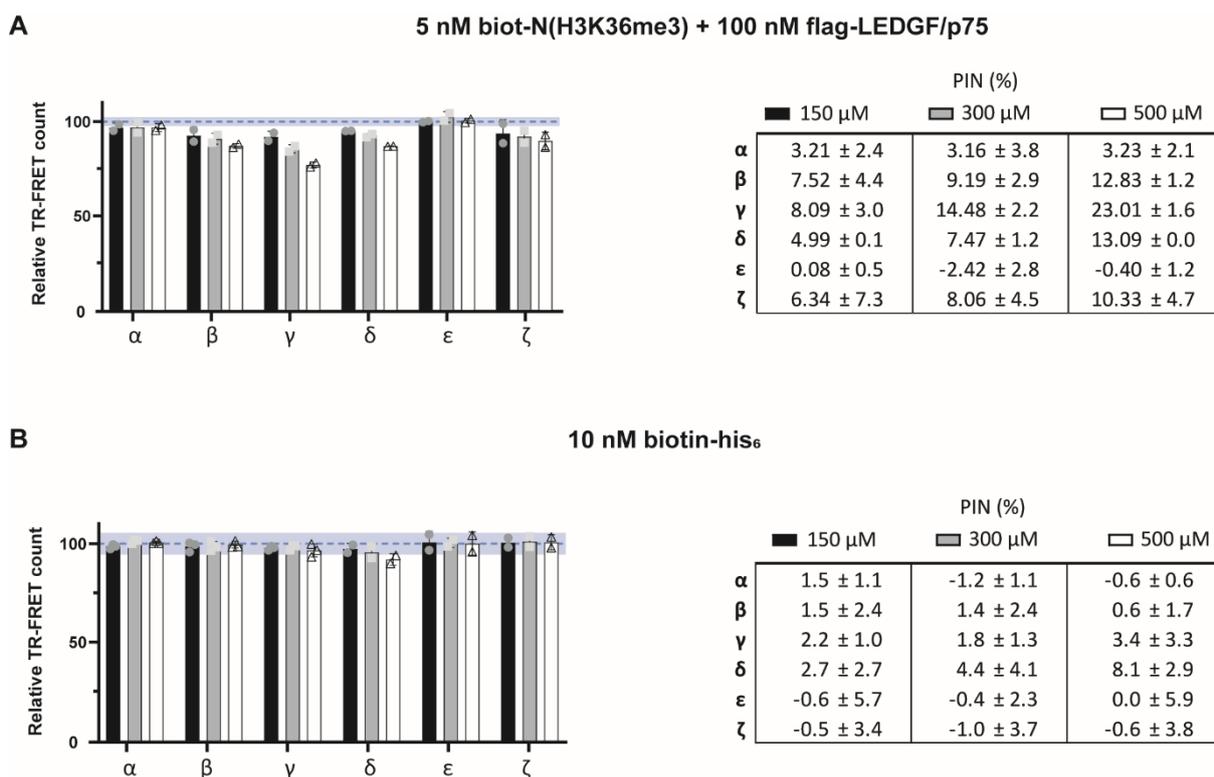


Figure 4.26: Dose dependent inhibition of the LEDGF/p75-nucleosome interaction in TR-FRET

The percentage of inhibition (PIN) for each fragment at different concentrations ($150 \mu\text{M}$, black; $300 \mu\text{M}$, grey; $500 \mu\text{M}$, white) or control (1% EG, blue) was determined by TR-FRET. TR-FRET measures the interaction between biotinylated and trimethylated nucleosome (biot-N(H3K36me3), 5 nM) and Flag-LEDGF/p75 (100 nM) ($n=2$) (A). Fragments were tested as well in a quenching assay (B) with biotinylated-His₆ (biotin-his, 10 nM) ($n=3$ for α , β and γ , $n=2$ for δ , ϵ and ζ). Average of the control condition is given by the dotted line and standard deviation is indicated with a blue area. Calculated PIN (%) of the detected TR-FRET signal is given in the tables on the right (PIN \pm SD).

Table 4.2. Summary of fragment characteristics

	ΔT_m at 500 μM					PIN at 500 μM									
	DSF					TR-FET					AlphaScreen				
	LEDGF _{PWWP}	HRP-2 _{PWWP}	BRPF2 _{PWWP}	WHSC1 _{PWWP1}		N - LEDGF/p75	biot-his ₆	N - LEDGF/p75	N - LEDGF _{PWWP}	biot-his ₆					
α	0.53 ± 0.51	0.77 ± 0.71	0.59 ± 1.22	0.08 ± 0.12		3.2 ± 2.1	-0.6 ± 0.6	19.4 ± 6.0	53.2 ± 39.7	2.9 ± 17.2					
β	1.43 ± 0.08	-0.03 ± 0.14	-0.46 ± 0.16	-0.52 ± 0.31		12.8 ± 1.2	0.6 ± 1.7	35.3 ± 12.6	83.7 ± 28.2	-22.7 ± 14.5					
γ	2.03 ± 0.08	3.27 ± 0.28	-0.16 ± 0.44	0.18 ± 0.26		23 ± 1.6	3.4 ± 3.3	31.4 ± 22.2	-15.9 ± 38.5	3.5 ± 9.9					
δ	0.43 ± 0.08	0.57 ± 0.42	-0.41 ± 0.09	-0.18 ± 0.11		13.1 ± 0.0	8.1 ± 2.9	13.5 ± 10.2	54.3 ± 9.7	3.6 ± 21.2					
ϵ	0.13 ± 0.06	-0.03 ± 0.42	-0.11 ± 0.23	0.28 ± 0.11		-0.4 ± 1.2	0.0 ± 5.9	3.3 ± 9.3	2.5 ± 10.5	8.2 ± 23.3					
ζ	3.23 ± 0.36	3.47 ± 0.28	-0.06 ± 0.02	0.03 ± 0.04		10.3 ± 4.7	-0.6 ± 3.8	29.5 ± 10.7	77.8 ± 2.7	3.6 ± 35.8					

N = biotinylated H3K36me3 nucleosome

Nevertheless, false positive results could be assessed using a similar quenching set-up as for the AlphaScreen assay in which all fragment concentrations were incubated with 10 nM of biotin-his₆ (Figure 4.26B). Interestingly, **fragment δ** induced a dose-dependent decrease in TR-FRET signal upon increasing concentration of fragment, with the highest inhibition of $8.1 \pm 2.9\%$ at 500 μM . This result suggests interference of the fragment with the assay and an overestimated PIN in the nucleosome-LEDGF interaction assay, which is most likely a combined effect of protein-protein interaction inhibition and interaction interference. Table 4.2 summarizes the performed experiments using these six fragments. Overall, fragments γ and ζ present as the most interesting and potent fragments with moderate activity over all different experiments.

5

Discussion

Part of the discussion has previously been published in S. Van Belle et al., "Unlike Its Paralog LEDGF/p75, HRP-2 Is Dispensable for MLL-r Leukemogenesis but Important for Leukemic Cell Survival," Cells, 2021 [272].

No specific treatment options are available for the genetically distinct and very aggressive form of leukemia characterized by *KMT2A*-rearrangements, called mixed lineage leukemia-rearranged (MLL-r) leukemia. The mechanism by which MLL-r leukemia is driven is relatively well investigated for a majority of the cases, and several experimental setups have resulted in a block of the leukemic phenotype in cells and in *in vivo* mouse models. These findings indicate that the development of small molecules that specifically target mechanistic pathways that underly MLL-r leukemia is a valid strategy to develop specific treatments.

LEDGF/p75 is a key player in the MLL-r leukemia disease mechanism. LEDGF/p75 tethers the MLL1 protein complex to the chromatin and the MLL1 target genes by forming an MLL1-menin-LEDGF/p75 ternary complex. We were interested to investigate whether the LEDGF/p75 paralog HRP-2 plays a role in MLL-r leukemia and aimed to develop small molecules targeting the LEDGF_{PWWP} domain.

5.1. The function of HRP-2 in normal hematopoiesis

To date, little is known about the function of HRP-2 in cell biology and oncogenesis. As paralog of LEDGF/p75, HRP-2 encompasses both a PWWP domain and IBD domain, two protein domains known to fold in full length LEDGF/p75. While the structures of both LEDGF/p75 domains are described in literature, only the PWWP domain of HRP-2 was crystallized and analyzed so far. We reported the NMR solution structure of the HRP-2_{IBD} domain and revealed a similar fold as the LEDGF_{IBD} domain consisting of 2 hairpin-alpha helices, connected by a fifth, smaller alpha helix (Figure 4.2). Superposition and alignment of both solution structures revealed a high similarity supporting the analogy between both proteins. Previously, co-immunoprecipitation experiments indicated PogZ, JPO2, and IWS1 as HRP-2 binding partners [199], [250]. In addition, we showed for the first time that MLL1 interacts with the IBD of HRP-2 by NMR (Figure 4.2). By comparing the amino acids involved in the IBD-MLL1 interaction for both LEDGF/p75 and HRP-2, an almost identical pattern was observed, suggesting that corresponding amino acids are responsible for interaction with MLL1. Next to these NMR experiments, we confirmed the interaction between MLL1 and HRP-2 in AlphaScreen and co-immunoprecipitation experiments (Figure 4.1 and Figure 4.2). In the context of MLL-r leukemia, it was of interest to find out whether HRP-2 is present in the ternary complex with MLL1 and menin. Whereas menin was required for the stabilization of the MLL1-LEDGF/p75 interaction as shown before [94], co-IP experiments with selective MLL-menin interaction inhibitors (Figure 4.1) and menin binding deficient MLL1 mutants (Figure 4.1) revealed that HRP-2 is less dependent on menin. Despite similar binding affinities as measured by NMR and AlphaScreen for the direct binding of HRP-2 or LEDGF/p75 to MLL1 (Figure 4.2), co-IP experiments indicate that menin modulates the interaction between the IBD and MLL1 in favor of LEDGF/p75 (Figure 4.1). We hypothesize that either low (undetectable) menin levels are sufficient

to support this interaction or addition of menin increases the binding affinity of LEDGF/p75 for MLL1 at the expense of HRP-2. Alternatively, it is possible that the HRP-2-MLL1 interaction is differentially regulated by PTMs or another cellular factor as compared to the LEDGF/p75-MLL1 interaction.

As HRP-2 is interacting with MLL1, the master regulator of normal development and adult hematopoiesis, a systemic *Hrp-2* knockout mouse model was used to study the role of HRP-2 in hematopoiesis. In contrast to the results of Wang et al. [287], the offspring from a heterozygous *Hrp-2* (*Hrp-2*^{+/-}) breeding couple deviated from the expected Mendelian inheritance pattern (Table 4.1). As discussed in section 4.1.3., we hypothesize that our mouse model presents with a more pronounced reduction of *Hrp-2* expression as no residual mRNA levels were detected in lineage depleted bone marrow cells. The mouse model by Wang et al. presented with up to 20% mRNA expression levels, suggesting that 20% gene expression is sufficient for normal development. Genotyping new born pups revealed a distribution more in line with the Mendelian theory, suggesting that HRP-2 is important for postnatal survival. So far, we did not explore the possible causes of this early death in more detail. Recent studies indicated a crucial function of HRP-2 during mRNA transcription [219] and differentiation [248] of muscle cells, suggesting *Hrp-2*^{-/-} muscles fail to support vital functions. Of note, high prenatal lethality was observed in systemic *Psip1* knockout mice with less than 1% reaching weaning age [206]. These surviving *Psip1* null mice presented with severe phenotypic abnormalities [294]. Unlike the phenotype observed in *Psip1* knockout mice, we could only distinguish adult HRP-2 knockout mice from their wild type and heterozygous littermates by small differences in the hematopoietic system (Figure 4.3 and Figure 4.4). The limited group size of *Hrp2*^{-/-} mice and high biological variability within groups might have disguised the detection of statistically significant differences. In fact, we only found a significant increase in the number of neutrophils. A more detailed analysis of the hematopoietic stem cells by colony formation experiments (Figure 4.3) and Gene Set Enrichment Analysis (Figure 4.4) hinted towards a stem-like state supported by HRP-2.

5.2. The function of HRP-2 in cell survival and MLL-r development

Hrp-2 depletion had minimal effects on the hematopoietic system in adult mice. Yet, because of the interaction with MLL1, we evaluated how HRP-2 affected growth and survival of leukemic cells *in vitro*. To allow an extensive analysis, a broad panel of leukemic cell lines was analyzed in parallel, including acute myeloid leukemia cell lines THP1 (MLL-AF9) and Kasumi1 (MLL1 wt), acute lymphoblastic leukemia cell lines SEM (MLL-AF4) and Nalm6 (MLL1 wt) and the chronic myeloid leukemia cell line K562.

Analysis of cell survival and clonal expansion potency of these different types of leukemia (Figure 4.5 and Figure 4.7) indicated that HRP-2 depletion impairs cellular growth independently of the presence of MLL1 fusion proteins. Of interest, HRP-2 knockdown was reported to reduce growth in hepatocellular carcinoma cells and induce cell death in U2OS cells [249], suggesting a more general pro-survival role of HRP-2. Of note, we observed that colony formation of the MLL-r driven cell lines THP1 and SEM at high vector titer were less affected by HRP-2 depletion than the MLL1 wild-type cell lines K562, Kasumi1 and Nalm6 (Figure 4.5). We hypothesize that HRP-2 and LEDGF/p75 compete for MLL1 binding, implying that reduced HRP-2 levels facilitate binding between LEDGF/p75, MLL1 and menin supporting MLL-r-induced leukemogenesis. However, since LEDGF/p75 is present in the human MLL-r driven cell lines when growth reduction is observed by HRP-2 depletion, we hypothesize that the effect of HRP-2 depletion, which seems independent of the MLL1 genotype, appears dominant over the LEDGF/p75 driven leukemogenesis (Figure 4.6). In line with the finding that HRP-2 supports a stem-like state in primary cells (Figure 4.3 and Figure 4.4), the effect on cellular growth in *Hrp2* knockdown cells in differentiation-stimulating medium of the colony formation was more pronounced than the reduction in growth in regular medium (Figure 4.5 and Figure 4.7).

As previously mentioned, LEDGF/p75 is important for the viral replication cycle of HIV-1 by affecting viral integration in the host genome [238], [240], [242]. Detailed analysis LEDGF/p75 depletion by Schrijvers R. *et al.* revealed redundancy by ectopic expression of the HRP-2 protein, thus supporting the theory that both proteins can fulfil similar functions [251], [252]. A similar experiment was performed in the MLL-r leukemic cell line THP1 and the MLL1 wild type cell line Nalm6 (Figure 4.8 and Figure 4.9). As expected, knockdown of LEDGF/p75 only affected the clonogenic growth of the MLL-r cell line, indicating that endogenous HRP-2 levels do not rescue colony formation. In contrast, a rescue of the phenotype was observed upon overexpression of HRP-2 (Figure 4.8), supporting the notion that HRP-2 can function as a tether between MLL-FPs and its target genes. This rescue phenotype is reminiscent of that of HRP-2 in HIV infection. However, as HRP-2 overexpression also promoted cell growth in hepatocellular carcinoma [250], we cannot exclude that HRP-2 rescues the cellular growth via a more general pathway.

Because of this functional interplay between HRP-2 and LEDGF/p75 and in light of the strategy to target the structural domains of LEDGF/p75, we further investigated the role of HRP-2 in the development and maintenance of MLL-r leukemia. Although HRP-2 seems to play an important but nonspecific role in leukemic cell survival, our results using primary murine cells indicate that HRP-2 is not important for the initial transformation of hematopoietic stem cells by MLL1 fusions (Figure 4.10). MLL-ENL transduced lineage depleted BMCs from HRP-2 knockout, heterozygous or wild-type mice transformed

irrespective of the genotype in a colony forming assay, indicating that HRP-2 knockdown, nor knockout is impairing MLL-r driven leukemogenesis in the presence of LEDGF/p75.

In lethally irradiated mice, engrafted MLL-ENL transduced cells induced leukemia with a survival time of 6 weeks post transplantation (Figure 4.10), equal to previously published experiments [179]. Transplantation of MLL-ENL transduced HRP-2^{-/-} or HRP-2^{+/-} cells also resulted in leukemogenesis and the life span of these mice was even shorter. The shorter life span of mice in the absence of HRP-2 may be due to a loss of competition between HRP-2 and LEDGF/p75 for binding to MLL1. By competing with LEDGF/p75, HRP-2 could act as a tumor suppressor. Alternatively, although HRP-2 might be dispensable for leukemogenesis, its role in hematopoiesis may affect the survival of mice.

The limited number of HRP-2^{-/-} animals in this project made it more challenging to study the role of HRP-2 in normal hematopoiesis and MLL-r leukemia *in vivo*. Understanding the high lethality of HRP-2^{-/-} pups could give insight in the possible risks of targeting HRP-2 in the early stages of life. In case of LEDGF/p75, a conditional knockout mouse model where LEDGF/p75 was depleted in the hematopoietic system turned out to be viable and a useful tool for understanding the importance of LEDGF/p75 in hematopoiesis. It would be interesting to study a similar mouse model in which HRP-2 is selectively knocked out in the hematopoietic system.

5.3. HRP-2 as target for MLL-r leukemia?

Our findings demonstrate that HRP-2 is able to interact with MLL1, but that the interaction is not important for MLL-ENL leukemogenesis. We hypothesize that LEDGF/p75 is the main driver of MLL-r leukemia. In line with previous reports, our results indicate that by replacing LEDGF/p75 from the nucleosomes, endogenous HRP-2 expression cannot substitute for this lost interaction. This indicates that LEDGF/p75 is the main target for MLL-r treatment. Since both PWWP domains of LEDGF/p75 and HRP-2 are highly similar, obtaining specificity for LEDGF_{PWWP} can be challenging. Interestingly, we have indicated that overexpressed HRP-2 is able to tether MLL-FP in the absence of LEDGF/p75 to the chromatin by observing an increase in number of colonies (Figure 4.8). Whether this observation is biological relevant and whether HRP-2 is upregulated in MLL-r leukemia after depletion of LEDGF/p75 has not been investigated. Together with our cell culture data that suggest a more general role for HRP-2 in the survival of leukemic cells, the prediction of potential toxicity induced by affecting HRP-2_{PWWP} with small molecules remains elusive. For a better understanding we lack mechanistical insight on how HRP-2 affects general cell growth and leukemic cell survival.

5.4. Validating PWWP as target for MLL-r leukemia

Currently, many efforts are ongoing to specifically target the oncogenic multi-protein complex involved in MLL-r leukemia. DOT1L and BET inhibitors target the MLL1 fusion partners [36], [39], whereas menin inhibitors are directly inhibiting menin binding to MLL1 [158], [160]–[162]. The importance of LEDGF/p75 in MLL-r leukemia was described by Yokoyama *et al.* [94] and later LEDGF/p75 was validated as potential target to treat MLL-r leukemia (reported in section 1.3.1.2., page 22). Yokoyama *et al.* defined that the LEDGF_{PWWP} domain alone, when fused to the MLL1 fusion protein, is sufficient for leukemic transformation. In line with those findings, we show here that expressing the LEDGF_{PWWP} domain can induce competition with endogenous LEDGF by occupying the nucleosomes (Figure 4.12). This results in a decreased number of colony formed units, suggesting that displacing LEDGF/p75 from the nucleosome is affecting the proliferation of the cells. The fact that this effect was less pronounced when using the nucleosome-binding deficient LEDGF_{PWWP} W21A, F44A mutant supports this theory. Although this experiment was controlled by an MLL1 wild type cell line in which no effect was observed, it would be of interest to study the HoxA9 expression levels upon competition, to support the hypothesis that this reduction in colony formation is MLL-r specific.

To detect the interaction between recombinant proteins *in vitro* (Figure 4.14 and Figure 4.18) and ectopically expressed proteins in cells (Figure 4.12 and Figure 4.20), an N-terminally tagged full length LEDGF/p75 or LEDGF_{PWWP} domain was used. Tagging these proteins N-terminally may appear contra intuitively, as the interaction surface with nucleosomes is located at this site. Hendrix J *et al.* showed that N-terminally tagged LEDGF is functionally active in cells [295] and our results indicate that these proteins can still interact with the nucleosomes, suggesting that the tag does not interfere with the LEDGF/p75 function.

5.5. Assay development to detect the PWWP-nucleosome interaction

For screening purposes, several *in vitro* assays were set up to detect the interaction between various LEDGF constructs and nucleosomes. At first, we performed AlphaScreen and TR-FRET experiments in which tagged, recombinant LEDGF/p75 or LEDGF_{PWWP} proteins were used. Both assays can detect a protein-protein interaction, however the detection method is substantially different, resulting in two complementary assays. As the name suggests, the Amplified Luminescent Proximity Homogenous Assay (AlphaScreen) signal is amplified by a chemical reaction in the acceptor bead. In combination with the avidity of the AlphaScreen beads, meaning that multiple proteins can interact with one bead, this assay is useful to detect weak protein-protein interactions. On the contrary, more potent inhibitors might be necessary to inhibit the interaction to non-detectable levels. In TR-FRET, tagged proteins are

recognized by donor and acceptor fluorophores conjugated to tag-recognizing moieties. In equilibrium, a one to one ratio of protein and fluorophore is achieved and no cascade enhances signal emission, making it easier to inhibit the protein interaction. To best mimic the cellular context, recombinant nucleosomes consisting of a histone octamer, wrapped in the 147 bp Widom sequence were preferred over the use of histone-mimicking peptides. Interestingly, an extension of 20 bp to the DNA Widom sequence was needed to stabilize the LEDGF_{PWWP} domain on the nucleosomes to obtain good cryo-EM structures [186]. Nonetheless, we were able to detect the interaction between the biotinylated nucleosome and Flag-LEDGF/p75 or His₆-LEDGF_{PWWP} (Figure 4.14). To further validate the interaction interface at the LEDGF_{PWWP} domain, we were able to displace His₆-LEDGF_{PWWP} from the nucleosome by introducing Flag-LEDGF/p75 (Figure 4.15). This suggested that both Flag-LEDGF/p75 and His₆-LEDGF_{PWWP} competed for nucleosome binding. *Vice versa*, His₆-LEDGF_{PWWP} could displace Flag-LEDGF/p75 from the nucleosome when present in excess. As these results support the interface between recombinant nucleosome and LEDGF at the LEDGF_{PWWP} domain, we used the LEDGF_{PWWP} W21A, F44A mutant that was described to have a reduced affinity for nucleosome binding [183]. In line with our hypothesis, the competition by the His₆-LEDGF_{PWWP} W21A, F44A mutant was less pronounced (Figure 4.15). Of note, the direct interaction of Flag-LEDGF/p75 wild type or W21A, F44A mutant was not strikingly different (Figure 4.15B), possibly because not only the PWWP domain, but also the following CR1 and AT hook motifs are important for DNA binding [189], [190], [192]. As reported in literature [184], [185], we were able to detect the interaction between His₆-LEDGF_{PWWP} and both di- and trimethylated nucleosomes. To a lesser extent, we observe an interaction between non and monomethylated nucleosomes and His₆-LEDGF_{PWWP} (Figure 4.16).

In TR-FRET, the interaction between biotinylated nucleosome and Flag-LEDGF/p75 could be observed, however we were not able to detect any interaction with the His₆-LEDGF_{PWWP} protein (Figure 4.18). This is likely due to the combination of the earlier explained assay technologies and the lack of CR1 and AT hook motifs that follow the PWWP domain and are important for DNA binding. However, the importance of the LEDGF_{PWWP} domain was shown by a decrease in TR-FRET signal when the W21A, F44A mutant was compared to a wild type protein. This specificity of the interaction was further confirmed by out competition with His₆-LEDGF/p52, the shorter isoform of LEDGF/p75 which encompasses an identical PWWP domain. Of interest, also here the interaction of Flag-LEDGF/p75 could be observed with all methylation states of H3K36. The decrease in signal in correspondence with the level of methylation suggests that LEDGF/p75 binds best to trimethylated nucleosomes and that the interaction with non-methylated nucleosomes is less strong. Comparing these results to the W21A, F44A mutant protein a similar pattern could be observed. Additionally, the ratio between wild type and mutant protein for their binding to biotinylated nucleosome decreases with a lower methylation

state, supporting the importance of both residues in the PWWP pocket for binding to the methyl groups of lysine 36 (Figure 4.18).

5.6. Discovery of small ligands targeting LEDGF_{PWWP} to treat MLL-r leukemia

To obtain small molecules that inhibit the interaction between nucleosome and LEDGF/p75-bound MLL-FP we followed a crystallography-based drug discovery strategy using crystals of the HRP-2_{PWWP} domain. This structural information on fragment binding is a very powerful and essential tool during drug design and development. However, the biological relevance of soaking these fragments in the presence of HRP-2_{PWWP} crystals is arguable. Unfortunately, when crystallizing LEDGF_{PWWP}, the PWWP pocket is occupied by an adjacent PWWP molecule, blocking the residues important for nucleosome interaction (data not shown). The HRP-2_{PWWP} domain, with its 94% sequence homology (Figure 4.21B) and highly related structure (Figure 1.9 and Figure 1.11), is the closest related PWWP domain of LEDGF_{PWWP}. In more detail, both PWWP pockets of LEDGF and HRP-2 show an identical fold, which is the rationale for using the HRP-2_{PWWP} domain for screening binding of fragments to the PWWP pocket residues.

The screening was initiated using small, commercially available fragments. We discriminate fragments from compounds on basis of their molecular weight (<300 Dalton). In a parallel approach, a medium throughput screen of 600 fragments was performed at the X-chem facility in Oxford, providing 64 additional hits. Based on these primary hits, commercially available new fragments and larger de novo synthesized molecules were analyzed. To compensate for the limits of crystallography and the lack of LEDGF_{PWWP} structures, fragment hits were tested and validated against LEDGF_{PWWP} in NMR and molecular interaction assays *in vitro*.

Here, we reported on six hits of the in-house fragment screen (Figure 4.22). At first, the fragments were tested for their binding to either the His₆-LEDGF_{PWWP} or the HRP-2_{PWWP} domain in the DSF assay (Figure 4.23). Interestingly, the same two fragments induced the most pronounced shift in melting temperature for both domains, in a concentration dependent manner. However, it should be noted that both the LEDGF_{PWWP} and HRP-2_{PWWP} domain present with a different intrinsic melting temperature (43.0 and 56.6°C, respectively, Figure 4.13). It should be investigated whether this difference depends on the presence of the 6x histidine tag N-terminally of LEDGF_{PWWP} or salt concentrations. Despite the high sequence similarity and apparent similar PWWP domain fold, both proteins seem to differ in thermostability. Additionally, fragment β showed moderate temperature shifts for the His₆-LEDGF_{PWWP} domain but not for HRP-2_{PWWP} domain, although the fragment was a binder in crystallography. This highlights the importance of screening fragments not only with the HRP-2_{PWWP} domain, but also

LEDGF_{PWWP} in a parallel approach. DSF does not provide any information on the interface where fragments are binding. To investigate whether they are interfering with the LEDGF_{PWWP} interaction activity, the fragments were tested for their potency to inhibit the interaction between LEDGF/p75 and nucleosomes in the optimized interaction assays. Results indicate a modest inhibition for the same two fragments γ and ζ , but due to a high variability between AlphaScreen repeats, interpretation of the assay is difficult. The compounds may contain too limited potency to inhibit multiple interactions that arise from the avidity of AlphaScreen beads. In TR-FRET, the fragment inhibition is less pronounced, but due to a much lower variability more reliable. We aim to design a labelled fragment to facilitate screening of newly designed fragments for their binding to the LEDGF/p75 or LEDGF_{PWWP} proteins by out-competition with the labelled fragment.

Results of DSF, AlphaScreen and TR-FRET (Table 4.2) support fragment ζ as the most potent fragment at this moment. Although these inhibitory effects are small and at not biological relevant concentrations, the fact that we measure activity for such small fragments, divided over different scaffolds is promising for the continuation of the project where we aim to grow potent fragments into compounds with low nanomolar activity. To obtain this, our assay results will be compared to the obtained K_d values by NMR and available crystal structures from soaking. Correlating the results from different groups will enforce the developmental process and will help to exclude fragments that interfere with an assay rather than induce an inhibitory effect (false positives). So far, the fragments were not tested in nanoBRET and available cellular assays such as MTT and CFU because of their limited potency.

5.7. The challenge to avoid toxicity and obtain LEDGF_{PWWP} specific activity

Drug discovery and development can be very challenging and only a limited number of compounds reach the market for a clinical application. One of the major hurdles in drug development is toxicity. As LEDGF is a ubiquitously expressed protein and involved in many cellular processes such as transcription elongation and DNA damage repair, control experiments and awareness during our drug discovery project are of great importance. In fact, a systemic LEDGF KO mouse is not viable, suggesting that LEDGF is important for the prenatal development [206], [287]. In contrast, the absence of LEDGF in the hematopoietic system from very early on did not affect viability and reproducibility [179]. However, a systemic and inducible mouse model for LEDGF, which would mimic the addition of LEDGF inhibitors, has not been described so far. On the contrary, our colony forming experiment (Figure 4.11), as well as previous publications [95], [97] indicate that targeting LEDGF by either overexpression of the PWWP domain or IBD domain alone, respectively, is sufficient to reduce the number of colonies formed in *KMT2A*-rearranged cells specifically. Compounds with cellular activity are needed to further

address this question. Lead compounds will be tested in primary cells and *in vivo* mouse models. In addition, we aim to reduce the risk of toxicity by developing fragments that specifically bind the LEDGF_{PWWP} pocket. However, 26 cellular proteins are described to harbor a PWWP domain and all proteins are involved in epigenetic reading of methyl-lysine marks. In particular, the HRP-2_{PWWP} domain shows high sequence similarity to the LEDGF_{PWWP} domain. As discussed previously a toolbox is available to study fragment specificity for either the HRP-2_{PWWP} or LEDGF_{PWWP} domain, as well as other PWWP domain proteins. Indeed, we do not limit our focus to the extensively discussed paralog, but investigate binding of fragments to other PWWP domain proteins as well. Interestingly, from our results in DSF (Figure 4.13 and Figure 4.24), we noticed different melting characteristics for each PWWP domain and subsequently different binding patterns for the tested fragments, suggesting that fragments bind differently to the distinct PWWP domain proteins. The induced temperature shift indicates that the fragment affects the denaturation of our protein domain. Based on available crystallography data of HRP-2, it is likely that the fragments bind at the desired HRP-2_{PWWP} pocket-site, but it remains elusive for other PWWP domain proteins whether binding occurs at this preferred PWWP pocket-site. To investigate in more detail whether fragment-binding to the other PWWP domain proteins is also interfering with the epigenetic reading function, we should optimize an interaction assay with the recombinant protein domain of interest and corresponding post-translationally modified nucleosomes.

Of note, despite this functional similarity between different PWWP-domain containing proteins, small differences in the structured domains have been reported. Obtaining specificity might be supported by a slightly longer sequence between two beta sheets or extra alpha helices at the C-terminal end of the PWWP domain. Additionally, two other groups have already succeeded in obtaining probes, or small molecules, that specifically bind one PWWP domain-containing protein. On the one hand, Böttcher J. *et al.* developed a NSD3_{PWWP1} (or WHSC1L1_{PWWP1}) inhibitor with an affinity of 166 nM in surface plasmon resonance and 445 nM in isothermal titration calorimetry [271]. In cells, the compound showed cell line-specific activity at $26.8 \pm 4.4 \mu\text{M}$ in Molm-13 cells and $13 \pm 2 \mu\text{M}$ in RN2 cells [271]. By DSF, the authors indicated a clear shift in the melting temperature of NSD3_{PWWP1} only, and not for the closest family member NSD2 and 12 other tested PWWP-domain proteins, neither for the second PWWP domain present in NSD3 [271]. The compound was not tested against NSD1. Not much later, very modest inhibitors of the NSD2_{PWWP1} molecule were reported [268], [269].

Theoretically, preventing the LEDGF_{PWWP} interaction with nucleosomes will not only affect MLL1 tethering to the nucleosomes, but also affect the other previously mentioned IBD interacting proteins that use LEDGF/p75 as guide to find actively transcribed genes labelled with the H3K36me3 mark. Some of these cellular proteins are associated with cancer or neurological disorders. Noteworthy, the

viral protein interacting with LEDGF/p75, HIV-1 integrase, could also be affected by LEDGF_{PWWP} inhibitors, suggesting our compounds could serve as treatment for HIV-1. In a later stage, the effect of LEDGF_{PWWP} inhibitors will also be assessed in this context.

In addition, the existing menin inhibitors have shown promising activity not only for MLL-r leukemia, but are also tested for NPM1-mutated leukemia [296], [297]. Both aggressive leukemias are characterized by elevated HOX levels and the VTP-50469 directly affects the cofactor MEIS1, which results in differentiation [296]. Promising *in vivo* results were also obtained using the SNDX-5613 compound [297], which are currently in clinical trial for both MLL-r and NPM1-mutated leukemia [167]. Whether hampering the LEDGF/p75-mediated tethering of MLL1, by blocking the LEDGF_{PWWP} domain interactions with chromatin, would be considered a potential therapeutic strategy for NPM1-mutated leukemia or other leukemias characterized by elevated HoxA9/Meis1 levels [298] has to be investigated.

5.8. Conclusion and further perspectives

Overall, we have concluded that HRP-2 is not essential for MLL-r leukemogenesis, but suggest that HRP-2 has a more general effect on the growth of leukemic cells. The mechanism by which growth is inhibited in the absence of HRP-2 was not elucidated. To further investigate, an RNA sequencing analysis of leukemic cells depleted for HRP-2 in comparison to leukemic wild type cells would be of interest. To study whether the reduced growth effect upon HRP-2 depletion is specific for leukemic cells rather than affecting all cancer cells or normal tissue, the different cell lines tested should be expanded to non-leukemic cancer cells and human primary cells. Nonetheless, we cannot predict at this moment whether small molecules targeting the LEDGF_{PWWP} domain will affect the HRP-2_{PWWP} function.

We aim to develop a compound class (i) specifically binding to the LEDGF_{PWWP} pocket, (ii) interfering with the nucleosome interaction so that (iii) the MLL-FP is displaced from its target genes and (iv) MLL-r leukemic phenotype is impaired. The road to potent, specific compounds is challenging, but the moderate activity obtained in this thesis with small, primary hits from different structural scaffolds is promising for further development. First line screening of fragments will be performed in crystallography and DSF in parallel, a variable set of additional experiments is present in the lab to further investigate the inhibitory effect of fragments. The iterative process between fragment growth and activity measurement will reveal the structure-activity relationship to further build fragments into compounds. Based on the NSD2_{PWWP1} and NSD3_{PWWP1} compounds [269], [271], were an *in vitro* affinity in the low nanomolar range resulted in cellular activity, our fragments with a similar *in vitro* activity

will be tested in cellular context. First, the compound will be tested for their effect on the viability of the cells and their potency to interfere with the nucleosome-LEDGF/p75 interaction in a nanoBRET experiment. The gene expression profile of treated cells will be compared to untreated controls and the effect on leukemic growth will be tested in a colony formation assay. After optimization to a lead compound, ADME and animal experiments will be performed. Next to the lead optimization, it would be interesting to use molecules with high affinity for the LEDGF_{PWWP} domain to investigate the biological functions of LEDGF not only in leukemia, but also in other cancers such as prostate and breast cancer [221]–[224].

Acknowledgements, personal contribution and conflict of interest statements

Scientific acknowledgements

The HRP-2 project ([Section 4.1](#)) was initiated by Jan De Rijck (J.D.R.) and Sara El Ashkar (S.E.A.) who designed experiments and wrote the manuscript together with Siska Van Belle (S.V.B.) and Zeger Debyser (Z.D.). I thank Katerina Cermakova (K.C.) for the provision of structural data and support in writing the manuscript. Additionally, K.C. and Courtney H. Hodghe (C.H.H.) supported the RNA analysis. I want to thank J.D.R., S.E.A. and Alessandro Canella (A.C.) for their contribution on generating data and Filip Mattijsens (F.M.) and Steven Goossens (S.G.) for the processing of RNA sequencing samples. I thank Martine Michiels for excellent technical assistance. The project was supervised by Frauke Christ (F.C.), Pieter Van Vlierberghe (P.V.V.), Vaclav Veverka and Zeger Debyser.

The drug discovery project ([Section 4.2](#)) is supervised by Frauke Christ, Zeger Debyser, Vaclav Veverka and Sergei Strelkov (S.S.). I want to thank Evgenii Osipov and Nayyar Ahmad Aslam, under supervision of Sergei Strelkov, Arthur Van Aerschot and Peter Verwilst, for the design, development and synthesis of fragments. I also thank Thibault Vantieghem (T.V.) for generating crystallography data, Vanda Lux for generating NMR data and Thatcher Akele for performing biochemical interaction assays. I thank Martine Michiels for excellent technical assistance.

Funding

This work was supported by KU Leuven Interdisciplinair onderzoeksprogramma (IDO) Program Grant IDO/12/008-3E130241, Flanders FWO Grant G.0595.13 (Z.D.), the Flemish SBO 140038 (ChromaTarget), the Belgian National Foundation Against Cancer, GACR 16-06357S (V.V.), the Ministry of Education of the Czech Republic LO1304 (V.V.), the European Regional Development Fund; OP RDE; the project ChemBioDrug CZ.02.1.01/0.0/0.0/16_019/0000729 (V.V.), the ERC Starting Grant 639784 (P.V.V.), the Fund for Scientific Research Flanders (FWO) (P.V.V.), Stand up to Cancer (the Flemish cancer society) (P.V.V.), the CPRIT RR170036 grant (H.C.H.), the V Foundation grant V2018-003 (H.C.H.), Gabrielle's Angel Foundation for Cancer Research (H.C.H.), NCI R00CA187565 (H.C.H.), TReaT MLL C32_18_038 KUL funding (S.S. and F.C.) and FWO project grant G053822N (S.S., Z.D.). F.C. is supported by the KU Leuven Industrial Research Fund. My contributions to this section were funded by a fellowship of FWO.

Personal contribution

All chapters were written by Siska Van Belle and proof-read by promotor Prof. Zeger Debyser and co-promotor Dr. Frauke Christ. All figures were generated by Siska Van Belle using Adobe Illustrator and Biorender, except Figure 1.5 (copied with permission from [97]), Figure 4.2 (K.C.) and Figure 4.21 (T.V.).

Experimentally, section 4.1. was partly performed by Siska Van Belle, in collaboration with Sara El Ashkar, Katerina Cermakova and Jan De Rijck, under supervision of Prof. Zeger Debyser and Prof. Vaclav Veverka. In this chapter, I performed part of the immunoprecipitation experiments, part of the genotyping, RNA sequencing, cellular experiments and assisted with *in vivo* experiments. For section 4.2., I performed all experiments published, except for the HRP-2_{PWWP} purification and generation of crystals.

Conflict of Interest

There are no relevant financial or non-financial competing interests to report.

List of references

- [1] T. Kouzarides, "Chromatin Modifications and Their Function," *Cell*, vol. 128, no. 4, pp. 693–705, Feb. 2007.
- [2] Q. Yao, Y. Chen, and X. Zhou, "The roles of microRNAs in epigenetic regulation," *Curr. Opin. Chem. Biol.*, vol. 51, pp. 11–17, Aug. 2019.
- [3] T. B. Miranda and P. A. Jones, "DNA methylation: the nuts and bolts of repression," *J. Cell. Physiol.*, vol. 213, no. 2, pp. 384–390, Nov. 2007.
- [4] S. Saxonov, P. Berg, and D. L. Brutlag, "A genome-wide analysis of CpG dinucleotides in the human genome distinguishes two distinct classes of promoters," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 103, no. 5, pp. 1412–1417, Jan. 2006.
- [5] H. Y. Ryu and M. Hochstrasser, "Histone sumoylation and chromatin dynamics," *Nucleic Acids Res.*, vol. 49, no. 11, pp. 6043–6052, Jun. 2021.
- [6] C. J. Nelson, H. Santos-Rosa, and T. Kouzarides, "Proline Isomerization of Histone H3 Regulates Lysine Methylation and Gene Expression," *Cell*, vol. 126, no. 5, pp. 905–916, Sep. 2006.
- [7] F. S. Howe and J. Mellor, "Proline cis-trans isomerization is influenced by local lysine acetylation-deacetylation," *Microb. Cell*, vol. 1, no. 11, p. 390, Nov. 2014.
- [8] J. Du, L. M. Johnson, S. E. Jacobsen, and D. J. Patel, "DNA methylation pathways and their crosstalk with histone methylation," *Nat. Rev. Mol. Cell Biol.* 2015 169, vol. 16, no. 9, pp. 519–532, Aug. 2015.
- [9] L. Li and Y. Wang, "Cross-talk between the H3K36me3 and H4K16ac histone epigenetic marks in DNA double-strand break repair," *J. Biol. Chem.*, vol. 292, no. 28, pp. 11951–11959, Jul. 2017.
- [10] G. L. Cuthbert *et al.*, "Histone deimination antagonizes arginine methylation," *Cell*, vol. 118, no. 5, pp. 545–553, Sep. 2004.
- [11] M. Altaf, N. Saksouk, and J. Côté, "Histone modifications in response to DNA damage," *Mutat. Res.*, vol. 618, no. 1–2, pp. 81–90, May 2007.
- [12] R. P. Halley-Stott and J. B. Gurdon, "Epigenetic memory in the context of nuclear reprogramming and cancer," *Brief. Funct. Genomics*, vol. 12, no. 3, p. 164, May 2013.
- [13] A. Sriraman, T. K. Debnath, B. Xhemalce, and K. M. Miller, "Making it or breaking it: DNA methylation and genome integrity," *Essays Biochem.*, vol. 64, no. 5, p. 687, Oct. 2020.
- [14] L. D. Moore, T. Le, and G. Fan, "DNA Methylation and Its Basic Function," *Neuropsychopharmacol.* 2013 381, vol. 38, no. 1, pp. 23–38, Jul. 2012.
- [15] S. Aricthota, P. P. Rana, and D. Haldar, "Histone acetylation dynamics in repair of DNA double-strand breaks," *Front. Genet.*, vol. 13, Sep. 2022.
- [16] J. Gräff and L. H. Tsai, "Histone acetylation: molecular mnemonics on the chromatin," *Nat. Rev. Neurosci.*, vol. 14, no. 2, pp. 97–111, Feb. 2013.
- [17] F. Gong and K. M. Miller, "Histone methylation and the DNA damage response," *Mutat. Res. Rev. Mutat. Res.*, vol. 780, pp. 37–47, Apr. 2019.

- [18] E. L. Greer and Y. Shi, "Histone methylation: a dynamic mark in health, disease and inheritance," *Nat. Rev. Genet.*, vol. 13, no. 5, pp. 343–357, May 2012.
- [19] T. Riedl and J. M. Egly, "Phosphorylation in Transcription: The CTD and More," *Gene Expr.*, vol. 9, no. 1–2, p. 3, 2000.
- [20] D. Rossetto, N. Avvakumov, and J. Côté, "Histone phosphorylation: a chromatin modification involved in diverse nuclear events," *Epigenetics*, vol. 7, no. 10, pp. 1098–1108, 2012.
- [21] N. Rahmanian, M. Shokrzadeh, and M. Eskandani, "Recent advances in γ H2AX biomarker-based genotoxicity assays: A marker of DNA damage and repair," *DNA Repair (Amst.)*, vol. 108, Dec. 2021.
- [22] I. Ouni, K. Flick, and P. Kaiser, "Ubiquitin and transcription: The SCFMet30/Met4 pathway, a (protein-) complex issue," *Transcription*, vol. 2, no. 3, p. 135, 2011.
- [23] J. Yu, B. Qin, and Z. Lou, "Ubiquitin and ubiquitin-like molecules in DNA double strand break repair," *Cell Biosci. 2020 101*, vol. 10, no. 1, pp. 1–10, Feb. 2020.
- [24] F. Mattioli and L. Penengo, "Histone Ubiquitination: An Integrative Signaling Platform in Genome Stability," *Trends Genet.*, vol. 37, no. 6, pp. 566–581, Jun. 2021.
- [25] V. Dehennaut, D. Leprince, and T. Lefebvre, "O-GlcNAcylation, an Epigenetic Mark. Focus on the Histone Code, TET Family Proteins, and Polycomb Group Proteins," *Front. Endocrinol. (Lausanne)*, vol. 5, no. SEP, 2014.
- [26] C. Liu and J. Li, "O-GlcNAc: A Sweetheart of the Cell Cycle and DNA Damage Response," *Front. Endocrinol. (Lausanne)*, vol. 9, p. 415, Jul. 2018.
- [27] K. Li and Z. Wang, "Histone crotonylation-centric gene regulation," *Epigenetics Chromatin*, vol. 14, no. 1, Dec. 2021.
- [28] A. Ntorla and J. R. Burgoyne, "The Regulation and Function of Histone Crotonylation," *Front. Cell Dev. Biol.*, vol. 9, p. 729, Apr. 2021.
- [29] R. Martinez-Zamudio and H. C. Ha, "Histone ADP-Ribosylation Facilitates Gene Transcription by Directly Remodeling Nucleosomes," *Mol. Cell. Biol.*, vol. 32, no. 13, p. 2490, Jul. 2012.
- [30] N. Li and J. Chen, "ADP-Ribosylation: Activation, Recognition, and Removal," *Mol. Cells*, vol. 37, no. 1, p. 9, Jan. 2014.
- [31] X. Zhang *et al.*, "Peptidylarginine deiminase 2-catalyzed histone H3 arginine 26 citrullination facilitates estrogen receptor α target gene activation," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 109, no. 33, pp. 13331–13336, Aug. 2012.
- [32] D. Zhu, Y. Zhang, and S. Wang, "Histone citrullination: a new target for tumors," *Mol. Cancer*, vol. 20, no. 1, Dec. 2021.
- [33] C. R. Vakoc, S. A. Mandat, B. A. Olenchok, and G. A. Blobel, "Histone H3 lysine 9 methylation and HP1 γ are associated with transcription elongation through mammalian chromatin," *Mol. Cell*, vol. 19, no. 3, pp. 381–391, Aug. 2005.
- [34] S. Bender *et al.*, "Reduced H3K27me3 and DNA Hypomethylation Are Major Drivers of Gene Expression in K27M Mutant Pediatric High-Grade Gliomas," *Cancer Cell*, vol. 24, no. 5, pp. 660–672, Nov. 2013.
- [35] K. M. Chan *et al.*, "The histone H3.3K27M mutation in pediatric glioma reprograms H3K27 methylation and gene expression," *Genes Dev.*, vol. 27, no. 9, pp. 985–990, May 2013.

- [36] S. R. Daigle *et al.*, "Selective killing of mixed lineage leukemia cells by a potent small-molecule DOT1L inhibitor.," *Cancer Cell*, vol. 20, no. 1, pp. 53–65, Jul. 2011.
- [37] S. R. Daigle *et al.*, "Potent inhibition of DOT1L as treatment of MLL-fusion leukemia.," *Blood*, vol. 122, no. 6, pp. 1017–1025, Aug. 2013.
- [38] E. Alexandrova *et al.*, "Histone Methyltransferase DOT1L as a Promising Epigenetic Target for Treatment of Solid Tumors," *Front. Genet.*, vol. 13, Apr. 2022.
- [39] M. A. Dawson *et al.*, "Inhibition of BET recruitment to chromatin as an effective treatment for MLL-fusion leukaemia," *Nature*, vol. 478, no. 7370, pp. 529–533, 2011.
- [40] H. G. Ozer *et al.*, "BRD4 profiling identifies critical chronic lymphocytic leukemia oncogenic circuits and reveals sensitivity to PLX51107, a novel structurally distinct bet inhibitor.," *Cancer Discov.*, vol. 8, no. 4, pp. 458–477, Apr. 2018.
- [41] M. P. Schwalm and S. Knapp, "BET bromodomain inhibitors," *Curr. Opin. Chem. Biol.*, vol. 68, p. 102148, Jun. 2022.
- [42] S. Ropero and M. Esteller, "The role of histone deacetylases (HDACs) in human cancer," *Mol. Oncol.*, vol. 1, no. 1, pp. 19–25, Jun. 2007.
- [43] J. L. Guerriero *et al.*, "Class IIa HDAC inhibition reduces breast tumours and metastases through anti-tumour macrophages," *Nat. 2017 5437645*, vol. 543, no. 7645, pp. 428–432, Mar. 2017.
- [44] S. Das Pramanik, A. Kumar Halder, U. Mukherjee, D. Kumar, Y. N. Dey, and R. Mogana, "Potential of histone deacetylase inhibitors in the control and regulation of prostate, breast and ovarian cancer," *Front. Chem.*, vol. 10, Aug. 2022.
- [45] M. Esteller, "Cancer epigenomics: DNA methylomes and histone-modification maps," *Nat. Rev. Genet.* 2007 84, vol. 8, no. 4, pp. 286–298, Mar. 2007.
- [46] K. K. Wong, C. H. Lawrie, and T. M. Green, "Oncogenic Roles and Inhibitors of DNMT1, DNMT3A, and DNMT3B in Acute Myeloid Leukaemia," *Biomark. Insights*, vol. 14, May 2019.
- [47] H. Zhang, H. Ying, and X. Wang, "Methyltransferase DNMT3B in leukemia," <https://doi.org/10.1080/10428194.2019.1666377>, vol. 61, no. 2, pp. 263–273, Jan. 2019.
- [48] A. Martisova, J. Holcakova, N. Izadi, R. Sebuyoya, R. Hrstka, and M. Bartosik, "DNA Methylation in Solid Tumors: Functions and Methods of Detection," *Int. J. Mol. Sci.*, vol. 22, no. 8, Apr. 2021.
- [49] A. Lazarenkov and J. L. Sardina, "Dissecting TET2 Regulatory Networks in Blood Differentiation and Cancer," *Cancers 2022, Vol. 14, Page 830*, vol. 14, no. 3, p. 830, Feb. 2022.
- [50] P. V. Peña *et al.*, "Histone H3K4me3 Binding Is Required for the DNA Repair and Apoptotic Activities of ING1 Tumor Suppressor," *J. Mol. Biol.*, vol. 380, no. 2, pp. 303–312, Jul. 2008.
- [51] S. Wei *et al.*, "Histone methylation in DNA repair and clinical practice: new findings during the past 5-years," *J. Cancer*, vol. 9, no. 12, pp. 2072–2081, 2018.
- [52] A. Kumar, N. Kumari, N. Nallabelli, and R. Prasad, "Pathogenic and Therapeutic Role of H3K4 Family of Methylases and Demethylases in Cancers," *Indian J. Clin. Biochem.*, vol. 34, no. 2, p. 123, Apr. 2019.
- [53] K. Park, J. A. Kim, and J. Kim, "Transcriptional regulation by the KMT2 histone H3K4 methyltransferases," *Biochim. Biophys. Acta - Gene Regul. Mech.*, vol. 1863, no. 7, p. 194545, Jul. 2020.

- [54] Y. Ohguchi and H. Ohguchi, "Diverse Functions of KDM5 in Cancer: Transcriptional Repressor or Activator?," *Cancers* 2022, Vol. 14, Page 3270, vol. 14, no. 13, p. 3270, Jul. 2022.
- [55] F. Lohmann *et al.*, "KMT1E mediated H3K9 methylation is required for the maintenance of embryonic stem cells by repressing trophoctoderm differentiation," *Stem Cells*, vol. 28, no. 2, pp. 201–212, Feb. 2010.
- [56] M. Udugama *et al.*, "Histone variant H3.3 provides the heterochromatic H3 lysine 9 trimethylation mark at telomeres," *Nucleic Acids Res.*, vol. 43, no. 21, p. 10227, Dec. 2015.
- [57] D. Nicetto *et al.*, "H3K9me3-heterochromatin loss at protein-coding genes enables developmental lineage specification," *Science*, vol. 363, no. 6424, p. 294, Jan. 2019.
- [58] N. Saha and A. G. Muntean, "Insight into the multi-faceted role of the SUV family of H3K9 methyltransferases in carcinogenesis and cancer progression," *Biochim. Biophys. Acta. Rev. cancer*, vol. 1875, no. 1, Jan. 2021.
- [59] E. Madrazo *et al.*, "Fast H3K9 methylation promoted by CXCL12 contributes to nuclear changes and invasiveness of T-acute lymphoblastic leukemia cells," vol. 41, no. 9, pp. 1324–1336, 2022.
- [60] D. Pasini, A. P. Bracken, M. R. Jensen, E. L. Denchi, and K. Helin, "Suz12 is essential for mouse development and for EZH2 histone methyltransferase activity," *EMBO J.*, vol. 23, no. 20, pp. 4061–4071, Oct. 2004.
- [61] E. Conway, E. Healy, and A. P. Bracken, "PRC2 mediated H3K27 methylations in cellular identity and cancer," *Curr. Opin. Cell Biol.*, vol. 37, pp. 42–48, Dec. 2015.
- [62] J. N. Nichol, D. Dupéré-Richer, T. Ezponda, J. D. Licht, and W. H. Miller, "H3K27 Methylation: A Focal Point of Epigenetic Deregulation in Cancer," *Adv. Cancer Res.*, vol. 131, p. 59, 2016.
- [63] E. T. Wiles and E. U. Selker, "H3K27 methylation: a promiscuous repressive chromatin mark," *Curr. Opin. Genet. Dev.*, vol. 43, pp. 31–37, Apr. 2017.
- [64] X. Zhu *et al.*, "Identification of functional cooperative mutations of SETD2 in human acute leukemia," *Nat. Genet.*, vol. 46, no. 3, p. 287, 2014.
- [65] C. C. Fahey and I. J. Davis, "SETting the Stage for Cancer Development: SETD2 and the Consequences of Lost Methylation," *Cold Spring Harb. Perspect. Med.*, vol. 7, no. 5, p. a026468, May 2017.
- [66] B. J. Klein, K. Krajewski, S. Restrepo, P. W. Lewis, B. D. Strahl, and T. G. Kutateladze, "Recognition of cancer mutations in histone H3K36 by epigenetic writers and readers," *Epigenetics*, vol. 13, no. 7, p. 683, Jul. 2018.
- [67] C. Xiao *et al.*, "H3K36 trimethylation-mediated biological functions in cancer," *Clin. Epigenetics*, vol. 13, no. 1, Dec. 2021.
- [68] U. T. F. Lam, B. K. Y. Tan, J. J. X. Poh, and E. S. Chen, "Structural and functional specificity of H3K36 methylation," *Epigenetics Chromatin* 2022 151, vol. 15, no. 1, pp. 1–20, May 2022.
- [69] Z. Farooq, S. Banday, T. K. Pandita, and M. Altaf, "The many faces of histone H3K79 methylation," *Mutat. Res. Rev. Mutat. Res.*, vol. 768, p. 46, Apr. 2016.
- [70] K. Wood, M. Tellier, and S. Murphy, "DOT1L and H3K79 Methylation in Transcription and Genomic Stability," *Biomolecules*, vol. 8, no. 1, Mar. 2018.
- [71] M. Ljungman, L. Parks, R. Hulbatte, and K. Bedi, "The role of H3K79 methylation in transcription and the DNA damage response," *Mutat. Res. Mutat. Res.*, vol. 780, pp. 48–54,

- Apr. 2019.
- [72] K. Cao *et al.*, “DOT1L-controlled cell-fate determination and transcription elongation are independent of H3K79 methylation,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 117, no. 44, pp. 27365–27373, Nov. 2020.
- [73] O. Arnold, K. Barbosa, A. J. Deshpande, and N. Zhu, “The Role of DOT1L in Normal and Malignant Hematopoiesis,” *Front. Cell Dev. Biol.*, vol. 10, May 2022.
- [74] S. Jørgensen, G. Schotta, and C. S. Sørensen, “Histone H4 Lysine 20 methylation: key player in epigenetic regulation of genomic integrity,” *Nucleic Acids Res.*, vol. 41, no. 5, p. 2797, Mar. 2013.
- [75] K. L. Paquin and N. G. Howlett, “The histone DNA repair code: H4K20me2 makes its mark,” *Mol. Cancer Res.*, vol. 16, no. 9, p. 1335, Sep. 2018.
- [76] A. Z. Corvalan and H. A. Collier, “Physiological Genomics of Cell States and Their Regulation and Single Cell Genomics: Methylation of histone 4’s lysine 20: a critical analysis of the state of the field,” *Physiol. Genomics*, vol. 53, no. 1, p. 22, Jan. 2021.
- [77] E. T. B. Antunes and K. Ottersbach, “The MLL/SET family and haematopoiesis,” *Biochim. Biophys. Acta - Gene Regul. Mech.*, vol. 1863, no. 8, p. 194579, Aug. 2020.
- [78] X. Zhang, W. Novera, Y. Zhang, and L. W. Deng, “MLL5 (KMT2E): structure, function, and clinical relevance,” *Cell. Mol. Life Sci.*, vol. 74, no. 13, pp. 2333–2344, Jul. 2017.
- [79] H. M. Herz, A. Garruss, and A. Shilatifard, “SET for life: biochemical activities and biological functions of SET domain-containing proteins,” *Trends Biochem. Sci.*, vol. 38, no. 12, p. 621, Dec. 2013.
- [80] B. E. Collins, C. B. Greer, B. C. Coleman, and J. D. Sweatt, “Histone H3 lysine K4 methylation and its role in learning and memory,” *Epigenetics Chromatin 2019 121*, vol. 12, no. 1, pp. 1–16, Jan. 2019.
- [81] R. C. Rao and Y. Dou, “Hijacked in cancer: the KMT2 (MLL) family of methyltransferases,” *Nat. Rev. Cancer 2015 156*, vol. 15, no. 6, pp. 334–346, May 2015.
- [82] S. der Poel *et al.*, “Identification of a gene, MLL, that spans the breakpoint in 11q23 translocations associated with human leukemias,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 88, no. 23, pp. 10735–10739, Dec. 1991.
- [83] G. Cimino *et al.*, “Cloning of ALL-1, the locus involved in leukemias with the t(4;11)(q21;q23), t(9;11)(p22;q23), and t(11;19)(q23;p13) chromosome translocations,” *Cancer Res.*, vol. 51, no. 24, pp. 6712–6714, Dec. 1991.
- [84] M. Djabali, L. Selleri, P. Parry, M. Bower, B. D. Young, and G. A. Evans, “A trithorax-like gene is interrupted by chromosome 11q23 translocations in acute leukaemias,” *Nat. Genet.*, vol. 2, no. 2, pp. 113–118, 1992.
- [85] I. Nilson *et al.*, “Exon/intron structure of the human ALL-1 (MLL) gene involved in translocations to chromosomal region 11q23 and acute leukaemias,” *Br. J. Haematol.*, vol. 93, no. 4, pp. 966–972, 1996.
- [86] L. H. Butler, R. Slany, X. Cui, M. L. Cleary, and D. Y. Mason, “The HRX Proto-oncogene Product Is Widely Expressed in Human Tissues and Localizes to Nuclear Structures,” *Blood*, vol. 89, no. 9, pp. 3361–3370, May 1997.
- [87] J. J.-D. Hsieh, P. Ernst, H. Erdjument-Bromage, P. Tempst, and S. J. Korsmeyer, “Proteolytic

- Cleavage of MLL Generates a Complex of N- and C-Terminal Fragments That Confers Protein Stability and Subnuclear Localization,” *Mol. Cell. Biol.*, vol. 23, no. 1, 2003.
- [88] J. J. D. Hsieh, E. H. Y. Cheng, and S. J. Korsmeyer, “Taspase1: A threonine aspartase required for cleavage of MLL and proper HOX gene expression,” *Cell*, vol. 115, no. 3, 2003.
- [89] A. Patel, V. Dharmarajan, V. E. Vought, and M. S. Cosgrove, “On the Mechanism of Multiple Lysine Methylation by the Human Mixed Lineage Leukemia Protein-1 (MLL1) Core Complex,” *J. Biol. Chem.*, vol. 284, no. 36, p. 24242, Sep. 2009.
- [90] Y. Dou *et al.*, “Regulation of MLL1 H3K4 methyltransferase activity by its core components,” *Nat. Struct. Mol. Biol.*, vol. 13, no. 8, pp. 713–719, 2006.
- [91] P. Ernst, J. Wang, M. Huang, R. H. Goodman, and S. J. Korsmeyer, “MLL and CREB bind cooperatively to the nuclear coactivator CREB-binding protein,” *Mol. Cell. Biol.*, vol. 21, no. 7, pp. 2249–2258, Apr. 2001.
- [92] N. K. Goto, T. Zor, M. Martinez-Yamout, H. J. Dyson, and P. E. Wright, “Cooperativity in transcription factor binding to the coactivator CREB-binding protein (CBP). The mixed lineage leukemia protein (MLL) activation domain binds to an allosteric site on the KIX domain,” *J. Biol. Chem.*, vol. 277, no. 45, pp. 43168–43174, Nov. 2002.
- [93] J. Grembecka, A. M. Belcher, T. Hartley, and T. Cierpicki, “Molecular Basis of the Mixed Lineage Leukemia-Menin Interaction: IMPLICATIONS FOR TARGETING MIXED LINEAGE LEUKEMIAS,” *J. Biol. Chem.*, vol. 285, no. 52, pp. 40690–40698, Dec. 2010.
- [94] A. Yokoyama and M. L. Cleary, “Menin Critically Links MLL Proteins with LEDGF on Cancer-Associated Target Genes,” *Cancer Cell*, vol. 14, no. 1, pp. 36–46, Jul. 2008.
- [95] H. Méreau *et al.*, “Impairing MLL-fusion gene-mediated transformation by dissecting critical interactions with the lens epithelium-derived growth factor (LEDGF/p75),” *Leukemia*, vol. 27, no. 6, pp. 1245–1253, 2013.
- [96] J. Huang *et al.*, “The same pocket in menin binds both MLL and JUND but has opposite effects on transcription,” *Nature*, vol. 482, no. 7386, pp. 542–546, 2012.
- [97] K. Cermakova *et al.*, “Validation and Structural Characterisation of the LEDGF/p75-MLL Interface as a New Target for the Treatment of MLL-Dependent Leukaemia,” *Cancer Res.*, no. 9, pp. 5139–5152, 2014.
- [98] R. K. Slany, C. Lavau, and M. L. Cleary, “The oncogenic capacity of HRX-ENL requires the transcriptional transactivation activity of ENL and the DNA binding motifs of HRX,” *Mol. Cell. Biol.*, vol. 18, no. 1, pp. 122–129, 1998.
- [99] K. Fair, M. Anderson, E. Bulanova, H. Mi, M. Tropschug, and M. O. Diaz, “Protein Interactions of the MLL PHD Fingers Modulate MLL Target Gene Regulation in Human Cells,” *Mol. Cell. Biol.*, vol. 21, no. 10, p. 3589, May 2001.
- [100] T. Cierpicki *et al.*, “Structure of the MLL CXXC domain-DNA complex and its functional role in MLL-AF9 leukemia,” *Nat. Struct. & Mol. Biol.*, vol. 17, no. 1, pp. 62–68, 2010.
- [101] P.-Y. Chang *et al.*, “Binding of the MLL PHD3 finger to histone H3K4me3 is required for MLL-dependent gene transcription,” *J. Mol. Biol.*, vol. 400, no. 2, pp. 137–144, 2010.
- [102] A. G. Muntean *et al.*, “The PAF complex synergizes with MLL fusion proteins at HOX loci to promote leukemogenesis,” *Cancer Cell*, vol. 17, no. 6, pp. 609–621, 2010.
- [103] T. A. Milne *et al.*, “Multiple interactions recruit MLL1 and MLL1 fusion proteins to the HOXA9

- locus in leukemogenesis.," *Mol. Cell*, vol. 38, no. 6, pp. 853–863, 2010.
- [104] J. Tan, A. G. Muntean, J. L. Hess, J. Tan, A. G. Muntean, and J. L. Hess, "PAFc, a Key Player in MLL-rearranged Leukemogenesis," *Oncotarget*, vol. 1, no. 6, pp. 461–465, Oct. 2010.
- [105] B. D. Yu, J. L. Hess, S. E. Horning, G. A. J. Brown, and S. J. Korsmeyer, "Altered Hox expression and segmental identity in Mll-mutant mice," *Nature*, vol. 378, no. 6556, pp. 505–508, Nov. 1995.
- [106] K. A. McMahon *et al.*, "Mll Has a Critical Role in Fetal and Adult Hematopoietic Stem Cell Self-Renewal," *Cell Stem Cell*, vol. 1, no. 3, pp. 338–345, Sep. 2007.
- [107] T. Gan, C. D. Jude, K. Zaffuto, and P. Ernst, "Developmentally induced Mll1 loss reveals defects in postnatal haematopoiesis," *Leukemia*, 2010.
- [108] T. A. Milne *et al.*, "MLL Targets SET Domain Methyltransferase Activity to Hox Gene Promoters," *Mol. Cell*, vol. 10, no. 5, pp. 1107–1117, Nov. 2002.
- [109] E. L. Artinger *et al.*, "An MLL-dependent network sustains hematopoiesis.," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 110, no. 29, pp. 12000–5, 2013.
- [110] C. N. Vallianatos and S. Iwase, "Disrupted intricacy of histone H3K4 methylation in neurodevelopmental disorders," <http://dx.doi.org/10.2217/epi.15.1>, vol. 7, no. 3, pp. 503–518, Jun. 2015.
- [111] W. D. Jones *et al.*, "De Novo Mutations in MLL Cause Wiedemann-Steiner Syndrome," *Am. J. Hum. Genet.*, vol. 91, no. 2, p. 358, Aug. 2012.
- [112] S. P. Strom *et al.*, "De Novo variants in the KMT2A (MLL) gene causing atypical Wiedemann-Steiner syndrome in two unrelated individuals identified by clinical exome sequencing," *BMC Med. Genet.*, vol. 15, no. 1, pp. 1–8, May 2014.
- [113] S. Kudithipudi and A. Jeltsch, "Role of somatic cancer mutations in human protein lysine methyltransferases," *Biochim. Biophys. Acta - Rev. Cancer*, vol. 1846, no. 2, pp. 366–379, Dec. 2014.
- [114] M. C. Béné, "Biphenotypic, bilineal, ambiguous or mixed lineage: strange leukemias!," *Haematologica*, vol. 94, no. 7, p. 891, Jul. 2009.
- [115] J. L. Abkowitz, S. N. Catlin, and P. Gutterop, "Evidence that hematopoiesis may be a stochastic process in vivo," *Nat. Med.* 1996 22, vol. 2, no. 2, pp. 190–197, 1996.
- [116] J. Xu, Y. Wang, P. Gutterop, and J. L. Abkowitz, "Visualizing hematopoiesis as a stochastic process," *Blood Adv.*, vol. 2, no. 20, p. 2637, Oct. 2018.
- [117] H. M. McRae, A. K. Voss, and T. Thomas, "Are transplantable stem cells required for adult hematopoiesis?," *Exp. Hematol.*, vol. 75, pp. 1–10, Jul. 2019.
- [118] C. E. Muller-Sieburg, R. H. Cho, L. Karlsson, J. F. Huang, and H. B. Sieburg, "Myeloid-biased hematopoietic stem cells have extensive self-renewal capacity but generate diminished lymphoid progeny with impaired IL-7 responsiveness," *Blood*, vol. 103, no. 11, pp. 4111–4118, Jun. 2004.
- [119] Y. Morita, H. Ema, and H. Nakauchi, "Heterogeneity and hierarchy within the most primitive hematopoietic stem cell compartment," *J. Exp. Med.*, vol. 207, no. 6, p. 1173, Jun. 2010.
- [120] R. Lu, A. Czechowicz, J. Seita, D. Jiang, and I. L. Weissman, "Clonal-level lineage commitment pathways of hematopoietic stem cells in vivo," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 116, no. 4, pp. 1447–1456, Jan. 2019.

- [121] R. Yamamoto *et al.*, "Clonal Analysis Unveils Self-Renewing Lineage-Restricted Progenitors Generated Directly from Hematopoietic Stem Cells," *Cell*, vol. 154, no. 5, pp. 1112–1126, Aug. 2013.
- [122] H. Iwasaki *et al.*, "The order of expression of transcription factors directs hierarchical specification of hematopoietic lineages," *Genes Dev.*, vol. 20, no. 21, pp. 3010–3021, Nov. 2006.
- [123] C. L. Hsu, A. G. King-Fleischman, A. Y. Lai, Y. Matsumoto, I. L. Weissman, and M. Kondo, "Antagonistic effect of CCAAT enhancer-binding protein- α and Pax5 in myeloid or lymphoid lineage choice in common lymphoid progenitors," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 103, no. 3, pp. 672–677, Jan. 2006.
- [124] B. V. Balgobind *et al.*, "Novel prognostic subgroups in childhood 11q23/MLL-rearranged acute myeloid leukemia: results of an international retrospective study," *Blood*, vol. 114, no. 12, pp. 2489–2496, Sep. 2009.
- [125] T. Szczepański, C. J. Harrison, and J. J. M. van Dongen, "Genetic aberrations in paediatric acute leukaemias and implications for management of patients," *Lancet Oncol.*, vol. 11, no. 9, pp. 880–889, Sep. 2010.
- [126] M. Khan, R. Siddiqi, and K. Naqvi, "An update on classification, genetics, and clinical approach to mixed phenotype acute leukemia (MPAL)," *Ann. Hematol.*, vol. 97, no. 6, pp. 945–953, Jun. 2018.
- [127] A. Schmitt-Gräff, H. Jürgens, A. Reifenhäuser, D. Schwamborn, and U. Göbel, "Childhood biphenotypic leukemia: Detection of mixed lymphoid and myeloid populations in bone marrow specimens," *Hum. Pathol.*, vol. 19, no. 6, pp. 651–656, Jun. 1988.
- [128] C. -H Pui *et al.*, "Acute leukaemia with mixed lymphoid and myeloid phenotype," *Br. J. Haematol.*, vol. 56, no. 1, pp. 121–130, 1984.
- [129] S. Stass, J. Mirro, S. Melvin, C. H. Pui, S. B. Murphy, and D. WILLIAMS, "Lineage switch in acute leukemia.," *Blood*, vol. 64, no. 3, pp. 701–706, Sep. 1984.
- [130] C. Constantin *et al.*, "ALL-503 Lineage Switch Upon Relapse of Acute Leukemia: A Rare Event Which Heralds Poor Clinical Outcomes," *Clin. Lymphoma. Myeloma Leuk.*, vol. 22 Suppl 2, pp. S206–S207, Oct. 2022.
- [131] J. K. Choi, "Acute Undifferentiated Leukemia and Mixed-Phenotype Acute Leukemias," *Hematop. A Vol. Ser. Found. Diagnostic Pathol.*, pp. 481-487.e1, Jan. 2018.
- [132] C. Meyer *et al.*, "The MLL recombinome of acute leukemias in 2017," *Leukemia*, vol. 32, no. 2, pp. 273–284, Feb. 2018.
- [133] N. R. McCabe *et al.*, "Cloning of cDNAs of the MLL gene that detect DNA rearrangements and altered RNA transcripts in human leukemic cells with 11q23 translocations," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 89, no. 24, 1992.
- [134] C. Meyer *et al.*, "Human MLL/KMT2A gene exhibits a second breakpoint cluster region for recurrent MLL–USP2 fusions," *Leukemia*, vol. 33, no. 9. 2019.
- [135] S. Takahashi and A. Yokoyama, "The molecular functions of common and atypical MLL fusion protein complexes," *Biochim. Biophys. Acta - Gene Regul. Mech.*, vol. 1863, no. 7, p. 194548, Jul. 2020.
- [136] B. M. Peterlin and D. H. Price, "Controlling the elongation phase of transcription with P-TEFb.," *Mol. Cell*, vol. 23, no. 3, pp. 297–305, Aug. 2006.

- [137] A. Yokoyama, M. Lin, A. Naresh, I. Kitabayashi, and M. L. Cleary, "A higher-order complex containing AF4 and ENL family proteins with P-TEFb facilitates oncogenic and physiologic MLL-dependent transcription.," *Cancer Cell*, vol. 17, no. 2, pp. 198–212, Feb. 2010.
- [138] C. Lin *et al.*, "AFF4, a component of the ELL/P-TEFb elongation complex and a shared subunit of MLL chimeras, can link transcription elongation to leukemia.," *Mol. Cell*, vol. 37, no. 3, pp. 429–437, Feb. 2010.
- [139] Q. Feng *et al.*, "Methylation of H3-Lysine 79 Is Mediated by a New Family of HMTases without a SET Domain," *Curr. Biol.*, vol. 12, no. 12, pp. 1052–1058, Jun. 2002.
- [140] F. Sarno, A. Nebbioso, and L. Altucci, "DOT1L: a key target in normal chromatin remodelling and in mixed-lineage leukaemia treatment," *Epigenetics*, vol. 15, no. 5, p. 439, May 2020.
- [141] R. K. Slany, "The molecular mechanics of mixed lineage leukemia," *Oncogene*, vol. 35, no. 40, pp. 5215–5223, Oct. 2016.
- [142] B. Argiropoulos and R. K. Humphries, "Hox genes in hematopoiesis and leukemogenesis," *Oncogene*, vol. 26, no. 47, pp. 6766–6776, 2007.
- [143] A. Yokoyama, T. C. P. Somervaille, K. S. Smith, O. Rozenblatt-Rosen, M. Meyerson, and M. L. Cleary, "The menin tumor suppressor protein is an essential oncogenic cofactor for MLL-associated leukemogenesis.," *Cell*, vol. 123, no. 2, pp. 207–218, 2005.
- [144] C. Caslini, Z. Yang, M. El-Osta, T. A. Milne, R. K. Slany, and J. L. Hess, "Interaction of MLL Amino Terminal Sequences with Menin Is Required for Transformation," *Cancer Res.*, vol. 67, no. 15, pp. 7275–7283, Aug. 2007.
- [145] C. Steudel *et al.*, "Comparative analysis of MLL partial tandem duplication and FLT3 internal tandem duplication mutations in 956 adult patients with acute myeloid leukemia," *Genes Chromosom. Cancer*, vol. 37, no. 3, pp. 237–251, Jul. 2003.
- [146] A. Sakhdari, Z. Tang, C. Y. Ok, C. E. Bueso-Ramos, L. J. Medeiros, and Y. O. Huh, "Homogeneously staining region (hsr) on chromosome 11 is highly specific for KMT2A amplification in acute myeloid leukemia (AML) and myelodysplastic syndrome (MDS)," *Cancer Genet.*, vol. 238, pp. 18–22, Oct. 2019.
- [147] G. Tang *et al.*, "MLL gene amplification in acute myeloid leukemia and myelodysplastic syndromes is associated with characteristic clinicopathological findings and TP53 gene mutation.," *Hum. Pathol.*, vol. 46, no. 1, pp. 65–73, Jan. 2015.
- [148] R. W. Maitta, L. A. Cannizzaro, and K. H. Ramesh, "Association of MLL amplification with poor outcome in acute myeloid leukemia," *Cancer Genet. Cytogenet.*, vol. 192, no. 1, pp. 40–43, Jul. 2009.
- [149] A. K. Andersson *et al.*, "The landscape of somatic mutations in infant MLL-rearranged acute lymphoblastic leukemias," *Nat. Genet.*, vol. 47, no. 4, pp. 330–337, 2015.
- [150] C. G. Mullighan *et al.*, "Genome-wide analysis of genetic alterations in acute lymphoblastic leukaemia," *Nat. 2007 4467137*, vol. 446, no. 7137, pp. 758–764, Mar. 2007.
- [151] I. Radtke *et al.*, "Genomic analysis reveals few genetic alterations in pediatric acute myeloid leukemia," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 106, no. 31, pp. 12944–12949, Aug. 2009.
- [152] O. Britten, D. Ragusa, S. Tosi, and Y. M. Kamel, "MLL-Rearranged Acute Leukemia with t(4;11)(q21;q23)—Current Treatment Options. Is There a Role for CAR-T Cell Therapy?," *Cells 2019, Vol. 8, Page 1341*, vol. 8, no. 11, p. 1341, Oct. 2019.

- [153] O. V. Zotova *et al.*, “11q23/MLL rearrangements in adult acute leukemia,” *Exp. Oncol.*, vol. 43, no. 3, pp. 229–233, 2021.
- [154] H. Super *et al.*, “Rearrangements of the MLL Gene in Therapy-Related Acute Myeloid Leukemia in Patients Previously Treated With Agents Targeting DNA-Topoisomerase II,” *Blood*, vol. 82, no. 12, pp. 3705–3711, Dec. 1993.
- [155] J. M. Hilden *et al.*, “Analysis of prognostic factors of acute lymphoblastic leukemia in infants: report on CCG 1953 from the Children’s Oncology Group,” *Blood*, vol. 108, no. 2, pp. 441–451, Jul. 2006.
- [156] “INTERFANT-06 INTERNATIONAL COLLABORATIVE TREATMENT PROTOCOL FOR INFANTS UNDER ONE YEAR WITH ACUTE LYMPHOBLASTIC OR BIPHENOTYPIC LEUKEMIA,” 2015.
- [157] R. Pieters *et al.*, “Outcome of infants younger than 1 year with acute lymphoblastic leukemia treated with the interfant-06 protocol: Results from an international phase III randomized study,” *J. Clin. Oncol.*, vol. 37, no. 25, pp. 2246–2256, Sep. 2019.
- [158] A. Shi *et al.*, “Structural insights into inhibition of the bivalent menin-MLL interaction by small molecules in leukemia,” *Blood*, vol. 120, no. 23, pp. 4461–4469, Nov. 2012.
- [159] G. F., M. F., and B. ML, “Multiple Endocrine Neoplasia Type 1,” 1993.
- [160] S. Klossowski *et al.*, “Menin inhibitor MI-3454 induces remission in MLL1-rearranged and NPM1-mutated models of leukemia,” *J. Clin. Invest.*, vol. 130, no. 2, pp. 981–997, Feb. 2020.
- [161] A. V. Krivtsov *et al.*, “A Menin-MLL Inhibitor Induces Specific Chromatin Changes and Eradicates Disease in Models of MLL-Rearranged Leukemia,” *Cancer Cell*, vol. 36, no. 6, pp. 660–673.e11, Dec. 2019.
- [162] M. Zhang *et al.*, “Discovery of M-1121 as an Orally Active Covalent Inhibitor of Menin-MLL Interaction Capable of Achieving Complete and Long-Lasting Tumor Regression,” *J. Med. Chem.*, vol. 64, no. 14, pp. 10333–10349, Jul. 2021.
- [163] C. Dafflon *et al.*, “Complementary activities of DOT1L and Menin inhibitors in MLL-rearranged leukemia,” *Leukemia*, vol. 31, no. 6, pp. 1269–1277, Jun. 2017.
- [164] B. W. Katona *et al.*, “Combined Menin and EGFR Inhibitors Synergize to Suppress Colorectal Cancer via EGFR-Independent and Calcium-Mediated Repression of SKP2 Transcription,” *Cancer Res.*, vol. 79, no. 9, pp. 2195–2207, 2019.
- [165] “First in Human Study of KO-539 in Relapsed or Refractory Acute Myeloid Leukemia - Full Text View - ClinicalTrials.gov.” [Online]. Available: <https://www.clinicaltrials.gov/ct2/show/NCT04067336?cond=menin&draw=2&rank=1>. [Accessed: 28-Mar-2022].
- [166] “A Study of JNJ-75276617 in Participants With Acute Leukemia - Full Text View - ClinicalTrials.gov.” [Online]. Available: <https://www.clinicaltrials.gov/ct2/show/NCT04811560?cond=menin&draw=2&rank=2>. [Accessed: 28-Mar-2022].
- [167] “A Study of SNDX-5613 in R/R Leukemias Including Those With an MLLr/KMT2A Gene Rearrangement or NPM1 Mutation - Full Text View - ClinicalTrials.gov.” [Online]. Available: <https://www.clinicaltrials.gov/ct2/show/NCT04065399?term=sndx-5613&draw=2&rank=1>. [Accessed: 28-Mar-2022].
- [168] “Study of BMF-219, in Adult Patients With Acute Leukemia, Diffuse Large B-Cell Lymphoma and Multiple Myeloma - Full Text View - ClinicalTrials.gov.” [Online]. Available:

- <https://www.clinicaltrials.gov/ct2/show/NCT05153330?term=bmf-219&draw=2&rank=1>. [Accessed: 28-Mar-2022].
- [169] “A Study of DSP-5336 in Relapsed/Refractory AML/ ALL With or Without MLL Rearrangement or NPM1 Mutation - Full Text View - ClinicalTrials.gov.” [Online]. Available: <https://www.clinicaltrials.gov/ct2/show/NCT04988555>. [Accessed: 27-May-2022].
- [170] K. M. Bernt *et al.*, “MLL-rearranged leukemia is dependent on aberrant H3K79 methylation by DOT1L,” *Cancer Cell*, vol. 20, no. 1, pp. 66–78, 2011.
- [171] F. Perner *et al.*, “Novel inhibitors of the histone methyltransferase DOT1L show potent antileukemic activity in patient-derived xenografts,” *Blood*, vol. 136, no. 17, pp. 1983–1988, Oct. 2020.
- [172] E. M. Stein *et al.*, “The DOT1L inhibitor pinometostat reduces H3K79 methylation and has modest clinical activity in adult acute leukemia,” *Blood*, vol. 131, no. 24, pp. 2662–2669, Jun. 2018.
- [173] Z. Yang *et al.*, “Recruitment of P-TEFb for stimulation of transcriptional elongation by the bromodomain protein Brd4,” *Mol. Cell*, vol. 19, no. 4, pp. 535–545, Aug. 2005.
- [174] K. J. Moon, K. Mochizuki, M. Zhou, H. S. Jeong, J. N. Brady, and K. Ozato, “The bromodomain protein Brd4 is a positive regulatory component of P-TEFb and stimulates RNA polymerase II-dependent transcription,” *Mol. Cell*, vol. 19, no. 4, pp. 523–534, Aug. 2005.
- [175] C. Y. Fong *et al.*, “BET inhibitor resistance emerges from leukaemia stem cells,” *Nature*, vol. 525, no. 7570, pp. 538–542, Sep. 2015.
- [176] K. Amari *et al.*, “The CDK4/6-UCHL5-BRD4 axis confers resistance to BET inhibitors in MLL-rearranged leukemia cells by suppressing BRD4 protein degradation,” *Biochem. Biophys. Res. Commun.*, vol. 588, pp. 147–153, Jan. 2022.
- [177] K. Liang *et al.*, “Therapeutic targeting of MLL degradation pathways in MLL-rearranged leukemia,” *Cell*, vol. 168, no. 1–2, p. 59, Jan. 2017.
- [178] Z. Zhao *et al.*, “Regulation of MLL/COMPASS stability through its proteolytic cleavage by *taspase1* as a possible approach for clinical therapy of leukemia,” *Genes Dev.*, vol. 33, no. 1–2, pp. 61–74, Jan. 2019.
- [179] S. El Ashkar *et al.*, “LEDGF/p75 is dispensable for hematopoiesis but essential for MLL-rearranged leukemogenesis,” *Blood*, vol. 131, no. 1, pp. 95–107, Jan. 2018.
- [180] F. Dietz, S. Franken, K. Yoshida, H. Nakamura, J. Kappler, and V. Gieselmann, “The family of hepatoma-derived growth factor proteins: characterization of a new member HRP-4 and classification of its subfamilies.,” *Biochem. J.*, vol. 366, no. Pt 2, pp. 491–500, Sep. 2002.
- [181] S. Desfarges, A. Abderrahmani, B. Hernández-Novoa, M. Munoz, and A. Ciuffi, “LEDGF/p75 TATA-Less Promoter Is Driven by the Transcription Factor Sp1,” *J. Mol. Biol.*, vol. 414, no. 2, pp. 177–193, Nov. 2011.
- [182] J. O. Eidahl *et al.*, “Structural basis for high-affinity binding of LEDGF PWWP to mononucleosomes.,” *Nucleic Acids Res.*, vol. 41, no. 6, pp. 3924–3936, Apr. 2013.
- [183] R. Van Nuland *et al.*, “Nucleosomal DNA binding drives the recognition of H3K36-methylated nucleosomes by the PSIP1-PWWP domain,” *Epigenetics and Chromatin*, vol. 6, no. 1, p. 12, May 2013.
- [184] L. Zhu *et al.*, “ASH1L Links Histone H3 Lysine 36 Dimethylation to MLL Leukemia.,” *Cancer*

- Discov.*, vol. 6, no. 7, pp. 770–783, 2016.
- [185] E. J. Wagner and P. B. Carpenter, “Understanding the language of Lys36 methylation at histone H3,” *Nat. Rev. Mol. Cell Biol.*, vol. 13, no. 2, pp. 115–126, Feb. 2012.
- [186] H. Wang, L. Farnung, C. Dienemann, and P. Cramer, “Structure of H3K36-methylated nucleosome-PWWP complex reveals multivalent cross-gyre binding,” *Nat. Struct. Mol. Biol.*, vol. 27, no. 1, pp. 8–13, Jan. 2020.
- [187] M.-C. Shun *et al.*, “Identification and characterization of PWWP domain residues critical for LEDGF/p75 chromatin binding and human immunodeficiency virus type 1 infectivity,” *J. Virol.*, vol. 82, no. 23, pp. 11555–11567, Dec. 2008.
- [188] G. Maertens, P. Cherepanov, Z. Debyser, Y. Engelborghs, and A. Engelman, “Identification and characterization of a functional nuclear localization signal in the HIV-1 integrase interactor LEDGF/p75,” *J. Biol. Chem.*, vol. 279, no. 32, pp. 33421–33429, Aug. 2004.
- [189] F. Turlure, G. Maertens, S. Rahman, P. Cherepanov, and A. Engelman, “A tripartite DNA-binding element, comprised of the nuclear localization signal and two AT-hook motifs, mediates the association of LEDGF/p75 with chromatin in vivo,” *Nucleic Acids Res.*, vol. 34, no. 5, p. 1653, 2006.
- [190] M. McNeely *et al.*, “In Vitro DNA Tethering of HIV-1 Integrase by the Transcriptional Coactivator LEDGF/p75,” *J. Mol. Biol.*, vol. 410, no. 5, pp. 811–830, Jul. 2011.
- [191] K. M. Tsutsui, K. Sano, O. Hosoya, T. Miyamoto, and K. Tsutsui, “Nuclear protein LEDGF/p75 recognizes supercoiled DNA by a novel DNA-binding domain,” *Nucleic Acids Res.*, vol. 39, no. 12, pp. 5067–5081, Jul. 2011.
- [192] M. Llano, M. Vanegas, N. Hutchins, D. Thompson, S. Delgado, and E. M. Poeschla, “Identification and Characterization of the Chromatin-binding Domains of the HIV-1 Integrase Interactor LEDGF/p75,” *J. Mol. Biol.*, vol. 360, no. 4, pp. 760–773, Jul. 2006.
- [193] L. S. Leoh *et al.*, “The stress oncoprotein LEDGF/p75 interacts with the methyl CpG binding protein MeCP2 and influences its transcriptional activity,” *Mol. Cancer Res.*, vol. 10, no. 3, pp. 378–391, Mar. 2012.
- [194] R. Li *et al.*, “Misregulation of Alternative Splicing in a Mouse Model of Rett Syndrome,” *PLOS Genet.*, vol. 12, no. 6, p. e1006129, Jun. 2016.
- [195] P. Cherepanov, Z. Y. J. Sun, S. Rahman, G. Maertens, G. Wagner, and A. Engelman, “Solution structure of the HIV-1 integrase-binding domain in LEDGF/p75,” *Nat. Struct. Mol. Biol.*, vol. 12, no. 6, pp. 526–532, Jun. 2005.
- [196] P. Tesina *et al.*, “Multiple cellular proteins interact with LEDGF/p75 through a conserved unstructured consensus motif,” *Nat. Commun.*, vol. 6, p. 7968, 2015.
- [197] G. N. Maertens, P. Cherepanov, and A. Engelman, “Transcriptional co-activator p75 binds and tethers the Myc-interacting protein JPO2 to chromatin,” *J. Cell Sci.*, vol. 119, no. 12, pp. 2563–2571, Jun. 2006.
- [198] K. Bartholomeeusen *et al.*, “Differential Interaction of HIV-1 Integrase and JPO2 with the C Terminus of LEDGF/p75,” *J. Mol. Biol.*, vol. 372, no. 2, pp. 407–421, Sep. 2007.
- [199] K. Bartholomeeusen *et al.*, “Lens epithelium-derived growth factor/p75 interacts with the transposase-derived DDE domain of PogZ,” *J. Biol. Chem.*, vol. 284, no. 17, pp. 11467–11477, Apr. 2009.

- [200] S. Hughes, V. Jenkins, M. J. Dar, A. Engelman, and P. Cherepanov, "Transcriptional co-activator LEDGF interacts with Cdc7-activator of S-phase kinase (ASK) and stimulates its enzymatic activity.," *J. Biol. Chem.*, vol. 285, no. 1, pp. 541–54, Jan. 2010.
- [201] S. Sharma *et al.*, "Affinity switching of the LEDGF/p75 interactome is governed by kinase-dependent phosphorylation.," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 115, no. 30, pp. E7053–E7062, Jul. 2018.
- [202] K. Cermakova *et al.*, "Validation and structural characterization of the LEDGF/p75-MLL interface as a new target for the treatment of MLL-dependent leukemia.," *Cancer Res.*, vol. 74, no. 18, pp. 5139–5151, Sep. 2014.
- [203] K. Cermakova *et al.*, "A ubiquitous disordered protein interaction module orchestrates transcription elongation," *Science (80-.)*, vol. 374, no. 6571, pp. 1113–1121, Nov. 2021.
- [204] W. Vanderlinden, J. Lipfert, J. Demeulemeester, Z. Debyser, and S. De Feyter, "Structure, mechanics, and binding mode heterogeneity of LEDGF/p75–DNA nucleoprotein complexes revealed by scanning force microscopy," *Nanoscale*, vol. 6, no. 9, pp. 4611–4619, Apr. 2014.
- [205] V. Lux *et al.*, "Molecular Mechanism of LEDGF/p75 Dimerization," *Structure*, vol. 28, no. 12, pp. 1288–1299.e7, Dec. 2020.
- [206] H. G. Sutherland *et al.*, "Disruption of Ledgef/Psip1 Results in Perinatal Mortality and Homeotic Skeletal Transformations," *Mol. Cell. Biol.*, vol. 26, no. 19, pp. 7201–7210, Oct. 2006.
- [207] H. Ge, Y. Si, and R. G. Roeder, "Isolation of cDNAs encoding novel transcription coactivators p52 and p75 reveals an alternate regulatory mechanism of transcriptional activation," *EMBO J.*, vol. 17, no. 22, pp. 6723–6729, Nov. 1998.
- [208] M. Nakamura, D. P. Singh, E. Kubo, L. T. Chylack Jr, and T. Shinohara, "LEDGF: Survival of Embryonic Chick Retinal Photoreceptor Cells," *Invest. Ophthalmol. Vis. Sci.*, vol. 41, no. 5, pp. 1168–1175, Apr. 2000.
- [209] D. P. Singh, N. Fatma, A. Kimura, L. T. Chylack, and T. Shinohara, "LEDGF binds to heat shock and stress-related element to activate the expression of stress-related genes," *Biochem. ...*, vol. 283, no. 4, pp. 943–955, 2001.
- [210] J. H. Shin, C. S. Piao, C. M. Lim, and J. K. Lee, "LEDGF binding to stress response element increases alphaB-crystallin expression in astrocytes with oxidative stress," *Neurosci. Lett.*, vol. 435, no. 2, pp. 131–136, Apr. 2008.
- [211] D. P. Singh, N. Ohguro, L. T. Chylack, and T. Shinohara, "Lens epithelium-derived growth factor: Increased resistance to thermal and oxidative stresses," *Investig. Ophthalmol. Vis. Sci.*, vol. 40, no. 7, pp. 1444–1451, 1999.
- [212] J. H. Shin, S. W. Kim, C. M. Lim, J. Y. Jeong, C. S. Piao, and J. K. Lee, "alphaB-crystallin suppresses oxidative stress-induced astrocyte apoptosis by inhibiting caspase-3 activation," *Neurosci. Res.*, vol. 64, no. 4, pp. 355–361, Aug. 2009.
- [213] N. Fatma, D. P. Singh, T. Shinohara, and L. T. Chylack, "Transcriptional regulation of the antioxidant protein 2 gene, a thiol-specific antioxidant, by lens epithelium-derived growth factor to protect cells from oxidative stress," *J. Biol. Chem.*, vol. 276, no. 52, pp. 48899–48907, Dec. 2001.
- [214] E. Kubo *et al.*, "Transactivation of involucrin, a marker of differentiation in keratinocytes, by lens epithelium-derived growth factor (LEDGF)," *J. Mol. Biol.*, vol. 320, no. 5, pp. 1053–1063, 2002.

- [215] N. Fatma, E. Kubo, L. T. Chylack, T. Shinohara, Y. Akagi, and D. P. Singh, "LEDGF regulation of alcohol and aldehyde dehydrogenases in lens epithelial cells: stimulation of retinoic acid production and protection from ethanol toxicity," *Am. J. Physiol. Cell Physiol.*, vol. 287, no. 2, Aug. 2004.
- [216] X. Wu, T. Daniels, C. Molinaro, M. B. Lilly, and C. A. Casiano, "Caspase cleavage of the nuclear autoantigen LEDGF/p75 abrogates its pro-survival function: Implications for autoimmunity in atopic disorders," *Cell Death Differ.*, vol. 9, no. 9, pp. 915–925, Sep. 2002.
- [217] M. Daugaard *et al.*, "LEDGF (p75) promotes DNA-end resection and homologous recombination," *Nat. Struct. Mol. Biol.*, vol. 19, no. 8, pp. 803–810, Aug. 2012.
- [218] V. Liedtke *et al.*, "LEDGF/p75 is required for an efficient dna damage response," *Int. J. Mol. Sci.*, vol. 22, no. 11, Jun. 2021.
- [219] G. LeRoy *et al.*, "LEDGF and HDGF2 relieve the nucleosome-induced barrier to transcription in differentiated cells," *Sci. Adv.*, vol. 5, no. 10, p. eaay3068, Oct. 2019.
- [220] B. A. Desimmie *et al.*, "Phage Display-directed Discovery of LEDGF/p75 Binding Cyclic Peptide Inhibitors of HIV Replication," *Mol. Ther.*, vol. 20, no. 11, pp. 2064–2075, 2012.
- [221] T. Daniels *et al.*, "Antinuclear autoantibodies in prostate cancer: immunity to LEDGF/p75, a survival protein highly expressed in prostate tumors and cleaved during apoptosis," *Prostate*, vol. 62, no. 1, pp. 14–26, Jan. 2005.
- [222] L. Ríos-Colón *et al.*, "Targeting the stress oncoprotein LEDGF/p75 to sensitize chemoresistant prostate cancer cells to taxanes," *Oncotarget*, vol. 8, no. 15, pp. 24915–24931, 2017.
- [223] G. L. Ortiz-Hernandez *et al.*, "The LEDGF/p75 Integrase Binding Domain Interactome Contributes to the Survival, Clonogenicity, and Tumorsphere Formation of Docetaxel-Resistant Prostate Cancer Cells," *Cells*, vol. 10, no. 10, Oct. 2021.
- [224] D. K. Singh *et al.*, "PSIP1/p75 promotes tumorigenicity in breast cancer cells by promoting the transcription of cell cycle genes," *Carcinogenesis*, vol. 38, no. 10, pp. 966–975, Oct. 2017.
- [225] X. H. Yin, Z. Q. Wang, Q. H. Guo, H. Wu, and M. Shi, "Overexpressed LEDGF is a novel biomarker of poor prognosis in patients with cervical cancer," *Eur. J. Gynaecol. Oncol.*, vol. 38, no. 2, pp. 245–250, Apr. 2017.
- [226] A. Basu *et al.*, "Expression of the stress response oncoprotein LEDGF/p75 in human cancer: A study of 21 tumor types," *PLoS One*, vol. 7, no. 1, Jan. 2012.
- [227] T. S. T. sheng Huang *et al.*, "LEDGF/p75 has increased expression in blasts from chemotherapy-resistant human acute myelogenous leukemia patients and protects leukemia cells from apoptosis in vitro," *Mol. Cancer*, vol. 6, Apr. 2007.
- [228] A. Canella *et al.*, "LEDGF/p75-mediated chemoresistance of mixed-lineage leukemia involves cell survival pathways and super enhancer activators," *Cancer Gene Ther.*, vol. 29, no. 2, pp. 133–140, Feb. 2022.
- [229] T. S. Y. Chan, C. Hawkins, J. R. Krieger, C. J. McClade, and A. Huang, "JPO2/CDCA7L and LEDGF/p75 Are Novel Mediators of PI3K/AKT Signaling and Aggressive Phenotypes in Medulloblastoma," *Cancer Res.*, vol. 76, no. 9, pp. 2802–2812, May 2016.
- [230] H. A. F. Stessman *et al.*, "Disruption of POGZ Is Associated with Intellectual Disability and Autism Spectrum Disorders," *Am. J. Hum. Genet.*, vol. 98, no. 3, p. 541, Mar. 2016.
- [231] K. V. Good, J. B. Vincent, and J. Ausió, "MeCP2: The Genetic Driver of Rett Syndrome

- Epigenetics," *Front. Genet.*, vol. 12, Jan. 2021.
- [232] J. Leitz *et al.*, "Oncogenic human papillomaviruses activate the tumor-associated lens epithelial-derived growth factor (LEDGF) gene," *PLoS Pathog.*, vol. 10, no. 3, 2014.
- [233] R. L. Ochs, Y. Muro, Y. Si, H. Ge, E. K. L. Chan, and E. M. Tan, "Autoantibodies to DFS 70 kd/transcription coactivator p75 in atopic dermatitis and other conditions," *J. Allergy Clin. Immunol.*, vol. 105, pp. 1211–1220, Jun. 2000.
- [234] G. L. Ortiz-Hernandez, E. S. Sanchez-Hernandez, and C. A. Casiano, "Twenty years of research on the DFS70/LEDGF autoantibody-autoantigen system: many lessons learned but still many questions," *Autoimmun. Highlights*, vol. 11, no. 1, Feb. 2020.
- [235] K. Busschots, A. Voet, M. De Maeyer, and J. C. Rain, "Identification of the LEDGF/p75 binding site in HIV-1 integrase," *J. Mol. ...*, 2007.
- [236] M. Llano *et al.*, "An Essential Role for LEDGF/p75 in HIV Integration," *Science (80-.)*, vol. 314, no. 5798, pp. 461–464, Oct. 2006.
- [237] A. Ciuffi *et al.*, "A role for LEDGF/p75 in targeting HIV DNA integration," *Nat. Med.*, vol. 11, no. 12, pp. 1287–1289, Dec. 2005.
- [238] L. Vandekerckhove *et al.*, "Transient and stable knockdown of the integrase cofactor LEDGF/p75 reveals its role in the replication cycle of human immunodeficiency virus.," *J. Virol.*, vol. 80, no. 4, pp. 1886–1896, Feb. 2006.
- [239] J. De Rijck *et al.*, "Overexpression of the lens epithelium-derived growth factor/p75 integrase binding domain inhibits human immunodeficiency virus replication.," *J. Virol.*, vol. 80, no. 23, pp. 11498–11509, Dec. 2006.
- [240] L. S. Vranckx *et al.*, "LEDGIN-mediated Inhibition of Integrase-LEDGF/p75 Interaction Reduces Reactivation of Residual Latent HIV," *EBioMedicine*, vol. 8, pp. 248–264, Jun. 2016.
- [241] Z. Debyser, A. Bruggemans, S. Van Belle, J. Janssens, and F. Christ, "LEDGINs, Inhibitors of the Interaction Between HIV-1 Integrase and LEDGF/p75, Are Potent Antivirals with a Potential to Cure HIV Infection," *Adv. Exp. Med. Biol.*, vol. 1322, pp. 97–114, 2021.
- [242] G. Vansant *et al.*, "Impact of LEDGIN treatment during virus production on residual HIV-1 transcription," *Retrovirology*, vol. 16, no. 1, 2019.
- [243] P. Cherepanov, E. Devroe, P. A. Silver, and A. Engelman, "Identification of an evolutionarily conserved domain in human lens epithelium-derived growth factor/transcriptional co-activator p75 (LEDGF/p75) that binds HIV-1 integrase.," *J. Biol. Chem.*, vol. 279, no. 47, pp. 48883–48892, Nov. 2004.
- [244] Y. Izumoto, T. Kuroda, H. Harada, T. Kishimoto, and H. Nakamura, "Hepatoma-derived growth factor belongs to a gene family in mice showing significant homology in the amino terminus.," *Biochem. Biophys. Res. Commun.*, vol. 238, no. 1, pp. 26–32, Sep. 1997.
- [245] M. Vanegas, M. Llano, S. Delgado, D. Thompson, M. Peretz, and E. Poeschla, "Identification of the LEDGF/p75 HIV-1 integrase-interaction domain and NLS reveals NLS-independent chromatin tethering," *J. Cell Sci.*, vol. 118, no. 8, pp. 1733–1743, Apr. 2005.
- [246] K. Thakar *et al.*, "Interaction of HRP-2 isoforms with HDGF. Chromatin binding of a specific heteromer," vol. 279, no. 5, pp. 737–751, Mar. 2012.
- [247] H. Wu *et al.*, "Structural and histone binding ability characterizations of human PWWP domains.," *PLoS One*, vol. 6, no. 6, p. e18919, 2011.

- [248] X. Zhu *et al.*, “HRP2-DPF3a-BAF complex coordinates histone modification and chromatin remodeling to regulate myogenic gene transcription,” *Nucleic Acids Res.*, vol. 48, no. 12, pp. 6563–6582, Jul. 2020.
- [249] A. Baude *et al.*, “Hepatoma-derived growth factor-related protein 2 promotes DNA repair by homologous recombination,” *Nucleic Acids Res.*, vol. 44, no. 5, pp. 2214–2226, Mar. 2016.
- [250] K. Gao *et al.*, “HDGF-related protein-2 (HRP-2) acts as an oncogene to promote cell growth in hepatocellular carcinoma,” *Biochem. Biophys. Res. Commun.*, vol. 458, no. 4, pp. 849–855, Mar. 2015.
- [251] R. Schrijvers *et al.*, “LEDGF/p75-independent HIV-1 replication demonstrates a role for HRP-2 and remains sensitive to inhibition by LEDGINs,” *PLoS Pathog.*, vol. 8, no. 3, 2012.
- [252] R. Schrijvers *et al.*, “HRP-2 determines HIV-1 integration site selection in LEDGF/p75 depleted cells,” *Retrovirology*, vol. 9, no. 1, p. 84, Oct. 2012.
- [253] H. Wang *et al.*, “HRP2 determines the efficiency and specificity of HIV-1 integration in LEDGF/p75 knockout cells but does not contribute to the antiviral activity of a potent LEDGF/p75-binding site integrase inhibitor,” *Nucleic Acids Res.*, vol. 40, no. 22, pp. 11518–11530, Dec. 2012.
- [254] H. Nakamura *et al.*, “Molecular cloning of complementary DNA for a novel human hepatoma-derived growth factor. Its homology with high mobility group-1 protein,” *J. Biol. Chem.*, vol. 269, no. 40, pp. 25143–25149, 1994.
- [255] H. Enomoto, H. Nakamura, H. Nishikawa, S. Nishiguchi, and H. Iijima, “Hepatoma-Derived Growth Factor: An Overview and Its Role as a Potential Therapeutic Target Molecule for Digestive Malignancies,” *Int. J. Mol. Sci.*, vol. 21, no. 12, pp. 1–14, Jun. 2020.
- [256] C. Bao, J. Wang, W. Ma, X. Wang, and Y. Cheng, “HDGF: A novel jack-of-all-trades in cancer,” *Futur. Oncol.*, vol. 10, no. 16, pp. 2675–2685, Dec. 2014.
- [257] T. Kuroda, H. Tanaka, H. Nakamura, Y. Nishimune, and T. Kishimoto, “Hepatoma-derived growth factor-related protein (HRP)-1 gene in spermatogenesis in mice,” *Biochem. Biophys. Res. Commun.*, vol. 262, no. 2, pp. 433–437, Aug. 1999.
- [258] M. Sausen *et al.*, “Clinical implications of genomic alterations in the tumour and circulation of pancreatic cancer patients,” *Nat. Commun.*, p. 7686, 2015.
- [259] B. Aschebrook-Kilfoy *et al.*, “Genome-wide association study of parity in Bangladeshi women,” *PLoS One*, vol. 10, no. 3, Mar. 2015.
- [260] J. Bellmunt, “Stem-Like Signature Predicting Disease Progression in Early Stage Bladder Cancer. The Role of E2F3 and SOX4,” *Biomedicines*, vol. 6, no. 3, Sep. 2018.
- [261] K. Ikegame *et al.*, “A new member of a hepatoma-derived growth factor gene family can translocate to the nucleus,” *Biochem. Biophys. Res. Commun.*, vol. 266, no. 1, pp. 81–87, Dec. 1999.
- [262] M. M. Abouzied, S. L. Baader, F. Dietz, J. Kappler, V. Gieselmann, and S. Franken, “Expression patterns and different subcellular localization of the growth factors HDGF (hepatoma-derived growth factor) and HRP-3 (HDGF-related protein-3) suggest functions in addition to their mitogenic activity,” *Biochem. J.*, vol. 378, no. Pt 1, pp. 169–76, Feb. 2004.
- [263] M. E. LeBlanc *et al.*, “Hepatoma-Derived Growth Factor-Related Protein-3 Is a Novel Angiogenic Factor,” *PLoS One*, vol. 10, no. 5, May 2015.

- [264] Q. Xiao *et al.*, "HDGF-related protein-3 is required for anchorage-independent survival and chemoresistance in hepatocellular carcinomas," *Gut*, vol. 62, no. 3, pp. 440–451, Mar. 2013.
- [265] S. Qin and J. Min, "Structure and function of the nucleosome-binding PWWP domain," *Trends Biochem. Sci.*, vol. 39, no. 11, pp. 536–547, 2014.
- [266] R. Chen *et al.*, "The Role of Methyltransferase NSD2 as a Potential Oncogene in Human Solid Tumors," *Onco. Targets. Ther.*, vol. 13, p. 6837, 2020.
- [267] N. P. Coussens *et al.*, "High-throughput screening with nucleosome substrate identifies small-molecule inhibitors of the human histone lysine methyltransferase NSD2," *J. Biol. Chem.*, vol. 293, no. 35, p. 13750, Aug. 2018.
- [268] R. Ferreira De Freitas *et al.*, "Discovery of Small-Molecule Antagonists of the PWWP Domain of NSD2," *J. Med. Chem.*, vol. 64, no. 3, pp. 1584–1592, Feb. 2021.
- [269] D. Dilworth *et al.*, "A chemical probe targeting the PWWP domain alters NSD2 nucleolar localization," *Nat. Chem. Biol.*, vol. 18, no. 1, pp. 56–63, 2022.
- [270] I. Stec *et al.*, "WHSC1, a 90 kb SET domain-containing gene, expressed in early development and homologous to a Drosophila dysmorphism gene maps in the Wolf-Hirschhorn syndrome critical region and is fused to IgH in t(4;14) multiple myeloma," *Hum. Mol. Genet.*, vol. 7, no. 7, pp. 1071–1082, Jul. 1998.
- [271] J. Böttcher *et al.*, "Fragment-based discovery of a chemical probe for the PWWP1 domain of NSD3," *Nat. Chem. Biol.*, vol. 15, no. 8, pp. 822–829, Aug. 2019.
- [272] S. Van Belle *et al.*, "Unlike Its Paralog LEDGF/p75, HRP-2 Is Dispensable for MLL-R Leukemogenesis but Important for Leukemic Cell Survival," *Cells*, vol. 10, no. 1, pp. 1–19, Jan. 2021.
- [273] P. S. Renshaw *et al.*, "Sequence-specific assignment and secondary structure determination of the 195-residue complex formed by the Mycobacterium tuberculosis proteins CFP-10 and ESAT-6," *J. Biomol. NMR*, vol. 30, no. 2, pp. 225–6, Oct. 2004.
- [274] V. Veverka, G. Lennie, T. Crabbe, I. Bird, R. J. Taylor, and M. D. Carr, "NMR assignment of the mTOR domain responsible for rapamycin binding.," *J. Biomol. NMR*, vol. 36, no. S1, p. 3, Dec. 2006.
- [275] T. Herrmann, P. Güntert, and K. Wüthrich, "Protein NMR structure determination with automated NOE assignment using the new software CANDID and the torsion angle dynamics algorithm DYANA," *J. Mol. Biol.*, vol. 319, no. 1, pp. 209–27, May 2002.
- [276] Y. Shen, F. Delaglio, G. Cornilescu, and A. Bax, "TALOS+: a hybrid method for predicting protein backbone torsion angles from NMR chemical shifts," *J. Biomol. NMR*, vol. 44, no. 4, pp. 213–223, Aug. 2009.
- [277] E. Harjes *et al.*, "GTP-Ras Disrupts the Intramolecular Complex of C1 and RA Domains of Nore1," *Structure*, vol. 14, no. 5, pp. 881–888, May 2006.
- [278] M. Yokogawa, Y. Kobashigawa, N. Yoshida, K. Ogura, K. Harada, and F. Inagaki, "NMR Analyses of the Interaction between the FYVE Domain of Early Endosome Antigen 1 (EEA1) and Phosphoinositide Embedded in a Lipid Bilayer," *J. Biol. Chem.*, vol. 287, no. 42, pp. 34936–34945, Oct. 2012.
- [279] M. Geraerts, M. Michiels, V. Baekelandt, Z. Debyser, and R. Gijssbers, "Upscaling of lentiviral vector production by tangential flow filtration," *J. Gene Med.*, vol. 7, no. 10, pp. 1299–1310, 2005.

- [280] “NorCOMM2 - Phenotyping - Project.” [Online]. Available: <http://www.norcomm2.org/norcomm2/index.php>.
- [281] P. Moll, M. Ante, A. Seitz, and T. Reda, “QuantSeq 3’ mRNA sequencing for RNA quantification,” *Nat. Methods*, vol. 11, no. 12, pp. i–iii, Dec. 2014.
- [282] A. Dobin *et al.*, “STAR: ultrafast universal RNA-seq aligner,” *Bioinformatics*, vol. 29, no. 1, pp. 15–21, Jan. 2013.
- [283] M. I. Love, W. Huber, and S. Anders, “Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2,” *Genome Biol.*, vol. 15, no. 12, p. 550, Dec. 2014.
- [284] S. El ashkar, “The role of LEDGF/p75 and HRP2 in hematopoiesis and mixed lineage leukemia,” *KU Leuven, Fac. Med.*, 2017.
- [285] J. Grembecka *et al.*, “Menin-MLL inhibitors reverse oncogenic activity of MLL fusion proteins in leukemia,” *Nat. Chem. Biol.*, vol. 8, no. 3, pp. 277–284, 2012.
- [286] D. Borkin *et al.*, “Property Focused Structure-Based Optimization of Small Molecule Inhibitors of the Protein-Protein Interaction between Menin and Mixed Lineage Leukemia (MLL),” *J. Med. Chem.*, vol. 59, no. 3, pp. 892–913, Feb. 2016.
- [287] H. Wang, M.-C. Shun, A. K. Dickson, and A. N. Engelman, “Embryonic Lethality Due to Arrested Cardiac Development in Psp1/Hdgfrp2 Double-Deficient Mice,” *PLoS One*, vol. 10, no. 9, p. e0137797, Sep. 2015.
- [288] P. T. Lowary and J. Widom, “New DNA sequence rules for high affinity binding to histone octamer and sequence-directed nucleosome positioning,” *J. Mol. Biol.*, vol. 276, no. 1, pp. 19–42, Feb. 1998.
- [289] M. M. Pradeepa, H. G. Sutherland, J. Ule, G. R. Grimes, and W. A. Bickmore, “Psp1/Ledgf p52 binds methylated histone H3K36 and splicing factors and contributes to the regulation of alternative splicing,” *PLoS Genet.*, vol. 8, no. 5, p. e1002717, 2012.
- [290] N. C. Dale, E. K. M. Johnstone, C. W. White, and K. D. G. Pflieger, “NanoBRET: The bright future of proximity-based assays,” *Front. Bioeng. Biotechnol.*, vol. 7, no. MAR, p. 56, 2019.
- [291] F. Weihs, J. Wang, K. D. G. Pflieger, and H. Dacres, “Experimental determination of the bioluminescence resonance energy transfer (BRET) Förster distances of NanoBRET and red-shifted BRET pairs,” *Anal. Chim. Acta X*, vol. 6, p. 100059, Nov. 2020.
- [292] D. J. Wood *et al.*, “FragLites[®] Minimal, Halogenated Fragments Displaying Pharmacophore Doublets. An Efficient Approach to Druggability Assessment and Hit Generation,” 2019.
- [293] “Fragment Libraries - Diamond Light Source, Oxford.” [Online]. Available: <https://www.diamond.ac.uk/Instruments/Mx/Fragment-Screening/Fragment-Libraries0.html>. [Accessed: 10-Jan-2022].
- [294] H. G. Sutherland *et al.*, “Disruption of Ledgf/Psp1 results in perinatal mortality and homeotic skeletal transformations,” *Mol. Cell. Biol.*, vol. 26, no. 19, pp. 7201–7210, 2006.
- [295] J. Hendrix *et al.*, “The transcriptional co-activator LEDGF/p75 displays a dynamic scan-and-lock mechanism for chromatin tethering,” *Nucleic Acids Res.*, vol. 39, no. 4, p. 1310, Mar. 2011.
- [296] M. C. Gundry, M. A. Goodell, and L. Brunetti, “It’s All About MEIs: Menin-MLL Inhibition Eradicates NPM1-Mutated and MLL-Rearranged Acute Leukemias in Mice,” *Cancer Cell*, vol. 37, no. 3, pp. 267–269, Mar. 2020.
- [297] W. Fiskus *et al.*, “Effective Menin inhibitor-based combinations against AML with MLL

rearrangement or NPM1 mutation (NPM1c),” *Blood Cancer J.* 2022 121, vol. 12, no. 1, pp. 1–11, Jan. 2022.

- [298] C. T. Collins and J. L. Hess, “Deregulation of the HOXA9/MEIS1 axis in acute leukemia,” *Curr. Opin. Hematol.*, vol. 23, no. 4, pp. 354–61, 2016.

Curriculum vitae

SISKA VAN BELLE

12 aug. 1993, Belgian

Haaggatstraat 15, 3071 Kortenberg

siskavanbelle@outlook.com

+32 479 632 552

EDUCATION AND WORKEXPERIENCE

2022 – to date	Lab coordinator Laboratory of Molecular Virology and Gene Therapy, KU Leuven
2016 – 2022	PhD Researcher Laboratory of Molecular Virology and Gene Therapy, KU Leuven
2014 – 2016	Master in biomedische wetenschappen, minor onderzoek, KU Leuven
2011 – 2014	Bachelor in biomedische wetenschappen, KU Leuven

GRANT

2018 – 2021	FWO SB grant
-------------	--------------

MEETINGS AND CONGRESSES

2020	EHA, virtual Poster presentation
2019	EHA, Amsterdam Poster presentation
2019	BACR, Antwerp Poster presentation

TRAININGS

Felasa B, Certificate Laboratory Animal Science, KU Leuven

Protein Purification School, Univ. Liège

Research integrity seminars, Doctoral program, KU Leuven

TEACHING & SUPERVISION

2017 – 2021	Practical guidance of first year students of medicine and biomedical sciences B-KUL-U02A4a, B-KUL-U01E3a, B-KUL-U03D6a, KULAK
2021	Research internship students Catherine Willems and Amber De Moor
2020	Research internship student Guillem Wetherell
2019 – 2020	Master thesis supervision of Joke Seuntjens and Michiel De Coster
2019	Research internship student Nienke Meeus
2018	Research internship students Deirdre Van Esse and Janne Huysmans

PUBLICATIONS

Presented meeting abstracts

Alessandro Canella, [Siska Van Belle](#), Tine Brouns, Giovanni Nigita, Marianne Carlon, Frauke Christ and Zeger Debyser. LEDGF/p75-mediated chemoresistance of Mixed-Lineage Leukemia involves cell survival pathways and super enhancer activators. European Hematology Association. Virtual (2020)

[Siska Van Belle](#)*, Sara El Ashkar*, Katerina Cermáková, Filip Matthijssens, Steven Goossens, Alessandro Canella, Frauke Christ, Pieter Van Vlierberghe, Václav Veverka, Jan De Rijck and Zeger Debyser. The LEDGF/p75 homolog HRP2 is not important for the initiation of MLL-rearranged leukemia. European Hematology Association. Amsterdam (2019).

Sara El Ashkar*, [Siska Van Belle](#)*, Katerina Cermáková, Filip Matthijssens, Steven Goossens, Frauke Christ, Pieter Van Vlierberghe, Václav Veverka, Jan De Rijck and Zeger Debyser. HRP2 plays a role in the maintenance of MLL-rearranged leukemia. Belgian Association for Cancer Research (BACR). Antwerp (2019)

Journal articles

Nombela I, Michiels M, Van Looveren D, Marcelis L, El Ashkar S, [Van Belle S](#), Bruggemans A, Tousseyn T, Schwaller J, Christ F, Gijsbers R, De Rijck J, Debyser Z. BET-Independent Murine Leukemia Virus Integration Is Retargeted *In Vivo* and Selects Distinct Genomic Elements for Lymphomagenesis. *Microbiol Spectr*. 2022 Aug 31. PMID: 35852337.

Acke A, [Van Belle S](#), Louis B, Vitale R, Rocha S, Voet T, Debyser Z, Hofkens J. Expansion microscopy allows high resolution single cell analysis of epigenetic readers. *Nucleic Acids Res*. 2022 Jun 18. PMID: 35716125.

Shirvani P, Fayyazi N, [Van Belle S](#), Debyser Z, Christ F, Saghiaie L, Fassihi A. Design, synthesis, in silico studies, and antiproliferative evaluations of novel indolin-2-one derivatives containing 3-hydroxy-4-pyridinone fragment. *Bioorg Med Chem Lett*. 2022 May 12:128784. PMID: 35569690.

Canella A, Van Belle S, Brouns T, Nigita G, Carlon MS, Christ F, Debyser Z. LEDGF/p75-mediated chemoresistance of mixed-lineage leukemia involves cell survival pathways and super enhancer activators. *Cancer Gene Ther.* 2022 Feb;29(2):133-140. PMID: 33795806.

Van Belle S, El Ashkar S, Čermáková K, Matthijssens F, Goossens S, Canella A, Hodges CH, Christ F, De Rijck J, Van Vlierberghe P, Veverka V, Debyser Z. Unlike Its Paralog LEDGF/p75, HRP-2 Is Dispensable for MLL-R Leukemogenesis but Important for Leukemic Cell Survival. *Cells.* 2021 Jan 19;10(1):192. PMID: 33477970.

Debyser Z, Bruggemans A, Van Belle S, Janssens J, Christ F. LEDGINs, Inhibitors of the Interaction Between HIV-1 Integrase and LEDGF/p75, Are Potent Antivirals with a Potential to Cure HIV Infection. *Adv Exp Med Biol.* 2021;1322:97-114. PMID: 34258738.

Shirvani P, Fassihi A, Saghaie L, Van Belle S, Debyser Z, Christ F. Synthesis, anti-HIV-1 and antiproliferative evaluation of novel 4-nitroimidazole derivatives combined with 5-hydroxy-4-pyridinone moiety. *Journal of Molecular Structure*, volume 1202, 2020, 127344, ISSN 0022-2860.

Reynhout S, Jansen S, Haesen D, van Belle S, de Munnik SA, Bongers EMHF, Schieving JH, Marcelis C, Amiel J, Rio M, Mclaughlin H, Ladda R, Sell S, Kriek M, Peeters-Scholte CMPCD, Terhal PA, van Gassen KL, Verbeek N, Henry S, Scott Schwoerer J, Malik S, Revencu N, Ferreira CR, Macnamara E, Braakman HMH, Brimble E, Ruzhnikov MRZ, Wagner M, Harrer P, Wieczorek D, Kuechler A, Tziperman B, Barel O, de Vries BBA, Gordon CT, Janssens V, Vissers LELM. De Novo Mutations Affecting the Catalytic α Subunit of PP2A, PPP2CA, Cause Syndromic Intellectual Disability Resembling Other PP2A-Related Neurodevelopmental Disorders. *Am J Hum Genet.* 2019 Jan 3;104(1):139-156. PMID: 30595372.

Sharma S, Čermáková K, De Rijck J, Demeulemeester J, Fábry M, El Ashkar S, Van Belle S, Lepšík M, Tesina P, Duchoslav V, Novák P, Hubálek M, Srb P, Christ F, Řezáčová P, Hodges HC, Debyser Z, Veverka V. Affinity switching of the LEDGF/p75 IBD interactome is governed by kinase-dependent phosphorylation. *Proc Natl Acad Sci U S A.* 2018 Jul 24;115(30):E7053-E7062. PMID: 29997176.

El Ashkar S, Schwaller J, Pieters T, Goossens S, Demeulemeester J, Christ F, Van Belle S, Juge S, Boeckx N, Engelman A, Van Vlierberghe P, Debyser Z, De Rijck J. LEDGF/p75 is dispensable for hematopoiesis but essential for MLL-rearranged leukemogenesis. *Blood.* 2018 Jan 4;131(1):95-107. PMID: 29084774.