

PhD degree in Systems Medicine (curriculum in Molecular Oncology)

European School of Molecular Medicine (SEMM),

University of Milan and University of Naples “Federico II”

Settore disciplinare: BIO/10

**Chromatin Proteomics dissects the mechanism of
action of LSD1 inhibitors in Acute Promyelocytic
Leukemia**

Luciano Nicosia

European Institute of Oncology (IEO), Milan

Matricola n. R11481

Supervisor: Dr. Tiziana Bonaldi

European Institute of Oncology (IEO), Milan

Anno accademico 2018-2019

Table of contents

List of abbreviations:	7
Figure and Table Index:	13
1. ABSTRACT	16
2. INTRODUCTION	18
2.1. LSD1: function and structure	18
2.2. LSD1 function as transcriptional regulator	19
2.3. Non-histone LSD1 substrates	22
2.4. LSD1 in cancer	22
2.4.1. Acute Myeloid Leukemia	24
2.4.2. LSD1 in Acute Myeloid Leukemia.....	26
2.5. State-of-the-art on the LSD1 inhibitors.....	27
2.5.1. The LSD1 inhibitors MC2580 and DDP-38003.....	30
2.6. Effects of pharmacological inhibition of LSD1 in NB4 and UF-1 APL cells	31
2.6.1. Combinatorial effect of LSD1 inhibitors with physiological doses of ATRA in NB4 cells.....	32
2.7. Mass spectrometry (MS)-based proteomics	36
2.8. The mass spectrometer	38
2.9. Quantitative proteomics	40
2.10. Immuno-precipitation coupled with MS to study protein-protein interactions	44
2.10.1. Quantitative proteomics to study PPIs.....	47
2.10.2. Interaction proteomics of chromatin-associated proteins	48
2.11. MS-based proteomics analysis of histone post-translational modifications (PTMs)	50
2.11.1. Identification of histone modifications	50
2.11.2. MS-based quantitative strategies of histone PTMs.....	51
3. AIM OF THE WORK	56
4. MATERIALS AND METHODS	58
4.1. Cell culture	58
4.1.1. SILAC labelling of NB4 cells for the LSD1 Interactome analysis.....	58
4.2. Compounds.....	59
4.3. LSD1 co-IP for mass-spectrometry analysis of protein-protein interactions	59
4.3.1. Subcellular fractionation.....	59
4.3.2. LSD1 protein co-immunoprecipitation (co-IP).....	59
4.4. In-gel digestion of immuno-precipitated proteins	60

4.5. Liquid Chromatography and Tandem Mass Spectrometry (LC-MS/MS)	61
4.6. Data analysis for SILAC- based protein quantitation	63
4.7. LSD1 interactome data analysis.....	63
4.8. Immuno-precipitation analysis for Western Blot validation.....	64
4.9. Retroviral and lentiviral cell transduction.....	65
4.9.1. Retroviral and lentiviral constructs	65
4.9.2. NB4 cell transduction.....	66
4.10. Western blot analysis	66
4.11. RNA extraction and RT-qPCR	67
4.12. Cell growth analysis.....	68
4.13. May Grunwald-Giemsa staining	69
4.14. Chromatin Immuno-Precipitation-sequencing (ChIP-seq) in NB4 cells	69
4.14.1. ChIP-seq analysis of LSD1 and GFI1	69
4.14.2. ChIP-seq analysis	70
4.15. Cell cycle analysis by flow cytometry	71
4.16. RNA-sequencing.....	71
4.17. MS-based profiling of histone PTMs.....	72
4.17.1 Histones extraction and digestion.....	72
4.17.2 LC-MS/MS and data analysis of histone PTMs.....	73
4.18. Statistical analysis	74
5. RESULTS	75
5.1. MS-based analysis of the basal LSD1-interactome in NB4 cells	75
5.1.1. SILAC labelling of NB4 cells	76
5.1.2. Optimization of the LSD1-IP conditions for the characterization of the LSD1 interactors in NB4 cells	77
5.1.3. SILAC proteomics in combination with LSD1-IP for the characterization of the basal LSD1-interactome in NB4 cells	80
5.2. MS-based analysis of the changes in the LSD1-interaction network upon treatment with LSD1 inhibitors.....	87
5.3. Impact of the LSD1 inhibition on the binding of the catalytic inactive LSD1 mutant to GFI1 and GSE1.....	91
5.4. Phenotypic and molecular effects of the inhibition of LSD1-GFI1 interaction in NB4 cells	92
5.5. Mechanistic analysis of the alteration of LSD1-GSE1 binding by LSD1 inhibitors	97
5.6. Dissecting the effects of GSE1 depletion in NB4 cells	101
5.6.1. Characterization of the phenotypic effects of GSE1 knock-down in NB4 cells	102

5.6.2. RNA-sequencing analysis upon GSE1 depletion in NB4 cells	107
5.7. Comparison of the transcriptomic data upon GSE1-KD and LSD1 inhibition in NB4 cells.....	110
5.8. MS-based profiling of histone PTMs in NB4 cells upon LSD1 inhibition	113
5.9. MS-based analysis of histone PTM changes in UF-1 cells upon LSD1 inhibition .	116
5.10. Comparative analysis of histone PTMs in APL cells with different sensitivity to LSD1 inhibitors	118
5.11. Combinatorial effect of LSD1 and EZH2 inhibitors on NB4 cell proliferation....	119
6. DISCUSSION	121
7. REFERENCES	132
8. APPENDIX	151
Appendix 1: List of putative LSD1 interactors in NB4 cells	151
Appendix 2: A Chromatin Proteomics (ChroP) approach dissects the histone post-translational modifications (PTMs) enriched at enhancers of mouse macrophages	155
BACKGROUND	155
MATERIALS AND METHODS.....	156
RESULTS AND DISCUSSION.....	161
REFERENCES	169
9. List of publications achieved during PhD.....	173
10. Acknowledgments	175

List of abbreviations:

μ	Mean
3D-LSD1	LSD1 mutant D553, 555, 556A
AABG	Advanced active beam guide
ACN	Acetonitrile
AML	Acute Myeloid Leukemia
AO	Amino oxidase
AOL	Amino oxidase-like
AP	Affinity-purification
APL	Acute Promyelocytic Leukemia
AR	Androgen receptor
ATRA	All-trans retinoic acid
AUC	Area under the curve
BP	Biological process
BSA	Bovine serum albumin
CAM	Cell adhesion molecule
CD	Cluster of differentiation
CFA	Colony forming activity
CFU	Colony forming unit
ChEP	Chromatin enrichment for proteomics
ChIP-seq	Chromatin immuno-precipitation-sequencing
ChroP	Chromatin proteomics
CID	Collision-induced dissociation
CKI	Cyclin-dependent kinase inhibitor
CMP	Common myeloid progenitor
CoREST	REST corepressor 1
CRPC	Castrate-resistant prostate cancer
D-CAP	Differential chromatin-associated proteins
DCIS	Ductal carcinoma in situ
DDA	Data-dependent acquisition

DEG	Differentially expressed genes
DIR	Forward replicate
DMSO	Dimethyl sulfoxide
DNase I	Deoxyribonuclease I
DSG	Disuccinimidyl glutarate
DTT	Dithiothreitol
ECD	Electron capture dissociation
ECL	Enhanced chemio luminescence
EMT	Epithelial to mesenchymal transition
ER	Estrogen receptor
ESI	Electrospray ionization
ESR1	Estrogen receptor 1
ETD	Electron transfer dissociation
EV	Empty vector
FA	Formic acid
FAB	French-American British
FACS	Fluorescence activated cell sorting
FAD	Flavin adenine dinucleotide
FBS	Fetal bovine serum
FC	Fold change
FT	Flow-through
FT-ICR	Fourier-transform ion cyclotron resonance
GAPDH	Glyceraldehyde 3-phosphate dehydrogenase
GBM	Glioblastoma
GFI1	Growth factor independent 1
GO	Gene ontology
GSE1	Genetic suppressor element 1
H	Heavy
H3K27me2	Histone H3 lysine 27 di-methylation
H3K27me3	Histone H3 lysine 27 tri-methylation
H3K4me1	Histone H3 lysine 4 mono-methylation

H3K4me2	Histone H3 lysine 4 di-methylation
H3K9me1	Histone H3 lysine 9 mono-methylation
H3K9me2	Histone H3 lysine 9 di-methylation
HAT	Histone acetyl-transferase
HCD	Higher energy collision dissociation
HCl	Hydrochloric acid
HDAC1	Histone deacetylase 1
HPLC	High-pressure liquid chromatography
IgG	Immunoglobulin G
IP	Immuno-precipitation
IPA	Ingenuity pathway analysis
iTRAQ	Isobaric tags for relative and absolute quantitation
K	Lysine
K661A-LSD1	LSD1 catalytic inactive
KD	Knock-down
KO	Knock-out
L	Light
LC-MS/MS	Liquid-chromatography tandem mass spectrometry
LFQ	Label-free quantitation
LIT	Linear ion trap
LSC	Leukemic stem cell
LSD1	Lysine-specific demethylase 1
M	Median
m/z	mass-to-charge ratio
MALDI	Matrix-assisted laser desorption ionization
MAO	Monoamine oxidase
me1	Mono-methylation
me2	Di-methylation
me3	Tri-methylation
MNase	Micrococcal nuclease
MQ	Max quant

MRM	Multiple reaction monitoring
MS	Mass-spectrometry
MTA	Metastasis tumour-antigen
NCOR	Nuclear receptor co-repressor
NES	Neuro-epithelial stem cells
NSCLC	Non-small cell lung cancer
NuRD	Nucleosome remodelling and deacetylase
P/S	Penicillin/Streptomycin
P-adj	Adjusted p-value
PBS	Phosphate-buffered saline
PCA	Principal component analysis
PI	Propidium iodide
PIC	Phenyl-isocyanate
PKC	Protein kinase C
PML	Promyelocytic Leukemia
PMSF	Phenylmethylsulfonyl fluoride
PPI	Protein-protein interaction
pRb	Phosphorylated retinoblastoma
PTM	Post-translational modification
Q	Quadrupole
R	Arginine
RA	Relative abundance
RARA	Retinoic acid receptor alpha
RC	Ratio count
REV	Reverse replicate
RF	Radio frequency
RNaseA	Ribonuclease A
RNA-seq	RNA sequencing
RP	Reversed-phase
SDS-PAGE	Sodium dodecyl sulphate - polyacrylamide gel electrophoresis
SEM	Standard error of the mean

shRNA	Short hairpin RNA
SILAC	Stable isotope labelling by amino acids in cell culture
siRNA	Short interfering RNA
SRM	Selected reaction monitoring
T-ALL	T-cell lymphoblastic leukemia
TBP	Tata binding protein
TBS	Tris-buffered saline
TCP	Tranlycypromine
TFA	Trifluoroacetic acid
TLR	Toll-like receptor
TMT	Tandem mass tag
TNBC	Triple-negative breast cancer
TNF	Tumour necrosis factor
TOF	Time-of-flight
WB	Western blot
WT	Wild-type
XIC	Extracted ion chromatogram
σ	Standard deviation

Figure and Table Index:

Figure 1: Enzymatic reaction of histone demethylation mediated by LSD1	18
Figure 2: LSD1 protein structure	19
Figure 3: The dual role of LSD1 as transcriptional co-regulator	21
Figure 4: Current model of APL pathogenesis and therapeutic treatment with Retinoic Acid	26
Figure 5: LSD1 scaffolding function	27
Figure 6: Effects of MC2580 in NB4 and UF-1 cells	32
Figure 7: LSD1 inhibitors sensitize NB4 cells to physiological doses of all-trans retinoic acid (ATRA) irrespective of the inhibition of the LSD1 catalytic activity	35
Figure 8: Schematic workflow of a standard “bottom-up” approach for MS analysis	37
Figure 9: Electrospray ionization (ESI) process	39
Figure 10: Quantitative MS-based proteomics approaches	41
Figure 11: Workflow of affinity-purification (AP) strategy with antibody targeting an endogenous protein coupled with MS-analysis	45
Figure 12: SILAC-MS analysis in interaction proteomics experiments	48
Figure 13: MS-based strategies to identify histone PTMs	51
Figure 14: LF-based strategy to quantify histone modifications	52
Figure 15: Super-SILAC approach for histone PTMs quantitation	54
Figure 16: Schematic representation of a Q Exactive HF mass spectrometer	62
Figure 17: Experimental design of the SILAC/co-IP to characterize LSD1 specific interactors	76
Figure 18: SILAC labelling of NB4 cells	77
Figure 19: Optimization of the LSD1-IP protocol	79
Figure 20: Optimization of peptide fold excess for the LSD1-IP	80
Figure 21: Western Blot quality control assessment of the SILAC LSD1-IP experiments prior to MS	81
Figure 22: MS-based identification of the LSD1 Interactors in NB4 cells	83
Figure 23: Characterization of the complexes enriched in the LSD1 interactome	86
Figure 24: Experimental design of SILAC/co-IP strategy to study the changes in the LSD1 interactome upon drug treatment	87
Figure 25: Western Blot validation of the SILAC dynamic LSD1-IP replicates	88
Figure 26: MS-based characterization of the dynamic LSD1 interactome upon drug treatment	89
Figure 27: LSD1 pharmacological inhibition alters the interaction of LSD1 with EDC4, GFI1 and GSE1	91
Figure 28: LSD1 inhibitor alters the interaction of LSD1 with GFI1 and GSE1 irrespective of the inhibition of the LSD1 enzymatic activity	92
Figure 29: Analysis of the LSD1-GFI1 binding upon expression of the LSD1-3D mutant in NB4 KO-LSD1 cells	93
Figure 30: LSD1 interaction with GFI1 serves to self-renewal and proliferation of NB4 cells	95
Figure 31: MC2580 displaces LSD1 from chromatin-bound GFI1 in cis-regulatory regions of genes involved in hematopoietic cell differentiation	96

Figure 32: LSD1 inhibitor does not inhibit LSD1-GSE1 physical interaction, but reduces GSE1 protein levels	98
Figure 33: LSD1 inhibitors regulate GSE1 expression at the transcriptional level.....	99
Figure 34: GSE1 protein level is reduced in NB4 KO-LSD1 cells compared to the WT	100
Figure 35: Experimental approach used to deplete GSE1 in NB4 cells	101
Figure 36: GSE1-KD induces reduction of cell proliferation and apoptosis in NB4 cells	103
Figure 37: GSE1-KD induces the expression of differentiation markers in NB4 cells....	104
Figure 38: GSE1-KD determines increase in G1-phase and induction of p21.....	106
Figure 39: Differential gene expression analysis upon GSE1-KD in NB4 cells.....	108
Figure 40: Gene Ontology (GO) analysis of the up-regulated genes upon GSE1-KD in NB4 cells	110
Figure 41: Analysis of the common DEG upon LSD1 inhibition and GSE1-KD in NB4 cells	112
Figure 42: Experimental strategy to analyse histone PTMs in NB4 cells following 24h treatment with the inhibitor DDP-38003	114
Figure 43: MS-based profiling of histone PTMs upon genetic and pharmacological LSD1 inhibition.....	115
Figure 44: MS-based analysis of histone PTMs upon 24 hours treatment with MC2580 in UF-1 cells.....	117
Figure 45: Comparative analysis of histone PTMs in sensitive and resistant APL and GBM models.....	119
Figure 46: Combination of LSD1 and EZH1/2 inhibition strongly affects NB4 cell proliferation	120
Figure 47: MC2580 but not ATRA increases the levels of the monocyte/macrophage markers IRF8 and CD14.....	125
Figure 48: Novel possible mechanisms of action of the LSD1 inhibitors in APL	130
Table 1: Most studied LSD1 inhibitors in cancer	28
Table 2: Biochemical IC50 values of MC2580 and DDP-38003	31

1. ABSTRACT

Lysine-Specific Demethylase 1 (LSD1) is a protein lysine demethylase that catalyzes the removal of methyl-groups from the mono- and di- methylated forms of lysine 4 and lysine 9 of histone H3 (H3K4me1/me2 and H3K9me1/me2), thus playing an important role in gene expression modulation. This enzyme is aberrantly expressed in various types of cancer and has emerged, in particular, as a promising target for the epigenetic therapy of Acute Myeloid Leukemia (AML). In my thesis project, I implemented and applied a panel of quantitative mass spectrometry (MS)-based approaches to characterize the molecular effects of two potent and selective LSD1 inhibitors under development in our Institute both on the LSD1 interaction network and on global histone post-translational modifications (PTMs), in Acute Promyelocytic Leukemia (APL) cell models, a subtype of AML. Preliminary results acquired by the group of Prof. Saverio Minucci had previously demonstrated that these inhibitors sensitize NB4-APL cells to physiological doses of all-trans retinoic acid (ATRA) independently from the inhibition of LSD1 catalytic activity. These data suggested that the phenotypic effects elicited by these compounds could be due to an effect on the LSD1-interaction network.

By SILAC-based quantitative MS-interactomics we first identified the complete set of LSD1 interactors, most of which involved in chromatin remodelling and transcription regulation activities. The subsequent analysis of the dynamic changes of the LSD1-interactome upon pharmacological inhibition of the enzyme with the drugs under focus led us to the identification of GFI1 and GSE1, two interactors that displayed a decreased binding to LSD1 upon drug treatment, independently from the specific inhibition of the catalytic activity of the enzyme. Functional and mechanistic follow-up experiments allowed demonstrating that the inhibition of the LSD1-GFI1 interaction sensitizes NB4 cells to physiological doses of ATRA, thus promoting reduction of cell proliferation and differentiation of the cells.

The LSD1-GSE1 interaction was basically uncharacterized before this study; the experiments carried out during my PhD project demonstrated that the two drugs lead to the down-regulation of GSE1 expression, both at the transcript and protein levels, rather than disrupting the physical interaction with LSD1. To mimic the effect of the drugs, we performed GSE1 knock-down (KD) in NB4 cells and analysed the consequence of this depletion on both cellular phenotype and transcription. We found that GSE1-KD reduced NB4 cell proliferation by triggering apoptosis. Before the initiation of cell death, GSE1 down-regulation also induces the expression of various differentiation markers and leads to cell cycle arrest by enhancing p21 expression. The comparative analysis of transcriptomic changes upon GSE1-KD and LSD1 pharmacological inhibition in NB4 cells unveiled a set of genes in common, which are involved in “cytokine-mediated signalling” and “regulation of cysteine-endopeptidase activity involved in the apoptotic process”. This result suggests the existence of a regulatory LSD1-GSE1 axis controlling the transcription of these genes. To complete the molecular inspection on the effects of LSD1 inhibition in APL, we carried out a systematic MS-based profiling of histone post-translational modifications (PTMs) changes upon drug treatment in NB4 cells. We observed not only the increase of the well-known LSD1 target H3K4me₂, but also a reproducible increase of H3K27me₂ and H3K27me₃, paralleled by the decrease of the H3K27me₁ in combination with H3K36me₁. UF-1 cells, an APL cell model more sensitive to pharmacological inhibition of LSD1 than NB4, did not present any drug-induced change in H3K27 methylation levels. Interestingly, the basal levels of H3K27me₂/me₃ measured by quantitative MS are much higher in NB4 than UF-1, which may suggest that the bulk levels of these modifications may be critical for the sensitivity of the cells to the LSD1 inhibitors. In line with this hypothesis, we found that pre-treating the NB4 with EZH2/1 inhibitor sensitizes the cells to LSD1 inhibition. Collectively the data presented in this thesis provide novel insights into the molecular activity of LSD1 and its inhibitors in APL cells.

binding and the substrate-binding subdomains^{5,6} (Figure 2). The SWIRM domain interacts with the AOL region to form an overall globular structure that is fundamental for protein function and stability^{5,7}. The Tower domain, instead, mediates the interaction of LSD1 with the subunits of specific chromatin complexes such as the REST corepressor 1 (CoREST) and the metastasis tumour-antigens (MTA1-3) of the CoREST complex and the Mi-2/nucleosomes remodelling and deacetylase (NuRD) complex, respectively. The interaction involves the SANT domain of the aforementioned LSD1-binding partners^{8,9}.

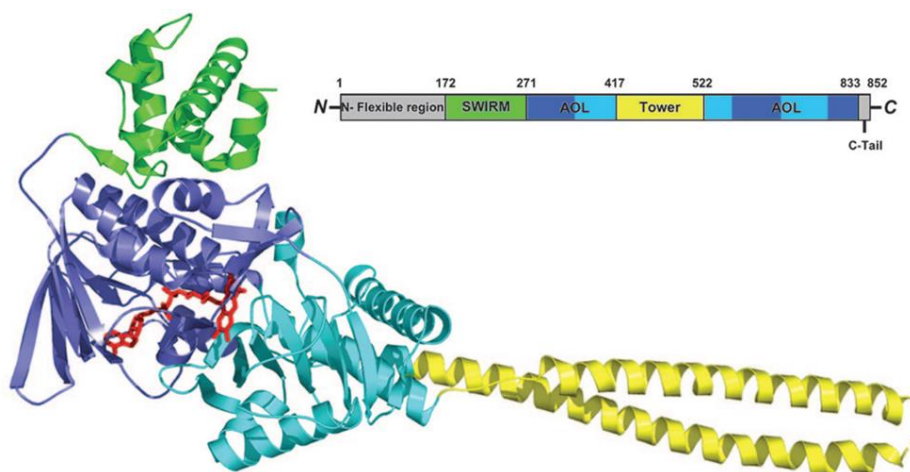


Figure 2: LSD1 protein structure. LSD1 structure is mainly composed of three domains: the SWIRM domain (in green), the Tower domain (in yellow) and the Amino-Oxidase Like (AOL) domain containing the FAD-binding subdomain (in blue) and a substrate-binding subdomain (in cyan). Between these two subdomains, FAD is represented in ball-and-stick and painted in red. The N-terminal region and the C-terminal tail are not present in the structure determination and are marked in grey. Adapted from⁵.

2.2. LSD1 function as transcriptional regulator

LSD1 plays a key role in the regulation of several cellular processes of normal and cancer cells including cell differentiation¹⁰, cell cycle¹¹, cell metabolism^{12,13}, cell motility¹⁴ and epithelial to mesenchymal transition (EMT)¹⁵⁻¹⁷. The regulation of these processes occurs through its presence in multi-protein complexes that may elicit opposite function in the regulation of gene transcription. In fact, depending on the distinct set of interaction established, LSD1 can act as either a transcriptional co-repressor or co-activator.

Epigenetically this is reflected in the de-methylation of different histone targets, in particular the H3K4me1/me2 in the case of gene silencing and the H3K9me1/me2 during gene activation ¹⁸.

The main LSD1 complexes with transcription repressive function are CoREST and NuRD ¹⁹. Components of these complexes have a fundamental role in regulating the LSD1 activity, for instance CoREST confers to LSD1 the ability to bind nucleosomes and activate its demethylase activity as well as protects it from proteasomal degradation ^{8,9}. HDACs, instead, create an hypo-acetylated chromatin environment that activates the LSD1 enzymatic activity ^{9,20}. On the other hand, PHF21A has an important role in stabilizing LSD1 on its target regions and promoting the de-methylation of the H3K4me1/me2 in the surrounding nucleosomes ^{21,22}.

The recruitment of LSD1 and its repressive CoREST complex to the target genes is mediated by the interaction of the enzyme with the SNAG domain of different transcription factors, such as Snail1, GFI1 and GFI1b: in the case of the binding to Snail1, the LSD1-CoREST complex is recruited to the promoter of the E-cadherin ¹⁵, while the LSD1-CoREST-GFI1 complexes target the promoters of several genes involved in hematopoietic differentiation ²³. The association of LSD1 with the NuRD complex is mediated by its interaction with other transcription factors, such as Pax2 and Prox1 ^{24,25}. In addition to CoREST and NuRD, LSD1 is embedded in other less-known repressive complexes such as the BHC ²⁶, the SIN3A/HDAC/LSD1 ²⁷ and the Notch/PRC2/LSD1 complexes ²⁸.

LSD1 has, instead, a role as transcriptional co-activator in the context of androgen (AR) and estrogen (ER)-receptor dependent transcription. Upon activation of these signalling pathways, LSD1 de-methylates H3K9me1/me2 thus activating downstream the expression of the AR- and ER-target genes ^{29,30}. In the case of AR-signalling, the change of the target specificity from H3K4 to H3K9 is dependent on the activity of a kinase called protein kinase C beta I (PKCbeta I) that, following hormone treatment, is recruited to the AR target

promoters and mediates the phosphorylation of H3 on threonine 6 (H3T6ph), a modification that directs the LSD1 catalytic activity towards the H3K9^{18,31}. In the context of ER-signalling this switch of substrate specificity seems instead to be driven by PELP1, a protein that recognizes H3K4me2 and H3K9me2 (histone reader)³². In addition to AR- and ER-target genes, the LSD1 enzymatic activity directed on the H3K9 was also detected at the promoter of some cell cycle-related genes^{11,33}, as well as at regulatory regions of genes involved in myogenesis³⁴, adipogenesis³⁵ and inner ear development³⁶ (Figure 3).

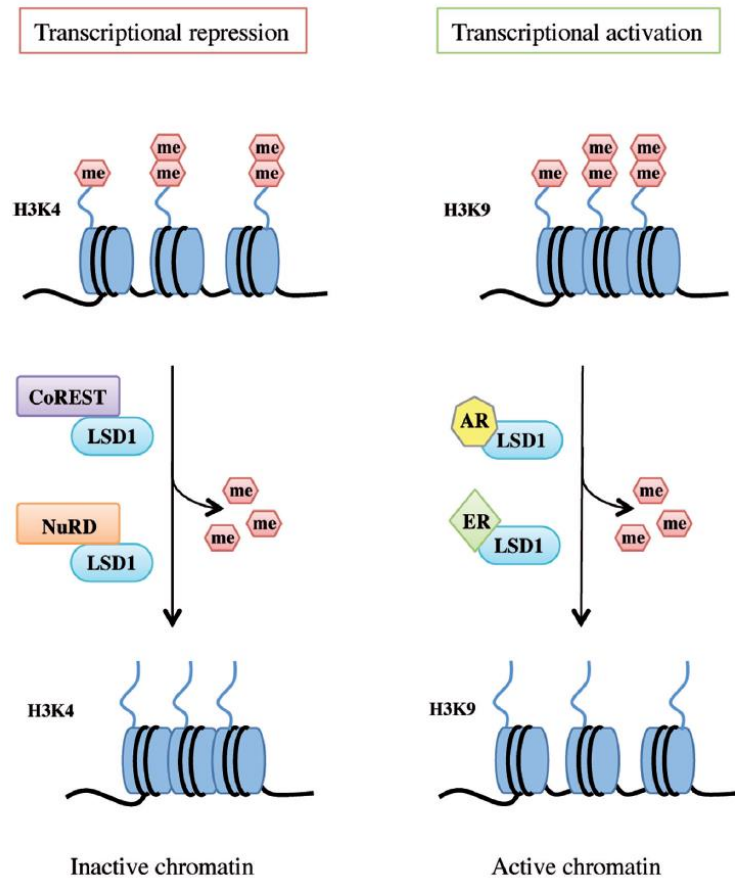


Figure 3: The dual role of LSD1 as transcriptional co-regulator. LSD1 functions as transcriptional repressor or activator by controlling chromatin accessibility through its demethylase activity on K4 or K9 of histone H3. On the left, LSD1 binds to the CoREST or NuRD complexes and catalyzes the de-methylation of H3K4me1/me2, thus allowing gene transcriptional repression. On the right, following AR or ER receptor binding, LSD1 activates transcription of the target genes by de-methylating H3K9me1/me2. Taken from²².

2.3. Non-histone LSD1 substrates

LSD1 has also different non-histone substrates such as p53³⁷, E2F1³⁸, DNMT1³⁹, MYPT1⁴⁰ and HIF1 α ^{41,42}. By targeting these proteins, LSD1 regulates various processes including apoptosis, DNA damage, cell cycle, DNA replication and cell metabolism. In particular, LSD1 blocks the pro-apoptotic activity of p53 by de-methylating its K370me2 residue³⁷. The di-methylated form of p53 serves to bind its co-activator 53BP1 and enhance p53-dependent transcriptional program^{37,43}.

LSD1 controls DNA damage-induced cell death by increasing the protein stability of the transcription factor E2F1. This occurs through the de-methylation of its K185 residue. Indeed, methylation of this residue stimulates other modifications such as ubiquitination that, in turn, induce E2F1 degradation³⁸.

Lysine methylation of the enzymes DNMT1 and MYPT1 and of the transcription factor HIF1 α regulates their protein stability; by de-methylating these proteins, LSD1 can affect their stability and thus can influence global levels of DNA methylation, cell cycle and glycolysis respectively. In the case of MYPT1, its de-methylation increases the amount of phosphorylated Retinoblastoma (pRb), thus promoting cell cycle progression^{39,40}.

2.4. LSD1 in cancer

LSD1 is overexpressed in several cancers, including solid and haematological tumours. Its higher levels usually correlate with aggressive cancer and poor prognosis⁴⁴⁻⁴⁷. Among the different cancer types, a wide number of studies investigated the involvement of this enzyme in breast cancer, prostate cancer, lung cancer, colon cancer and especially leukemia.

In breast cancer, upregulation of the LSD1 protein level correlates with the progression of the tumour grade. As a matter of fact, LSD1 expression is higher in invasive ductal carcinoma and high grade ductal carcinoma in situ (DCIS) compared to the low grade DCIS

⁴⁸. Despite its well-known function in the context of ER-signalling (as explained in the paragraph 2.2.), LSD1 has also an important role in ER- unresponsive tumours ⁴⁹, thus making it a potential target especially for triple-negative breast cancers (TNBCs), for which the therapeutic possibilities are limited. In ER-negative breast cancer, LSD1 regulates the expression of several proliferation-related genes such as *p21*, *ERBB2* and *CCNA2* ⁴⁷ and increases the progression and invasiveness of the tumour. These enhanced capacities are, at least in part, due to its interaction with the transcription factor Slug. Indeed, this binding is fundamental both to activate EMT-programs ⁵⁰ and suppress the expression of the estrogen receptor 1 gene (*ESR1*), thus making the tumour insensitive to hormone therapy ⁵¹.

In prostate cancer, LSD1 promotes cell proliferation, angiogenesis, migration and invasion ^{52,53}, including in castrate-resistant prostate cancer (CRPC), a type of tumour that grows also in the presence of low androgen levels ⁵⁴. In CRPC, LSD1 acts as an important epigenetic regulator by controlling the expression of genes involving in androgen synthesis, DNA synthesis and cell cycle, through its histone de-methylase activity ⁵⁵⁻⁵⁷. However, a recent study published by Sehrawat *et al.* ⁵⁸ unveiled a novel LSD1 function, independent from its catalytic activity, whereby the protein can promote survival and progression of prostate cancer cells by triggering a transcriptional program specifically in patients affected by a lethal type of prostate cancer. In particular, the authors demonstrated that the interaction of LSD1 with the transcription factor ZNF217 activated this transcriptional program, independently from its histone demethylase activity.

In non-small cell lung cancer (NSCLC), LSD1 up-regulation is associated with short survival of the patients, as a consequence of the increase in cell proliferation and activation of the EMT program ⁵⁹. Interestingly, from a clinical proteomics study by Doll *et al.*, LSD1 emerged as one of the most upregulated proteins in lung metastases, thus suggesting that this protein may be a potential therapeutic option for this terminal illness ⁶⁰.

LSD1 has also a relevant role in colon cancer, since it contributes to the tumorigenesis and metastasis formation. Mechanistically, these processes occur through the activation of the Wnt/ β catenin signalling pathway ⁶¹, as well as by the down-regulation of the E-cadherin gene and promotion of the EMT programs ^{62,63}.

In addition to the above-mentioned cancer, LSD1 has also a remarkable function in many other solid tumours such as neuroblastoma, pancreatic cancer, glioblastoma, medulloblastoma, hepatocarcinoma, bladder and ovarian cancer ^{7,46}.

Regarding the haematological tumours, an abnormal LSD1 activity correlates with a variety of myeloproliferative disorders, since this enzyme has a critical function in the process of terminal haematopoietic differentiation ⁶⁴. Among these malignancies, several studies investigated the oncogenic role of LSD1 in Acute Myeloid Leukemia (AML) and proposed this enzyme as prospective treatment target for different AML subtypes.

In this thesis I will describe the investigation of the mechanism of action and pharmacological modulation of LSD1 in AML, with a specific focus on the subtype Acute Promyelocytic Leukemia (APL).

2.4.1. Acute Myeloid Leukemia

AML is a tumour of the myeloid line of blood cells, characterized by the rapid growth of undifferentiated myeloid precursors that accumulate in the bone marrow and inhibit the growth of normal blood cells ⁶⁵. AML is the most frequent type of acute leukemia in adults, representing the 80% of acute leukemia cases ⁶⁶. It is a heterogeneous group of leukemias characterized by specific and distinctive cytogenetic features. According to the French-American British (FAB) classification, which integrates morphological and cytogenetic characteristics, nine different subtypes of AML can be distinguished ⁶⁷⁻⁶⁹. Among these, APL corresponds to the M3 subtype of the FAB classification and is characterized by the

accumulation of immature granulocytes called promyelocytes. From a genetic point of view, APL cells show the chromosomal translocation involving the promyelocytic leukemia gene (*PML*) -located on the chromosome 15- and the retinoic acid receptor alpha (*RARA*) in chromosome 17⁷⁰. Nowadays, the therapy adopted for this type of leukemia consists in the use of the all-trans retinoic acid (ATRA), the ligand of the receptor $RAR\alpha$, that induces the reactivation of the APL cell differentiation into mature granulocytes⁷¹. The other classical therapy for this pathology is characterized by the use of the arsenic trioxide, which mainly promotes apoptosis, but in combination with cytokines and cyclic AMP (cAMP) also terminal myeloid differentiation⁷²⁻⁷⁵. With these strong advances in the therapy, APL has become one of the most curable forms of acute leukemia, with cure rates of about 80%⁷⁶.

The molecular mechanisms involving APL pathogenesis are well characterized and involve directly the activity of the fusion protein PML- $RAR\alpha$. Specifically, this onco-protein has a dual function: on the one hand, it disrupts the PML nuclear bodies^{71,77} which contain stress-sensitive nuclear domains involving in the activation of p53 and the regulation of senescence^{78,79}; on other hand, it controls the transcription of $RAR\alpha$ target-genes implicated in self-renewal and differentiation. According to the classical model of APL pathogenesis, this latter function is achieved by the recruitment of the nuclear receptor co-repressor complex (NCOR) containing also several histone de-acetylases (HDACs) that facilitate the repression of the target genes⁸⁰⁻⁸². This leads to the increase of self-renewal and the block of differentiation. High concentration of ATRA reactivates the transcription of the downstream target genes through the recruitment of epigenetic co-activators, such as the histone acetyltransferases (HATs)^{80,83}. In addition, ATRA induces the degradation of PML- $RAR\alpha$ fusion gene by activation of the proteasome-dependent pathway and restoration of the PML nuclear bodies^{84,85} (Figure 4). This model demonstrates the important role of numerous epigenetic modifiers for the development of the APL disease and its response to treatment.

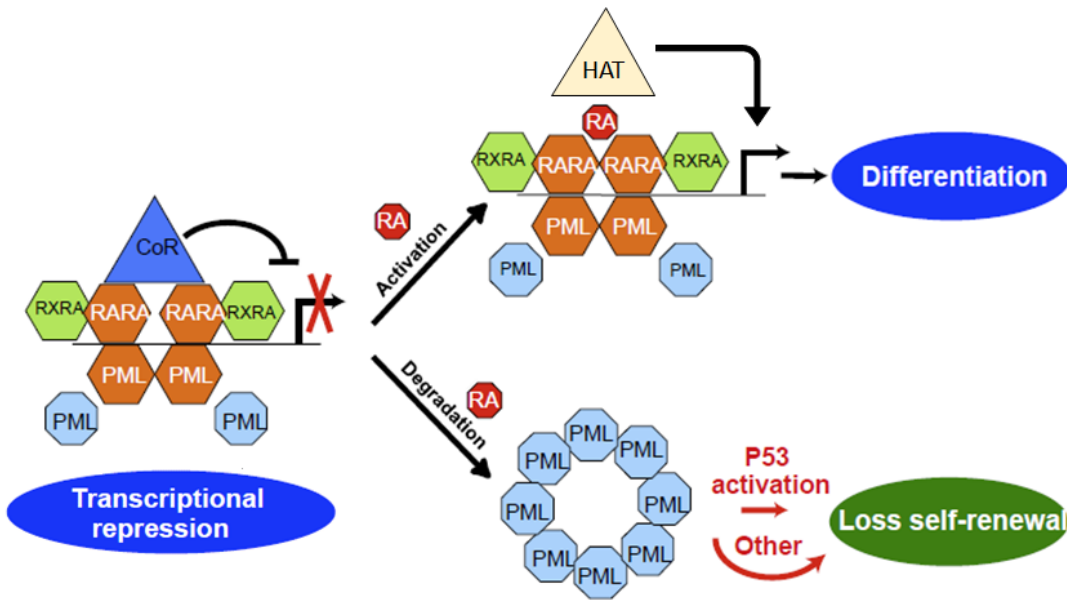


Figure 4: Current model of APL pathogenesis and therapeutic treatment with Retinoic Acid. In the absence of retinoic acid, PML-RAR α fusion protein binds and represses RAR α target genes through recruitment of the nuclear receptor corepressor complex NCOR. In the presence of retinoic acid, a reactivation of RAR α target genes is achieved by recruitment of several co-activators such as histone acetyl-transferases (HATs). Furthermore, in this condition PML-RAR α is degraded and PML can restore its nuclear bodies. This determines the reactivation of p53 and the loss of self-renewal. Adapted from ⁸⁵.

2.4.2. LSD1 in Acute Myeloid Leukemia

Unlike other epigenetic modifiers such as the histone methyl-transferase mixed-lineage leukemia 1 (MLL1), LSD1 is not frequently mutated, but rather over-expressed in about 60% of AML patients ^{22,86}. This over-expression is widespread among all the AML subtypes, therefore is not dependent on specific cytogenetic characteristics ²².

LSD1 has been described as an important regulator of the Leukemic Stem Cells (LSCs) potential, whereby it supports the clonogenic activity of LSCs as well as sustains their oncogenic gene expression programs ⁸⁷. LSD1 was also described as a negative regulator of cell differentiation programs, while it promotes the self-renewal and proliferation of leukemic cells. Indeed, at the transcriptional level, LSD1 inhibition causes the activation of different myeloid lineage genes, such as CD11b and CD86 together with a remarkable decrease of AML cell proliferation and clonogenic capacity ^{88,89}.

The molecular mechanisms underpinning these phenotypic effects are various and largely dependent on the cellular context. For instance, Fang *et al.* and Schenk *et al.* detected a strong increase of the LSD1-histone target H3K4me2 at the promoters of several myeloid differentiation genes following LSD1 inhibition, further corroborating the importance of the LSD1 catalytic activity to drive the re-activation of the differentiation pathway ^{88,90}. However, additional results indicated that LSD1 can control AML progression independently of its de-methylase activity, through its protein-scaffolding function, whereby LSD1 recruits members of the CoREST complex on the target genes of the transcription factors GFI1 or its isoform GFI1b, thus promoting their repression (Figure 5) ^{56,89}.

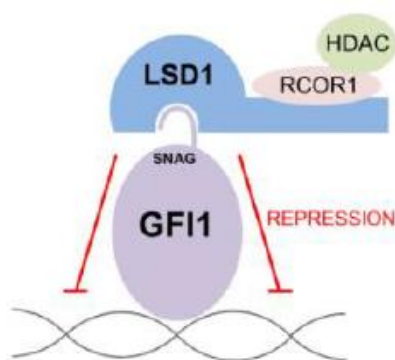


Figure 5: LSD1 scaffolding function. The role of LSD1 as scaffolding protein is exemplified here: LSD1 can recruit the transcription factor GFI1 and other members of the CoREST complex on GFI1-target genes, to maintain them in a repressive state. Adapted from ⁸⁹.

2.5. State-of-the-art on the LSD1 inhibitors

Given the relevant role of LSD1 in cancer, the development of pharmacological inhibitors targeting LSD1 has gained strong interest. Since the enzyme is a member of the monoamine oxidase (MAO) family, to which MAO-A and MAO-B belong, known MAO inhibitors were tested for their efficacy in inhibiting also LSD1 catalytic activity. Among them, the nonselective monoamine oxidase inhibitor tranylcypromine (TCP) was the first compound described to inhibit LSD1 quite efficiently ⁹¹. This drug acts through a suicide-inactivation mechanism: specifically, it binds covalently to FAD forming a covalent adduct that blocks

the LSD1 enzymatic reaction⁹². TCP was described to inhibit colony forming activity (CFA) and to induce differentiation of AML cells derived from MLL-AF9 leukemia mouse models⁸⁷. However, TCP showed low potency and specificity versus LSD1, therefore a series of tranylcypromine derivatives with enhanced LSD1 specificity were developed. The increase of specificity was achieved by modifications of the TCP phenyl-ring, in particular by the addition of side groups or *N*-alkylations. These changes increased the specificity and affinity towards LSD1 since its catalytic site is much larger than that of MAO-A and MAO-B⁹³. Currently, the most efficient LSD1 inhibitors developed in the last years are the irreversible compounds ORY-1001, GSK2879552, OG86, T-3775440, RN-1, NCD38 and the reversible ones GSK690 and SP2509 (Table 1). Apart from the reversible drugs, all the irreversible ones are derived from TCP.

Table 1: Most studied LSD1 inhibitors in cancer.				
Compound	Drug type	Tumour setting	Manufacturer	Ref.
ORY-1001	Tranylcypromine derivative	AML, SCLC	Oryzon Genomics	94-97
GSK2879552	Tranylcypromine derivative	AML, SCLC, T-ALL, Hepatocellular carcinoma	GlaxoSmithKline	98-103
OG86	Tranylcypromine derivative	AML	Oryzon Genomics	89
T-3775440	Tranylcypromine derivative	AML, SCLCL	Takeda Pharmaceutical	104,105
RN-1	Tranylcypromine derivative	AML	Athinoula A. Martinos Center for Biomedical Imaging	106
NCD38	Tranylcypromine derivative	AML, erythroleukemia, megakaryoblastic leukemia	Kyoto University	107
SP2509	Benzohydrazide	AML, Ewing sarcoma	Salarius Pharmaceutical	108
GSK690	Benzohydrazide	AML, rhabdomyosarcoma	GlaxoSmithKline	109

The compound ORY-1001, produced by the Oryzon Genomics, was studied in the context of AML and SCLC ^{94,96}. In AML, this compound induces differentiation and reduces colony-forming potential of leukemic cells. Furthermore, it showed great synergy with standard-of-care drugs, such as cytarabine (ara-C) as well as with other specific inhibitors like quizartinib ¹¹⁰ (a FLT3 inhibitor) and the DOT1L inhibitors ^{94,111,112}. Currently, ORY-1001 is in early phases clinical trials and results are expected in refractory and relapsed AML ¹¹³.

Also the compound GSK2879552 was in clinical trials for the treatment of relapsed and refractory AML and SCLC, but the risk/benefit analysis did not favour the continuation of the study. This drug has been used in different types of tumours: in addition to AML and SCLC also in hepatocellular carcinoma and T-cell lymphoblastic leukemia (T-ALL) ⁹⁹⁻¹⁰². In AML, GSK2879552 reduced cell viability of MLL-AF9 cells ¹⁰¹ and, in combination with ATRA, it gained cytotoxic effects across different AML subtypes ¹⁰³.

OG86, another compound produced by the Oryzon Genomics, was one of the first tranlycypromine derivatives that was demonstrated to alter the LSD1 scaffolding function by displacing LSD1 itself from its interaction with GFI1 and inducing, in this manner, differentiation of THP-1 AML cells ⁸⁹. A similar mechanism of action was also proved in both SCLC and AML for the inhibitor T-3775440. In these two cancer cell models, T-3775440 exerted its anti-proliferative effects by altering the interaction of LSD1 with the SNAG domain proteins GFI1b and INSM1 ^{104,105}.

RN-1 and NCD38 displayed strong effects in impairing cell growth of different AML subtypes ^{106,107}. In particular, NCD38 was shown to inhibit leukemogenic programs by activating the super-enhancers of different hematopoietic regulators like *GFI1* and *ERG* ¹⁰⁷.

The reversible LSD1 inhibitors, instead, compete with the LSD1 substrate for the binding to the enzyme. Despite most of these drugs are in early phase of development, they have the potentiality to reduce some of the side-effects that TCP derivatives showed on erythropoiesis ¹¹³. In most of the studies, these drugs were tested in combination with other compounds,

like HDAC inhibitors^{108,109}. This co-treatment significantly reduced cultured and primary AML blasts and enhanced the survival of mice engrafted with human AML cells, without eliciting evident side-effects¹⁰⁸.

In some studies, LSD1 inhibitors were tested in combination with ATRA as a new “differentiation therapy” for the treatment of several AML subtypes¹¹⁴. In these studies, LSD1 inhibitors greatly potentiated the ATRA-driven differentiation process of non-APL AML cells by increasing the expression of myeloid-associated genes^{90,115}. These results strongly suggest that this combined treatment may be used for the AML treatment and could be explored in more details, especially from a molecular point of view.

2.5.1. The LSD1 inhibitors MC2580 and DDP-38003

Between 2010 and 2016 the Experimental Therapeutic Unit of our Campus, in collaboration with the University of Pavia and La Sapienza University in Rome developed a series of irreversible anti-LSD1 compounds showing higher selectivity towards LSD1 compared to other structurally and functionally similar enzymes such as MAO-A, MAO-B and LSD2^{116,117}. Among these, MC2580 and DDP-38003 have been used and studied in this thesis.

The first set of inhibitors including the MC2580 were produced in 2010 and developed using TCP as chemical scaffold. As in the case of TCP, these compounds bind covalently to the co-factor FAD, thus impeding the LSD1 catalytic reaction⁹². MC2580 was the most selective of these drugs, even though it maintained some inhibition activity towards MAO-A. Specifically, its biochemical half maximal inhibitory concentration (IC₅₀) for LSD1 inhibition corresponds to 0.129 μM while for MAO-A to 0.31 μM (Table 2). Biochemical and phenotypic cellular assays proved the efficacy of this compound in inhibiting H3K4me2 demethylation and synergizing with ATRA to reduce cell growth and trigger differentiation of NB4 cells and primary murine APL blasts¹¹⁶.

After this initial study, in 2016 a set of novel drugs were developed based on the same TCP scaffold, but showing higher potency (Table 2) ¹¹⁷. Among them, the compound named DDP-38003 was one of the most potent and thus selected for the cellular and *in vivo* studies. Specifically, its biochemical IC₅₀ towards LSD1 is much lower than that of MC2580, while its inhibitory activity on MAO-A similar (Table 2). At cellular level, DDP-38003 displayed strong inhibitory effects in murine APL blasts, as demonstrated by the effective reduction of the colony forming unit (CFU) and the increase of LSD1 target genes expression upon treatment ¹¹⁷.

Table 2: Biochemical IC₅₀ values of MC2580 and DDP-38003.

Compound	LSD1 IC₅₀ (μM)	LSD2 IC₅₀ (μM)	MAO-A IC₅₀ (μM)	MAO-B IC₅₀ (μM)
MC2580	0.129	not tested	0.31	> 100
DDP-38003	0.06	36.6	0.29	84.59

2.6. Effects of pharmacological inhibition of LSD1 in NB4 and UF-1 APL cells

The group of Prof. Saverio Minucci at IEO, with whom we are collaborating, have been investigating the phenotypic and molecular effects of LSD1 inhibitors on APL cell lines for about ten years. Most of their studies focused initially on MC2580, but many of the assays and results collected were then validated with DDP-38003. The two cell lines employed in this project were NB4 and UF-1 cells. Unlike the NB4, UF-1 cells were established from a patient clinically resistant to ATRA treatment ¹¹⁸ and, therefore, represent a more interesting model to investigate from a therapeutic point of view, being representative of patients resistant to the treatment currently in use.

These cell lines displayed differential response to LSD1 inhibitors, both in terms of cell growth and differentiation. In particular, UF-1 cells were more sensitive to the MC2580 treatment compared to NB4, as displayed by the strong proliferation reduction following the treatment with the drug. Conversely, NB4 cells showed only a little effect on cell growth

after MC2580 treatment (Figure 6A-B). The observed phenotype was due to cell differentiation, as confirmed by the greater progression in the nuclear lobulation process (a typical process of myeloid cells associated with neutrophilic differentiation) in UF-1 cells compared to NB4 (Figure 6C).

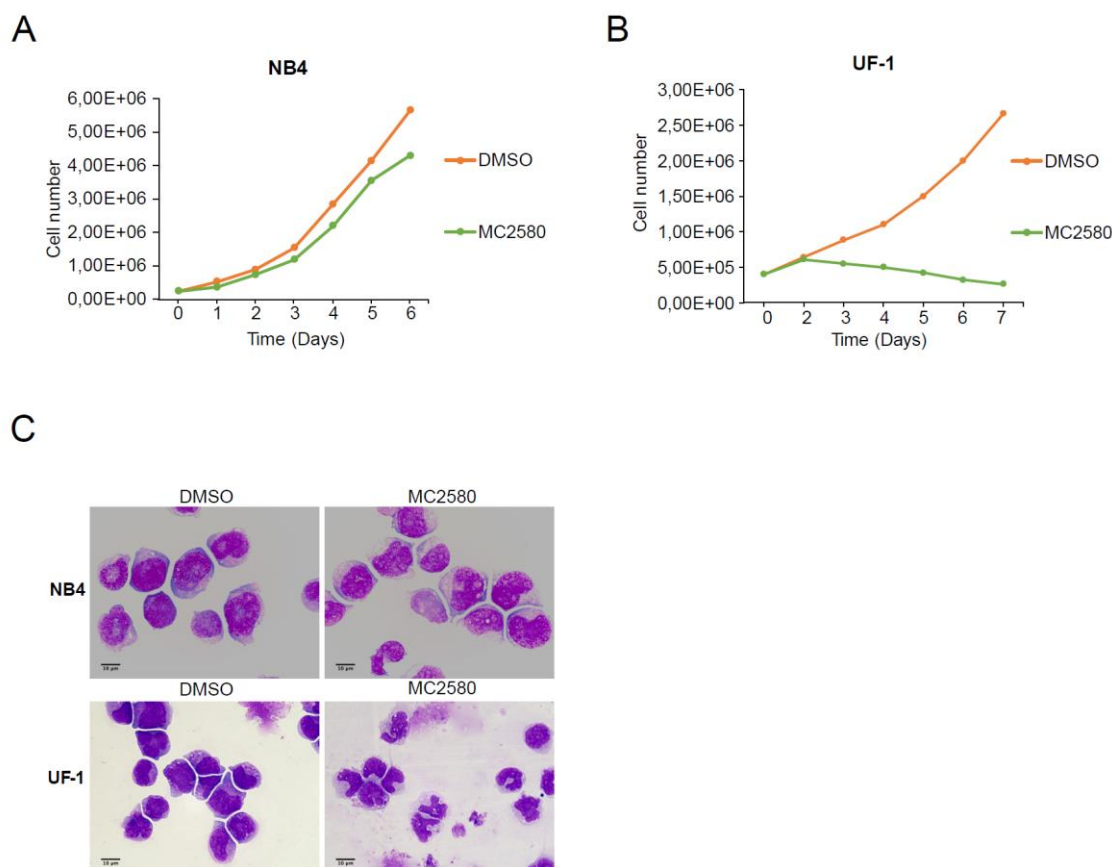


Figure 6: Effects of MC2580 in NB4 and UF-1 cells. A) Representative cell growth of NB4 cells treated with either MC2580 (2 μ M) or control dimethyl sulfoxide (DMSO) for 6 days. B) Representative cell growth of UF-1 cells treated with either MC2580 (2 μ M) or control DMSO for 7 days. C) Morphologic analysis through May Grünwald-Giemsa staining of NB4 and UF-1 cells treated with MC2580 (2 μ M) and DMSO for 96 hours in liquid culture.

2.6.1. Combinatorial effect of LSD1 inhibitors with physiological doses of ATRA in NB4 cells

Having observed that treatment with any LSD1 inhibitor elicited very limited response on NB4 cell when used alone, the Minucci's group investigated the effect of the combined

treatment with LSD1 inhibitors and physiological doses of ATRA, which are also ineffective *per se* (ATRA 0.01 μM)¹¹⁹. Previous studies had proved that NB4 cells undergo a complete differentiation process when 1 μM ATRA is used¹²⁰, which is however a concentration at which some adverse effects are observed. This led to the question whether the combination of the two drugs at doses which individually ineffective and thus not toxic could be a possible strategy to achieve leukemic cell differentiation without side effects.

Indeed, the combination of MC2580 2 μM + ATRA 0.01 μM turned out to be much more effective, with a strong reduction of cellular proliferation, to a similar extent than ATRA 1 μM and MC2580 2 μM + ATRA 1 μM (Figure 7A). Also in this case the reduced proliferation was a consequence of cell differentiation, as confirmed by the strong induction of the myeloid differentiation marker CD11b and by the observed morphological changes associated with nuclear lobulation (Figure 7B-C). This differentiation process was re-activated without PML-RAR α degradation, which instead occurs when high doses of ATRA are used (Figure 7D). This is particularly interesting, because it suggests that this novel combinatorial drug treatment bypasses the action of the main oncogene and therefore might be applied also to other less curable AML subtypes, not based on PML-RAR α fusion.

Another interesting observation was that the re-expression of either the wild-type (WT) or the catalytically-inactive (K661A) form of LSD1 in LSD1 knock-out (KO) NB4 cells, previously generated by the Minucci's group through the CRISPR/Cas9 technology, induced a similar molecular and cellular effect. Indeed, while NB4 LSD1-KO cells treated with ATRA showed a strong reduction of cell growth and higher expression of differentiation markers such as CD11b, both the wild-type (WT) and the catalytic inactive LSD1 (K661A) were able to rescue the proliferative and clonogenic potential of the cells (Figure 7E-F). These results demonstrated that LSD1 catalytic activity is dispensable in these cells to promote self-renewal and proliferation.

These results indicate that the effect of the LSD1 inhibitors in triggering cell differentiation is not due to the inhibition of the enzymatic activity, but is achieved through a different mechanisms. Elaborating on the published evidence which showed that some LSD1 inhibitors induce differentiation of leukemic cells by altering the interaction of LSD1 with some other proteins ^{89,104,105} (as described in the paragraph 2.5.), we set to investigate the effect of these drugs on the LSD1 interactome by using quantitative mass-spectrometry (MS)-based approaches.

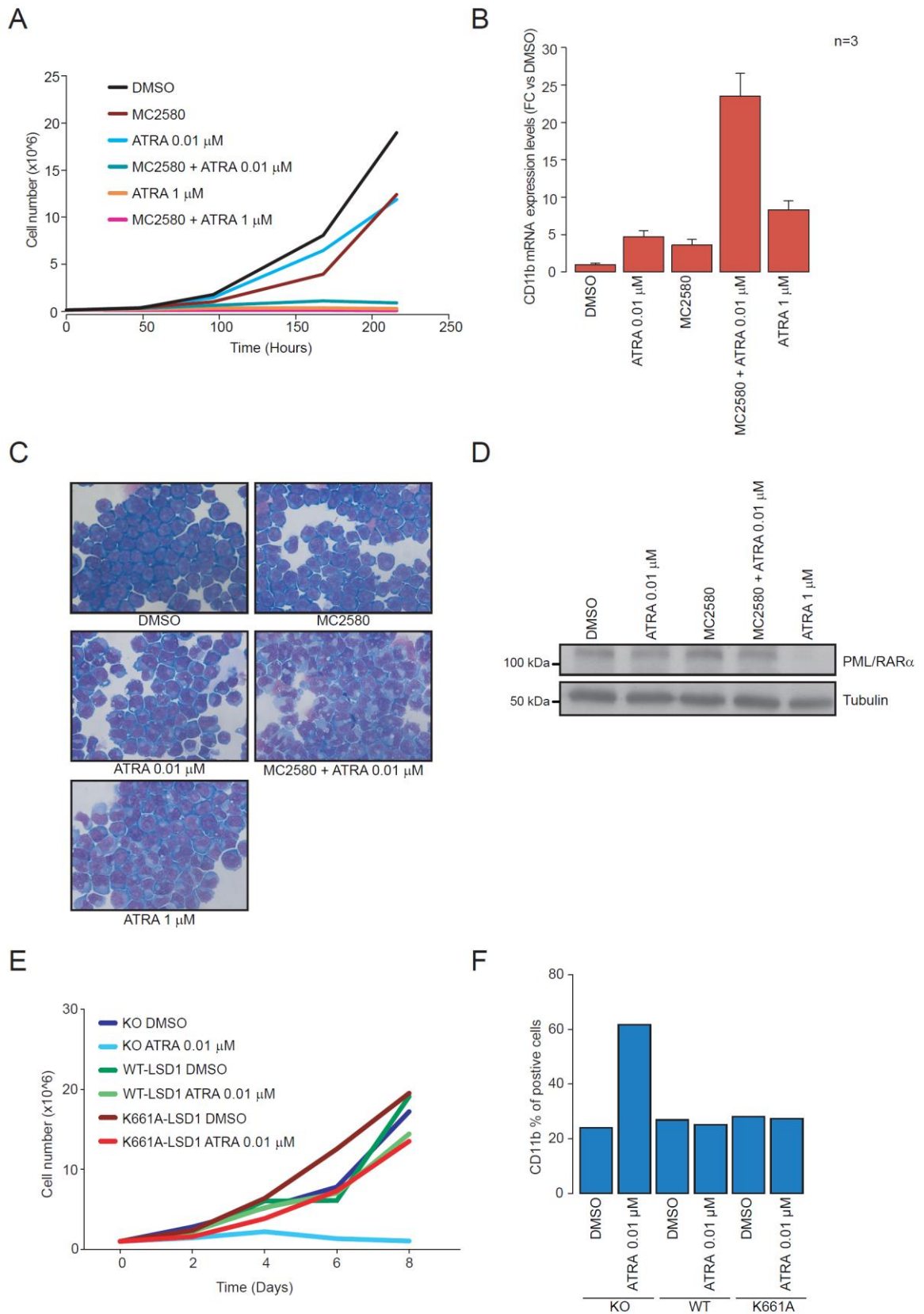


Figure 7: LSD1 inhibitors sensitize NB4 cells to physiological doses of all-trans retinoic acid (ATRA) irrespective of the inhibition of the LSD1 catalytic activity. A) Growth curve of NB4 cells treated with the LSD1 inhibitor MC2580 2 μ M, ATRA 0.01 μ M, ATRA 1 μ M and the combination of drugs for 6 days. B) Analysis of CD11b mRNA levels in NB4 cells treated with the LSD1 inhibitor MC2580 2 μ M, ATRA 0.01 μ M, ATRA 1 μ M and

combination of MC2580 2 μM plus ATRA 0.01 μM for 96 hours in liquid culture. Ct values are normalized against glyceraldehyde 3-phosphate dehydrogenase (GAPDH) and referred to DMSO. Graph represents the mean and standard deviation of three independent experiments. C) Morphologic analysis through May Grünwald-Giemsa staining of NB4 cells treated with the LSD1 inhibitor MC2580 2 μM , ATRA 0.01 μM , ATRA 1 μM and combination of MC2580 2 μM plus ATRA 0.01 μM for 96 hours in liquid culture. D) Western blot analysis of PML-RAR α levels in NB4 cells treated with the LSD1 inhibitor MC2580 2 μM , ATRA 0.01 μM , ATRA 1 μM and combination of MC2580 2 μM plus ATRA 0.01 μM for 24 hours. Tubulin is used as loading control. E) Cell proliferation assay of LSD1-KO NB4 cells infected with WT-LSD1, K661A-LSD1 and empty vector (KO). Cells are treated with either DMSO or ATRA 0.01 μM up to 8 days. (D) FACS analysis of CD11b in LSD1-KO NB4 cells infected with WT-LSD1, K661A-LSD1 and empty vector (KO). Cells are treated with either DMSO or ATRA 0.01 μM for 24 hours.

2.7. Mass spectrometry (MS)-based proteomics

Mass-spectrometry (MS) is the main analytical technique adopted for the identification and quantitation of proteins in biological samples¹²¹. In the last decades, several technological improvements led to the development of mass spectrometers characterized by progressively higher resolution, sensitivity and performance. Nowadays, the most common analytical setup used in proteomics analysis is the liquid-chromatography coupled with tandem mass spectrometry (LC-MS/MS). In its classical workflow, known as “bottom-up” approach, proteins are first separated within a complex mixture, then digested with specific proteases into peptides prior to the MS analysis. The most widely used protease in these experiments is trypsin, since it produces peptides about 10-12 amino acids long that are more prone to be detected in MS. To reduce the complexity of the mixture, tryptic peptides are then separated by nano-liquid chromatography, usually in reversed-phase setup (RP-nLC) according to their hydrophobicity. In the mass spectrometer, precursor peptides are isolated and scanned (full scan MS or MS1) to detect their mass-to-charge ratios (m/z). The most intense precursor peptides are, then, fragmented into their constituent fragments by collisions with rare gas atoms (tandem MS or MS2) to retrieve the information of the amino acid sequence of the peptides. Specific search engines allow the identification of peptides analyzed in MS by

comparing the experimental MS/MS spectra with a theoretical one produced by the *in silico* digestion of a protein database. The output of this search consists in a list of fragment ion spectra matched to the peptide sequence and ranked according to the search score. The score is a measure of the similarity between the experimental spectrum and the theoretical one. This value allows discriminating the true positives from the false identifications. In this manner, peptide sequences are assigned to specific proteins thus permitting their identification (Figure 8). Different programs are available and used for MS-protein identification, such as Andromeda within the MaxQuant suite ¹²², MASCOT ¹²³ and SEQUEST ¹²⁴.

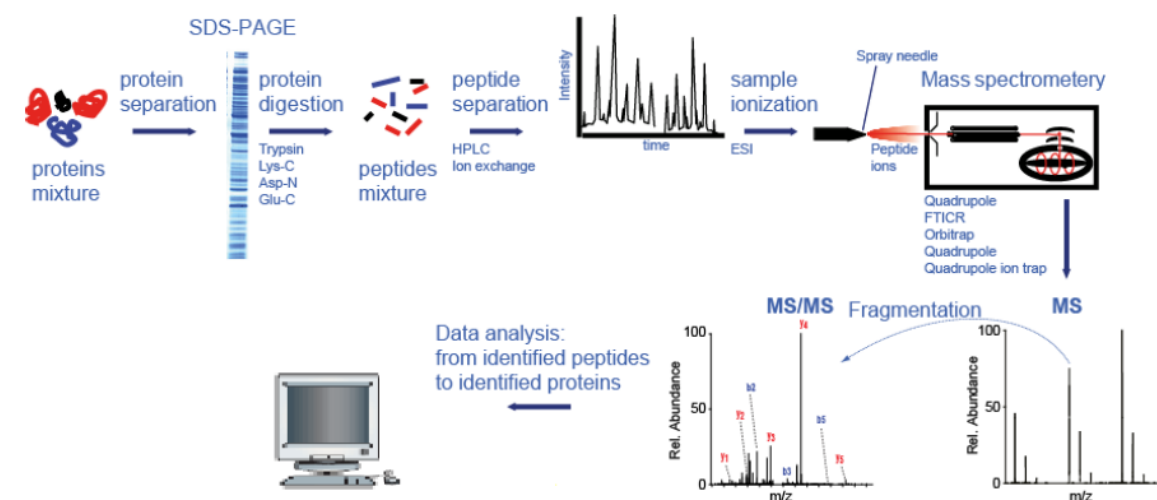


Figure 8: Schematic workflow of a standard “bottom-up” approach for MS analysis.

A protein mixture extracted from a biological sample is separated by sodium dodecyl sulphate - polyacrylamide gel electrophoresis (SDS-PAGE) and in-gel digested into peptides by proteases (most commonly trypsin). Peptides are, then, separated by high-performance nano-liquid chromatography (nHPLC), ionized through the ion source into the mass spectrometer and analysed by tandem mass spectrometry (MS/MS), which produces both MS and MS/MS spectra. Specific search engines compare the experimental MS/MS spectra with the theoretical ones extrapolated from a protein database to identify proteins within the mixture. Taken from www.biochem.mpg.de.

2.8. The mass spectrometer

A mass spectrometer is composed of three main elements: 1) the ion source that transforms the peptides into gas-phase ions, 2) the mass analyzer that distinguishes the ions according to their m/z and 3) the detector that records the number of ions per each m/z value.

I. Ion Source

Since mass spectrometers measure m/z values of ionized molecules in gas-phase, peptides must be converted into the gas-phase for the subsequent MS-analysis. Two different ionization techniques are commonly used: matrix-assisted laser desorption ionization (MALDI) ^{125,126} and electrospray ionization (ESI) ¹²⁷. In the context of this thesis, I will explain the principle of the ESI since it is the ion source installed in the mass spectrometer used for the MS-experiments.

ESI source generates ions in a gas phase from peptides in aqueous solution and their charge is controlled by the pH of the solution. At acidic pH values, protonation of the amines gives positive charge to peptides, while at high (basic) pH, de-protonation of the amines and carboxyl groups confers an overall negative charge. Fragmentation of the peptide ions is favoured by positive charges and, for this reason, ESI of peptides is commonly done in the positive ion mode. During the ESI process, the application of a high voltage (2–6 kV) between the end of the LC-column and the entrance of the mass spectrometer forms an electrically charged spray that causes de-solvation of peptide droplets and formation of ions with a certain charge. Heated capillary and the sheath gas flow present at the mass spectrometer inlet produces a high temperature that helps this process (Figure 9). ESI sources are, usually, combined “on-line” with the LC instruments, especially reverse phase (RP)-LC, to obtain continuous analysis of the samples. An important development of ESI includes nano-ESI sources, where the flow rates are lowered to nanoliter/min to improve the sensitivity ^{128,129} and increase the concentration of the analyte.

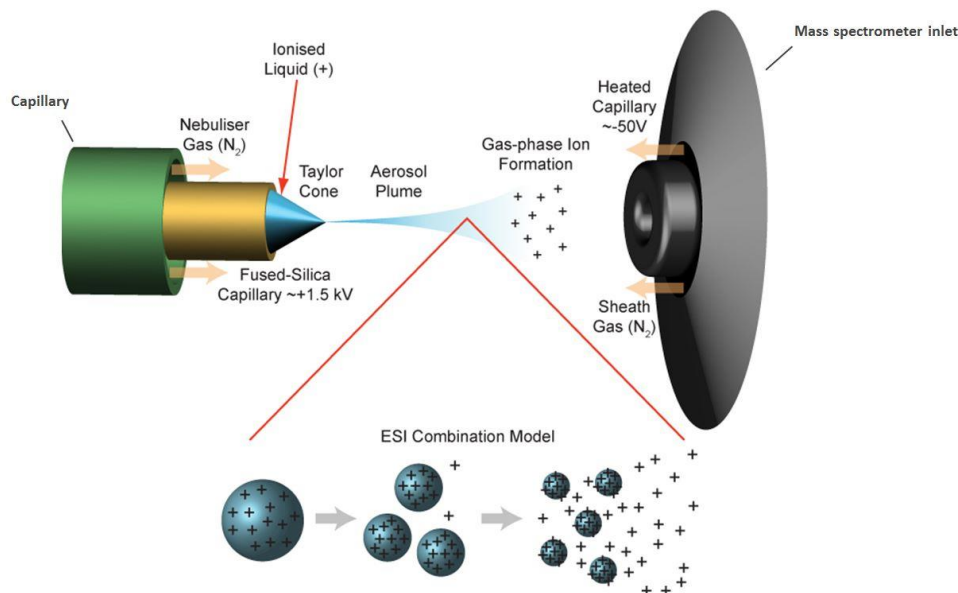


Figure 9: Electro spray ionization (ESI) process. Peptides eluted from the chromatographic column are ionized by a high voltage applied between the end of the column and the mass spectrometer. Charged liquid generates a cone shape, known as the Taylor cone, where the analyte in the form of solvent droplets burst away into a spray. Taken from <http://www.lamondlab.com>.

II. Mass analyzer and fragmentation methods

The mass analyser is the central core of the mass spectrometer and its main role consists in the storage and separation of ions based on their m/z . The most common mass analysers in proteomics are the linear ion trap (LIT), the Orbitrap, the Fourier-transform ion cyclotron resonance (FT-ICR), the quadrupole (Q) and the time-of-flight (TOF). Despite these analysers differ in their mode of operation, all of them have the function to select a single m/z species from a mixture of peptide ions and lead them to be fragmented for generating the MS/MS spectrum. This allows extrapolating the primary sequence of the peptides and assuming the presence and position of post-translational modifications.

Different fragmentation techniques can be used, but the most common are the collision-induced dissociation (CID)¹³⁰ and the higher energy collision dissociation (HCD) methods. In these types of fragmentation, protonated ion peptides underwent multiple collisions with rare gas atoms, thus generating the breakage of the peptide backbone at $-CO-NH-$ bonds¹³¹ and the formation of the characteristic b- and y-ions (at N- and C-terminus, respectively)

which are used for the univocal peptide identification. CID and HCD are suitable for the analysis of post-translational modifications (PTMs) in “bottom-up” experiments since they increase the frequency of modification-specific fragmentation events in high resolution MS/MS spectra ¹³². However, these methods produce only limited information for peptides longer than 15 amino acids. In these cases, electron capture dissociation (ECD) ¹³³ and electron transfer dissociation (ETD) ¹³⁴ can be used. These approaches induce fragmentation of the peptide backbone based on gas-phase reactions using either thermal electrons or formation of radical ions ¹³⁵.

In a classical shotgun proteomics workflow, peptides are analysed in a data-dependent acquisition setup (DDA) ¹³⁶, where the mass analyser selects the most intense ions in a certain time window and lead them to fragmentation. Afterwards, the analyser automatically returns to the MS1 mode and selects the next most intense ions to be fragmented. Thus, DDA is biased towards the most abundant species.

III. Detector

The detector is at the end of the mass spectrometer. It has the function to record the number of ions at each m/z value and convert them into mass spectra. Most frequently, the detectors are electron multipliers or micro-channel plates that produce an electronic cascade when each ion hits the detector plate. This process is performed under vacuum to eliminate gas molecules and non-sample ions that can collide with the sample ions generating non-specific reaction products.

2.9. Quantitative proteomics

In the last decade, it has become apparent that the sole qualitative identification of proteins is not enough to describe the dynamicity of the proteome at different functional states and

that a quantitative information about global changes in protein levels could be much more informative. Yet, mass spectrometry is not an inherently quantitative technique, for various reasons. First, the intensity of each ion is proportional to its quantity but is also dependent on the chemical-physical properties of the corresponding peptide, such as the charge, the hydrophobicity and the molecular composition. Second, each LC-MS/MS run is influenced by external variations such as the temperature and the chromatography reproducibility thus rendering difficult the comparison of different samples. In order to overcome these limitations, different quantitative strategies have been developed and can be divided in two distinct groups: label-free quantitation (LFQ) and isotope-labelling approaches (Figure 10).

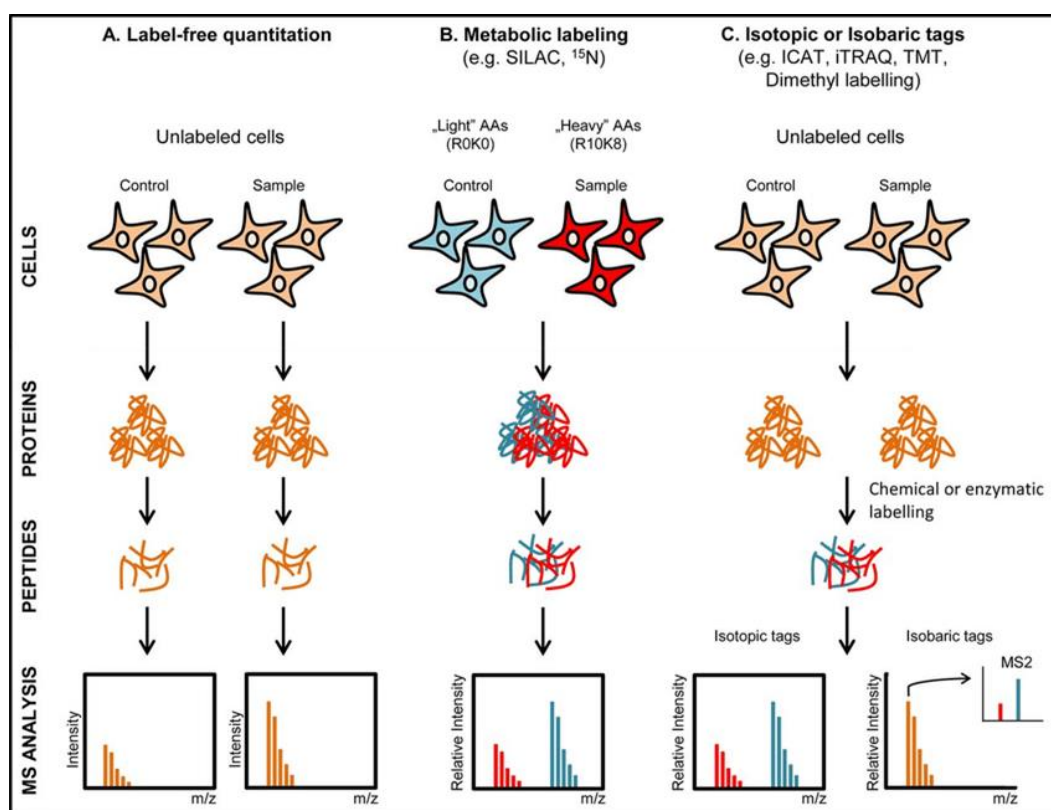


Figure 10: Quantitative MS-based proteomics approaches. Three main available strategies for protein quantification in MS proteomics are: label-free quantitation (LFQ), metabolic labelling and chemical labelling. In LFQ (A), control and sample are kept separated during the entire workflow of sample preparation and run separately in MS. In metabolic labelling workflows such as stable isotope labelling by amino acids in cell culture (SILAC) (B), the different biological conditions are labelled with light (K0, R0) and heavy (K8, R10) isotopes of lysine and arginine and, then, cells or protein extracts are mixed in equal amount before sample preparation, protein digestion and MS-analysis. In chemical labelling approaches (C), isotopic or isobaric tags are added to purified protein or peptides and the resulting differentially labelled conditions are pooled together before MS-analysis.

Quantitation can be performed at the MS1 level or and at the MS2 as in the case of the isobaric tags for relative and absolute quantitation (iTRAQ) and the tandem mass tag (TMT) strategies. Adapted from ¹³⁷.

Label-free strategies take advantage of the intrinsic characteristics of the polypeptides. These strategies can be divided, in turn, in two groups: 1) intensity-based strategies, where one or more peptide intensities of the same protein are quantified and compared among different samples and 2) spectra counting strategies, where the quantitation is based on the number of MS/MS spectra for the same protein in different samples. The first approach is based on the rationale that extracted ion chromatogram (XIC, or area under the curve, AUC) of a given peptide linearly correlates with its abundance. Therefore, XIC of each peptide from a protein is quantified and compared among the different samples. These strategies need extremely reproducible chromatography among multiple runs and *ad hoc* software capable to perform retention time realignment and peptide intensity normalization over the global intensity. Spectral counting strategies are, instead, based on the concept that the abundance of a protein correlates linearly with the number of MS/MS spectra generated by its peptides. Relative quantitation is, therefore, obtained by comparing the number of these spectra between different functional states ^{138,139}. To be accurate, a high number of spectra must be acquired since the physical and chemical properties of the peptides influence the above-mentioned linear correlation between number of spectra and abundance. Theoretically, label-free strategies can be applied to an unlimited number of sample types, from tissues to cell lines and primary cells. However, they require high technical and experimental reproducibility and generate an overall lower accuracy in protein quantitation compared to isotope-based techniques.

Isotope-labelling methods can be divided into metabolic and chemical strategies depending on how proteins or peptides are labelled. All these approaches are based on the same concept: producing a mass shift that differentiates proteins belonging to different biological

conditions within a single MS analysis. Despite the isotope could be introduced in the protein at different stages of the sample preparation, earlier this step is performed and less are the variations introduced between the samples, thus generating a more accurate quantitation.

In the metabolic labeling strategies, the isotope is added to culturing cells as a metabolic precursor in order to be incorporated in the proteome during protein biosynthesis. These approaches are very advantageous because they can be applied to *in vivo* studies and the sample mixing is performed at the beginning of the workflow, thus minimizing the introduction of processing errors. The most successful metabolic labelling strategy was named stable isotope labelling by amino acids in cell culture (SILAC) and was introduced by Mann and *et al.* in 2002¹⁴⁰. This strategy, which was applied in this thesis in different MS-experiments, is based on the use of essential amino acids, typically lysine and arginine, where the naturally isotopes ¹²C, ¹H and ¹⁴N (or light, L) are substituted with heavy (or H) isotopes (¹³C, ²H and ¹⁵N, respectively). In a standard SILAC workflow, two cell populations representing two different biological states are grown in medium containing either the light or the heavy isotope encoded version of the essential amino acids for a certain number of cell doublings to guarantee the full incorporation of the isotopes in the newly synthesized proteins. Upon complete labelling, cells grown in the two media have identical proteomes whereby each protein differs only in its respective isotopic composition. At this step, samples can be mixed in equal amounts and further processed for the MS-analysis. In a MS1 spectrum, each peptide is represented as a pair of chemically identical peptides with different stable isotope composition, where the Δ -mass shift is dependent on both the number and the type of H and L amino acids incorporated in the sequence¹⁴⁰. Metabolic labelling with H and L arginine and lysine in combination with trypsin digestion ensures that all peptides of a protein (except the C-terminal one) contain at least one labelled amino acid that is therefore able to produce a Δ -mass. SILAC metabolic labelling has been successfully used in a variety

of studies: from protein expression to global PTM analysis, protein-protein and DNA-protein interactions ¹⁴¹⁻¹⁴⁴.

In chemical-labelling approaches, a specific tag is bound covalently to the reactive side chains of amino acids through a chemical reaction that is carried out either before or after the proteolytic cleavage. The most known and used chemical labelling methods are: isobaric tags for relative and absolute quantitation (iTRAQ) ¹⁴⁵ and tandem mass tags (TMT) ¹⁴⁶. In iTRAQ, side chains of lysine residues and the N-terminus of peptides are labelled with an isobaric tag that allows the distinction and quantification of peptides belonging to different experimental states at the MS/MS level. The great advantage of this approach is the possibility of multiplexing, by using up to 8 different isobaric tags. Tandem mass tag (TMT) is a similar quantitative approach and is based on the same principle, namely the use of isobaric tags attached at the N- terminus of a peptide or to a lysine residue, with the option of multiplexing up to 10 different tags ¹⁴⁶. One limitation of these approaches is that the labelling is often performed at a later step of sample processing compared to SILAC, and this can result in a less accurate protein quantification.

2.10. Immuno-precipitation coupled with MS to study protein-protein interactions

One of the most common application of MS-based proteomics concerns the study of protein-protein interactions (PPIs). In this regard, this technique has been proved to be very efficient in dissecting in a global and unbiased manner protein-interaction networks, understanding the biological functions associated to distinct protein complexes and evaluating the plasticity of PPIs following various stimuli or perturbations ^{147,148}.

The first step of an interaction proteomics workflow is the purification of the protein of interest, defined as bait. This is obtained by affinity-purification (AP) strategies using either antibodies or tag-based approaches ^{149,150}. The AP method used in this thesis is based on

antibody purification of endogenous proteins (Figure 11), thus in this introductory section I will focus on this specific approach explaining its advantages and disadvantages.

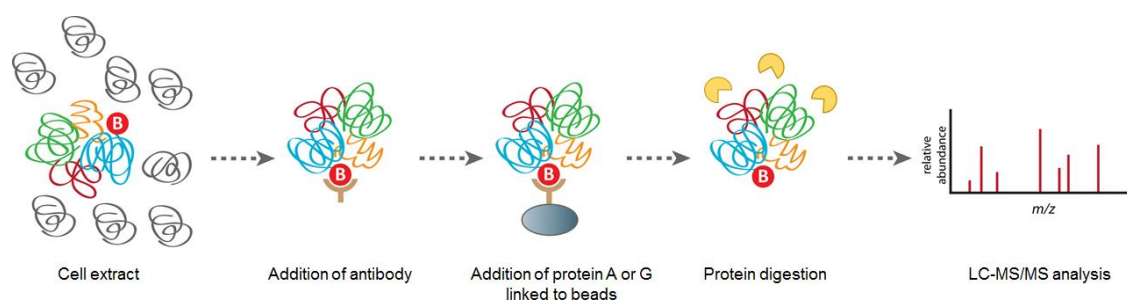


Figure 11: Workflow of affinity-purification (AP) strategy with antibody targeting an endogenous protein coupled with MS-analysis. In this workflow, the antibody binds to the protein bait together with its associated binders and the immune-complex is then purified by protein A or G linked to commercial beads. Protein complexes are subsequently digested with proteases such as trypsin and subjected to LC-MS/MS analysis. Adapted from: www.uu.nl.

Immuno-precipitation (IP) with antibodies targeting endogenous proteins has the main advantage that the bait is purified from cells or tissue extracts in close to physiological conditions, so that its conformation, modification state and stoichiometry is preserved and representative of its status within the cellular context. Furthermore, multiple isoforms such as different splice variants can be in principle analysed simultaneously, if the antibody recognizes them ¹²¹. On the other hand, this strategy presents several issues: one major issue regards the amount of bait available in the cells, with a low expressed bait which typically leads to a reduced detection of its interactors by MS. Another problem concerns the high concentration of antibody added to the affinity purification reaction that can mask in MS the detection of low abundant peptides derived from specific interactors. A few strategies have been developed to overcome this problem, such as the use of antibodies cross-linked to the beads and of pH-based buffers to elute the immuno-precipitated proteins from the beads. Specifically, the antibody chains are denatured in low pH buffers with the consequent elution of the bait and its interactors; however the antibody chains will remain cross-linked to the beads, limiting the amount of antibody contamination compared to buffers based on

denaturing agents ¹⁵¹. Another method is based on the separation of the eluted proteins by SDS-PAGE, whereby the protein band of the gel containing light and heavy antibody chains are cut, digested and analysed by MS separately from the rest of the sample. Here, the problem remains the detection of proteins having the same size and running in the same position of the antibody H and L chains which still may be masked. An additional problem concerns the availability of antibodies against the endogenous proteins: as a matter of fact, despite the extensive production of antibodies against as many human and mouse proteins as possible in the last decade, only a portion of them is really suitable for efficient and specific IP experiments.

Last, one of the major problems in IP-MS studies is the identification of several false positives, namely proteins that are co-immuno-precipitated with the bait but are not true interactors. In the majority of cases, they correspond to sticky or high abundant proteins that bind aspecifically to the antibody or the beads ¹⁵². The false positive detection can be addressed biochemically by different strategies, like increasing the salt concentration or the amount of detergents in the washing solutions as well as reducing the incubation time of the antibody and the beads within the protein extract. In the case of the RNA-binding proteins, ribonuclease A (RNaseA) can be added to the protein extract to remove false positive interactions eventually derived from the binding of the bait to cellular RNAs ¹⁵³. However, these biochemical methods are in most cases insufficient to remove all the false positives derived from an interactomics experiment. Therefore, the use of a proper experimental control in parallel with the sample of interest is highly useful. The main controls adopted in these experiments include the use of isotype-matched antibodies (e.g. IgG), as well as an excess of soluble blocking peptide, which competes with the bait for the antibody binding allowing to better discriminate specific from unspecific binders ^{121,154}. The aforementioned experimental controls are, however, not suitable when antibodies show high cross-reactivity ¹⁵⁵. In this case, the most efficient negative control is given by the IP performed in the absence of the protein bait, which could be depleted by knock-down or knock-out ^{121,156}.

2.10.1. Quantitative proteomics to study PPIs

A better strategy to study PPIs and to discriminate, in a better manner, the true interactors from background proteins is based on quantitative proteomics¹²¹. In this regard, several studies combined biochemical approaches of protein complexes purification to label-free based quantitation. Here, quantitation is based on spectral counts or intensity-based measurements either at the MS1 or at the MS2 level (i.e. analysing intensity of the fragment ions). In all these cases, the idea is to distinguish the “signal” of the proteins identified by MS in the sample where the bait is purified in relation to the control IP. Different computational and statistical approaches were applied to score these quantitative values and establish a threshold that allows distinguishing the interactors from the background^{121,157,158}.

Isotope-labelling strategies have also been applied in MS-interactomics studies to increase the discrimination of specific versus unspecific interactors. One of the most commonly used is the SILAC approach. The use of SILAC in this type of studies is based on a simple principle: if the cells where the bait is purified are labelled with light isotopes and the control with heavy isotopes, a putative interactor should be enriched in the light-labelled condition compared to the heavy one while nonspecific binders should be present in equal amounts in the two SILAC conditions or more enriched in the negative control¹⁵⁹ (Figure 12). According to this assumption, several groups have successfully used this metabolic labelling strategy in quantitative MS-studies to analyse PPIs. For instance, Selbach et al. used SILAC-quantitative proteomics in combination with protein co-immunoprecipitation (co-IP) and RNA interference to detect, with high confidence, the interaction partners of β -catenin in mammalian cells¹⁵⁶. Moreover, SILAC approach was also used to analyse the dynamicity of some proteins in specific complexes, for instance Wang et al. found that 16 of 67 putative interactors of the human 26 S proteasome were dynamically and not statically bound to the proteasome¹⁶⁰. In addition to SILAC, chemical labelling approaches were also used in the

interaction proteomics field. Following the same principle of SILAC-based methods, a protein is identified as a true interactor if it is strongly enriched in the sample of interest compared to the control.

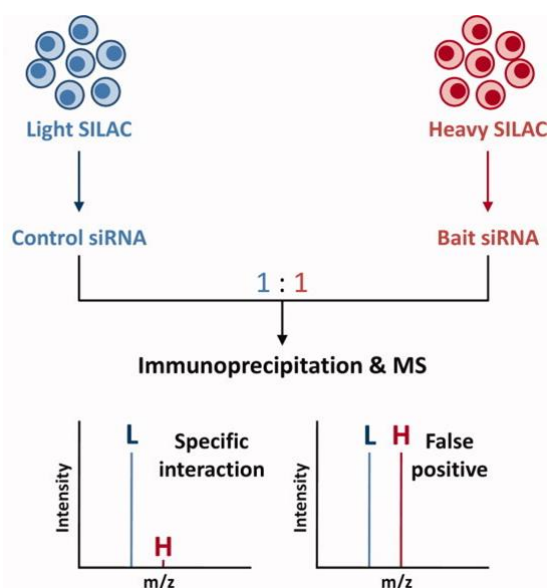


Figure 12: SILAC-MS analysis in interaction proteomics experiments. In this example of SILAC-based interaction proteomics experiment, light-labelled cells are transfected with control short interfering RNA (siRNA) while the heavy-labelled ones with siRNA targeting the bait. Cell extracts obtained from the two SILAC conditions are mixed 1:1 before IP and MS analysis. Specific interactors are much more enriched in the light condition compared to the heavy one while false positives show equal intensity in the two SILAC conditions. Adapted from ¹⁶¹.

2.10.2. Interaction proteomics of chromatin-associated proteins

Chromatin-associated complexes are difficult to characterize, both because many chromatin proteins are low abundant or only transiently associated to chromatin and also because, chromatin-associated proteins are difficult to extract from the cells ¹⁶². In the last decades, different protocols for the enrichment of nuclear proteins have been developed but a “gold standard” method to purify chromatin proteins does not exist and it depends on the downstream application ^{162,163}. One of this approach takes advantage of the insolubility of the chromatin in non-ionic detergents. The use of this type of detergents allows separating the chromatin fraction from the soluble one. Subsequently, chromatin proteins can be

extracted by using buffers containing ionic detergents. The use of this method allowed the MS-based identification and quantification of 282 differentially expressed transcription factors and chromatin modifiers in human-B lymphocytes expressing, or not, cMyc ¹⁶⁴. Several other procedures have been employed for the extraction and purification of chromatin proteins. Most of them are based on buffers having high salt concentration ^{147,165} or containing endonucleases such as deoxyribonuclease I (DNase I) ¹⁶⁶ or Benzonase ¹⁶⁷ to favour the release of proteins bound to DNA. For instance, a method called differential chromatin-associated proteins (D-CAP) is characterized by four consecutive low-salt buffer washes to separate proteins that are not strongly bound to chromatin followed by extraction of chromatin-associated proteins through buffers containing increased concentration of micrococcal nucleases (MNases). With this approach, Alajem *et al.* identified by MS analysis proteins differentially associated with chromatin during the differentiation process including SMARCD1, a protein that activates differentiation programs by regulating H3K27 methylation ¹⁶⁸.

In some chromatin purification solution, a step of protein cross-linking is added. This is achieved by using chemical agents such as formaldehyde or disuccinimidyl glutarate (DSG). This step allows stabilizing the interactions among the proteins and reducing the loss of proteins transiently bound to chromatin. As a further advantage, more stringent washes are permitted reducing, in this manner, the amount of contaminant proteins in MS ¹⁶³. Interaction proteomics methods based on chromatin cross-linking are, for instance, the chromatin enrichment for proteomics (ChEP) strategy developed by Kustatscher *et al.* ¹⁶⁹ as well as the cross-linking Chromatin Proteomics (X-ChroP) approach established by Soldi *et al.* ¹⁵⁴. These methods were used to comprehensively characterize the so-called “interphase chromatin” and to detect novel chromatin players within specific genomic regions, respectively. Problems related with the cross-linking step may concern the pull-down of proteins that are not directly bound to the bait but are, for instance, located on the same DNA portion. Furthermore, another issue is given by the MS-compatibility of the cross-linkers,

indeed some agents such as the glutaraldehyde can interfere with subsequent MS-analysis. Instead, the most common cross-linker formaldehyde has not been found to negatively influence enzymatic digestion, peptide fragmentation and MS-protein identification ¹⁷⁰.

2.11. MS-based proteomics analysis of histone post-translational modifications (PTMs)

2.11.1. Identification of histone modifications

In the last two decades, MS has shown to be an efficient technique for the identification of histone PTMs, given to the high mass accuracy and resolution achieved by the modern MS-instruments. The classical workflow of histone preparation prior to MS-analysis is very similar to what already described in Figure 8 with some expedients. Indeed, histones are very basic proteins and trypsin digestion generates extremely short peptides that are poorly detectable in standard RP chromatography. For this reason, the Arg-C protease which cleaves only at the C-terminus of arginine residues, is usually used for histone digestion in “bottom-up” experiments. When Arg-C does not work (e.g. in the case of in-gel digestions), lysine residues on histones are modified with chemical products, such as deuterated (D₆) or propionic anhydride. These chemical modifications alter trypsin digestion on the same sites, thus producing an “Arg-C like” digestion with peptides having an optimal length for the subsequent RP-LC-MS/MS analysis. Furthermore, chemical derivatization in particular propionylation of lysine residues enhances the hydrophobicity of short histone peptides, increasing in this manner their retention time in standard RP chromatography ^{171,172}. This allows to better distinguish the different isobaric forms of the histone modified peptides. More recently, other chemical agents used for histone derivatization are proved to further improve the detection of short and hydrophilic histone peptides, such as the H3 3-8 ¹⁷³. However, the “bottom-up” approach is not suitable for the study of the physical association of long-distance histone modifications, since it generates peptides of only 7-8 amino-acids. Co-existence of histone modifications can be instead better assessed by other MS-strategies

such as the “top-down” and “middle-down” MS-methods, where intact proteins or large peptide fragments are, respectively, ionized and analysed in MS ^{174,175}. Nevertheless, in top-down experiments the production of a high number of co-isolated and co-fragmented isobaric histone proteoforms generates very complex MS2 spectra that are difficult to interpret ¹⁷⁶. In this context, middle-down MS has gained an increased interest for the study of the histone PTM co-association since it is a good compromise between bottom-up and top-down methods. Classical middle-down MS workflow includes digestion with specific proteases such as Glu-C and Asp-N that cleave at less frequent amino-acid sites, namely C-terminus of glutamic acid and N-terminus of aspartic acid respectively. This allows studying association of histone modifications on peptides longer (> 5kDa) than those generated by bottom-up methods (Figure 13).

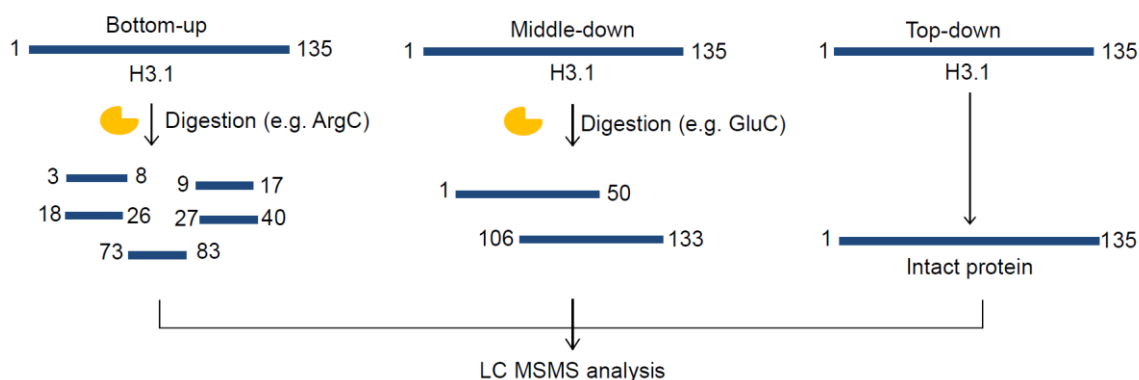


Figure 13: MS-based strategies to identify histone PTMs. The three main strategies adopted for the identification of histone modifications are the bottom-up, middle-down and top-down approaches. The first one is based on the use of proteases such as Arg-C that generate peptides of about 7-8 amino acids. Middle-down strategies instead take advantage of other proteases such as Glu-C and Asp-N that produce longer peptides around 50 amino-acids, while Top-down techniques involve the MS-analysis of intact histones. Adapted from ¹⁷⁷.

2.11.2. MS-based quantitative strategies of histone PTMs

Histone PTMs can be quantitatively analysed by means of different MS-approaches, including label-free methods, SILAC and SILAC-derived strategies as well as chemical

labelling and targeted-based techniques. The two quantitative strategies employed in this thesis for the histone PTM quantitation are based on label-free and spike-in SILAC approaches. The classical label-free method adopted for the quantitation of histone modifications involves the calculation of the relative abundance (RA) value, which corresponds to the ratio between the XIC of a specific histone modified peptide over the sum of the XICs of all detected unmodified and modified forms of the same peptide ¹⁵⁴ (Figure 14). Isobaric isoforms that cannot be separated and distinguished by LC can be quantified at the MS2 level on the basis of the relative ratios of their fragment ions ¹⁷⁸. LF-approaches were used to profile a panel of histone PTMs in various cancer cells ¹⁷⁹ as well as to assess the changes of histone acetylation in the absence of DNA glycosylase ¹⁸⁰.

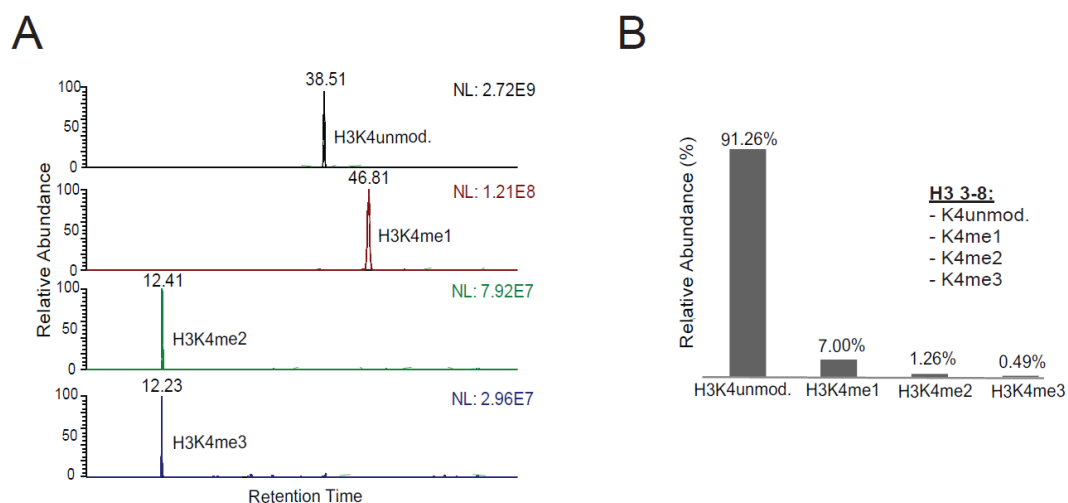


Figure 14: LF-based strategy to quantify histone modifications. A) Chromatographic peaks of unmodified and modified forms of the H3 3-8 peptide (H3K4 mono-, di- and tri-methylated), extracted from Thermo Xcalibur Qual Browser. The figure displays the retention time and the intensity values (NL) of each unmodified and modified histone peptide. B) Calculation of the percentage relative abundance (%RA) values of each form of the H3 3-8 peptide. %RA corresponds to the ratio of the MS-extracted ion chromatogram (XIC) of the unmodified or a specific modified form of the H3K4 over the sum of the XICs of all the isoforms of the H3 3-8 peptide. Adapted from ¹⁷⁷.

A common labelling approach for the relative quantification of histone modifications among different samples is based on SILAC-derived methods, such as the spike-in SILAC approaches. The rationale of these strategies consists in the use of an internal standard that

originates either from a single cell line or a mix of cell lines that represents a histone-focused version of the “super-SILAC” mix ¹⁸¹ (Figure 15). This method allows overcoming two main restraints of the standard SILAC, namely the incapacity to quantify more than three biological conditions and to use samples that cannot be metabolically labelled, such as the clinical samples. For instance, a super-SILAC approach was used to profile epigenetic variations among different breast cancer subtypes ¹⁸². Another labelling strategy used for the quantitation of histone PTMs is based on chemical derivatization of lysines using various isotope-labelled compounds, such as the propionic anhydride ¹⁸³. After the labelling with the different isotope-encoded chemicals, histones deriving from different biological conditions are mixed in equal amount and analysed by MS. Finally, histone PTMs can be quantified by targeted strategies, such as selected and multiple reaction monitoring (SRM and MRM, respectively). In this approach, single or multiple fragments ions of selected peptides are quantitatively measured. Through this method both relative and absolute quantitation can be performed, by using synthetic and isotopically-labelled peptides as internal spike-in standard. Relative quantitation is achieved by comparing the intensity of each modified peptide with that of the internal standard, which is added at constant concentration to each sample ¹⁸⁴. Absolute quantitation is instead obtained by the creation of a calibration curve of the ion intensity of the modified peptide with respect to the standard, injected at different concentrations ¹⁸⁵.

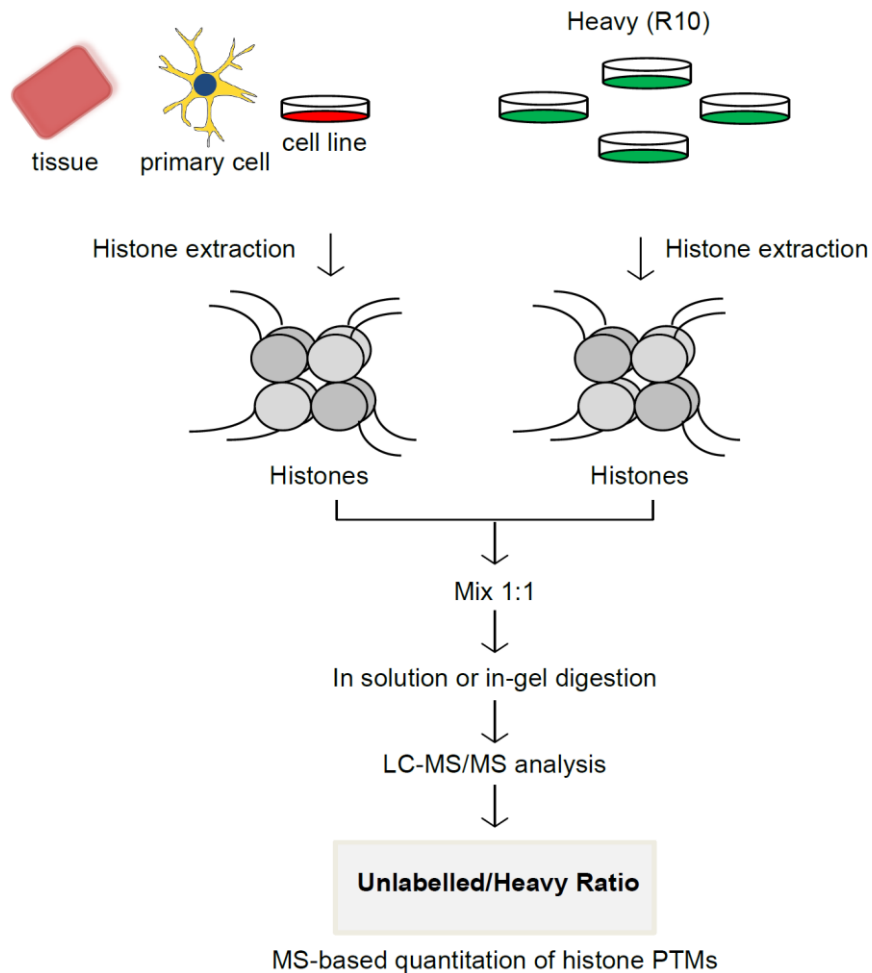


Figure 15: Super-SILAC approach for histone PTMs quantitation. Histones are extracted from either unlabeled or labelled samples with heavy-arginine (R10). Super-SILAC mix is usually obtained by combination of heavy histones extracted from different cell lines. Unlabelled and labelled SILAC histones are mixed in equal amount prior to histone digestion and LC-MS/MS analysis. To quantify histone PTMs, each histone modified peptide from unlabeled samples is normalized to the level of its internal heavy counterpart by calculating the unlabelled/heavy ratios. In this way, several unlabeled samples can be compared normalizing the levels of each histone modified peptide to that of the spike-in SILAC reference. Adapted from ¹⁷⁷.

3. AIM OF THE WORK

LSD1 has been widely studied in the last decade as a promising epigenetic target for different types of tumours, in particular AML. Despite the significant effort devoted to the development of selective LSD1 inhibitors, the thorough understanding of their mechanism of action is not yet fully clarified, also because it seems to depend on the different cellular context in which the compounds act. Until recently, the most widely accepted idea about the mode of action of LSD1 inhibitors was that they elicited their molecular and phenotypic effects on cancer cells by inhibiting the de-methylation activity of LSD1 on its own targets³⁷⁻³⁹. Recent studies have instead provided evidence that, in some circumstances, LSD1 inhibitors can reduce the proliferation and induce differentiation of cancer cells by altering the binding of LSD1 with some of its interactors rather than by inhibiting its catalytic activity^{58,89,104,105}. Preliminary data collected in our APL model system pointed towards this mode of action of the small molecules under investigation, as demonstrated by the fact that the LSD1 enzymatic activity was dispensable for promoting the differentiation of NB4 cells induced by the drugs.

The aim of my PhD project was therefore to dissect the mechanism of action of the MC2580 and DDP-38003 inhibitors in APL cell models by using *ad hoc* quantitative MS-based proteomics strategies. First, by combining biochemical approaches with MS-analysis, I examined the impact of these inhibitors on the LSD1 interactome. Follow-up studies on the two candidates emerged from this screening helped in reconstructing possible novel mechanisms of action of these drugs in APL. Second, I investigated the effects of these inhibitors on global histone modification levels of two APL cellular models that present different sensitivity to the drugs, with the aim of unravelling epigenetic mechanisms underpinning the different responsiveness to LSD1 inhibition.

4. MATERIALS AND METHODS

4.1. Cell culture

NB4 cell line was derived from the bone marrow of a patient with APL in second relapse in 1989^{186,187}. These cells were grown in RPMI plus 10% of fetal bovine serum (FBS), 2mM glutamine and 1% Penicillin/Streptomycin (P/S). Instead, UF-1 cell line was derived from the peripheral blood sample of a patient with APL in second relapse in 1994 and clinically resistant to ATRA¹¹⁸. These cells were grown in RPMI medium, complemented with 20% FBS, 2mM glutamine and 1% P/S. Cultures were maintained in a humidified tissue culture incubator at 37°C in 5% CO₂.

4.1.1. SILAC labelling of NB4 cells for the LSD1 Interactome analysis

For SILAC metabolic labelling, NB4 cells were grown in “light” and “heavy” SILAC RPMI 1640 (Thermo Fisher Scientific 89984) supplemented with either L-arginine and L-lysine, or with their heavy isotope-counterparts L-arginine-¹³C₆, ¹⁵N₄ hydrochloride (Arg10, Sigma 608033) and L-lysine-¹³C₆, ¹⁵N₂ hydrochloride (Lys 8, Sigma 608041), respectively¹⁸⁸. All media were supplemented with 10% dialyzed FBS (26400-044 Gibco, Life Technology), 2mM glutamine and 1% P/S. Cells were grown in SILAC media for up to 9-10 doublings, in order to ensure the complete incorporation of “heavy” amino acids. Monitoring of the cells’ growth rate, viability and morphology served to exclude effects of SILAC conditions on their physiology.

4.2. Compounds

The LSD1 inhibitors MC2580 and DDP-38003 were produced by the Experimental Therapeutic Unit at the IFOM-IEO Campus^{116,117} while the EZH1/2 inhibitor UNC1999¹⁸⁹ was purchased by the Sigma Aldrich (SML0778).

4.3. LSD1 co-IP for mass-spectrometry analysis of protein-protein interactions

4.3.1. Subcellular fractionation

For the preparation of cellular sub-fractions, NB4 cells were resuspended in Hypotonic Buffer (10 mM Tris HCl pH 7.6, 1.5 mM MgCl₂, 10 mM KCl, 1X EDTA-free Roche protease inhibitors, 0.5 mM phenylmethylsulfonyl fluoride). Upon 10 minutes of incubation in ice, 0.3% of Triton X-100 was added to the cell re-suspension and samples were incubated on a rotating wheel for 10 minutes at 4°C and then centrifuged for 1 minute at 11000 rpm^{165,190}. The pellet (corresponding to the fraction enriched of nuclei) was resuspended in Nuclear Extraction Buffer (50 mM Tris HCl pH 7.6, 150 mM NaCl, 0.5% NP-40, 20% Glycerol, 2 mM MgCl₂, 1X EDTA-free Roche protease inhibitors, 0.5 mM phenylmethylsulfonyl fluoride, 250U Merck Benzonase) for 1 hour at 4°C and, then, centrifuged at 13000 rpm for 30 minutes. The supernatant was collected and used for the subsequent immuno-precipitation experiments.

4.3.2. LSD1 protein co-immunoprecipitation (co-IP)

The content of nuclear proteins of light and heavy samples was quantified by Bradford assay and diluted to a concentration of 2 mg/ml in IP buffer (10 mM Tris HCl pH 7.6, 150 mM NaCl, 0.2% NP-40), supplemented with 1X protease inhibitors and 0.5 mM phenylmethylsulfonyl fluoride (PMSF). Preclearing of the lysates was achieved by incubation with protein G magnetic beads (Invitrogen 10004D) for 1 hour at 4°C. The lysates

were, then, re-quantified and diluted to a concentration of 1.3 mg/ml by using the IP buffer supplemented with protease inhibitors and 0.5 mM PMSF. A fraction (1/20) of the Input was collected before adding 10 µg of anti-LSD1 antibody to each sample. The samples were, subsequently, incubated on a rotating wheel at 4°C, overnight. For the acquisition of the basal LSD1 interactome, 120-fold molar excess (120X) of LSD1 blocking peptide was incubated together with the antibody as negative control. The peptide was added to the light channel in the forward experiment and to the heavy channel in the reverse one. On the following day, 100 µl of Dynabeads-protein G, pre-equilibrated in phosphate-buffered saline (PBS) supplemented with 0.5% bovine serum albumin (BSA), were added to each sample and incubated for 3 hours on a rotating wheel at 4°C. Beads were then washed 3 times with IP buffer and once with the Washing Buffer IP (10 mM Tris HCl pH 7.6, 250 mM NaCl, 0.2% NP-40), both supplemented with protease inhibitors. In the last washing step, light and heavy samples of each SILAC replicate were mixed and the co-immunoprecipitated proteins were eluted by incubation at 95°C for 5 minutes with the LDS Sample Buffer (NuPAGE-Invitrogen, NP0007), supplemented with 100 mM dithiothreitol (DTT). Samples were loaded on a SDS-PAGE gradient gel (Invitrogen, NP0335BOX) for subsequent protein separation and stained with colloidal coomassie staining (Invitrogen, LC6025). The antibody used for the preparative LSD1 co-immunoprecipitation was the anti-LSD1 antibody (AB17721, Abcam). The blocking peptide used as mock control was Human KDM1 / LSD1 peptide (AB17763).

4.4. In-gel digestion of immuno-precipitated proteins

Gel-separated proteins were digested with Trypsin prior to MS analysis, as described in ¹⁹¹. Each lane of the gel was cut into 8 fractions. After digestion and extraction from the gel pieces, the tryptic peptides were desalted and concentrated by reversed-phase chromatography onto micro-column C18 Stage Tips ¹⁹².

4.5. Liquid Chromatography and Tandem Mass Spectrometry (LC-MS/MS)

Peptides were eluted from the Stage Tips with buffer B (80% acetonitrile, 0.5% acetic acid), lyophilized, re-suspended in 0.1% Formic Acid (FA) and subjected to LC-MS/MS analysis. Peptide samples were analyzed by online nano-flow liquid chromatography tandem mass spectrometry using an EASY-nLC™ 1000 (Thermo Fisher Scientific, Odense, Denmark) coupled to a hybrid quadrupole/Orbitrap mass spectrometer (Q Exactive HF, Thermo Fisher Scientific) through a nanoelectrospray ion source. The nano-LC system worked with one column set up with an EASY-Spray™ LC Columns (50-cm length, 75- μ m inner diameter) packed with C18 reversed-phase resin (2 μ m). Solvent A was 0.1% FA in ddH₂O and solvent B was 80% acetonitrile (ACN) with 0.1% FA. Peptides were injected at a flow rate of 500 nL/min and separated with a gradient of 5-30% solvent B over 80 minutes, followed by a gradient of 30-60% for 10 minutes and 60-95% over 3 minutes at a flow rate of 250 nL/min. The Q Exactive HF Orbitrap mass spectrometer used for the analysis of the LSD1 co-IPs is composed of an ESI source, where ions are injected into the mass spectrometer and enter the radio frequency (RF)-lens. The lens focuses the ion beam before they pass through the Advanced Active Beam Guide (AABG). In the AABG, the path is curved of 90° allowing the removal of many uncharged/contaminants before entering the Quadrupole. Here, the most abundant ions are selected and transferred into the C-trap where they are then guided to the Orbitrap analyzer. This analyzer is characterized by a coaxial central spindle electrode surrounded by a barrel-like electrode¹⁹³. It catches ions in an electric field and makes them oscillate around a central electrode in ring shapes. Each ion has its own oscillation frequency according to its m/z value. This motion is collected by the detector and Fourier-transformed to generate high- resolution mass spectra¹⁹⁴. After the MS1 acquisition, ions are then led to the HCD cell to be fragmented. In the meantime, the following set of ions is ready to enter in the C-trap to start a new MS-MS/MS cycle. The fragment ions generated into the HCD

cell come back into the Orbitrap analyzer where also the MS/MS spectra are acquired at high resolution (Figure 16).

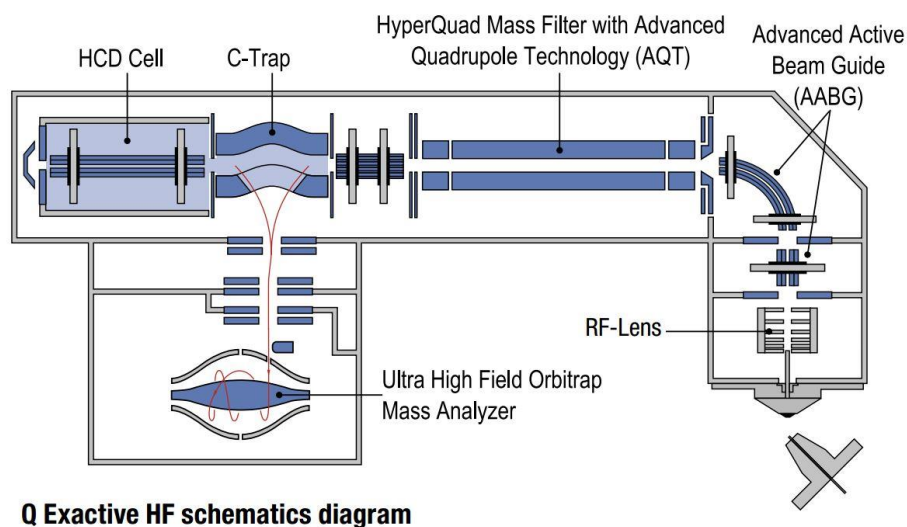


Figure 16: Schematic representation of a Q Exactive HF mass spectrometer. The main components of a Q Exactive HF mass spectrometer are: the RF-lens that focuses ions into a tight beam after they emerge from the ionization source, the AABG that decreases background noise removing many contaminants and ensuring more stability to the entire system, the quadrupole mass filter with AQT that selects ion precursors, the c-trap that regulates the flux of ion population that must enter in the Orbitrap analyzer, the Orbitrap mass analyzer that examines the m/z of ions as a function of their rotational frequencies and the HCD cell that is responsible of ions fragmentation. Taken from www.thermofisher.com.

The Q-Exactive was set in the data-dependent mode (DDA) to automatically switch between full scan MS and MS/MS acquisition. MS- experiments consisted of an Orbitrap full scan followed by the acquisition of the 15 most intense peptide ions (Top15) detected in the survey MS scan. Mass spectrometer conditions were as follows: full MS (AGC $3e^6$; resolution 60,000; m/z range 375-1650; maximum ion time 20 ms); MS/MS (AGC 15,000; maximum ion time 80 ms; isolation width 2 m/z with a dynamic exclusion time of 20 seconds). Singly charged ions and ions for which no charge state could be determined were excluded from the selection. Normalized collision energy was set to 28%; spray voltage was 1.9 kV; no sheath and auxiliary gas flow; heated capillary temperature was 275 °C; S-lens RF level of 50%.

4.6. Data analysis for SILAC- based protein quantitation

Acquired MS raw data were analyzed with the integrated MaxQuant (MQ) software v.1.6.0.1, using the Andromeda search engine ^{122,195}. The February 2018 version (UniProt Release 2018_02) of the Uniprot sequence was used for peptide identification. Enzyme specificity was set to Trypsin/P, meaning that trypsin cleavage occurs also in the presence of proline, at the C-terminal of lysine or arginine residues. A maximum of 3 missed cleavages were permitted, and the minimum peptide length was fixed at 7 amino acids. Carbamidomethylation of cysteine was set as a fixed modification. ‘Requantify’ and ‘Match between runs’ functions were enabled. ‘Multiplicity’ was set as 2 and in the heavy labels Arg10 and Lys8 were added.

4.7. LSD1 interactome data analysis

MQ outputs were manually filtered for the LSD1 interactome analysis according to the following parameters: proteins were considered in the analysis if identified with more than two peptides of which at least one unique and Ratio Count greater than 1 ($RC > 1$). To identify the putative LSD1 interactors from the SILAC-based co-IP, the “mixtools” R package was employed. This tool allowed defining for each replicate two populations of enriched and background proteins based on their respective SILAC ratio values. Mean (μ) and standard deviation (σ) of only the enriched protein population were calculated. Finally, proteins with a SILAC ratio higher than $\mu - \sigma$ of the enriched population were considered as putative LSD1 interactors. To identify the interactors whose binding was modulated upon treatment of the cells with the LSD1 inhibitor, median (M) and standard deviation (σ) of the total protein SILAC distributions were calculated: proteins displaying a SILAC ratio higher than the $M + 2\sigma$ in both replicates were defined as recruited while protein with a SILAC ratio lower than the $M - 2\sigma$ as evicted. Once analyzed the distributions of each replicate, the modulated

proteins were filtered on the basis of the results of the basal LSD1 interactome, which sets the list of specific binders. Analysis of the protein distribution of each SILAC replicate as well as Scatter Plots were obtained with Perseus Software. Protein-protein interaction analysis was achieved through STRING functional database and visualized by Cytoscape. Protein complexes were analysed by CORUM database contained within the web-based toolset g:Profiler¹⁹⁶.

4.8. Immuno-precipitation analysis for Western Blot validation

For Western Blot (WB) analysis of co-IP experiments, 600 µg of nuclear extract, prepared as explained in the paragraph 4.3.1., were diluted to 300 µl of IP buffer supplemented with 1X protease inhibitors (Roche) and 0.5 mM PMSF. Preclearing was performed by incubating the lysates with protein G magnetic beads for 1 hour on a rotating wheel at 4°C. Protein extracts were then re-quantified and diluted to a concentration of 1.3 µg/µl with the IP buffer supplemented with protease inhibitors and 0.5 mM PMSF. A minor fraction (1/20) of Input was collected before adding 3 µg of antibody to each sample followed by incubation on a rotating wheel at 4°C overnight. The following day 30 µl of Dynabeads-protein G, pre-equilibrated in PBS supplemented with 0.5% BSA, were added to the extracts and incubated for 3 hours on a rotating wheel at 4°C. Beads were, then, washed 3 times with IP buffer and once with Washing Buffer IP (10 mM Tris HCl pH 7.6, 250 mM NaCl, 0.2% NP-40), both supplemented with protease inhibitors. Co-immunoprecipitated proteins were eluted by incubation at 95°C for 5 minutes with the LSD Sample Buffer (NuPAGE-Invitrogen) supplemented with 100 mM DTT. Eluted proteins were separated on a SDS-PAGE for subsequent WB analysis. The antibodies used for small scale validation of basal and dynamic co-IP were: LSD1 (AB17721, Abcam), GFI-1 (sc-376949, Santa Cruz Technology) and GSE1 (24947-1-AP, Proteintech).

4.9. Retroviral and lentiviral cell transduction

4.9.1. Retroviral and lentiviral constructs

To perform rescue experiments in NB4 LSD1-KO cells, LSD1 N-terminal truncated (172-833) wild-type (WT-LSD1) and catalytic mutant (K661A-LSD1) constructs were kindly provided by Prof. Elena Battaglioli (University of Milan). Constructs were PCR-amplified from original vectors and cloned into pCR-TOPO 2.1 (Invitrogen) through the following primers:

LSD1 forward: ATGTCGGGTGTGGAGGGCGCAGCTTTC

LSD1 reverse: TCACATGCTTGGGGACTGCTGTGC.

PCR products were, then, sub-cloned into the EcoRI site of the retroviral PINCO vector for ectopic expression. To create the D553,555,556A triple mutant construct ¹⁵, the following primers were used in three sequential site-directed mutagenesis reactions using the pCR-TOPO-LSD1WT as template vector:

LSD1_D553A_forward: CTTAAGCACTGGGCTCAGGATGATGACTTTGAGTTC

LSD1_D553A_reverse: GAACTCAAAGTCATCATCCTGAGCCCAGTGCTTAAG

LSD1_D553,555A_forward: CTTAAGCACTGGGCTCAGGCTGATGACTTTGAGTTC

LSD1_D553,555A_reverse: GAACTCAAAGTCATCAGCCTGAGCCCAGTGCTTAAG

LSD1_D553,555,556A_forward:

CTTAAGCACTGGGCTCAGGCTGCTGACTTTGAGTTC

LSD1_D553,555,556A_reverse:

GAACTCAAAGTCAGCAGCCTGAGCCCAGTGCTTAAG

To knock-down (KD) GSE1, short hairpin RNA (shRNA) constructs for GSE1 depletion were cloned into pLKO.1 expression vector. The sequences of the shRNAs used in the experiments were:

shA1: GAACTCACCTTGACGTCAATG

shB2: CTGAGCATGCTTCACTATATC

4.9.2. NB4 cell transduction

Lentiviral constructs were transiently transfected in HEK-293T cells by using the calcium phosphate transfection method ¹⁹⁷ together with the packaging plasmid pCMV-DR8.74 and the envelope plasmid pMD2G-VSVG. The retroviral vectors were, instead, transiently transfected in phoenix-AMPHO cells, together with the packaging plasmid pKAT ¹⁹⁸. After 12 hours of transfection, the culturing medium was changed with fresh one. The day after, the medium was filtered and ultra-centrifuged for 2 hours at 24000 rpm at 4°C. The pellet enriched of viral particles was resuspended in 400µl of medium (RPMI plus 10% of FBS, 2mM glutamine and 1% P/S). 2 million of NB4 cells at a concentration of 1 million/ml were plated in 24-well plates (500 µl for each well). 4 µg of polybrene and 100µl of concentrated virus were added to each well. The plates were centrifuged for 1 hour at 1000 g at room temperature (Spin Infection) and, after 4 hours, 500 µl of medium were added to each well. Transfected HEK-293T or phoenix-AMPHO cells were replaced with fresh medium that was, again, filtered and ultra-centrifuged for the second round of infection executed in the same manner the day after. Lentiviral and retroviral transductions were performed in a Biosafety level 2 laboratory (BSL-2).

4.10. Western blot analysis

Cells were harvested, washed twice with PBS and lysed in RIPA buffer (10mM Tris pH 8, 150mM NaCl, 0.1% SDS, 1% Triton X-100, 1mM EDTA, 0.1% Sodium Deoxycholate), supplemented with 1X Protease Inhibitors (Roche) and 0.5 mM PMSF. The cell extract was, then, quantified by Bradford assay. Thirty micrograms of proteins were mixed with Laemmli

Buffer supplemented with 100 mM DTT, denaturated for 5 minutes at 95°C and then loaded on SDS-PAGE gel. The transfer to PVDF membranes was performed at 100V for 1 hour and 30 minutes at 4°C, or overnight at 30 V at 4°C in Transfer Buffer containing 10% methanol. Membranes were blocked in 10% BSA/tris buffered saline (TBS)-Tween for 1 hour at room temperature or overnight at 4°C. Primary antibodies were diluted in TBS-Tween + 5% BSA and incubated with the membrane overnight at 4°C or for 3 hours at room temperature. After three washes with TBS-Tween (5 minutes each), membranes were incubated with the proper secondary antibody in TBS-Tween + 5 % BSA for 1 hour at room temperature. After three more washes, signals were detected using the ECL (Enhanced Chemio Luminescence) method. For fluorescent WB, anti-rabbit Alexa680 and anti-rabbit Alexa800 were used as secondary antibodies and the signals were acquired using the LI-COR Odyssey v3.0. Quantitation of the blot bands was obtained using ImageJ software. The following antibodies were used, according to the company's instruction, at the specific dilutions: LSD1 (1:1000), GFI1 (1:100), GSE1 (1:600), ZMYM3 (AB106626, Abcam, 1:1000), HDAC1 (AB7028, Abcam, 1:2000), HMG20B (14582-1-AP, Proteintech, 1:500), Vinculin (Millipore 06-866, 1:10000), Lamin-B1 (AB16048, 1:10000), Cleaved caspase-3 (Cell Signalling Technology #9661, 1:1000), Caspase-3 (Cell Signalling Technology #9662, 1:1000), GAPDH (AB9484, 1:1000), H3 (AB1791, 1:5000), p21 (Cell Signalling Technology #2947, 1:1000), H3K4me2 (AB7766, 1:1000), H3K27me3 (Cell Signalling Technology, #9733, 1:1000).

4.11. RNA extraction and RT-qPCR

Total RNA was extracted by using the Quick-RNA™MiniPrep kit (Zymo Research) while reverse transcription was performed with OneScript® Plus cDNA Synthesis Kit (Abm), according to the manufacturer's protocols. Real time PCR was performed in triplicates in 20 µL of final reaction volume containing SYBR green buffer (Applied Biosystems), 100 ng of cDNA retro-transcribed from the RNA, and 0.5 µM of each primer mix. All the quantitative

PCR amplifications were performed in the CFX96 Real-Time System (BioRad) with this protocol: 1) 95 °C for 2 minutes, 2) 40 cycles at 95 °C for 10 seconds and 60 °C for 30 seconds. The primer sequences used for the quantitative PCR analysis are the following:

CD11b: Forward (Fwd) AACCCCTGGTTCACCTCCT

Reverse (Rev) CATGACATAAGGTCAAGGCTGT

GAPDH: Fwd TCAAGAAGGTGGTGAAGCAGG

Rev ACCAGGAAATGAGCTTGACAAA

GSE1: Fwd CAGGAGAAAGGGTACTACTA

Rev GGGACTGTTGTTTCATCTCAT

TBP: Fwd CGGCTGTTTAACTTCGC

Rev CACACGCCAAGAAACAGTGA

CD86: Fwd GTATTTTGGCAGGACCAGGA

Rev GCCGCTTCTTCTTCTTCCAT

IRF8: Fwd GAGGTGGTCCAGGTCTTCG

Rev CGGCCCTGGCTGTTATAG

p21: Fwd GTCACTGTCTTGTACCCTTG TG

Rev CGGCGTTTGGAGTGGTAGAAA

4.12. Cell growth analysis

For proliferation assays cells were counted in Trypan Blue (Sigma Aldrich, St. Louis, Missouri USA) by using the TC20 automated cell counter from Biorad (<http://www.bio->

rad.com/it-it/product/tc20-automated-cell-counter). Percentage of living cells was also obtained with this counting method.

4.13. May Grunwald-Giemsa staining

To prepare cells for May-Grunwald Giemsa staining, 200.000 NB4 cells were cytopinned at 250 rpm for 5 minutes. Then, slides were let to dry overnight before staining. At this point, cells were stained for 8 minutes with May-Grunwald solution, washed 3 times in deionized water and then incubated for 40 minutes with Giemsa. After three more washes with water, samples were air dried and evaluated with HistoFluo microscopy.

4.14. Chromatin Immuno-Precipitation-sequencing (ChIP-seq) in NB4 cells

4.14.1. ChIP-seq analysis of LSD1 and GFI1

Forty millions of cells were cross-linked in culture medium containing 1% of formaldehyde in PBS and the reaction was stopped after 10 minutes at room temperature by the addition of 0.125 M of glycine for 5 minutes at 4°C. Cells were, then, washed twice with PBS and lysed in SDS buffer (50 mM Tris-HCl pH 8.0, 0.5% SDS, 100 mM NaCl, 5 mM EDTA, 0.02% NaN₃ and 1X Roche Protease Inhibitors). At this point, fixed cell extracts were stored at -80° or directly processed. Then, extracts were diluted with Triton Dilution buffer (100 mM NaCl, 100 mM Tris-HCl pH 8.5, 5 mM EDTA, 5% Triton X-100, 0.02% NaN₃) supplemented with protease inhibitors, in order to obtain the IP buffer conditions (100 mM NaCl, 33 mM Tris-HCl pH 8.0, 5 mM EDTA, 0.02% NaN₃, 0.33% SDS, 1.7% Triton X-100, 33 mM Tris-HCl pH 8.5). Chromatin was sonicated using a Branson Sonifier 250 to obtain an average size length of about 300 bp and then pre-cleared by incubation with protein A-Sepharose beads (Amersham) for 2 hours on a rotating wheel. Precleared chromatin was centrifuged to discard the beads and the supernatant was immuno-precipitated overnight in

the presence of 10 µg of antibody and 50 µl of protein G magnetic beads. Before IP, 2.5% of Input was collected. The day after, the beads bound to the immuno-precipitated chromatin were washed twice with Mixed Micelle Buffer (150 mM NaCl, 20 mM Tris-HCl pH 8.0, 5 mM EDTA, 5.2% sucrose, 0.02% NaN₃, 1% Triton X-100, 0.2% SDS), twice with Buffer 500 (0.1% Deoxycholic acid, 1 mM EDTA, 50 mM HEPES pH 7.5, 500 mM NaCl, 1% Triton X-100, 0.02% NaN₃), twice with LiCl/detergent (0.5% deoxycholic acid, 1 mM EDTA, 250 mM LiCl, 0.5% NP-40, 10 mM Tris-Cl, pH 8.0, 0.02% NaN₃) and once with TE 1X before the de-crosslinking step. De-crosslinking was performed for all IP samples and the corresponding Inputs by an overnight incubation in 0.1% SDS and 0.1% NaHCO₃. The day after, the enriched DNA was treated with proteinase-K at 56°C for 40 min for protein removal and subsequently purified with a DNA purification kit (Qiagen). The antibodies used for ChIP were: LSD1 (AB17721, Abcam), GFI1 (AB21061, Abcam).

4.14.2. ChIP-seq analysis

Reads were obtained from Illumina Genome Analyzer II and quality-filtered according to the Illumina pipeline. Analysis of the samples was performed through the Fish the ChIPs pipeline¹⁹⁹ and included alignment to the hg18 reference genome using Bowtie v1.0.1²⁰⁰ and MACS version 1.4.1²⁰¹ as peak caller to identify regions of ChIP-seq enrichment over background. Only reads with a unique match to the genome and with two or fewer mismatches (-m 1 -v 2) were maintained. The MACS algorithm was used applying a p-value threshold of 10⁻⁵ for the GFI1-ChIP while for the LSD1 ChIP the threshold was set by RT-qPCR validation. Each sample was compared to the Input DNA coming from NB4 cells (DMSO). When calling the differentially enriched regions between treated and untreated samples, the resulting regions were further filtered, keeping only those found enriched against the Input as well. All the lists containing the differentially enriched regions were annotated over RefSeq genes according to GIN²⁰² while intergenic regions were considered

as those showing a distance higher than 22kb from the nearest gene. The bigwig files for UCSC browser visualization of genome profiles were normalized with the deepToos suite²⁰³ using RPCG. LSD1 and GFI1 ChIP were compared with ChIP of different histone modifications performed in the same model to distinguish specific genomic regions. Active enhancers were defined as regions with a coincident peak of H3K4me1 and H3K27ac, which fell into distal genomic regions (defined as -20/-5kb and +5/+20kb from TSS). Instead, active promoters were defined as regions with a coincident peak of H3K4me3 and H3K27ac falling into a 5kb neighbourhood of Refseq-annotated TSS. Pathway analysis was performed with QIAGEN's Ingenuity Pathway Analysis (IPA, QIAGEN Redwood City, www.qiagen.com/ingenuity).

4.15. Cell cycle analysis by flow cytometry

Cells were harvested, washed once with BSA 1% in PBS and resuspended in 250 µl of PBS. Fixing of the cells was achieved by adding 750 µl of pure ethanol dropwise, while vortexing. After 30 minutes, cells were washed again with BSA 1% in PBS and resuspended in Propidium Iodide (PI, 20 µg/ml) + RNase A (250 µg/ml). Stained cells were let on fluorescence activated cell sorting (FACS) tubes overnight at 4°C before analysing the samples with FACS Celesta 2. Analysis of the cell cycle was performed by FlowJo software.

4.16. RNA-sequencing

mRNA-sequencing (mRNA-seq) libraries were prepared according to the True-seq Low sample protocol (Illumina, San Diego, California USA), starting from 500 ng of total RNA per sample and sequencing was performed through the NovaSeq 6000 (Illumina) instrument. Raw reads were mapped to the human reference genome hg38 using STAR aligner²⁰⁴, while differentially expressed genes (DEG) were determined through DEseq2 package²⁰⁵. Genes

with an adjusted p-value less than 0.01 and \log_2 Fold Change (FC) greater than 1 and smaller than -1 were considered as upregulated and downregulated, respectively. Gene ontology analysis of the enriched biological processes (BP) was performed through EnrichR software²⁰⁶ while pathway analysis was executed by Reactome database within EnrichR. Significant BP and pathway terms displayed an adjusted p-value lower than 0.05. Treemap was obtained by using Revigo Database.

4.17. MS-based profiling of histone PTMs

4.17.1 Histones extraction and digestion

Cells were homogenized in lysis buffer (10% sucrose, 0.5 mM EGTA, 60 mM KCl, 15 mM NaCl, 15 mM HEPES, 0.5 mM PMSF, 5 μ g/ml Aprotinin, 5 μ g/ml Leupeptin, 1 mM DTT, 5 mM NaButyrate, 5 mM NaF, 30 μ g/ml Spermine, 30 μ g/ml Spermidine and 0.5% Triton X-100) and nuclei were separated from cytoplasm by centrifugation on sucrose cushions for 30 minutes at 3750 rpm. Then, histones were extracted through 0.4 N hydrochloric acid for 5 hours at 4°C^{207,208}. Extracted histones were lyophilized, resuspended in milliQ water and quantified by Bradford assay.

Five micrograms of histones were in-solution digested prior to LC-MS/MS analysis through the hybrid chemical labelling “Pro-PIC” method. This method is based on an initial conversion of free lysines to their propionylated forms under mild aqueous conditions followed by trypsin digestion and labelling of new peptide N-termini with phenyl isocyanate (PIC)¹⁷³. For the untreated and treated UF-1 cells with MC2580, 2.5 μ g of histones were mixed with 2.5 μ g of arginine heavy (R10)-labelled histones extracted from NB4 cells prior to the propionylation step. The digested peptides were desalted and concentrated by reversed-phase chromatography onto micro-column C18 Stage Tips¹⁹².

4.17.2 LC-MS/MS and data analysis of histone PTMs

Peptides were eluted from the stage tips with Elution Buffer (60% ACN, 0.1% trifluoroacetic acid), lyophilized, re-suspended in 1% trifluoroacetic acid (TFA) and subjected to LC-MS/MS analysis. The samples were analysed onto the Q Exactive HF Orbitrap mass spectrometer, upon separation with a gradient of 10-40% solvent B over 100 minutes, followed by a gradient of 40-60% for 10 minutes and 60-95% over 3 minutes at a flow rate of 250 nL/min using a Reversed Phase-UHPLC system.

MaxQuant software v.1.6.0.1 was used for the analysis of MS data, for both protein and peptides identification including as variable modifications propionylation, mono-, di- and tri-methyl lysine and lysine acetylation¹²². The Uniprot HUMAN histones 1502 database was used for histone peptide identification. Enzyme specificity was set to Arg-C, since the propionylation of lysine residues allows the trypsin cutting only at the C-terminus of arginine residues. A maximum of 3 missed cleavages were permitted, and the minimum peptide length was fixed at 6 amino acids. PIC at the N-terminus of each peptide was set as a fixed modification.

Quantitation of histone PTMs was performed by a label-free approach, in particular by calculating the extracted ion chromatogram (XIC) of each modified histone peptide and, then, extrapolating the percentage relative abundance (%RA) as explained in the paragraph 2.11.2 of the Introduction. To estimate the possible changes upon LSD1 inhibition/depletion, the ratio of the %RA of each modified peptide in the treated and KO sample over their corresponding values in the control were calculated. For the quantitation of histone PTM changes in UF-1 cells treated with MC2580, a spike-in SILAC strategy was instead employed, whereby the %RA of each histone modified peptide was normalized over the %RA of its heavy counterpart (L/H ratio) derived from NB4 heavy (R10)-labelled histones spiked-in at the same concentration across the samples¹⁷⁷. Then, to assess the changes upon

LSD1 inhibition, the L/H ratios of each modification in the treated samples over the control ones were calculated. Hierarchical clustering analysis was obtained with Perseus Software.

4.18. Statistical analysis

Most of the data represented in this thesis, such as those obtained from WB, RT-qPCR, FACS and MS-based histone PTM experiments, are shown as mean \pm standard error of the mean (SEM) and statistical analyses were carried out using two-tailed paired Student's t tests, unless otherwise specified. The number of biological (n) replicates, the specific type of statistical analyses performed, and statistical significance are reported in the corresponding Figures and Figure legends.

5. RESULTS

5.1. MS-based analysis of the basal LSD1-interactome in NB4 cells

In order to define the network of proteins specifically interacting with LSD1 in NB4 cells, we set-up a strategy based on the combination of SILAC-based protein profiling with LSD1 co-immuno-precipitation (co-IP) followed by mass-spectrometry (MS) analysis. To increase the confidence in the identification of “true” interactors from background, as mock control we carried out the LSD1 co-IP in the presence of an excess fold of the soluble LSD1-blocking peptide, which competes with the bait and all its co-associated factors for the antibody binding.

Briefly, NB4 cells were grown in medium containing either light or heavy isotopic variants of lysine (K) and arginine (R), up to full incorporation. Labelled cells were harvested and subjected to fractionation into nuclear and cytosol fractions. The nuclear fraction was, then, used as Input for the co-IP experiments, which were performed in both forward and reverse SILAC setups. In the forward experiment, the heavy (H)-labelled nuclear extract was incubated with the anti-LSD1 antibody and the light (L) one with the same antibody in the presence of the excess fold of soluble LSD1 blocking peptide. In the reverse experiment, the two experimental conditions were swapped. The two H and L co-IPs were performed in parallel and the proteins pulled-down were mixed 1:1 prior to LC-MS/MS analysis. Due to the competition given by the blocking peptide, specific LSD1 binders had a SILAC ratio different from 1 ($H/L > 1$ in the forward experiment) while background proteins were not affected by the competition generating a SILAC ratio similar to 1 ($H/L = 1$) (Figure 17).

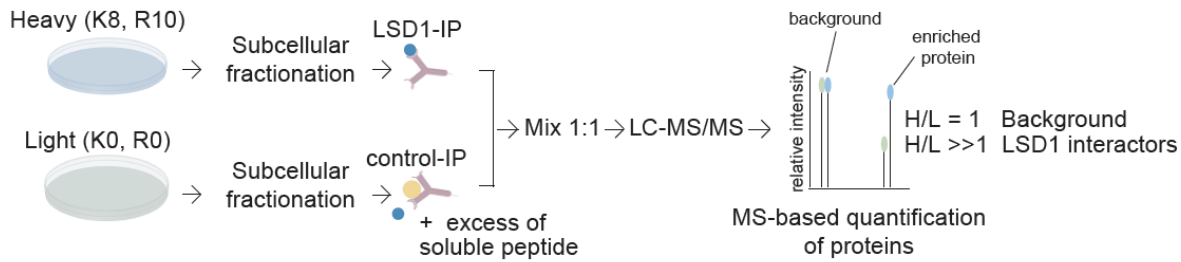


Figure 17: Experimental design of the SILAC/co-IP to characterize LSD1 specific interactors. NB4 cells are labelled in culture with both light and heavy arginine and lysine. Cells are harvested and cell extracts fractionated into cytosolic and nuclear fractions. Nuclear extracts are used as Input for the LSD1-co-IP. As mock co-IP, in one of the two channels the anti-LSD1 antibody is co-incubated with an excess fold of the corresponding LSD1-blocking peptide (light in the forward replicate). Upon carrying out the two co-IPs in parallel, the precipitated proteins are mixed 1:1 prior to LC-MS/MS analysis and quantitation of LSD1 interactors. In the forward experiment specific binders show a SILAC H/L ratio higher than 1, while background proteins a H/L ratio equal or similar to 1.

5.1.1. SILAC labelling of NB4 cells

NB4 cells were cultured in parallel in SILAC RPMI medium supplemented with either light (K0, R0) or heavy (K8, R10) lysine and arginine. The SILAC medium included heat-inactivated and dialyzed FBS, to guarantee that the isotope-coded amino acids represented the only source of amino acids employed for new protein synthesis. We performed growth curve analysis of NB4 cells cultured in light and heavy SILAC conditions up to 12 days of culturing. Light and heavy labelled cells showed similar and comparable growth (Figure 18A), suggesting that the viability was not affected by the different isotope-coded amino acid added to the media. We cultured NB4 cells for nine replications in heavy SILAC medium before assessing the incorporation rate of the heavy isotopic variants of lysine and arginine. To determine the incorporation rate, NB4 whole cell extract was subjected to in-solution trypsin digestion before analysing the sample in LC-MS/MS. We found that around 95% of peptides containing at least one lysine or arginine were present only in the heavy form, the 0.3%-0.4% only in the light one while 3% of peptides with lysine and arginine were detected in both light and heavy conditions. The analysis of the distribution of the H/L

ratios associated to these peptides allowed extrapolating an incorporation rate of 98% (Figure 18B). These results confirmed the almost complete incorporation of the heavy isotope-coded amino acids and that these cells were suitable for the following quantitative proteomics experiments.

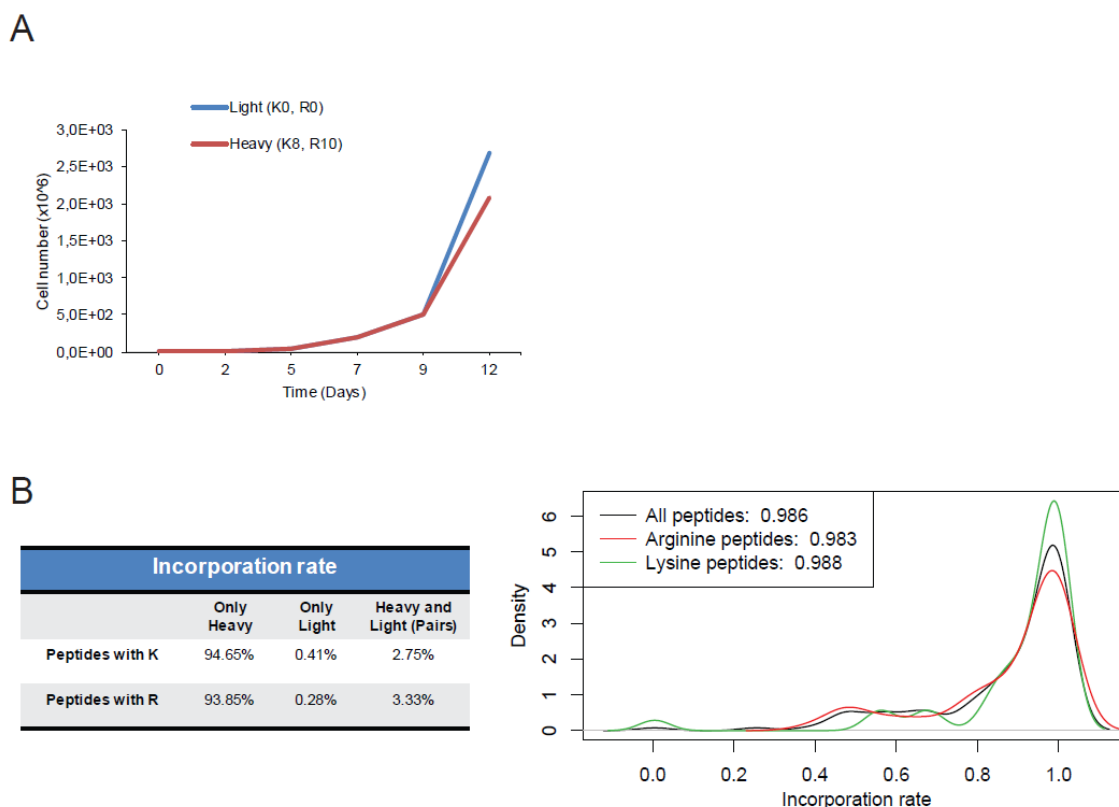


Figure 18: SILAC labelling of NB4 cells. A) Growth curve analysis of heavy (in red) and light (in blue) labelled NB4 cells. B) Analysis of the incorporation rate of heavy lysine (K8) and arginine (R10). The table (on the left) shows the percentage of peptides containing lysine and arginine identified in each SILAC form. The distribution of H/L ratios of peptides containing the intensity values in both light and heavy conditions is displayed on the right.

5.1.2. Optimization of the LSD1-IP conditions for the characterization of the LSD1 interactors in NB4 cells

We setup to determine the optimal conditions of LSD1-co-IP in unlabelled NB4 cells prior to move to the SILAC setup. We started by optimizing a cell fractionation protocol (described in details in Materials and Method, paragraph 4.3.1) that allowed the efficient separation of cytosol and nucleus, as assessed by the increased signals of Vinculin and

GAPDH in the cytosol, and of Lamin-B1 in the nuclear fraction. As expected, LSD1 and its interactor HDAC1 accumulated in the nucleus rather than in the cytosol, given their known function as chromatin modifying enzymes (Figure 19A). Next, we tested the LSD1-IP efficiency using the same experimental conditions already established in the lab (see Material and Methods, paragraph 4.3.2). In particular, we used 20% of the IPed sample for Western Blot analysis and the remaining 80% for Coomassie-staining followed by LC-MS/MS analysis. By Western Blot, we detected a strong enrichment of the bait in the IP sample (IP-LSD1) and a corresponding depletion in the flow-through (FT) (Figure 19B). Conversely, in the IgG control IP the signal of LSD1 was absent and all the bait remained in the FT. Moreover, two known LSD1 interactors -HDAC1 and HMG20B- were co-IPed efficiently with LSD1 and were not detected in the mock IP (Figure 19B). These results suggested that the IP worked efficiently and we could proceed with the analysis of the same IP samples in MS. After running the Input and IP samples in polyacrylamide gel (SDS-PAGE) and staining it with colloidal coomassie (Figure 19C), each lane was cut in 8 slices and in-gel digested with trypsin prior to MS. From MaxQuant Label-Free Quantitation (LFQ) analysis of MS raw data, we found that well-known LSD1 interactors RCOR1, HDAC1, HDAC2, ZNF217 and YBX3 were co-enriched with LSD1 in the LSD1-IP compared to the mock control IP (Figure 19D). Furthermore, a panel of other known LSD1 binders such as HMG20B, ZMYM3 and CHD4 were detected only in the LSD1-IP and did not show any LFQ value in the control, which indicates their unique enrichment in the LSD1-IP (Figure 19E).

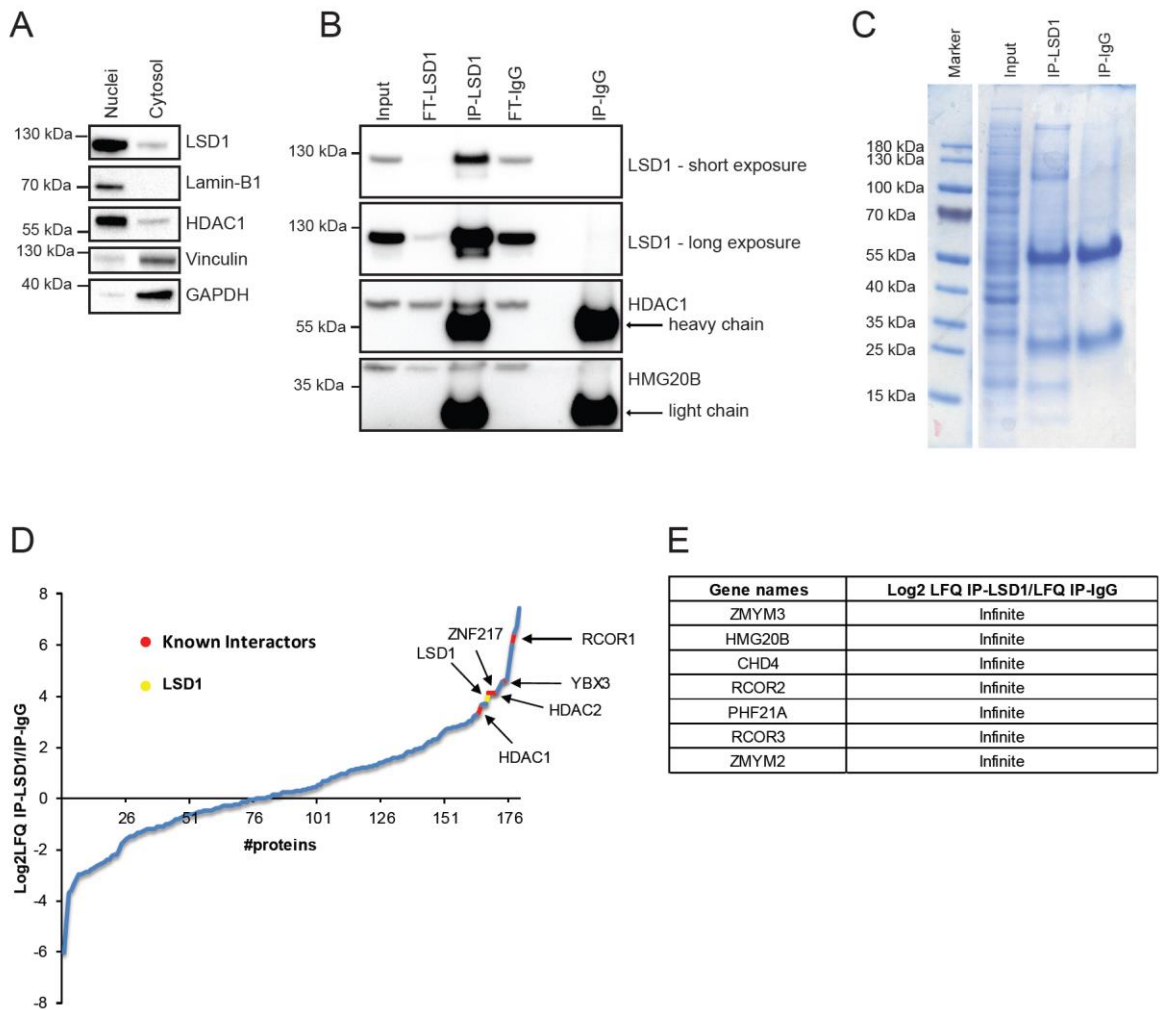


Figure 19: Optimization of the LSD1-IP protocol. A) Western Blot validation of the NB4 sub-cellular fractionation. Vinculin and GAPDH are used as markers of cytoplasm while Lamin-B1 as marker of nucleus. B) Western blot validation of the co-IP of some known LSD1 binders together with the bait, in NB4 cells. Immuno-precipitation with Immunoglobulin G (IgG) is used as negative control. FT = Flow-Through. C) Colloidal coomassie staining of Input and IPed samples from both LSD1-IP and IgG-IP. D) Distribution of proteins based on the log₂ ratio of the LFQ intensity in the IP-LSD1 over the IP-IgG. The bait is displayed in yellow while known LSD1 interactors in red. E) Table displaying the list of some known LSD1 binders detected by MS only in the LSD1-IP, not in the control.

As last step of optimisation of the MS-interactomics experiment, we determined the correct amount of blocking LSD1 peptide to be used in combination with the anti-LSD1 antibody as negative control to saturate the antibody binding and, therefore, evicts the bait together with all its specific interactors¹⁵⁴. First, we performed a small-scale LSD1-IP in the presence

of increasing doses (60X, 80X, 100X and 120X fold molar excess) of soluble LSD1 blocking peptide. The competition worked rather efficiently in all the chosen conditions. However, with 120X the blocking resulted slightly more efficient since the LSD1 signal was almost absent upon ECL detection (Figure 20A). We then performed the same experiment in a larger-scale setup, by choosing two different doses of peptide: 80X as intermediate condition and 120X as the strongest one. Here, the addition of 120X excess fold led to a slightly stronger reduction of the LSD1 signal in the IP sample and a corresponding increase in the FT without abolishing completely the binding (Figure 20B). Hence, we chose 120X as the amount of peptide to ensure stronger competition of the bait and produce an optimal H/L SILAC ratio of LSD1 and its co-associated proteins. Altogether, these results proved that we had established optimal conditions and could move to a SILAC experimental design to gain a more confident definition of LSD1 interactors.

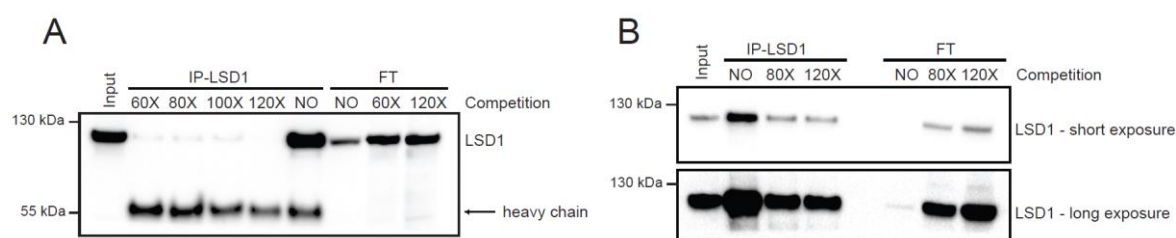


Figure 20: Optimization of peptide fold excess for the LSD1-IP. A) Analysis of the LSD1 levels in the LSD1-IP not competed (NO) or competed with 60, 80, 100 or 120 fold molar excess of blocking peptide in small-scale. B) Analysis of the LSD1 levels in the LSD1-IP not competed (NO) or competed with 80 and 120 fold molar excess of blocking peptide in a big-scale setup.

5.1.3. SILAC proteomics in combination with LSD1-IP for the characterization of the basal LSD1-interactome in NB4 cells

We carried out three SILAC co-IP experimental replicates, two in forward setup (DIR1 and DIR2) whereby the blocking peptide was added to the light channel and one in reverse mode (REV), where the peptide was added to the heavy-labelled samples (Figure 21A). We first

validated the efficiency of the IP and of the competition with the peptide by Western Blot on 10% of the IP samples. For both DIR1 and REV, we detected an enrichment of LSD1 in the IP channels without the peptide compared to the Input (i.e. heavy in the forward experiment and light in the reverse one) and a corresponding depletion in the FT. Furthermore, the competition worked efficiently as demonstrated by the lower amount of LSD1 in the IP carried out in the presence of the peptide and the corresponding increased level in the FT (Figure 21B-C).

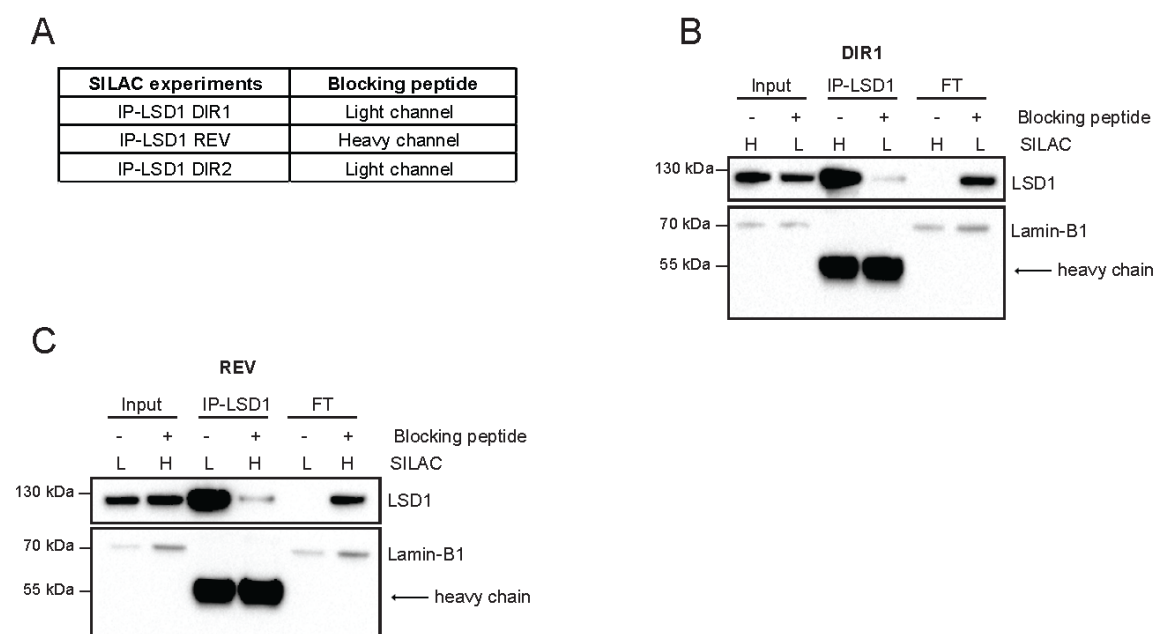


Figure 21: Western Blot quality control assessment of the SILAC LSD1-IP experiments prior to MS. A) The set of SILAC LSD1-co-IP replicates carried out in NB4 cells. B) Western blot validation of the LSD1-IP in the DIR1 replicate. IP with the LSD1 antibody in combination with the LSD1 soluble peptide is used as negative control. C) Western blot validation of the LSD1-IP in the REV replicate. IP with the LSD1 antibody in combination with the LSD1 soluble peptide is used as negative control.

The remaining 90% of samples were loaded on SDS-PAGE for MS-analysis. The three replicates showed a similar pattern of bands suggesting a good level of reproducibility within the experiment (Figure 22A). Each lane from Input and IP were cut in 8 slices and subjected to in-gel trypsin digestion (see Materials and Methods, paragraph 4.4) prior to LC-MS/MS.

Samples were acquired on a Orbitrap- QExactive HF mass spectrometer and MS raw data analysed by Max-Quant (MQ) algorithm for protein ID and quantification. Then, output files were manually filtered as follows: proteins were considered as confidently identified and quantified if at least two peptides, one of which unique, were assigned to them and if they have a SILAC Ratio Count greater than 1 ($RC > 1$). Upon this filtering, we identified 430 proteins in DIR1, 444 in REV and 343 in DIR2 (Figure 22B). To define the putative LSD1 interactors in NB4 cells, we used a mixing-model statistical approach applied to the protein SILAC \log_2 ratio distribution of each replicate to distinguish the background population of proteins -displaying a SILAC \log_2 ratio close to 0- from the specific interactors that displayed a SILAC \log_2 ratio much higher than 0. As such we discriminated, in each replicate, two subpopulations, one of which corresponded to the enriched proteins (Figure 22C; right-side populations, coloured in green). We calculated the mean (μ) and standard deviation (σ) of the enriched protein distribution and defined as putative LSD1 interactors those proteins showing a \log_2 SILAC ratio higher than $\mu - \sigma$. Thus, we identified 147 putative LSD1 interactors from the overlap of the putative binders of three biological replicates (Figure 22D).

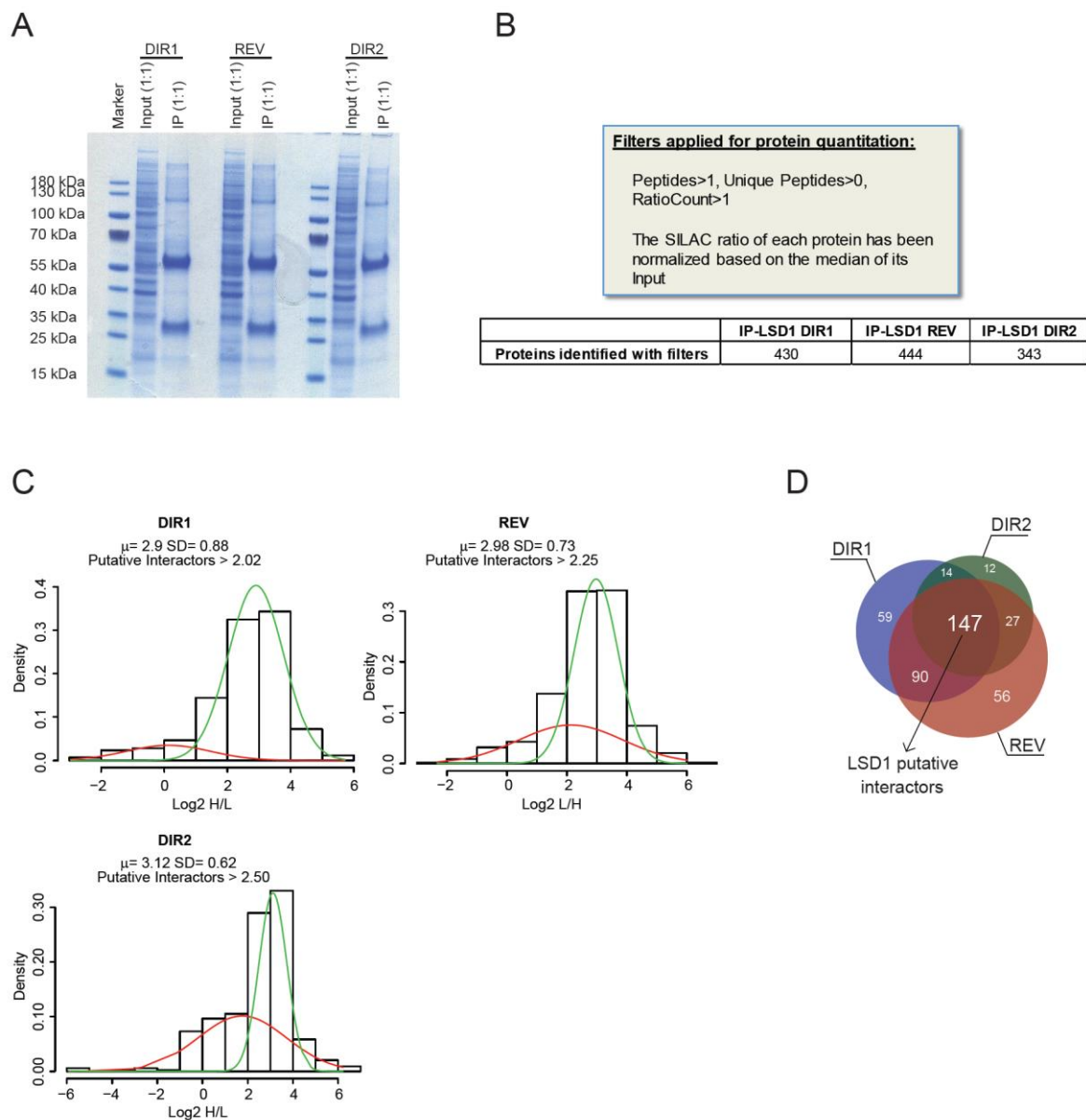


Figure 22: MS-based identification of the LSD1 Interactors in NB4 cells. A) Coomassie staining of Input and IP samples of each SILAC replicate. B) Description of the filtering criteria applied for protein identification and quantitation prior to the LSD1 interactome analysis. The table shows the number of quantified proteins in each replicate after the application of these criteria. C) Distribution of proteins based on their \log_2 SILAC ratio values. A mixing model approach allows distinguishing two different population of proteins. μ = average while SD = standard deviation (σ). For each replicate, the minimum \log_2 ratio value used to define a putative LSD1 interactor is indicated. D) Venn diagrams with number of individual and overlapping putative LSD1 interactors identified in three different SILAC replicates.

Among this list of putative LSD1 binders (Appendix 1) we detected almost all the proteins belonging to the CoREST complex such as HDAC1, HDAC2 and RCOR1 and other known

interactors, as represented in the right-upper quadrant of the scatterplot displaying the comparison between the DIR1 and REV replicates as well as in the table beside (Figure 23A-B). To understand which protein complexes were significantly enriched in our interactome data, we used the "g:Profiler" tool ¹⁹⁶, which generates functional descriptions of gene lists derived from large-scale experiments, including the presence of protein complexes described in the CORUM database ²⁰⁹. We observed a significant enrichment of several LSD1-containing complexes such as the NuRD, the BHC, the CtBP complex and the ALL-1 supercomplex (Figure 23C). These protein complexes had already been reported in several research studies to have important roles in the initiation and development of different leukemia types. For instance, in APL the NuRD complex is recruited by PML-RAR α to silence specific target genes and facilitate the blockade of cellular differentiation ²¹⁰. Instead, the CtBP complex had been described to interact with the leukemia-associated fusion gene AML1-ETO promoting, this way, growth increase and abnormal differentiation of murine bone marrow progenitors ²¹¹.

Importantly, this analysis also revealed novel interactions, which may suggest novel molecular processes in which LSD1 is involved in NB4 cells: for instance, we found the enrichment of the spliceosome complex represented by various subunits of the complex, such as PRPF6, TRA2B and SRRB as well as several serine/arginine-rich splicing factors. In line with this result, other members of the CoREST complex, such as HDAC1/2 have been previously showed to interact with different splicing factors and to regulate alternative splicing processes ²¹². So, this network may be the results of indirect, yet specific, interactions pulled down by our approach. Evidences suggest that the splicing regulation operated by HDACs occurs co-transcriptionally, when nascent RNA is still associated to chromatin, by controlling splice site selection ²¹³. Our results may suggest that other chromatin modifiers are also involved in regulating this process, such as LSD1 itself. In this regard, it is interesting to note that the increase of H3K9me2 (a known LSD1 substrate) at

the alternative exon's target site has been mechanistically linked to a diminished elongation rate and inclusion of exons ²¹⁴.

Our LSD1-interactome results also showed the enrichment of the RNA-Pol II core complex represented by the RPB1 and RPB2 subunits of the RNA-Polymerase II as well as different factors involving in the polyadenylation complex such as CPSF1, CPSF3, CPSF7, CSTF1, CSTF3, SYMPK. Altogether, these data suggested a strong link between LSD1 and the CoREST complex to the processes of transcription and mRNA processing.

Among the putative specific interactors we also identified MCM5 and MCM7, subunits of the MCM complex and RFC2 and RFC4 that belong to the RFC complex, all proteins involved in DNA replication (Figure 23C).

We corroborated the SILAC-MS readout for some of the putative binders, by LSD1 Co-IP followed by WB and confirmed that the signal of the interactors was increased in the IP compared to the Input and reduced in the IP with the excess of blocking peptide (Figure 23D).

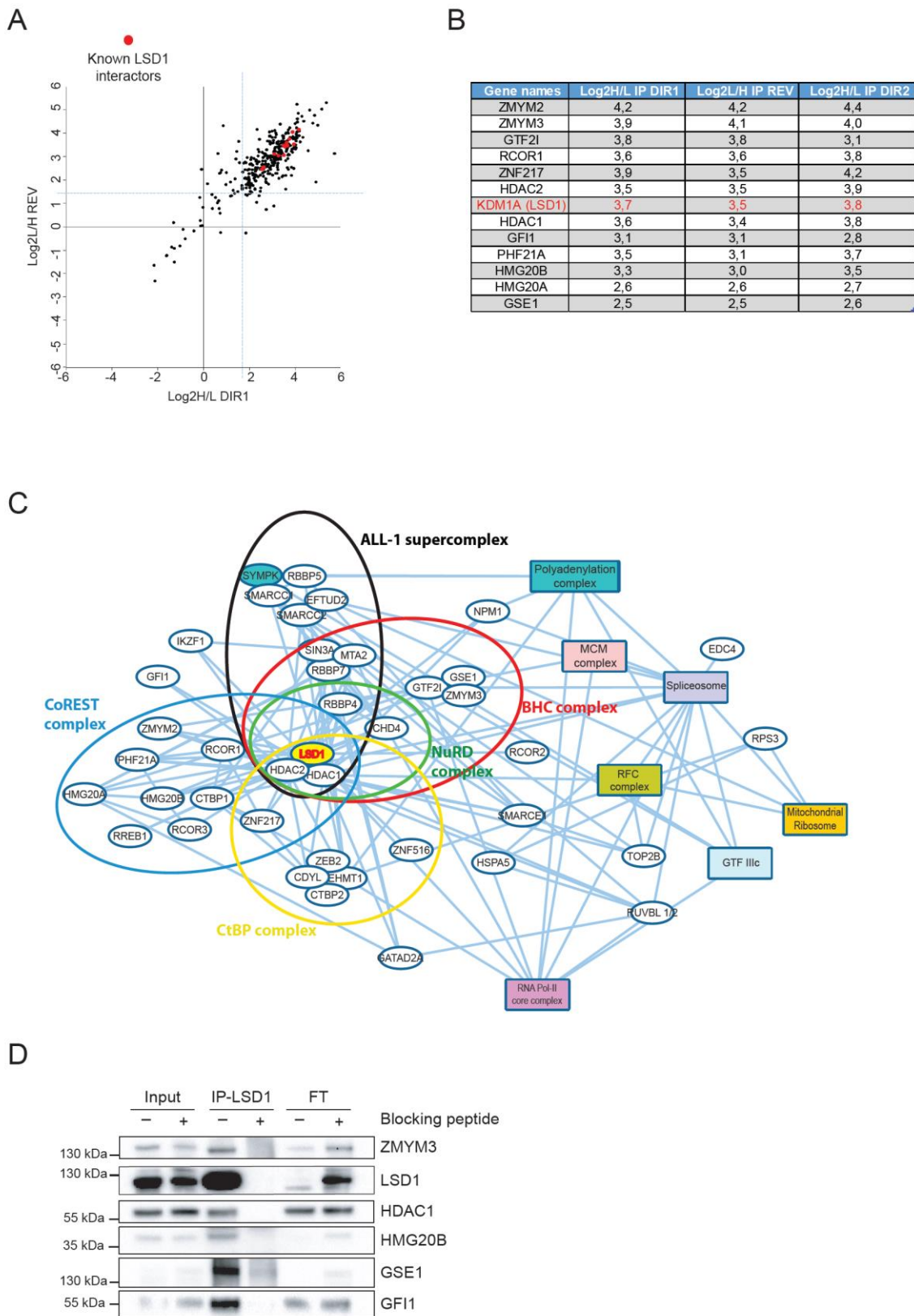


Figure 23: Characterization of the complexes enriched in the LSD1 interactome. A) Scatterplot showing the \log_2 (heavy/light) ratio of forward reaction (DIR1) on the x axis and the \log_2 (light/heavy) ratio of reverse reaction (REV) on the y axis. The LSD1 interactors are represented in the top-right quadrant. The blue dashed lines define the threshold used to discriminate the interactors from the background. In red dots known LSD1 interactors are shown. B) Table displaying some of the known LSD1 binders detected within the list of 147

putative interactors obtained from the LSD1 basal interactome. The \log_2 SILAC ratio of each interactor in each SILAC replicate is shown. C) Analysis of the complexes enriched in the LSD1 interactome. Protein-protein interaction analysis is obtained by using STRING functional database. Interactions are visualized by Cytoscape and the complexes analyzed by g:Profiler containing CORUM database. D) Western blot analysis of LSD1 and some identified interactors in LSD1 IP with or without the presence of the blocking peptide.

5.2. MS-based analysis of the changes in the LSD1-interaction network upon treatment with LSD1 inhibitors

With the basal LSD1 interactome at hands, we set to profile its dynamic changes upon pharmacological treatment of the cells with the LSD1 inhibitor MC2580. Again, we performed the experiment in SILAC, whereby in the forward setup the heavy-labelled NB4 cells were treated for 24 hours with the LSD1 inhibitor while in the reverse replicate the drug was added to the light cells. Upon LSD1-co-IP, the pulled-down proteins were mixed 1:1 prior to LC-MS/MS-analysis. Max-Quant protein quantitation allowed to define recruited, evicted and stably associated proteins to LSD1 upon drug treatment: in particular, in the forward replicate, proteins with a SILAC H/L ratio > 1 were defined as recruited, those with a SILAC ratio $H/L < 1$ as evicted, while proteins with a SILAC ratio $= 1$ as the interactors whose binding was not influenced by the drug treatment (Figure 24).

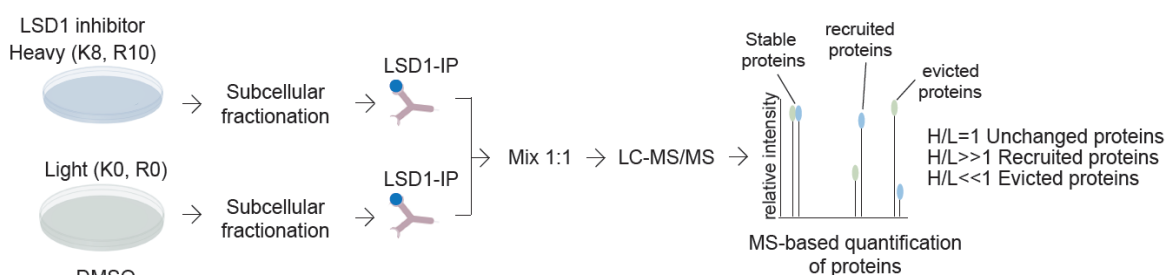


Figure 24: Experimental design of SILAC/co-IP strategy to study the changes in the LSD1 interactome upon drug treatment. Light and Heavy NB4 cells are treated with either DMSO or the LSD1 inhibitor MC2580 for 24 hours. Nuclear extracts are employed as Input for the LSD1 co-IP in both the channels. Upon IP, the precipitated proteins are mixed 1:1 prior to LC-MS/MS and quantitation of recruited, evicted and stably associated proteins to the LSD1 complexes that is determined on the bases of their respective SILAC

ratios. Proteins with a SILAC ratio > 1 are defined as recruited, while proteins with a SILAC ratio < 1 as evicted. Unchanging proteins display a SILAC ratio = 1.

Before MS analysis, we performed WB validation of the LSD1-IP using 10% of the samples. In both SILAC replicates, the LSD1 signal was increased in all IP samples and almost absent in the FT. Moreover, the LSD1 histone target H3K4me2 increased in the nuclear Input of the treated samples compared to the untreated ones, demonstrating that the drug MC2580 was active in inhibiting LSD1 enzyme (Figure 25A-B).

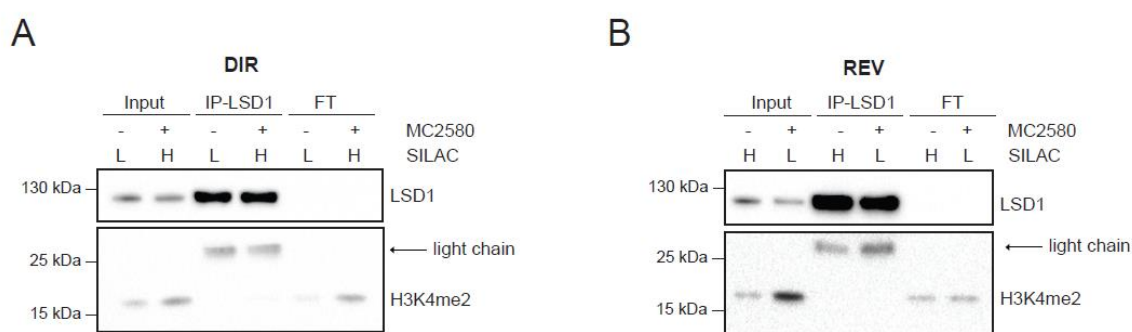


Figure 25: Western Blot validation of the SILAC dynamic LSD1-IP replicates. A) Western blot validation of the LSD1-IP of the DIR1 replicate in both untreated and treated NB4 cells with the inhibitor MC2580. H3K4me2 is used as a control of the drug treatment. B) Western blot validation of the LSD1-IP of the REV replicate in both untreated and treated NB4 cells with the inhibitor MC2580. H3K4me2 is used as a control of the drug treatment.

90% of the samples, together with the 1:1 mixed Input, were loaded on SDS-PAGE and coomassie-stained for subsequent sample preparation processing and MS-analysis. Also in this dynamic experiment, the IP samples displayed a comparable pattern of bands, suggesting a good level of reproducibility within the experiment (Figure 26A). Each gel lane was cut in 8 slices and in-gel digested (See Materials and Methods, paragraph 4.4) with trypsin prior to LC-MS/MS analysis. After analysis of the MS-data by MQ, the outputs were filtered following the same parameters used for the definition of the basal LSD1 interactome (See paragraph 5.1.3.). We identified 638 proteins in the forward (DIR) replicate and 538 in

the reverse (REV) one (Figure 26B). The distribution of proteins based on the \log_2 normalized SILAC ratios (H/L in the DIR replicate and L/H in the REV replicate) appeared normal (Gaussian) in both replicates (Figure 26C). Therefore, to identify the modulated proteins after LSD1 inhibition we used an approach based on the calculation of the median (M) and the standard deviation (σ) of the \log_2 SILAC ratios of each protein distribution. LSD1 interactors were grouped into three classes: 1) proteins with a \log_2 SILAC ratio lying between the $M \pm 2\sigma$ in both replicates are those stably associated to LSD1, 2) proteins with a \log_2 SILAC ratio $> M + 2\sigma$ were considered recruited to LSD1 binding upon MC2580 and 3) proteins with a \log_2 SILAC ratio $< M - 2\sigma$ in both replicates evicted from the LSD1 network (Figure 26D).

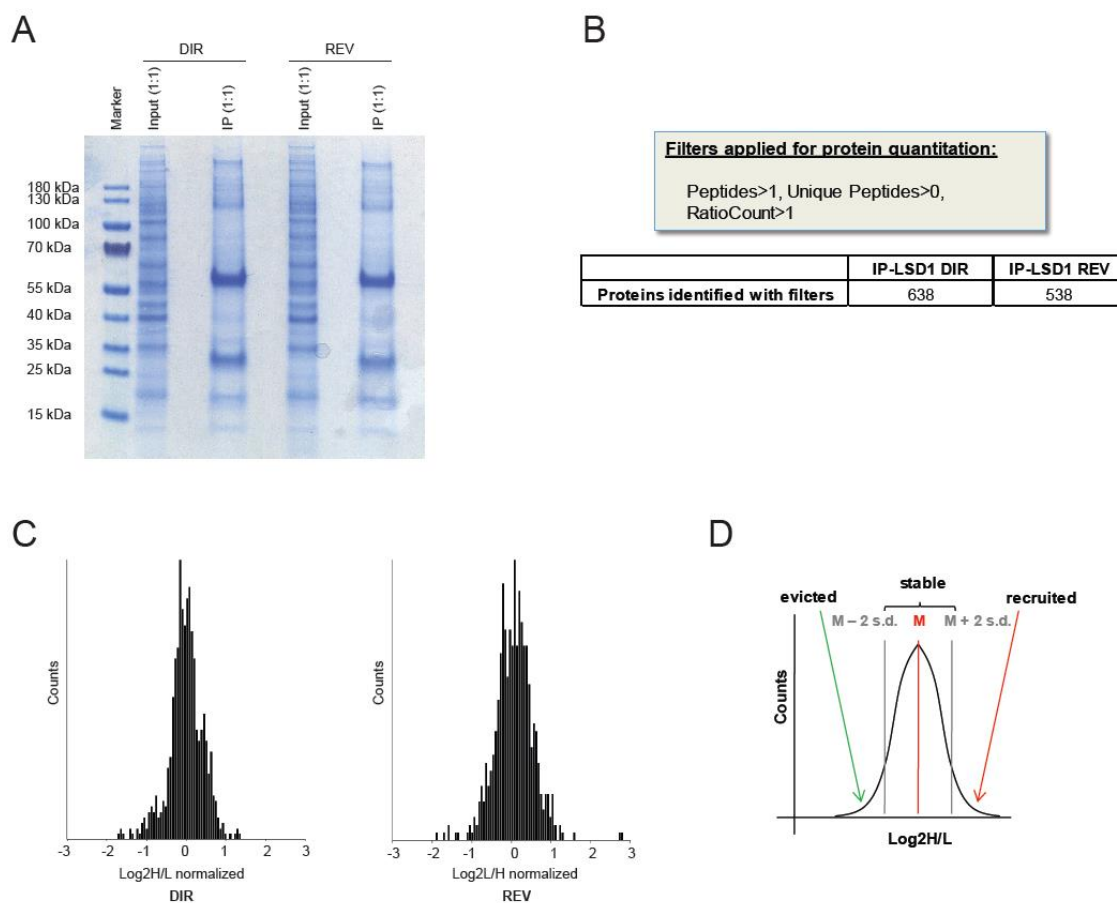


Figure 26: MS-based characterization of the dynamic LSD1 interactome upon drug treatment. A) Coomassie staining of both Input and IP samples of the two SILAC replicates. B) Number of proteins quantified in each replicate upon application of specific filtering criteria for protein quantitation. C) Protein distribution of each SILAC replicate based on the

log₂ normalized SILAC ratio values of each quantified protein. D) Statistical approach employed to define recruited, evicted and stable associated interactors upon treatment with the LSD1 inhibitor. M = median while SD = standard deviation.

Overall, the majority of interactors remained stably associated with LSD1 upon drug treatment, which is partially expected since we used a compound that binds to the catalytic domain of the enzyme, and not to the tower domain, which is currently considered the regions mostly involved in protein-protein interaction⁵. The only exceptions were represented by the proteins GFI1, GSE1 and EDC4, with EDC4 that is dynamically recruited, and GFI1 and GSE1 being evicted by the drug treatment (Figure 27A). EDC4 is a component of a complex containing DCP2 and DCP1A which functions in decapping of mRNAs²¹⁵. *GSE1* is an oncogene that is overexpressed in both breast and gastric cancer and whose increased expression correlates with increased cell proliferation, colony formation, cell migration and invasion^{216,217}. However, very recently, GSE1 was also described to have tumour-suppressor roles in neuro-epithelial stem (NES) cells²¹⁸. GFI1 is a transcription repressor essential for haematopoiesis. It regulates neutrophil differentiation, promotes proliferation of lymphoid cells, and is required for granulocyte development^{219,220}. Among them, we decided to follow-up GSE1 and GFI1 based on various consideration. In the case of GSE1, following our screening we discovered a number of studies which demonstrated that this protein binds to LSD1²²¹⁻²²³. However, despite the evidence of a role in some tumours, nobody have so far investigated its function in cancer in the context of its interaction with LSD1. The importance of the LSD1-GFI1 interaction was already reported in several works, with a paper showing that interaction of LSD1 with the N-terminal SNAG domain of GFI1 as well as its isoform GFI1b is fundamental for the GFI1 role as transcriptional repressor in haematopoietic cells²³. Interestingly, two research studies demonstrated that the LSD1 inhibitor T-3775440 was able to disrupt the interaction of LSD1 with GFI1b^{104,105}. Based on these works, we set to verify the effects of the disruption of the LSD1-GFI1 binding in our model system and with our inhibitor compounds. We first

validated the MS data by co-IP of LSD1 followed by Western Blot analysis in the presence and absence of both MC2580 and DDP-38003 compounds and could confirm that the interaction of GFI1 and GSE1 with LSD1 is indeed diminished by treatment of the cells with both drugs (Figure 27B).

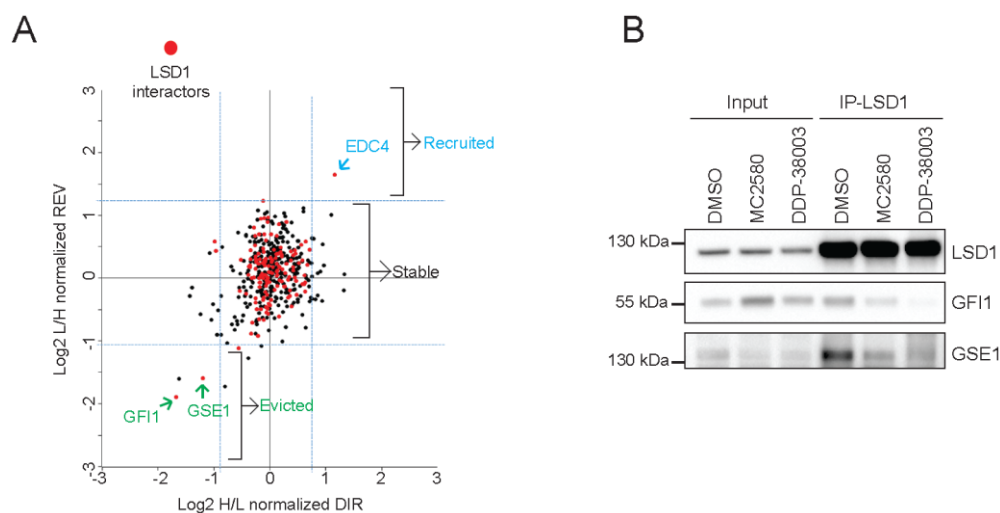


Figure 27: LSD1 pharmacological inhibition alters the interaction of LSD1 with EDC4, GFI1 and GSE1. A) Scatterplot showing the log₂ (heavy/light) ratio of forward reaction on the x axis and the log₂ (light/heavy) ratio of reverse reaction on the y axis. In the top-right quadrant are represented proteins recruited by LSD1 after its inhibition, while in the bottom left quadrant the evicted proteins. Proteins previously identified as interactors from the basal experiment are shown as red dots. The blue dashed lines delimitate the statistical threshold used to define recruited and evicted proteins. B) Western blot analysis of LSD1, GFI1 and GSE1 in LSD1-IPs in control and treated NB4 cells with MC2580 (2 μM) and DDP-38003 (2 μM) for 24 hours.

5.3. Impact of the LSD1 inhibition on the binding of the catalytic inactive LSD1 mutant to GFI1 and GSE1

To gain mechanistic insights on GFI1 and GSE1 interaction with LSD1, we investigated whether the reduced binding upon pharmacological treatment was truly dependent on the catalytic activity of the enzyme, also taking into account that the emerging evidence suggested that the catalytic activity was indeed not essential to sustain NB4 proliferation and differentiation (see Introduction paragraph 2.6.1.). We then performed the IP of GFI1 and LSD1 in NB4 where we first knocked-out (KO) LSD1 and, then, re-expressed either the

wild-type enzyme (WT), or its catalytic inactive form (K661A), at both the basal level and upon 24 hours treatment with MC2580. We observed that both GFI1 and GSE1 were able to bind similarly to WT and K661A-LSD1 at basal state (Figure 28A), and that their binding to both forms of the enzyme was reduced upon drug treatment (Figure 28B). This demonstrated that the inhibition of GFI1 and GSE1 interaction to LSD1 is independent from the catalytic activity of the enzyme.

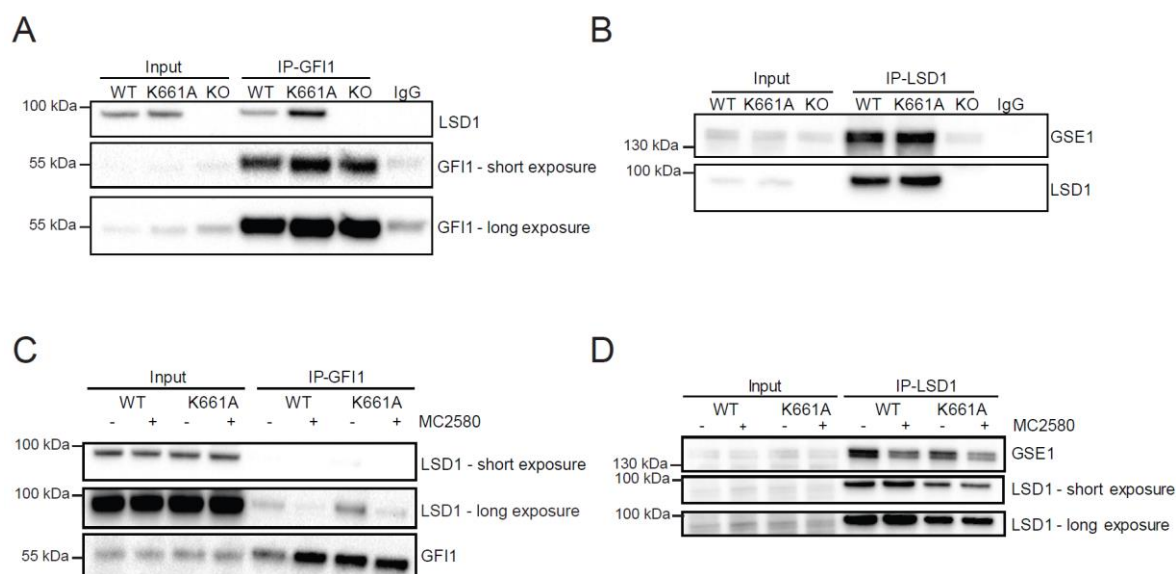


Figure 28: LSD1 inhibitor alters the interaction of LSD1 with GFI1 and GSE1 irrespective of the inhibition of the LSD1 enzymatic activity. A) Immuno-precipitation of endogenous GFI1 in NB4 KO-LSD1 cells transduced with empty vector (EV), wild-type (WT) or catalytic inactive LSD1 (K661A). B) Immuno-precipitation of exogenous LSD1 in NB4 KO-LSD1 cells infected with EV, WT or K661A-LSD1. C) Immuno-precipitation of endogenous GFI1 in NB4 KO-LSD1 cells transduced with WT or K661A-LSD1. The cells are treated with either DMSO or the inhibitor MC2580 (2 μ M) for 24 hours. D) Immuno-precipitation of exogenous LSD1 in NB4 KO-LSD1 cells transduced with WT or K661A-LSD1. The cells are treated with either DMSO or the inhibitor MC2580 (2 μ M) for 24 hours.

5.4. Phenotypic and molecular effects of the inhibition of LSD1-GFI1 interaction in NB4 cells

Given these results, we started a collaboration with the group of Prof. Saverio Minucci to follow-up functionally the impact of LSD1-GFI1 interaction in our model system. Different research studies suggested that this interaction regulates the balance between self-renewal

and differentiation in leukemic cells^{23,89}. Based on the evidence that GFI1 interacts with LSD1 through its SNAG domain²²⁴, we generated a LSD1 mutant (D553, 555, 556A) that was previously been shown to be unable to bind to the SNAG domain of Snail1, a protein involved in the epithelial-mesenchymal transition (EMT) which interaction with LSD1 was found in HEK293 cells¹⁵. We transduced this 3D-LSD1 mutant together with the WT form and the Empty Vector (EV) in LSD1-KO NB4 cells. After the infection, cells were sorted based on their GFP fluorescence to select only the cell population correctly transduced with the viral particles (Figure 29A). We assessed that the expression of the 3D-LSD1 was similar to that of the WT (Figure 29B), thus corroborating the possibility to compare the phenotypic effects of these different LSD1 forms in NB4 cells. We also evaluated the ability of the 3D mutant to disrupt the interaction with GFI1, as already proven with Snail1. To this aim, we performed the Co-IP of GFI1 in NB4 LSD1-KO cells transduced with WT-LSD1, 3D-LSD1 and EV and found that, unlike the WT, the 3D-LSD1 lost its binding to GFI1 (Figure 29C).

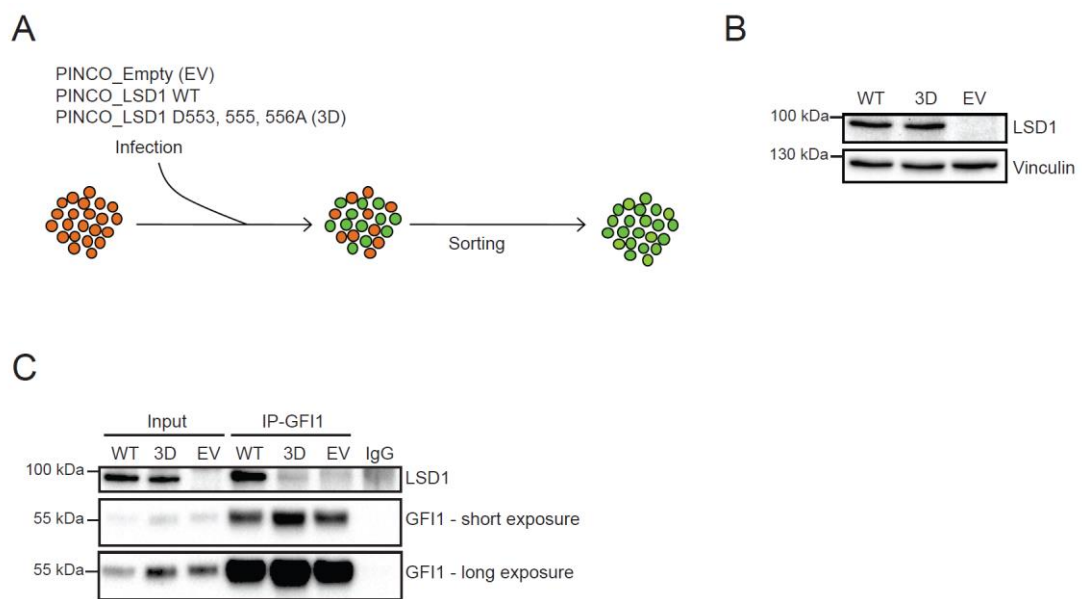


Figure 29: Analysis of the LSD1-GFI1 binding upon expression of the LSD1-3D mutant in NB4 KO-LSD1 cells. A) Schematic representation of the infection of NB4 LSD1-KO cells with either the Empty Vector (EV) or the WT-LSD1 or the LSD1 D553,555,556A (3D) mutant. B) Western blot analysis of exogenous LSD1 in NB4 LSD1-KO cells transduced with either EV or WT-LSD1 or 3D-LSD1 mutant. Vinculin is used as loading control. C) Immuno-precipitation of endogenous GFI1 in NB4 KO-LSD1 cells transduced with EV, WT-LSD1 or 3D-LSD1 mutant.

Since our preliminary data demonstrated that LSD1 inhibitors synergized with physiological doses of ATRA (ATRA 0.01 μ M) to reduce NB4 cell proliferation and induce differentiation (See Introduction, paragraph 2.6.1., Figure 7A-C), we assessed the proliferation of NB4 KO-LSD1 cells re-expressing either WT or 3D-mutant LSD1 upon treatment with either DMSO or ATRA 0.01 μ M. We found that the growth of NB4 cells re-expressing 3D-LSD1 was strongly affected by ATRA treatment, similarly to that of the EV-infected cells. Conversely, NB4 cells re-expressing WT-LSD1 did not show substantial changes in their proliferation in the presence of ATRA (Figure 30A). This indicates that the disruption of the LSD1-GFI1 binding when combined with physiological doses of ATRA causes a strong reduction of the NB4 cell growth. Then, we assessed if the cell proliferation decrease was due to the induction of a differentiation process: we analysed the expression of the myeloid differentiation marker CD11b in NB4-LSD1 KO cells transduced with EV and 3D-LSD1 upon 24, 48 and 72 hours treatment with either DMSO or ATRA 0.01 μ M. The expression of this marker after ATRA treatment was comparable between 3D-LSD1 and EV transduced cells in all the different time points (Figure 30B), suggesting that the differentiation process was re-activated. To confirm this result, we analysed by May Grümwald-Giemsa staining the morphological changes of NB4 KO-LSD1 cells infected with EV, 3D-LSD1 and WT-LSD1 and treated for 96 hours with physiological doses of ATRA. Unlike the WT, EV and 3D-LSD1-transduced cells treated with ATRA 0.01 μ M started to display some morphological changes associated with the re-activation of myeloid differentiation, like decrease of the nucleus-to-cytoplasm (N:C) ratio and increase in nuclear lobulation (Figure 30C).

These data indicate that, in NB4 cells, LSD1-GFI1 interaction may be essential for the establishment of a differentiation block by LSD1 and that the ATRA-sensitization of these cells relies on the specific inhibition of the LSD1-GFI1 interaction.

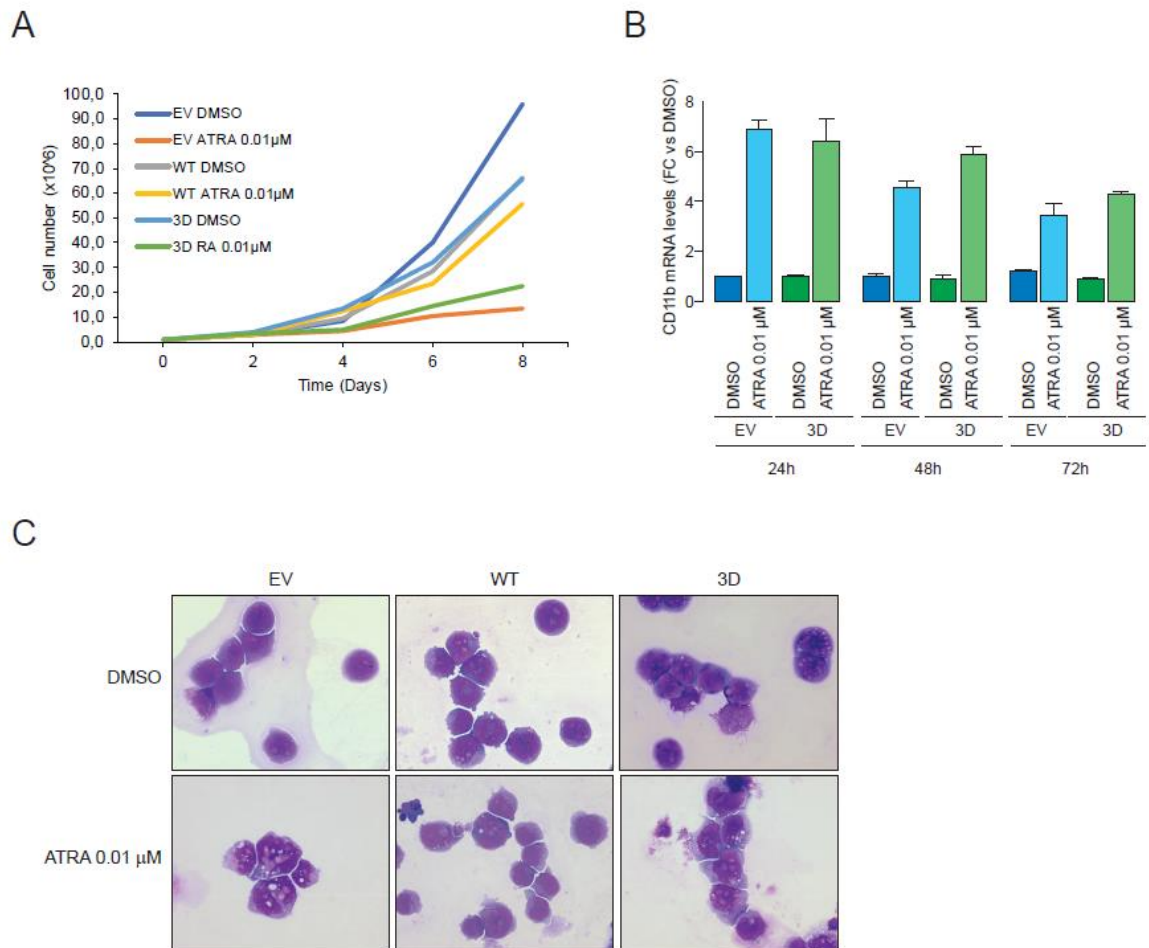


Figure 30: LSD1 interaction with GFI1 serves to self-renewal and proliferation of NB4 cells. A) Growth curve of NB4 LSD1-KO cells transduced with EV, LSD1-WT and LSD1 3D mutant. The cells are treated with ATRA 0.01 μ M or DMSO as control. B) RT-qPCR analysis of CD11b expression levels in NB4 KO-LSD1 cells transduced with EV and LSD1 3D mutant. The cells are treated for 24, 48 and 72 hours with ATRA 0.01 μ M or DMSO as control. Ct values are normalized against GAPDH and referred to DMSO. Graph represents the mean and standard deviation of three independent replicates. C) May Grünwald-Giemsa staining of NB4 cells transduced with EV, WT and 3D-LSD1 and treated for 96 hours with ATRA 0.01 μ M or DMSO as control.

To assess the potential effect on chromatin of the disruption of the LSD1-GFI1 interaction upon MC2580, we performed GFI1 ChIP-seq analysis in NB4 cells and compared with the LSD1-ChIP seq before and after treatment with the LSD1 inhibitor. Most of the LSD1 binding sites largely overlapped with those of GFI1 at the genome level, at basal condition (Figure 31A). However, upon MC2580 LSD1 was evicted from 732 GFI1-bound regions,

among which we found several *cis*-regulatory regions of LSD1-target genes, such as the promoter of PI16 and a putative IRF8 enhancer (Figure 31B). By applying the Ingenuity Pathway analysis (IPA) on the 732 GFI1-bound regions that lost LSD1 ChIP-seq signal after MC2580 treatment we found that the majority of these regions regulated the expression of genes with a relevant role in hematopoietic cell differentiation and function (Figure 31C), further confirming the importance of the LSD1-GFI1 interaction in the regulation of the myeloid differentiation process. Overall, our results suggest that the LSD1 inhibitor MC2580 is able to induce cell differentiation programs by disrupting LSD1-GFI1 interaction on chromatin and displacing LSD1 from a subset of GFI1-regulated genes.

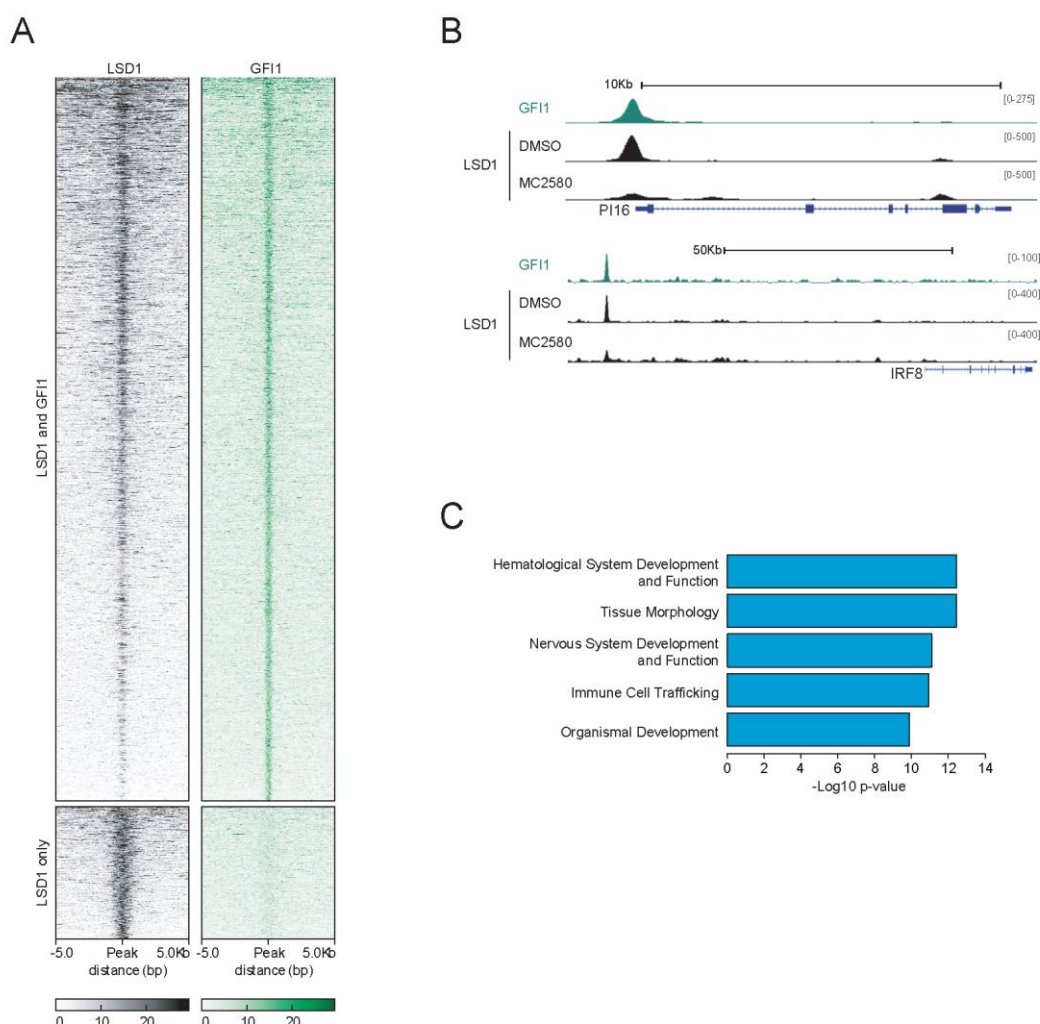


Figure 31: MC2580 displaces LSD1 from chromatin-bound GFI1 in *cis*-regulatory regions of genes involved in hematopoietic cell differentiation. A) Heatmaps displaying ChIP-Seq analysis of LSD1 and GFI1 with regions ranked according to decreasing LSD1 signal. B) Representative snapshots of overlapping LSD1 and GFI1 binding regions in NB4 treated with the LSD1 inhibitor MC2580 or control DMSO. UCSC Genome Browser profile

of LSD1 and GFI1 ChIP-seq on the P116 gene promoter and IRF8 putative enhancer. C) Ingenuity Pathway Analysis (IPA) of genes associated to the 732 GFI1-bound regions that missed the LSD1 ChIP-seq signal after treatment with MC2580. The x axis displays the $-\log_{10}$ of the adjusted p-value of each enriched biological process term.

5.5. Mechanistic analysis of the alteration of LSD1-GSE1 binding by LSD1 inhibitors

After investigating the molecular effects of the drugs on LSD1-GFI1 interaction, we started following-up GSE1, which also emerged as being displaced from LSD1 by the dynamic interactomics experiment. While some research studies already reported the importance of the LSD1-GFI1 binding²³, nothing is known about the possibility of modulating this protein-protein interaction pharmacologically with LSD1 inhibitors, similarly to what had been described for GFI1 or its isoform GFI1b^{89,104,105}. First, we validated the eviction of GSE1 from LSD1 upon drug treatment by performing the reciprocal IP using GSE1 as bait in NB4 cells and profiling its binding to LSD1 upon 24 hours of treatment with MC2580. This analysis allowed us to appreciate an unexpected decrease of expression of GSE1 protein in the Input sample upon LSD1 inhibition. The reduction of GSE1 expression was also reflected in the IPs, with a lower amount of IPed GSE1 in the treated condition than in the untreated one (Figure 32A). Then, we verified whether the LSD1-GSE1 interaction was specifically altered by the drug treatment and performed the LSD1-co-IP in NB4 lysates treated with increasing concentration of MC2580 (2, 10 and 20 μ M), with the expectation that if the drug had interfered directly with the protein-protein interaction, we should have expected a decrease of GSE1 pulled-down. This experiment, instead, showed that GSE1 remained bound to LSD1 with the same efficiency even at increasing doses of the drug, co-incubated *in vitro*. This led us to conclude that the results of the quantitative interactomics analysis which suggested a diminished GSE1 levels in LSD1 co-IP was not due to a physical interference of their interactions but to the reduced expression of GSE1. As positive control of this assay we profiled GFI1 and confirmed that its binding to LSD1 was effectively

competed by the compound (Figure 32B). Hence, the effect of how MC2580 acts on the interaction of LSD1 with these two protein is mechanistically completely different.

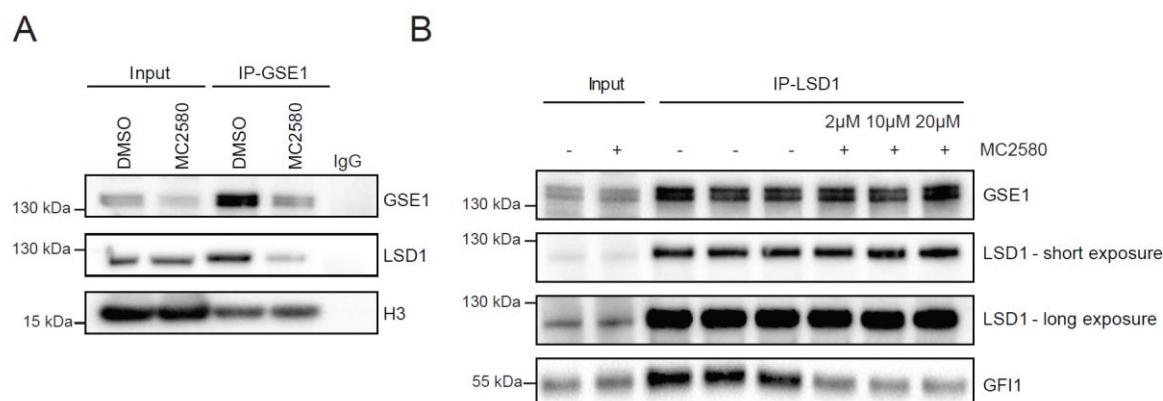


Figure 32: LSD1 inhibitor does not inhibit LSD1-GSE1 physical interaction, but reduces GSE1 protein levels. A) Western Blot analysis of GSE1, LSD1 and H3 in GSE1-IPs, both in control and treated cells with MC2580 (2 μM) for 24 hours. B) Western Blot analysis of GSE1, LSD1 and GFI1 in *in vitro* LSD1-IPs using NB4 nuclear cell extracts co-incubated with increasing doses of MC2580 (2, 10 and 20 μM).

We then assessed whether the reduction of GSE1 expression upon LSD1 inhibition occurred at the transcriptional or translational level. To do this, we treated NB4 cells with MC2580 (2 μM) for 12 and 24 hours and measured both GSE1 mRNA and protein levels by RT-qPCR and WB analysis, respectively. We found that GSE1 protein level was reduced at both time points (Figure 33A). Assembling the WB results of 4 different biological replicates, we showed that the LSD1 inhibitor MC2580, on average, diminished GSE1 protein level of about 30%-40%, with a slight more significant reduction after 24 hours treatment (Figure 33B). The GSE1 transcript was also diminished with the same trend, suggesting a direct regulation of GSE1 transcription by LSD1 and its inhibitor (Figure 33C). With the same experimental setup, we also tested the second LSD1 inhibitor DDP-38003 and confirmed the reduction of both GSE1 protein and mRNA levels at both 12 and 24 hours of treatment with these drugs (Figure 33D-E). These results allowed us to extrapolate that the reduction of

GSE1 is directly dependent on the pharmacological inhibition of LSD1 and not on mere off-targets effect by MC2580.

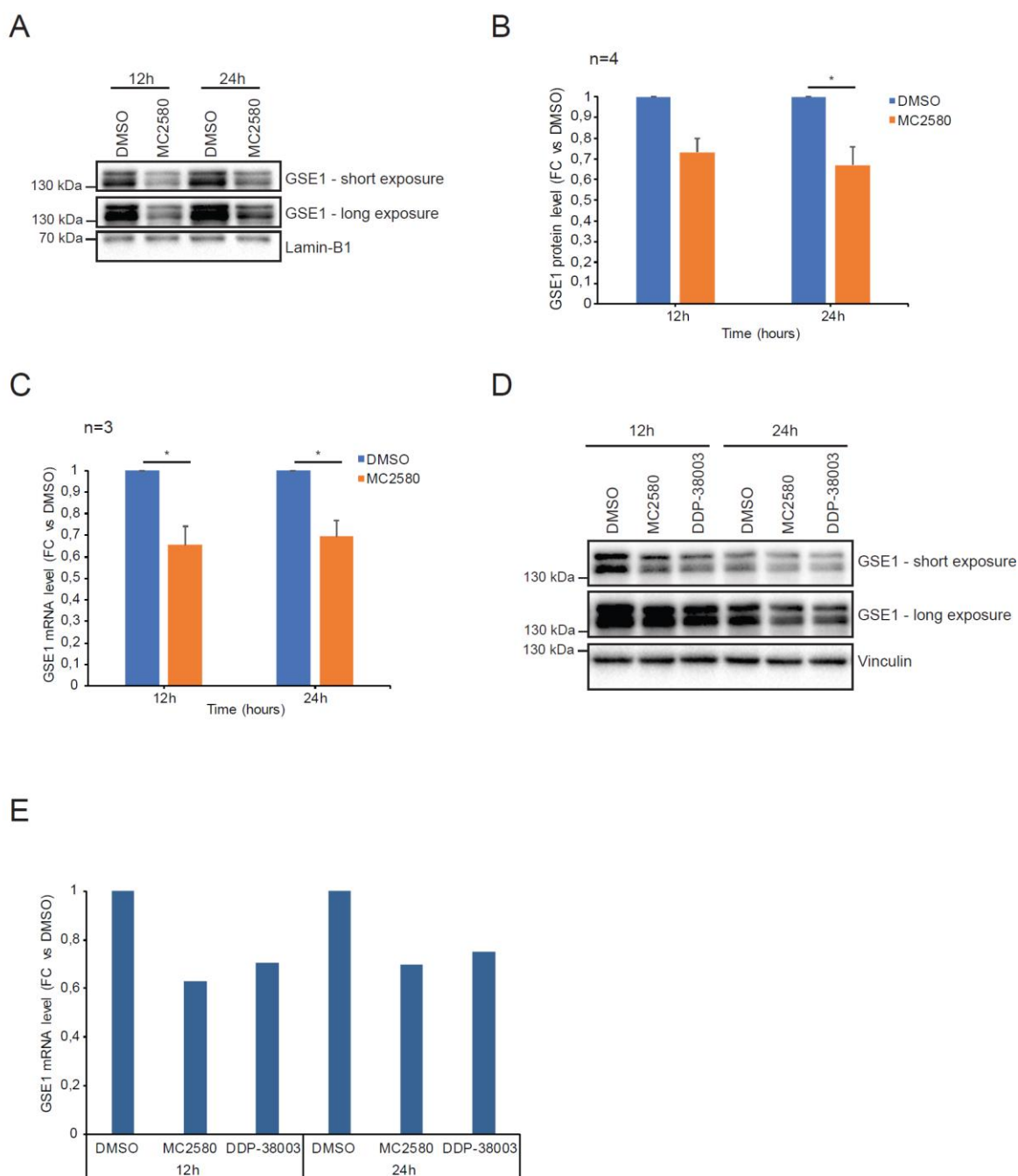


Figure 33: LSD1 inhibitors regulate GSE1 expression at the transcriptional level. A) Western Blot analysis of GSE1 in NB4 cells treated with MC2580 (2 μ M) and control DMSO for 12 and 24 hours. Laminin-B1 is used as loading control. B) Bar graph displaying the quantitation results of GSE1 protein level, normalized over the loading control in 4 different replicates of NB4 cells treated with either MC2580 (2 μ M) or DMSO. The results are plotted as fold change of GSE1 protein level in the treated sample compared to the control DMSO. The chart represents mean \pm SEM (n=4 biological replicates; Paired t-test, *p value<0.05). C) RT-qPCR profiling of GSE1 transcript level upon 12 and 24 hours treatment

with either MC2580 or DMSO (negative control). Ct values are normalized against TATA-binding protein (TBP). Results are plotted as fold change of GSE1 transcript level in the treated sample compared to its control DMSO. Chart represents mean \pm SEM (n=3 biological replicates; Paired t-test, *p value<0.05). D) Western Blot analysis of GSE1 in NB4 cells treated with MC2580 (2 μ M), DDP-38003 (2 μ M) and control DMSO for 12 and 24 hours. Vinculin is used as loading control. E) RT-qPCR analysis of GSE1 transcript level upon 12 and 24 hours of treatment with MC2580, DDP-38003 and the control DMSO. Ct values are normalized against TBP. The results are plotted as fold change of GSE1 transcript level in the treated samples compared to its control DMSO (n=1 biological replicate).

To corroborate the pharmacological data, we also profiled GSE1 protein level by WB in three different biological replicates of LSD1-KO and WT NB4 cells. From this experiment, we observed the reduction of GSE1 level in the KO samples compared to the WT, further demonstrating that the decrease of GSE1 expression is directly dependent on the LSD1 absence/inhibition (

Figure 34).

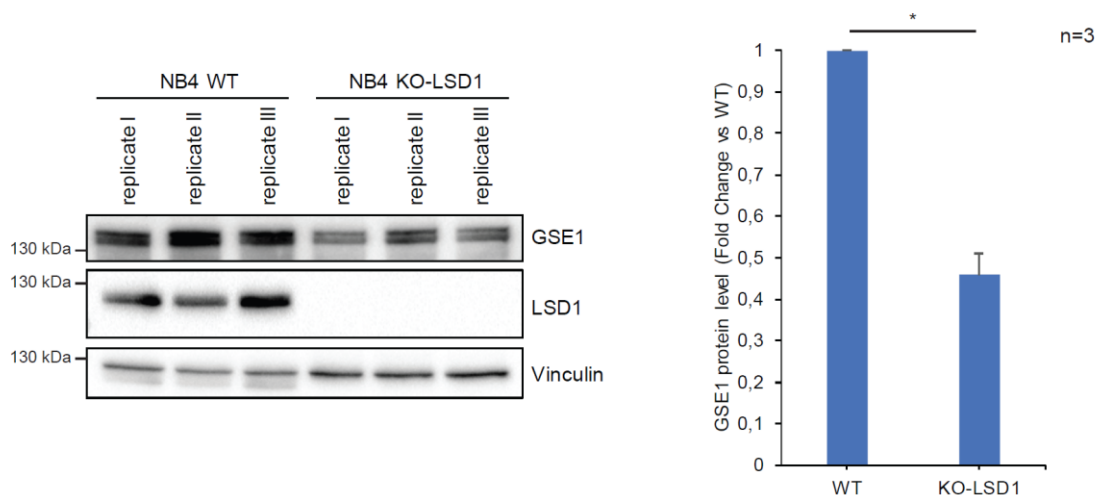


Figure 34: GSE1 protein level is reduced in NB4 KO-LSD1 cells compared to the WT. Left panel: Western Blot analysis of GSE1 in NB4-WT and LSD1-KO cells. Vinculin is used as loading control. Right Panel: The bar-graph shows GSE1 expression values normalized over the Vinculin in WT and LSD1-KO cells. The data are plotted as fold change of GSE1 protein level in the KO-samples compared to its control WT. Graph represents mean \pm SEM (n=3 biological replicates). Statistical significance is calculated by a paired-t test, *p value<0.05.

5.6. Dissecting the effects of GSE1 depletion in NB4 cells

In order to assess the effects of GSE1-depletion in NB4 cells and recapitulate, at least in part, the activity of the inhibitors, we knocked down GSE1 levels by using two different short hairpin RNAs (shRNAs). We transduced these shRNAs together with the negative control EV in NB4 cells and analysed both the phenotypic and molecular consequences of GSE1 depletion (Figure 35A). The two shRNAs modulated GSE1 with different efficiency. In particular, the shB2 reduced the amount of GSE1 more efficiently than the shA1 (Figure 35B). Collecting the WB results from 4 different biological replicates we obtained, on average, 80% of depletion with the stronger shRNA and 70% with the weaker one (Figure 35C).

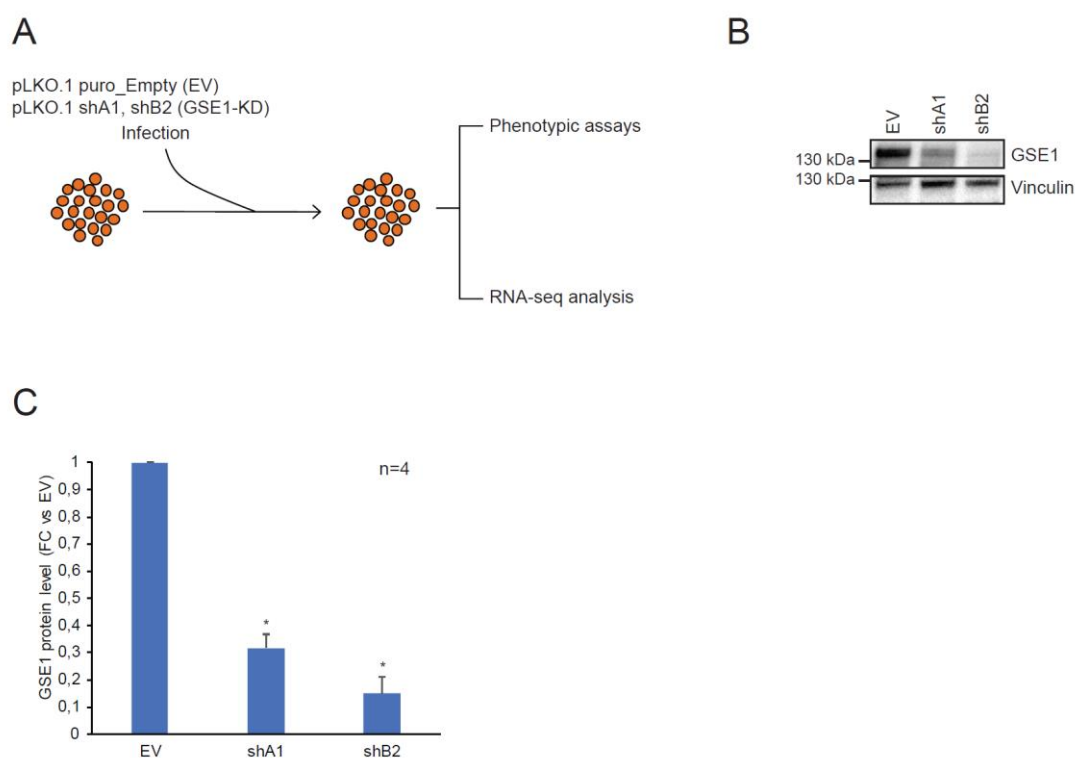


Figure 35: Experimental approach used to deplete GSE1 in NB4 cells. A) Schematic representation of the infection of NB4 cells with either the EV or two different shRNA targeting GSE1. After infection, phenotypic and molecular assays such as RNA-seq are performed. B) Western blot analysis of GSE1 in NB4 cells transduced with either EV or the shRNAs targeting GSE1. Vinculin is used as loading control. C) Bar graph displaying the quantitation results of GSE1 protein level normalized over the Vinculin in four different replicates, whereby the results are plotted as fold change of GSE1 protein level in the knock-

down (KD) samples compared to its control EV. Chart represents mean \pm SEM (n=4 biological replicates; Paired t-test, *p value<0.05).

5.6.1. Characterization of the phenotypic effects of GSE1 knock-down in NB4 cells

We then set to assess the effect of GSE1 knock-down (KD) in NB4 cell proliferation. We found that NB4 growth was strongly diminished by the reduction of GSE1 expression in a dose-dependent manner, with the shB2, which is the construct inducing the highest protein reduction, also displaying a stronger effect than the shA1 (Figure 36A). Alongside with cell growth reduction, cell death also displayed an increasing trend with time in GSE1-depleted cells compared to controls, with the death rate appearing as strictly related to the efficacy of the KD. Indeed, the percentage of cell death monitored after 72, 120 and 168 hours of infection was higher in cells transduced with the shB2 than the other construct. In particular, shB2 caused about 30% cell death 72 hours post infection and increased up to 60%-80% and 70%-90% at 120 and 168 hours, respectively. Instead, the shA1 led to only 10% cell death after 72 hours that tended increased up to 50%-60% at 168 hours (Figure 36B). These observations suggested that the decreased cell proliferation may be the consequence of an augmented cell death. To verify if cell mortality was due to the activation of the apoptotic cascade, we analysed by WB the levels of cleaved caspase-3 (activated caspase-3) ²²⁵ 96 hours post infection and detected the presence of cleaved-caspase 3 only in GSE1-KD cells. Moreover, cells transduced with shB2 also showed a reduction of the amount of total (not cleaved) caspase-3, indicating that the majority of this enzyme is in its active form in these cells (Figure 36C) and that prolonged GSE1 depletion impairs the NB4 cell vitality by activating the apoptotic process.

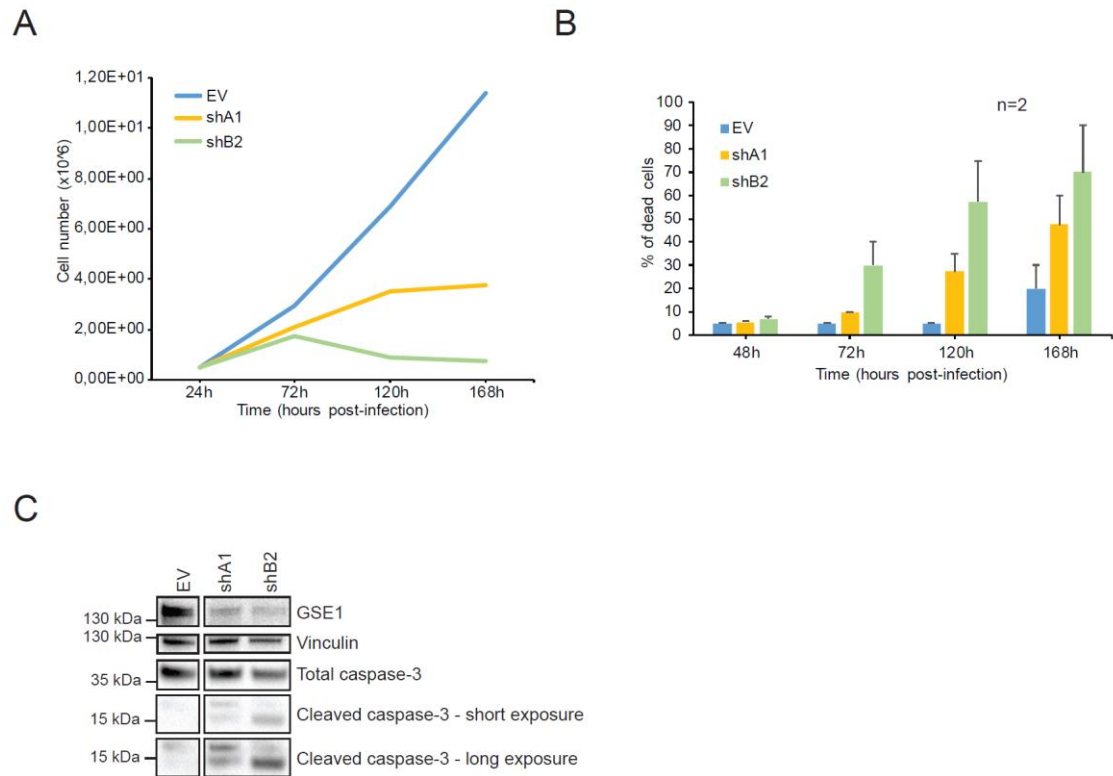


Figure 36: GSE1-KD induces reduction of cell proliferation and apoptosis in NB4 cells. A) Representative growth curve of NB4 cells transduced with either EV or 2 different shRNA targeting GSE1. B) Bar graph displaying the percentage of dead cells in control EV and GSE1-KD cells. The analysis is performed after 48, 72, 120 and 168 hours post-infection. Chart represents mean \pm SEM (n=2 biological replicates). C) Western Blot analysis of GSE1, cleaved caspase-3 and total caspase-3 in cells transduced with either control or shRNAs targeting GSE1 for 96 hours. Vinculin is used as loading control.

We know that both LSD1 inhibitors under study do not activate the apoptotic cascade in NB4 cells, but induce the cells to reactivate their differentiation process. Given the link between LSD1-inhibitor treatment and GSE1 gene expression reduction, we wanted to examine the effect of GSE1-KD in differentiation through the expression profiling of CD11b gene as a marker of myeloid differentiation. CD11b transcript was upregulated upon LSD1 inhibition and to a stronger extent by combination of MC2580 with physiological doses of ATRA (ATRA 0.01 μ M) (See Introduction, paragraph 2.6.1., Figure 7B). We investigated the induction of this marker at 48 and 72 hours of infection where the mortality was on average below 30% in all samples (Figure 36B) and observed a strong induction of CD11b gene in GSE1-KD cells compared to the EV, with a peak at 72 hours from shB2 transduction

(Figure 37A). We extended this analysis to other differentiation markers regulated by LSD1 inhibitors such as CD86⁸⁸ and IRF8²²⁶. Also these genes were upregulated in GSE1-depleted cells compared to EV after 48 hours from infection thus recapitulating the effect of the LSD1 inhibitors. The increased expression was stronger in cells transduced with shB2 compared with the shA1, as a probable consequence of the different reduction of GSE1 obtained with the two shRNAs (Figure 37B-C).

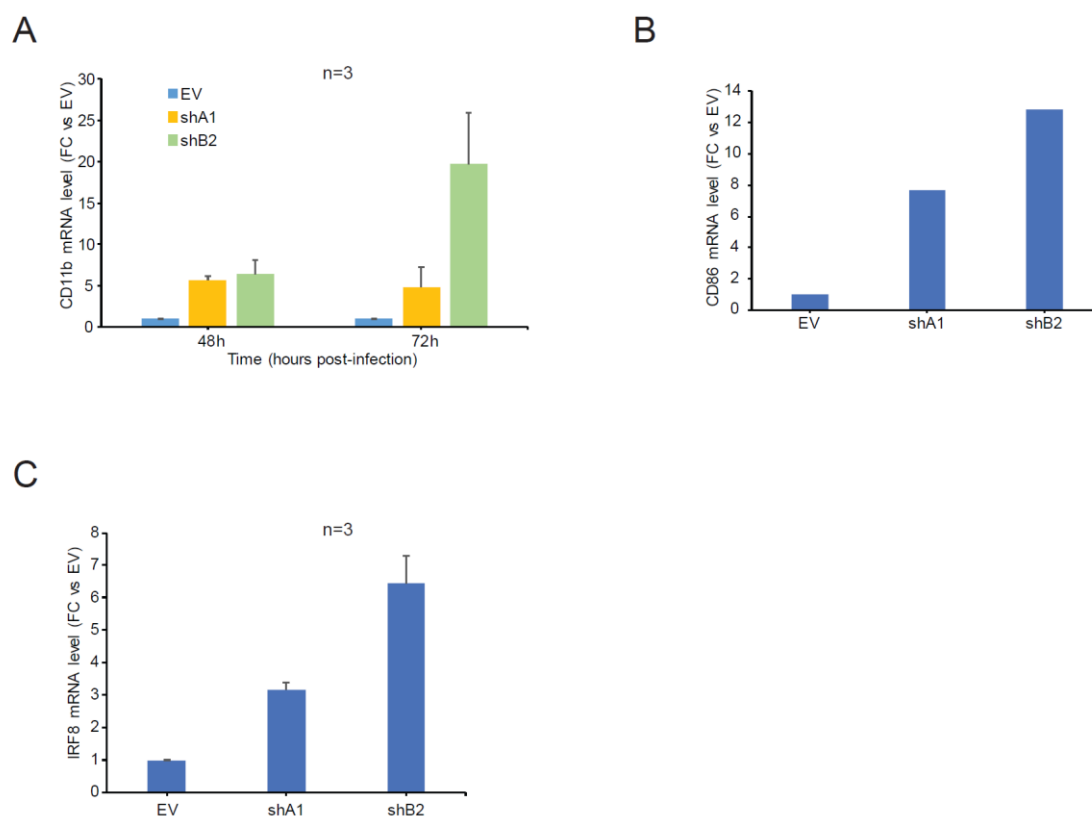


Figure 37: GSE1-KD induces the expression of differentiation markers in NB4 cells. A) RT-qPCR analysis of CD11b transcript level in NB4 cells upon 48 and 72 hours post-infection with shRNAs targeting GSE1 or EV as control. Ct values are normalized against GAPDH. The results are plotted as fold change of CD11b mRNA level in the KD samples compared to its control EV. Chart represents mean \pm SEM (n=3 biological replicates). B) Analysis of CD86 transcript in NB4 cells at 48 hours post-infection with shRNAs targeting GSE1 or EV as control. Ct values are normalized against GAPDH. The results are plotted as fold change of CD86 mRNA level in the KD samples compared to the control EV (n=1 biological replicate). C) RT-qPCR profiling of IRF8 transcript level in NB4 cells after 48 hours of infection with shRNAs targeting GSE1 or EV as control. Ct values are normalized against GAPDH. The data are plotted as fold change of IRF8 mRNA level in the KD samples compared to its control EV. Chart represents mean \pm SEM (n=3 biological replicates).

Myeloid differentiation process is usually associated with a reduction in the cell cycle progression^{227,228}. Hence, we assessed the possible effect of GSE1-KD on cell cycle progression by measurement of NB4 DNA content (with Propidium Iodide staining) by Fluorescence Activated Cell Sorting (FACS). Also in this case, we performed the experiments at earlier time points in respect to the progression of the apoptotic cascade and the increase of cell death, namely at 48h and 72h upon transduction. On average, shRNAs targeting GSE1 caused 10%-20% cell population increase in the G1-phase, followed by a corresponding decrease of their S-phase compared to the EV at both 48h and 72h post-infection (Figure 38A-B). We then assessed the expression of p21, the best-known cyclin-dependent kinase inhibitor (CKI)²²⁹ that controls the progression from G1 to S phase of the cell cycle^{230,231}, after 72 hours post-infection with shA1 and shB2 or the control EV. We found that the expression of p21 strongly increased in GSE1-KD cells compared to the control, corroborating the FACS data (Figure 38C). The increase of p21 was already evident at the transcript level upon 48 hours post transduction (Figure 38D), further suggesting that cell cycle arrest was induced by the activation of p21 signaling pathway. Overall, these results suggest that in myeloid blast cells GSE1 depletion at early time points causes differentiation process re-activation and reduction of cell cycle progression; at later time points induces apoptosis.

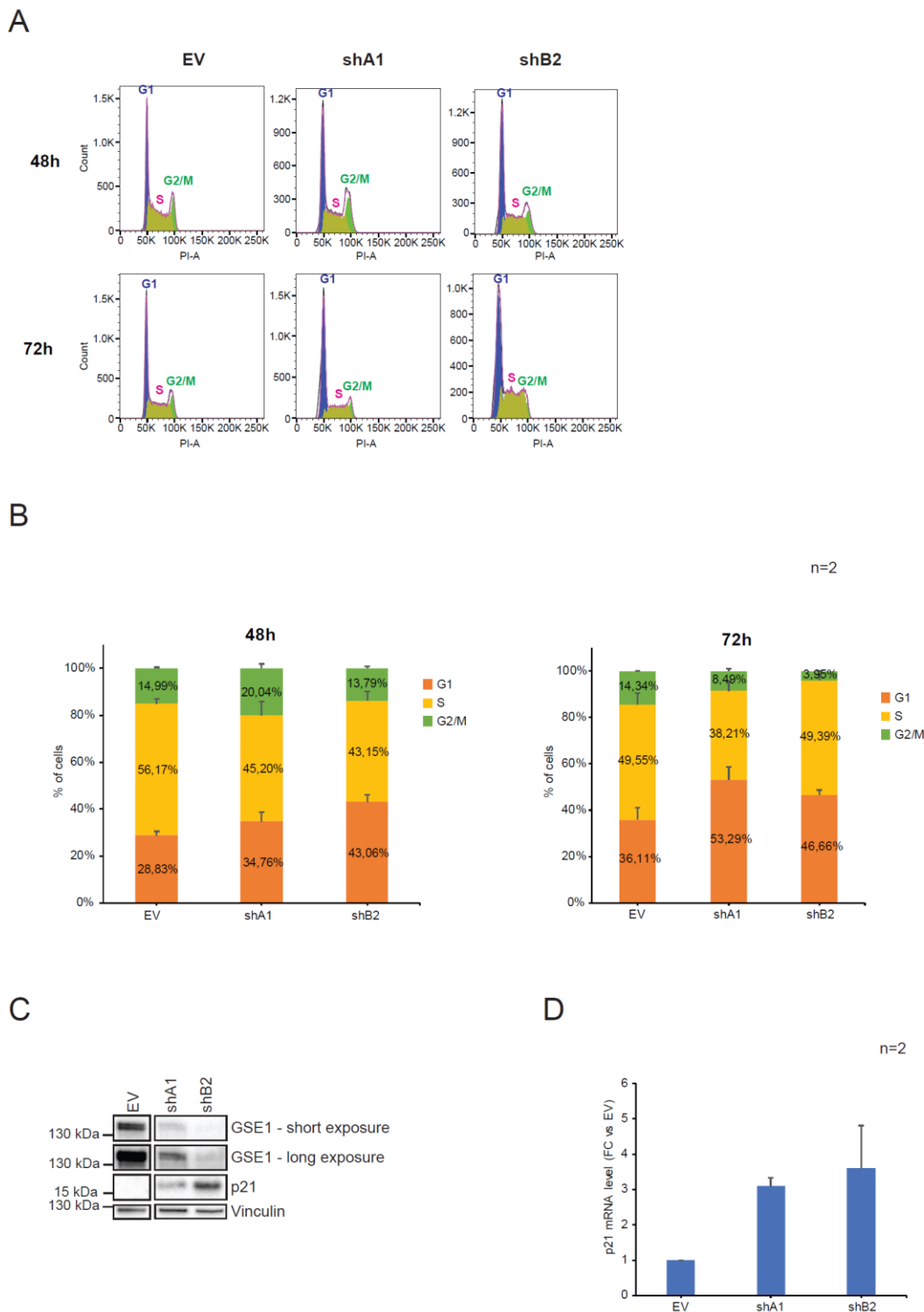


Figure 38: GSE1-KD determines increase in G1-phase and induction of p21. A) Representative FACS analysis of GSE1-KD and EV cells after 48 and 72 hours post-transduction. B) Percentage of GSE1-KD and control EV cells in each phase of the cell cycle at 48 and 72 hours upon infection. Chart represents mean \pm SEM (n=2 biological replicates). C) Western Blot analysis of GSE1 and p21 in NB4 GSE1-KD and EV samples after 72 hours of transduction with either shRNAs targeting GSE1 or EV as control. Vinculin is used as loading control. D) RT-qPCR profiling of p21 transcript level after 48 hours of GSE1-KD. Ct values are normalized against TBP. The data are plotted as fold change of p21 mRNA

level in the KD samples compared to its control EV. Chart represents mean \pm SEM (n=2 biological replicates).

5.6.2. RNA-sequencing analysis upon GSE1 depletion in NB4 cells

For a more in-depth evaluation of molecular mechanisms underlying the downregulation of GSE1 in NB4 cells, we performed RNA-seq analysis of cells transduced with the control EV and the shRNAs targeting GSE1 at 48 hours post-infection, in order to avoid potential side-effects due to the initiation of cell death.

NB4 GSE1-KD cells showed high differences in terms of gene expression compared to control cells, as visualized by their high distance in the first component (PC1) of the principal component analysis (PCA). Nevertheless, also cells transduced with shA1 and shB2 showed some differences among them, as depicted by their separation within the second component (PC2) of the PCA (Figure 39A). Given this difference, we analysed separately the differentially expressed genes (DEG) upon the individual shRNA infection and then extrapolated only the genes in common. Reassuringly, both shA1 and shB2 transduced cells displayed GSE1 down-regulation, with a \log_2 fold change (FC) value compared to the control EV of -0.77 and -0.78, respectively (Figure 39B). To extrapolate the DEG following GSE1-KD, we set the following filters: a gene was considered as differentially expressed when presenting an adjusted p-value ($p\text{-adj}$) < 0.01 and a \log_2 FC compared to EV more (in the case of up-regulation) or less (in the case of down-regulation) than 1. In both shA1 and shB2 transduced cells, we detected a greater number of up-regulated genes, which may be in line with the presence of this gene in transcriptional co-repressor complexes, like the BHC complex^{26,221}. In particular, in shA1- infected cells we found 720 up-regulated and 131 down-regulated genes, while in shB2- transduced cells we detected 999 up-regulated and 521 down-regulated genes. From the overlap of the DEG obtained from the two different GSE1-KD cells, we identified a total of 422 up-regulated and 47 down-regulated genes (Figure 39C).

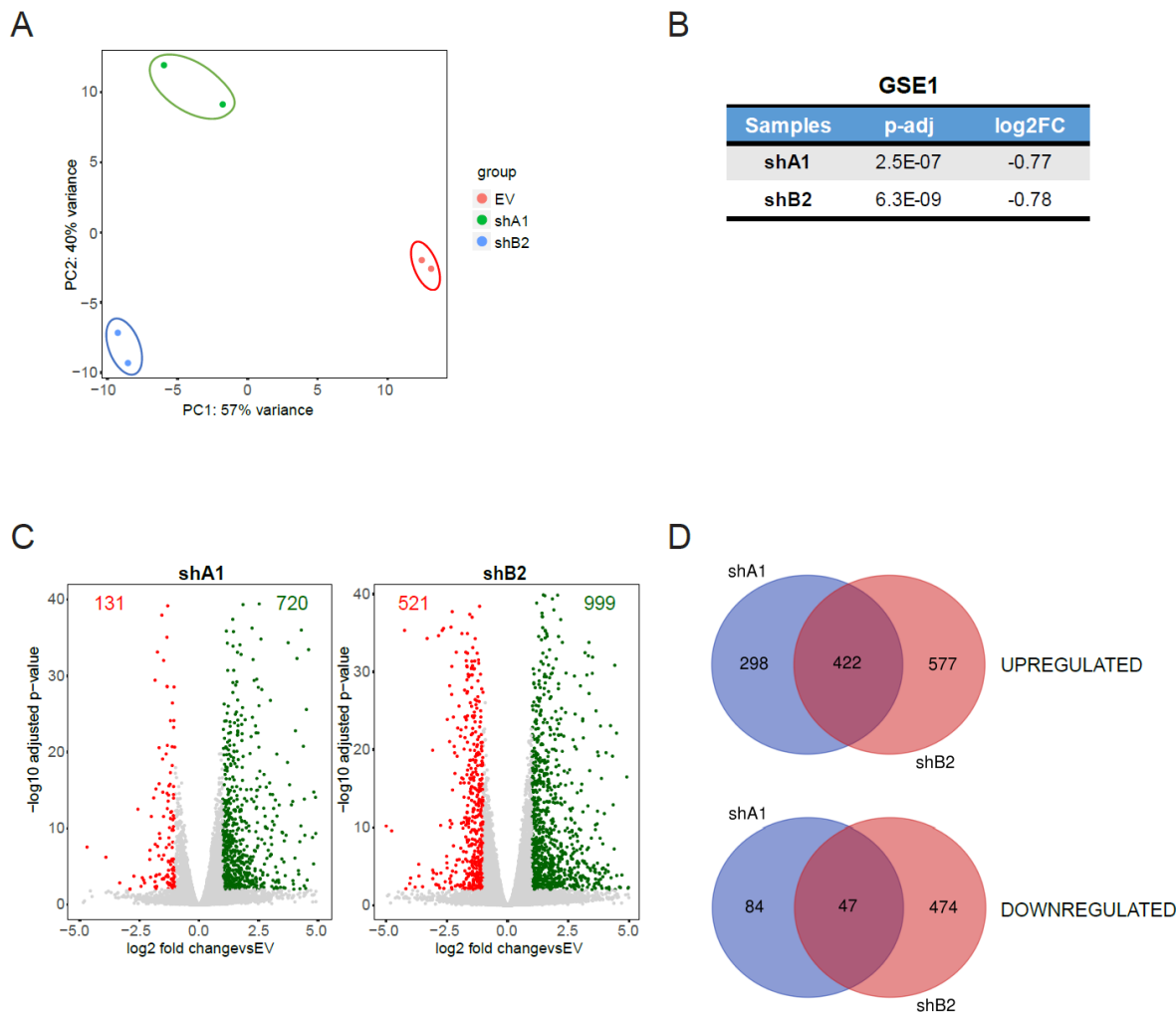


Figure 39: Differential gene expression analysis upon GSE1-KD in NB4 cells. A) Principal component analysis (PCA) of the EV, shA1 and shB2 samples. Each condition is characterized by two different biological replicates. B) Table displaying the \log_2 fold change (FC) and p-adjusted (p-adj) values for GSE1 transcript in shA1 and shB2 transduced cells. Fold change value is calculated with DEseq2 program using two biological replicates for each condition. C) Volcano plot displaying the up- and down- regulated genes upon shA1 and shB2 transduction. The x axis shows the \log_2 fold change values of each gene in the shRNAs-transduced cells compared to the control EV, while y-axis displays their $-\log_{10}$ p-adj values. Fold change value is calculated with DEseq2 program using two biological replicates for each condition. D) Venn diagrams with number of individual and overlapping putative up-regulated and down-regulated genes identified upon transduction with shA1 and shB2.

We performed a gene ontology (GO) analysis on both the common up-regulated and down-regulated genes extrapolated from the RNA-seq analysis. We did not obtain any statistically significant enrichment of biological processes in the down-regulated genes, while among the up-regulated genes we detected a significant enrichment of the following GO processes:

cytokine-mediated signalling, regulation of cysteine-endopeptidase activity involved in apoptotic process, activation of monocyte and dendritic differentiation, as well as of immune response (Figure 40A). Among the genes up-regulated within the category of cytokine-mediated signalling, we found some interferon-induced proteins (e.g. IFI35 and IFIT3), interferon-regulatory transcription factors (e.g. IRF5, IRF7 and IRF8), cluster of differentiation (CD) markers (e.g. CD11b, CD54 and CD86) as well as interleukin receptors (e.g. IL21R, IL12RB1).

Furthermore, we detected a substantial increase of some tumour necrosis factor (TNF) receptors, like TNFRSF10C, TNFRSF10B and TNFRSF10D as well as the TNF ligand TRAIL, all proteins involved in the activation of the apoptotic cascade.

Interestingly, we also found the up-regulation of a cluster of genes implicated in both monocytic differentiation such as CD74, IRF8 and CSF1 and dendritic differentiation like LILRB1 and LGALS9. In addition, other cellular processes like cell migration, angiogenesis and proliferation resulted as affected (Figure 40A).

The GO data were further corroborated by analysis on the enriched signalling pathway carried out by Reactome: we confirmed a substantial up-regulation of the cytokine- and interferon- signalling pathways. In addition, we detected the enrichment of integrin cell surface interactions mediated by cell adhesion molecules (CAMs), including CD11a, CD11b and CD31 that are associated to the process of myeloid differentiation. Moreover, other pathways associated to the activation of the innate immune system and apoptosis resulted up-regulated, such as the Toll Like Receptor 4 (TLR4) cascade and TRAIL signalling respectively. The first is an intracellular signalling mechanism that leads to activation of the transcription factor NF- κ B and inflammatory cytokine production ²³², while the second is based on the activity of the ligand TRAIL, which binds to the death receptors DR4 and DR5 thus activating a signalling cascade that induces apoptosis ²³³ (Figure 40B).

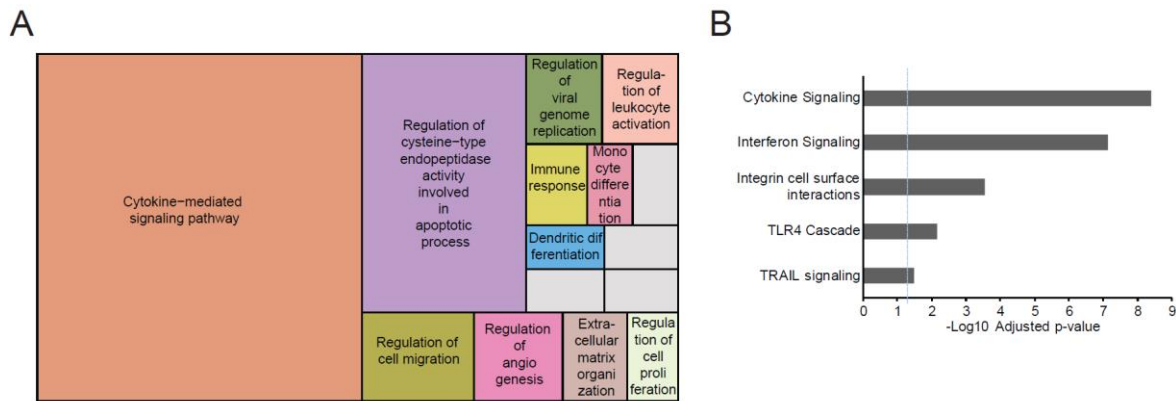


Figure 40: Gene Ontology (GO) analysis of the up-regulated genes upon GSE1-KD in NB4 cells. A) Tree-map displaying the statistically significant GO biological process (BP) terms of the common 422 up-regulated genes upon transduction with shA1 and shB2. GO analysis and calculation of the statistically significant BP is performed through EnrichR (Adjusted p-value < 0.05). Tree-map includes all BP terms mapped to a high hierarchical level. The size of the boxes corresponds to the number of genes in that category. The grey color in the plot include the following other BP terms: regulation of leukocyte degranulation, regulation of B cell proliferation, neutrophil-mediated immunity and regulation of leukocyte mediated cytotoxicity. B) Bar graph displaying the Reactome analysis of the significantly enriched pathways among the 422 up-regulated genes upon GSE1 depletion in NB4 cells. Reactome analysis and calculation of the statistically significant pathway terms is performed through EnrichR (Adjusted p-value < 0.05). The x axis shows the $-\log_{10}$ of the adjusted p-value for each significant term. The blue dashed line defines the threshold used to discriminate the significant enriched terms ($-\log_{10}$ p-adj > 1.3).

5.7. Comparison of the transcriptomic data upon GSE1-KD and LSD1 inhibition in NB4 cells

We also performed a global transcriptomic analysis at 24 hours of treatment with the LSD1 inhibitor MC2580. To analyse the DEG after drug treatment, we set the same filters applied to the previous RNA-seq experiment. Compared to GSE1-KD, the MC2580 produced much less variations in terms of gene expression. This was probably due to the stronger phenotypic effects associated with GSE1 down-regulation compared to LSD1 pharmacological inhibition in NB4 cells. Overall, MC2580 induced the up-regulation of 109 genes and the down-regulation of only 3 genes (Figure 41A). The higher number of up-regulated genes was probably associated with the main LSD1 function as transcriptional co-repressor in NB4 cells. Once established the DEG upon MC2580 treatment, we overlapped them with those

identified upon GSE1 depletion and found the common down-regulation of 1 gene and up-regulation of 45 genes among LSD1-inhibited and GSE1-KD cells (Figure 41B). GO analysis of the biological processes enriched among the 45 up-regulated genes allowed identifying two different functional categories: “cytokine-mediated signalling pathway” and “regulation of cysteine-endopeptidase activity involved in the apoptotic process” (Figure 41C). This result together with our discovery that LSD1 inhibition regulates the expression of GSE1 could indicate the presence of a LSD1-GSE1 regulatory axis controlling the expression of genes belonging to these pathways. An in-depth analysis of these 45 up-regulated genes indicated that most of them are implicated in immune response, such as ICAM1, ICAM4, IFI16 and IFIT3. This is in agreement with the significant enrichment of the “cytokine-mediated signalling pathway” category, as one of the mechanisms associated with immune system activation²³⁴ (Figure 41D). Some of the genes belonging to this group are also up-regulated during differentiation, like CD11b, IRF8 and CD86 thus confirming the previous results obtained by RT-qPCR regarding the increase of their expression upon GSE1 KD (Figure 37) and MC2580 treatment. We also detected the increase of genes implicated in other processes like cytoskeleton organization, intracellular signal transduction, ion transport and metabolism (Figure 41D). Overall, our transcriptomic results seem to suggest the presence of a set of genes/pathways regulated by the LSD1 inhibitor in NB4 cells through the reduction of GSE1 expression.

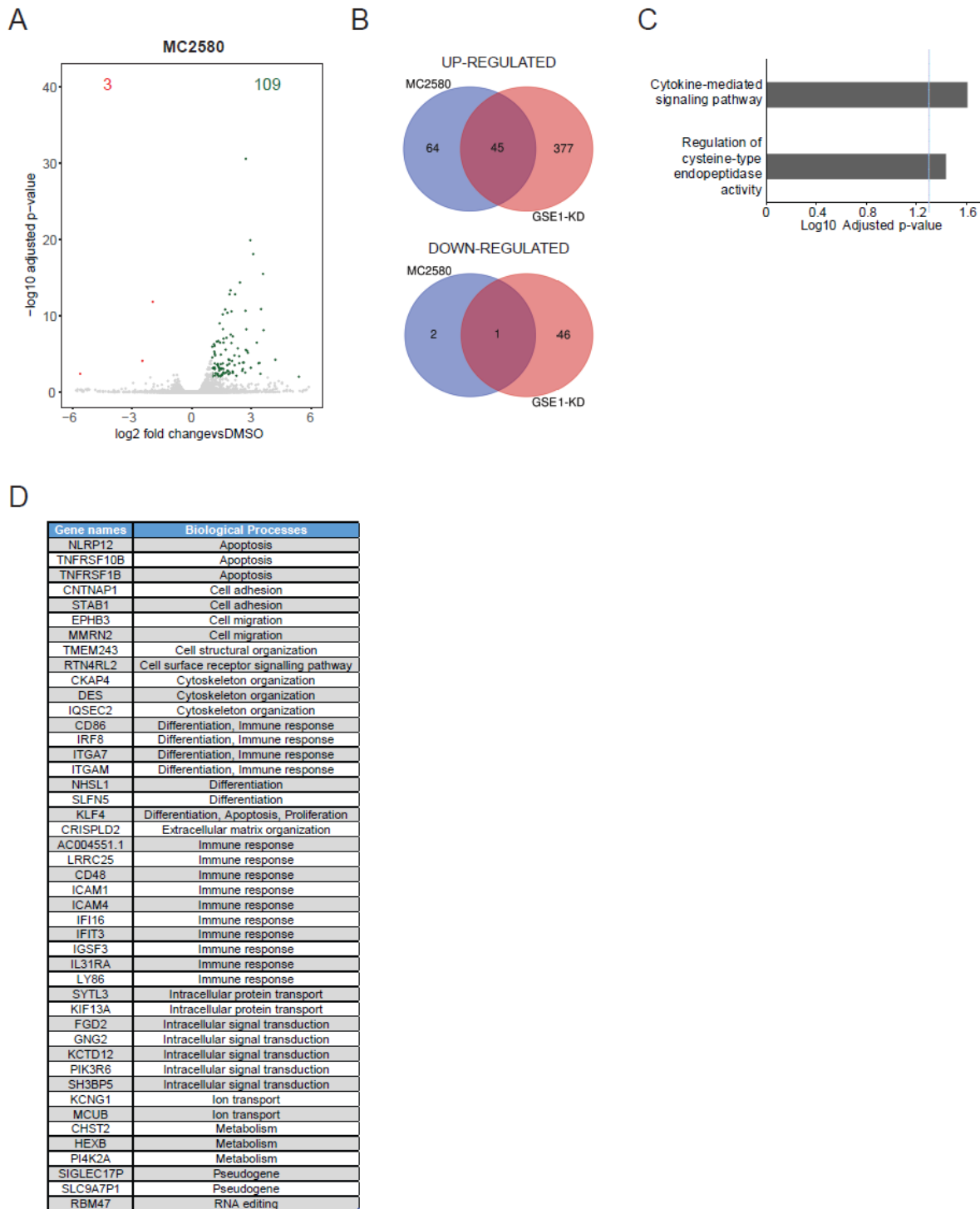


Figure 41: Analysis of the common DEG upon LSD1 inhibition and GSE1-KD in NB4 cells. A) Volcano plot displaying the up- and down- regulated genes upon 24 hours treatment with MC2580. The x axis shows the \log_2 fold change (FC) values in the LSD1-inhibited cells compared to the control DMSO, while y-axis displays the $-\log_{10}$ of the adjusted p-value (p-adj) of each gene. Fold change value is calculated with DEseq2 program using two biological replicates for each condition. B) Venn diagrams with number of individual and overlapping up- and down- regulated genes identified upon MC2580 treatment and transduction with shRNAs targeting GSE1. In the case of GSE1 KD, the common set of DEG genes with both shRNAs are overlapped with those extrapolated from MC2580 treatment. C) Bar graph displaying the GO analysis of the significantly enriched biological processes (BP) of the 45 commonly up-regulated genes upon GSE1-KD and LSD1 inhibition. GO analysis and

calculation of the statistically significant BP terms is performed through EnrichR (Adjusted p-value < 0.05). The x axis shows the $-\log_{10}$ of the p-adj for each significant term. The blue dashed line defines the threshold used to discriminate the significant enriched terms ($-\log_{10}$ p-adj > 1.3). D) Table displaying the gene names and associated biological processes of the 45 genes commonly up-regulated upon GSE1-KD and MC2580 treatment.

5.8. MS-based profiling of histone PTMs in NB4 cells upon LSD1 inhibition

Given the major role of LSD1 as histone-modifying enzyme, we set to investigate also the effect of the LSD1 inhibitors on bulk histone PTMs, including not only the known targets H3K4me1/me2 and H3K9me1/me2, but also all other major acetylations and methylations occurring on core histones. Although the de-methylase activity of the enzyme has emerged as dispensable in our model system, we reasoned that LSD1 can still maintain its physical and/or functional interactions with various histone modifiers, such as HDAC1 and 2; thus, the comprehensive profiling of bulk modification levels in response to the pharmacological inhibition of LSD1 could reveal novel downstream chromatin-linked processes mediated by the enzyme and modulated by the respective drugs. Hence, we decided to employ quantitative MS that has been shown to be an efficient approach for the unbiased and global analysis of combinatorial PTM patterns on histones^{235,236}. To this aim, we treated NB4 cells with the LSD1 inhibitor DDP-38003 for 24 hours, the same time interval that we had used in the interactomics experiments. After cell harvesting, histones were extracted with hydrochloric acid (HCl) prior to protease digestion and LC-MS/MS analysis. Histone modifications were quantified by a label-free approach by measuring the extracted ion intensity of each modified histone peptide identified and then calculating their percentage relative abundance (%RA)¹⁵⁴ (See Introduction, paragraph 2.11.2 and Figure 14) (Figure 42).

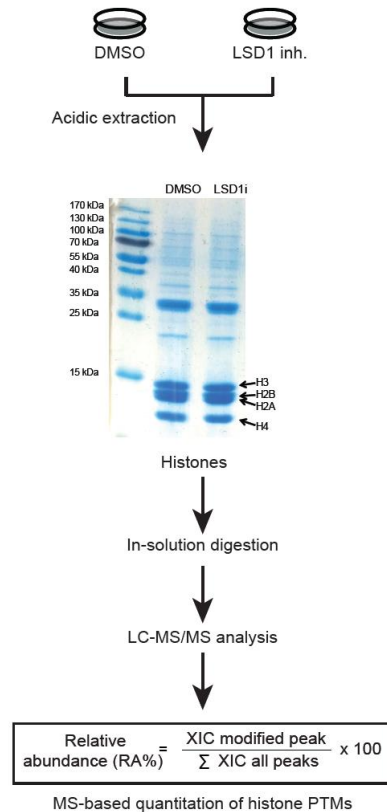


Figure 42: Experimental strategy to analyse histone PTMs in NB4 cells following 24h treatment with the inhibitor DDP-38003. NB4 cells are treated for 24 hours with either DMSO or DDP-38003 (2 μ M). Histones are then extracted by hydrochloric acid from nuclear lysates and in-solution digested with the “Pro-PIC” method (See Material and Methods, paragraph 4.17.1.) prior to LC-MS/MS analysis. To quantify each differentially modified histone peptide, calculation of the percentage relative abundance (%RA) is used.

The histone PTM profiling showed that the pharmacological inhibition of LSD1 caused a slight increase not only of H3K4me2, as expected, but also of H3K27me2, H3K27me3, H3K9me2 and H3K9me3, alone or in combination with H3K14ac, which were mirrored by a decrease of H3K27me1 in combination with H3K36me1. This indicates that, in NB4 cells, DDP-38003 determines an increase of histone PTMs typically associated with gene silencing, while histone marks associated with active chromatin regions are either unchanged or slightly reduced. In support of this evidence, we also detected a slight decrease of the histone marks H3K27ac, H3K18ac and H3K4ac.

To corroborate the evidence that these changes are due to the blockage of LSD1 activity, we compared these results with the data collected with a similar MS- analysis of bulk histone

modifications in two distinct NB4 KO clones of LSD1 (KO1 and KO2) and the corresponding WT cells. We observed that some histone marks displayed the same variations in both LSD1 drug-inhibited and KO cells, confirming that the molecular mechanism underpinning their changes may be indeed ascribed to LSD1 activity. Specifically, H3K27me3 and H3K4me2 showed reproducibly an increasing trend, while H3K27me1 combined with H3K36me1 decreased (Figure 43A). The same changes were further validated by using the LSD1 inhibitor MC2580 (Figure 43B).

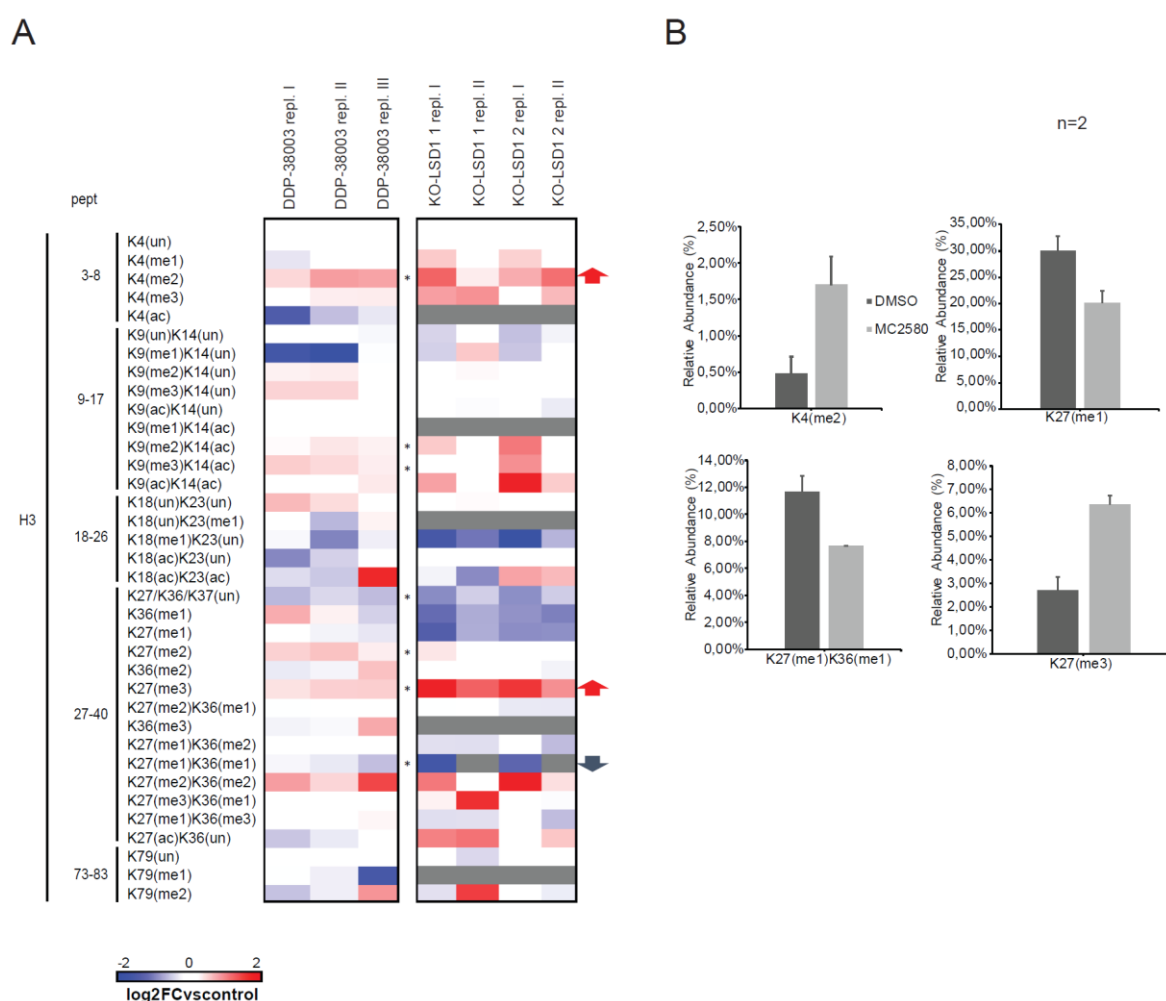


Figure 43: MS-based profiling of histone PTMs upon genetic and pharmacological LSD1 inhibition. A) Heatmap summarizing the changes in bulk histone modification levels obtained from 3 biological replicates of NB4-treated cells for 24 hours with the inhibitor DDP-38003 (2 μ M) (n=3), and two biological clones of NB4 LSD1-KO cells (n=2). For each modification, the results are plotted as log₂ fold change of the %RA of the treated and KO sample versus their respective controls: DMSO and WT. Statistical significance is calculated for the comparison between untreated and treated cells with the DDP-38003 by a paired t-test, with n=3 *p value<0.05. Asterisks indicate modifications significantly different in the treated samples compared to the control ones. Arrows point out modifications that show a

similar trend upon genetic depletion and pharmacological inhibition (red: up-regulation, blue: down-regulation). C) Bar graphs displaying the %RA of a panel of histone modifications in control and treated NB4 cells with the inhibitor MC2580 (2 μ M) for 24 hours. Graph represents mean \pm SEM (n=2, biological replicates).

5.9. MS-based analysis of histone PTM changes in UF-1 cells upon LSD1 inhibition

Once established the histone PTM changes upon treatment with the LSD1 inhibitors in NB4 cells, we investigated whether the same alterations were detected in UF-1 cells, another APL model that displays more sensitivity to these drugs than NB4 cells (See Introduction, paragraph 2.6., Figure 6). We treated UF-1 cells for 24 hours with either MC2580 or the control DMSO and, then, proceeded with the same workflow described in Figure 42. To improve the quantitation of each histone modifications, we employed a spike-in SILAC strategy, whereby heavy-labelled histones extracted from SILAC-labelled NB4 cells were spiked in at a fixed concentration in both treated and control samples and then the heavy-labelled counterpart of each modified histone peptide was used to normalize the levels of each modification across the samples (See Introduction, paragraph 2.11.2., Figure 15)¹⁷⁷.

This histone PTM analysis demonstrated that UF-1 had a different epigenetic response to the treatment with the LSD1 inhibitor compared to NB4, which could be linked to their different phenotypic responses to the drugs. Apart from the LSD1-histone target H3K4me2, the other histone PTMs that changed in NB4 cells were unchanged or slightly depleted in UF-1, such as the H3K27me2/me3 and the H3K9me2/me3 alone or in combination with the H3K14ac. On the contrary, we detected a very mild but significant increase of H3K36me3, a mark that is usually linked to active transcription (Figure 44A-B). These data are in line with the notion that the increase of H3K36 methylation is antagonistic to the activity of the PRC2 complex and its deposition of the H3K27me2/me3 marks, and vice-versa²³⁷⁻²⁴⁰.

On the one hand, the result that globally H3K4me2 increased to a similar extent both in cells that are sensitive and resistant to these inhibitors corroborates the concept that the phenotypic

and cellular effect of these drugs in APL models are not directly dependent on the catalytic activity of the enzyme, but rely on other functions associated to this protein. On the other hand, the differential response of these two cell lines to the drug may be interpreted epigenetically in light of the differential changes in other histone PTMs here observed.

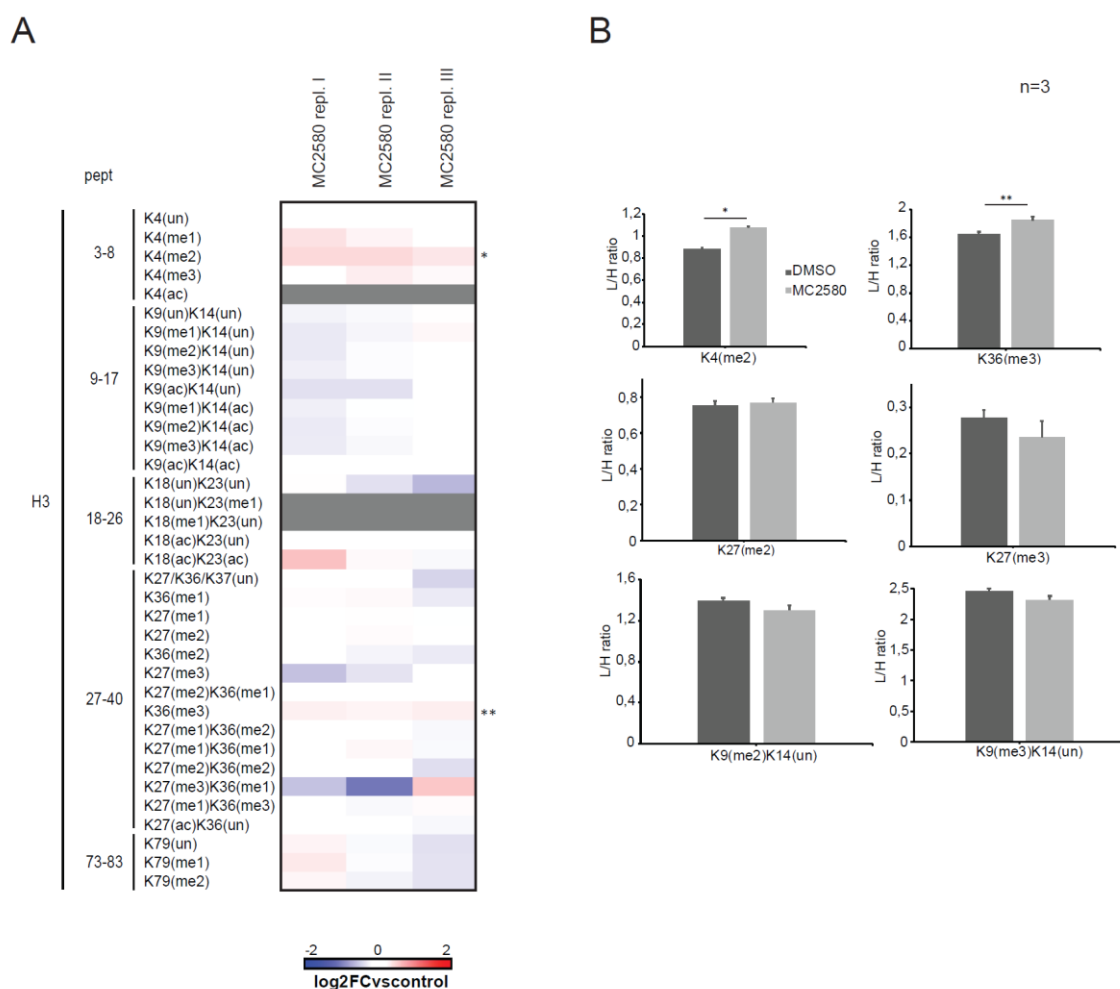


Figure 44: MS-based analysis of histone PTMs upon 24 hours treatment with MC2580 in UF-1 cells. A) Heatmap displaying the changes in histone PTMs obtained from three biological replicates of UF-1-treated cells for 24 hours with the inhibitor MC2580 (2 μ M) (n=3). Each modification is normalized over its heavy counterpart obtained from NB4 heavy-labelled histones. For each modification, the results are plotted as log₂ fold change of the L/H ratio in the treated samples over their corresponding controls. Statistical significance is calculated by a paired t-test, with n= 3 *p value<0.05, **p value<0.01. B) Bar graphs displaying the L/H ratio of a panel of histone modifications in control and treated NB4 cells with the inhibitor MC2580 (2 μ M) for 24 hours. Graph represents mean \pm SEM (n=3, biological replicates).

5.10. Comparative analysis of histone PTMs in APL cells with different sensitivity to LSD1 inhibitors

The previous observation that the histone PTM changes triggered by the drugs were different in NB4 and UF-1 cell lines prompted us to hypothesize that a different epigenetic background of the cells at the basal level could explain their differential response to the inhibitors. To this aim, we profiled a panel of histone PTMs and compared their abundance in untreated NB4 and UF-1 cells. We detected a substantial difference in the epigenetic patterns of the two cells: UF-1 cells displayed significantly lower levels of the H3K27me2/me3 and higher levels of H3K9me2/me3, H3K79me2 and H3K27me1, alone or in combination with H3K36me1 compared to NB4 cell lines (

Figure 45A).

Interestingly, a similar MS- analysis of histone PTMs, carried out (in collaboration with Prof. G. Pelicci's group) on a panel of glioblastoma (GBM) neurospheres displaying differential responsiveness to LSD1 inhibition, showed that H3K27 methylation levels correlated to the different sensitivity to the LSD1 inhibitors similarly to what we observed in APL cells. In particular, also GBM sensitive neurospheres showed higher levels of H3K27me1 in combination with H3K36me1, mirrored by lower levels of H3K27me2/me3 than GBM resistant neurospheres (

Figure 45B).

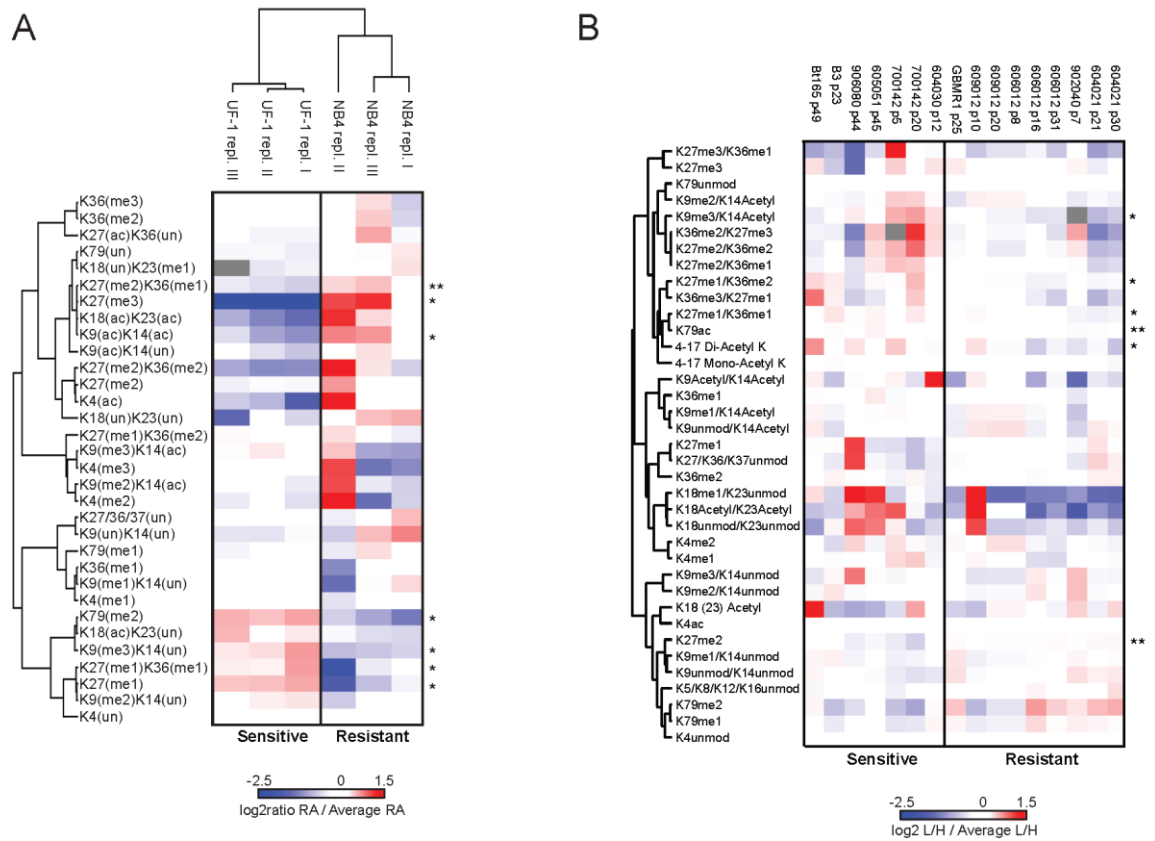


Figure 45: Comparative analysis of histone PTMs in sensitive and resistant APL and GBM models. A) Unsupervised hierarchical clustering displaying the log₂ of the %RA of a panel of histone PTMs obtained from three biological replicates of NB4 and UF-1 cells (n=3, biological replicates). %RA of each modification is normalized over the mean value across the samples are shown. Statistical significance is calculated by a paired t-test, *p value<0.05, **p value<0.01. B) Unsupervised hierarchical clustering showing the differences in the histone PTM abundance between sensitive and resistant glioblastoma neurospheres to LSD1 inhibition. RA of each modification is divided over the RA of its heavy counterpart obtained from a mix of heavy-labelled breast cell lines^{181,182}. L/H ratios are normalized over the mean value across the samples are displayed. Statistical significance is calculated by a paired t-test, *p value<0.05, **p value<0.01.

5.11. Combinatorial effect of LSD1 and EZH2 inhibitors on NB4 cell proliferation

The previous histone PTM analyses pointed towards a possible role of H3K27me2/me3 and hence the related enzymes in the response of the NB4 cells to LSD1 inhibition and in its resistance to pharmacological treatment. Therefore, in collaboration with Prof. S. Minucci's group, we tested the combination of LSD1 and EZH2/1 (the members of the Polycomb Repressive Complex 2 responsible for the deposition of the H3K27 methylation) inhibition in NB4 cells. NB4 cells were pre-treated with either DMSO or the EZH2/1 inhibitor

UNC1999¹⁸⁹ for 96 hours in order to obtain the complete removal of the H3K27me3 mark from the chromatin (Figure 46A). Afterwards, we added the MC2580 to both conditions and analysed cell proliferation for a total of 6 days. While the two epi-drugs alone had a minor effect on NB4 cell proliferation, their combination led to a strong reduction of cell growth. This result suggests that, in the absence of EZH2/1 activity, pharmacological inhibition of LSD1 has much stronger effect on NB4 cell growth, further supporting our hypothesis that the increase of this modification confers resistance to LSD1 inhibition (Figure 46B).

These results provide novel relevant data about the mechanisms of action of the LSD1 inhibitors in APL models, and possibly also in other tumour models, and pave the way to testing combinatorial treatments in those tumour types more resistant to LSD1 inhibition.

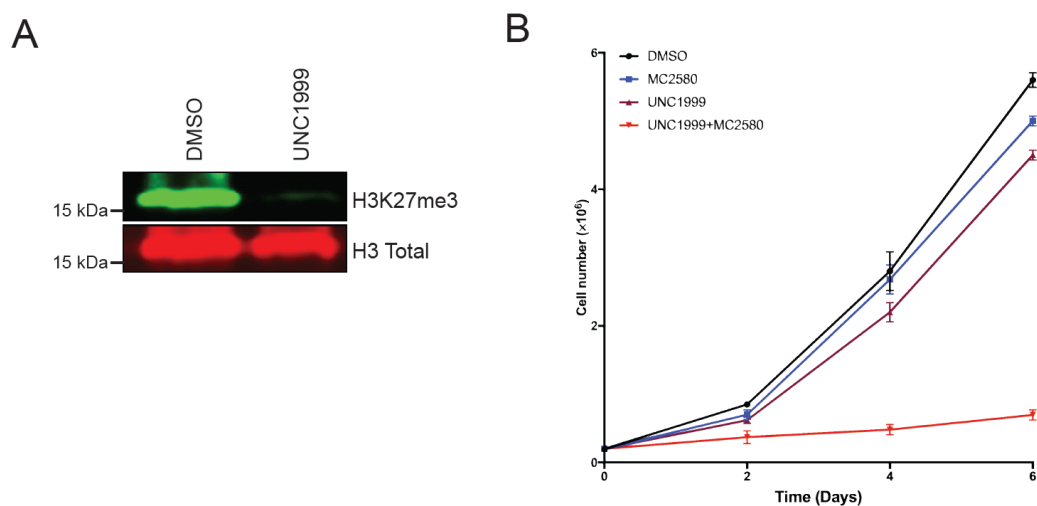


Figure 46: Combination of LSD1 and EZH1/2 inhibition strongly affects NB4 cell proliferation. A) Western Blot analysis of H3K27me3 in NB4 cells treated with 1 μ M of EZH2/1 inhibitor and control DMSO for 96 hours. H3 total is used as loading control. Anti-rabbit Alexa680 and anti-rabbit Alexa800 were used as secondary antibodies (i.e. fluorescent antibodies) and the signals were acquired using the LI-COR Odyssey v3.0. B) Cell growth of NB4 cells treated with MC2580 (2 μ M) following 96 hours of pre-treatment with either UNC1999 (1 μ M) or DMSO. As further experimental controls, cells are also maintained upon treatment with UNC1999 and its control DMSO for the analysis of cell proliferation. Chart represents the mean \pm standard deviation of two different biological replicates.

6. DISCUSSION

In this study, we employed different quantitative MS-based approaches to contribute to the understanding of the mechanism of action of the LSD1 inhibitors MC2580 and DDP-38003 in NB4 leukemia cells. Specifically, we assessed the effects of these drugs on both the LSD1 interactome and global histone PTM changes.

The preliminary observation that the catalytic activity of LSD1 is dispensable for the induction of the phenotypic effect of these drugs in the cellular system under investigation suggested that these molecules could act through a mechanism independent from the inhibition of the LSD1 enzymatic activity and could instead operate through the alteration of the LSD1 binding with some of its interactors. In support of this hypothesis, recent studies reported that some tranylcypromine-derivative LSD1 inhibitors can impair cell growth in SCLC and induce differentiation in AML by inhibiting the interaction between LSD1 and the interactors GFI1, GFI1B and INSM1^{89,104,105}. However, for the first time we assessed this mechanism with a global, and not candidate-based approach.

We initially combined SILAC-based proteomics with LSD1-IP to determine the list of LSD1 putative interactors and then, assess the changes of their binding to LSD1 upon drug treatment. The setup of the IP-MS strategy required the optimization of some steps, such as: the SILAC labelling of the cells, the appropriate IP conditions including the amount of blocking peptide to use in combination with the antibody as negative control for the IP. From the SILAC-LSD1 IP at basal state, we detected the enrichment of known LSD1-chromatin complexes, like the CoREST, the NuRD, the BHC complexes and the ALL-1 supercomplex. The CORUM analysis of our list of interactors revealed novel complexes in which LSD1 may be involved in NB4 cells. Most of them are implicated in transcription and mRNA-processing such as spliceosome, polyadenylation and RNA-Pol II complexes. Interestingly, HDAC1 and 2 (known LSD1 interactors) bind to some splicing factors, regulating directly this process^{212,213}. In addition, HDAC1 interacts with proteins of the polyadenylation

complex, thus controlling this other mechanism ²⁴¹. Since LSD1 is a direct component of various HDAC-containing complexes, the novel interactions found may suggest a functional role of this enzyme in these cellular processes. Nevertheless, the AP-MS approach does not immuno-precipitate only direct interaction of the bait but also various indirect ones; hence, we cannot exclude that the presence of these proteins in our interactomic list is a consequence of secondary interactions.

Instead, the presence of various subunits of the RNA-pol II complex may be in line with its function as transcriptional co-regulator. In fact, despite being mainly associated to co-repressor functions, LSD1 can act as transcriptional co-activator in some loci, by de-methylating H3K9me2/me1 and facilitating the recruitment of the RNA-pol II to activate transcription ²⁴². Interestingly, the MS-based histone PTM profiling upon drug treatment detected a slight increase of bulk H3K9me2/me3 levels, both alone and in combination with H3K14ac in LSD1-inhibited cells. This may suggest the presence of specific loci where LSD1 de-methylates a silencing mark and recruit some subunits of the transcriptional machinery, to start transcription.

By analysing the changes in the LSD1-interactome upon drug treatment, we found that the inhibitors alter the interaction of LSD1 with some of its associated proteins, in particular GFI1, GSE1 and EDC4, while most interactors were not affected by the drug treatment. According to the literature, most of the interactions are known to occur in the Tower Domain ⁵, while these inhibitors bind to the catalytic pocket of the enzyme, located within the Amino-Oxidase Like (AOL) domain; this may explain while only few interactors are indeed affected by the presence of the drugs. Also, we observed that the altered interaction of GFI1 and GSE1 to LSD1 is independent from the inhibition of the LSD1 catalytic activity, suggesting that other mechanisms are involved. The paper of Maiques-Diaz *et al.*, that was published when we had already acquired most of the data about the functional implication of the disruption of the LSD1-GFI1 interaction, demonstrated that the LSD1 inhibitor OG86

physically disrupts the LSD1-GFI1 interaction ⁸⁹. We showed that the inhibitor MC2580 behaves similarly to OG86, since it is able to block the LSD1-GFI1 interaction both *in-vitro* and *in-vivo*. However, unlike OG86, MC2580 can also dissociate GFI1 from the catalytic inactive form K661A-LSD1. This could be a relevant information from a therapeutic point of view, since the use of this drug would allow overcoming potential resistances due to mutations in the LSD1 catalytic site. This unique feature can be explained in light of the three-dimensional structure of the enzyme-inhibitor and enzyme-SNAIL1 complexes: SNAIL-1 binds to LSD1 via the SNAG domain, like GFI1. From the structural analysis emerged that the binding site to LSD1 of the inhibitor and the SNAIL1 protein fully overlap, so that the drug could cause steric hindrance preventing the binding to proteins containing the SNAG-domain ^{116,243}.

Regarding the phenotypic effects of the LSD1-GFI1 interaction, we demonstrated that the eviction of GFI1 sensitizes cells to physiological doses of ATRA by re-activating the myeloid differentiation process. Maiques-Diaz *et al.* found the presence of similar phenotypic effects in THP-1 cells (another AML model) following the disruption of this binding ⁸⁹. A study published in 2007 revealed that the physical association of LSD1 with the N-terminal SNAG domain of GFI1 is essential for the function of GFI1 as transcriptional repressor in haematopoietic cells ²³. Therefore, one of the mechanisms through which our LSD1 inhibitors favour the differentiation process is the block of the repressive activity of GFI1 by altering its binding to LSD1.

GSE1 was previously described to interact with LSD1 in both erythroleukemia ²²², ovarian ²²¹ and prostate cancer cells ⁵⁸. By both MS and WB analyses, we confirmed this binding in NB4 cells and found that it is altered upon treatment with the inhibitors MC2580 and DDP-38003. The mechanism through which the drugs modulate this interaction is not completely clarified. Currently, we demonstrated that the interaction between the two proteins is retained *in vitro* independently of the drug treatment and that levels of GSE1 are reduced

upon treatment at both the protein and mRNA levels. Altogether these experiments prove that LSD1 regulates, directly or indirectly, GSE1 transcription. ChIP-seq profiling of the histone modifications H3K4me1, H3K4me2, H3K4me3 and H3K27ac was carried out by the group of Prof. Saverio Minucci in NB4 cells and allowed to design primers amplifying genomic regions within *cis*-regulatory elements of GSE1 (i.e. promoter and putative enhancers). ChIP-qPCR of these histone modifications in cells untreated and treated with the LSD1 inhibitors may allow identifying potential epigenetic changes on GSE1 *cis*-regulatory regions following drug treatment. In addition, ChIP of the H3K27me3 may also be informative, given the relative global increase of this modification that we observed when LSD1 is inhibited.

To analyse the phenotypic and molecular effects of GSE1 down-regulation in NB4 cells, we transduced NB4 cells with two shRNAs targeting GSE1 and assessed cell proliferation and differentiation. We thus discovered that GSE1 depletion reduces NB4 cell proliferation by inducing apoptosis. However, at early time points, GSE1-KD induces the increase of various differentiation markers, such as CD11b, CD86 and IRF8, as well as cell cycle arrest by enhancing p21 expression. Increase in the expression of these genes was demonstrated by both RNA-seq and RT-qPCR analysis. Induction of p21 could cause the re-activation of the myeloid differentiation process: indeed, cell cycle arrest is the first step towards terminal differentiation²²⁸. In line with our results, Santos *et al.* found that increase of p21 upon DNA-damage leads to terminal differentiation of MLL-AF9 blasts²⁴⁴. The amount of DNA damage is critical for the choice between DNA repair or apoptosis activation. Low DNA damage leads to p21-dependent inhibition of apoptosis allowing DNA repair, while elevated DNA damage determines apoptosis²⁴⁵, probably by cleavage of p21 by caspase-3²⁴⁶. In our context, the apoptotic process observed at later time points upon GSE1-KD could be triggered by the presence of high levels DNA damage that cannot be repaired. Expression of some DNA damage markers like gamma-H2AX as well as comet assay will allow assessing

the potential presence and amount of DNA damage upon GSE1-KD and eventually providing answers to this hypothesis.

Functional analysis of RNA-seq data upon GSE1-KD confirmed the previous observation on the induction of apoptosis and differentiation markers when GSE1 is depleted. Indeed, we detected the upregulation of genes involved in different pathways associated with these processes, such as “TRAIL signalling” and “integrin cell surface interactions”. Furthermore, we found the up-regulation of genes specifically involved in monocytic differentiation lineage, like CD74, CSF1, CD86 and IRF8. The last two genes were also up-regulated upon LSD1 inhibition. Preliminary data collected by the group of Prof. S. Minucci demonstrated that myeloid markers associated with the monocytic/macrophagic lineage like IRF8 and CD14 were specifically up-regulated in NB4 cells upon treatment with MC2580 alone and in combination with physiological doses of ATRA (0.01 μ M), but not with ATRA alone (Figure 47).

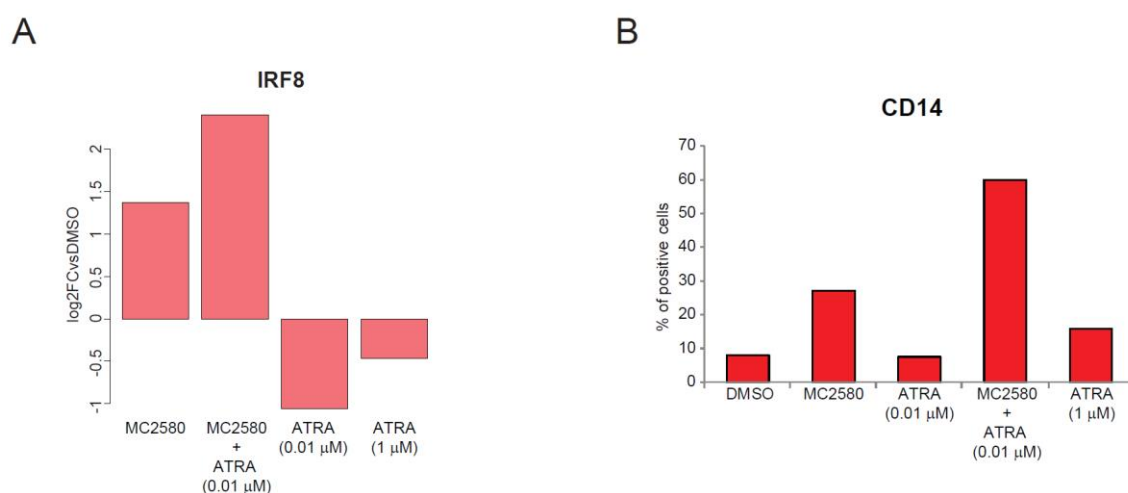


Figure 47: MC2580 but not ATRA increases the levels of the monocyte/macrophage markers IRF8 and CD14. A) IRF8 transcript level measured by RNA-seq analysis in NB4 cells treated for 24 hours with MC2580 (2 μ M), ATRA (0.01 μ M), MC2580 (2 μ M) + ATRA (0.01 μ M) and ATRA (1 μ M). Bar graph represents the log₂ fold change (FC) vs DMSO for IRF8 transcript in each biological condition. B) FACS analysis of the percentage of CD14 positive NB4 cells upon 24 hours treatment with MC2580 (2 μ M), ATRA (0.01 μ M), MC2580 (2 μ M) + ATRA (0.01 μ M) and ATRA (1 μ M).

NB4 cells are promyelocytes, namely cells that are already committed towards the granulocytic lineage; nevertheless they are prone to activate a transcriptional program associated with the monocytic/macrophagic lineage when LSD1 is inhibited. The ability of these cells to differentiate in this specific lineage was also proved by other research studies^{247,248}. Interestingly, our results indicate that GFI1 binds to a putative enhancer of IRF8, the main transcription factor regulating the differentiation process of a myeloid progenitor into a monocyte precursor cell²⁴⁹. Upon MC2580 treatment, LSD1 is evicted from this region, thus probably inhibiting GFI1 repressive function of IRF8 transcription^{23,89}. Overall, our results suggest that LSD1 inhibitors could induce the activation of monocytic/macrophagic differentiation program by both modulating the LSD1-GFI1 interaction and reducing GSE1 expression.

In addition to the above-mentioned pathways, we observed that GSE1-KD induces a strong up-regulation of the “cytokine-mediated signalling”. This process, together with the “regulation of cysteine-endopeptidase activity involved in the apoptotic process” were enriched both upon GSE1 depletion and MC2580 treatment, thus suggesting that they may be regulated by the LSD1 inhibitor through the reduction of GSE1 level. In line with these results, other studies have demonstrated the role of LSD1 in mediating the transcriptional regulation of genes involved in interferon signalling^{250,251}. GSE1 overexpression in NB4 cells followed by treatment with the LSD1 inhibitors and RT-qPCR analysis of genes commonly modulated by both GSE1-KD and MC2580 will allow assessing our hypothesis of the existence of a LSD1-GSE1 regulatory axis controlling these processes.

The observation that the GSE1-KD induced a stronger phenotypic and transcriptional effects than the pharmacological inhibition of the enzyme could be explained by the different efficiency of the two stimuli in decreasing GSE1 levels. As a matter of fact, MC2580 leads to 30%-40% of depletion of GSE1 while with the shRNAs targeting to 70%-80% reduction. GSE1 protein level seems, therefore, to be critical for NB4 proliferation and initiation of cell

death. In agreement with this, we also found that the shB2 construct –which leads to a higher GSE1 protein reduction – elicited a stronger effect on NB4 cell proliferation and apoptosis. These data are *per se* interesting, not only because allow unveiling a potential novel mechanism through which the LSD1 inhibitors act in APL cells, but also because they provide suggestions on other targetable strategies for AML treatment. Indeed GSE1 has already proved to be an important oncogene in other tumours, mainly breast and gastric cancers ^{216,217}. Chai *et al.* ²¹⁶ and Ding *et al.* ²¹⁷ demonstrated that GSE1 downregulation reduced cell proliferation, colony formation, cell migration and invasion in breast and gastric cancer, respectively. Moreover, GSE1 levels are much more elevated in some breast and gastric cancer tissues compared to normal tissues. In this context, analysing the expression of GSE1 and the effects of its down-regulation in other less curable AML subtypes or in other solid tumours where GSE1 is upregulated may provide new ideas on the impact of this protein in other malignancies.

The MS-based profiling of bulk histone PTM levels upon treatment with LSD1 inhibitors in NB4 cells allowed discovering that alterations occur in the H3K27 methylation status, in particular H3K27me2 and H3K27me3 levels increase, while H3K27me1 -combined with H3K36me1- decreases. We observed the same changes upon LSD1 KO, thus proving they are directly linked to LSD1 presence. The molecular mechanisms underpinning the changes of these histone marks have not been clarified, yet. RNA-seq analysis upon MC2580 treatment does not highlight significant changes in the expression of methyltransferases/demethylases acting on this site. Interestingly, also Jin *et al.* observed an increase of H3K27 methylation upon siRNA targeting LSD1 in MCF-7 ²⁵². The authors claimed that this change could depend on two different mechanisms: 1) the alteration in the status of H3K4 methylation (the main LSD1 target) may change the conformation of histone, which recruits other histone-modifying enzymes that modify the methylation of other lysines on the same protein, 2) LSD1 inhibition might alter the activity of other histone methyltransferases and de-methylases interacting physically or functionally with LSD1. For

instance, LSD1 interacts with the H3K27me3 methyl-transferase EZH2 through the long non-coding RNA HOTAIR²⁵³. Besides, LSD1 also binds to different subunits of the MLL2/3 complex containing the tri-methyl H3K27 demethylase UTX²⁵⁴. In our interactome, we found some subunits of this complex, such as RBBP5 and PELP1. However, we did not detect interactions with the enzymes responsible for histone H3K27 methylation/demethylation. Affinity purification approaches involving cross-linking to stabilize protein-protein interaction, such as the ChroP¹⁵⁴ or ChIP-SICAP²⁵⁵ may help to investigate the effect of the LSD1 inhibitors on the chromatin-related interactions that may involve LSD1 and H3K27 methyl-transferases/demethylases. In addition, ChIP-seq profiling of H3K27me1/me2/me3 in the presence and absence of LSD1 inhibitor would also help to identify the genomic regions implicated in these changes. Compared to NB4, UF-1 cells (APL cells showing more sensitivity to LSD1 inhibition than NB4 cells) displayed a different epigenetic response to LSD1 inhibition. In these cells, H3K36me3 increases upon MC2580 treatment while H3K27 methylation is not altered. As in the case of NB4, the histone target H3K4me2 shows an increasing trend upon drug treatment, thus corroborating the finding that phenotypic and cellular effects of LSD1 inhibition in APL models are not directly dependent on the catalytic activity of the enzyme. An additional MS-analysis of histone PTMs allowed detecting higher levels of H3K27me2 and H3K27me3 in NB4 than UF-1 at basal state and a similar result was also found when we analysed patient-derived human glioblastoma cells displaying different sensitivity to LSD1 inhibition. Altogether these data suggest that basal bulk levels of H3K27me2/me3 may be critical to determine (or at least predict) the tumour cell sensitivity to LSD1 inhibition. To verify this hypothesis, we treated NB4 cells with the EZH2/1 inhibitor UNC1999 prior to MC2580 incubation. Reduction of H3K27me3 made NB4 cells much more sensitive to MC2580 treatment. In line with this result, Wen *et al.* demonstrated the synergistic activity of LSD1 and EZH2 inhibitors against AML, both *in vitro* and *in vivo*. This synergy was explained by the decrease of the levels of the anti-apoptotic protein Bcl-2 and by alterations of cell metabolism²⁵⁶. These data could

pave the way to novel combinatorial treatments in those leukemia subtypes showing resistance to LSD1 inhibitors alone. In this context, the group of Prof. S. Minucci tested the effect of the LSD1 inhibitors in a panel of AML cellular subtypes that were resistant to the sole LSD1 inhibition. Assessment of the H3K27me2/me3 levels in these cells in combination with the evaluation of the phenotypic effect of the dual drug treatment could corroborate this preliminary interesting finding.

Taken together, the results I obtained during my PhD provide new insights into the mechanism of action of the LSD1 inhibitors in APL cells. The quantitative MS-interactomics approach allowed detecting the eviction of two LSD1 interactors from the LSD1 complexes upon drug treatment: GFI1 and GSE1. As regards GFI1, we proved that the drugs disrupt its interaction with LSD1, thus sensitizing the cells to physiological doses of ATRA and the consequent re-activation of the myeloid differentiation process. On the other hand, the molecular mechanisms underpinning the regulation of GSE1 by LSD1 inhibitors remain to be fully dissected. The data collected until now allow hypothesizing a model whereby pharmacological inhibition of LSD1 with our small molecules modulate GSE1 transcription which in turns leads to the up-regulation of genes involved in cytokine-signalling and regulation of apoptosis. MS-profiling of histone modifications upon drug treatment revealed that H3K27me2/me3 could be potential markers associated with the resistance to LSD1 inhibition (Figure 48) and demonstrated the efficacy of the combinatorial inhibition of LSD1 and EZH2/1 in reducing NB4 cell proliferation.

While the effect of the LSD1 inhibitors on the LSD1-GFI1 binding had been previously described ⁸⁹, the observations regarding the other two mechanisms are completely novel and may also be relevant from a translational point of view. Indeed, continuing the investigation of the above-mentioned processes may allow optimizing more potent and selective compounds as well as establishing the efficacy of novel combinatorial treatments. In this context, it would be interesting to assess whether the mechanisms detected in this thesis are

also triggered by other LSD1 inhibitors currently in pre-clinical and clinical trials. This information is *per se* useful to assess potential differences and/or similarities among them from a molecular and phenotypic point of view, which in turn can also be reflected in a different therapeutic efficacy.

More in general, my PhD thesis is an ideal example of how MS-proteomics combined with other -omics approaches like ChIP-seq and RNA-seq can push the understanding of the mechanisms of action of various anti-cancer drugs under development, thus offering valuable insights towards more translational application.

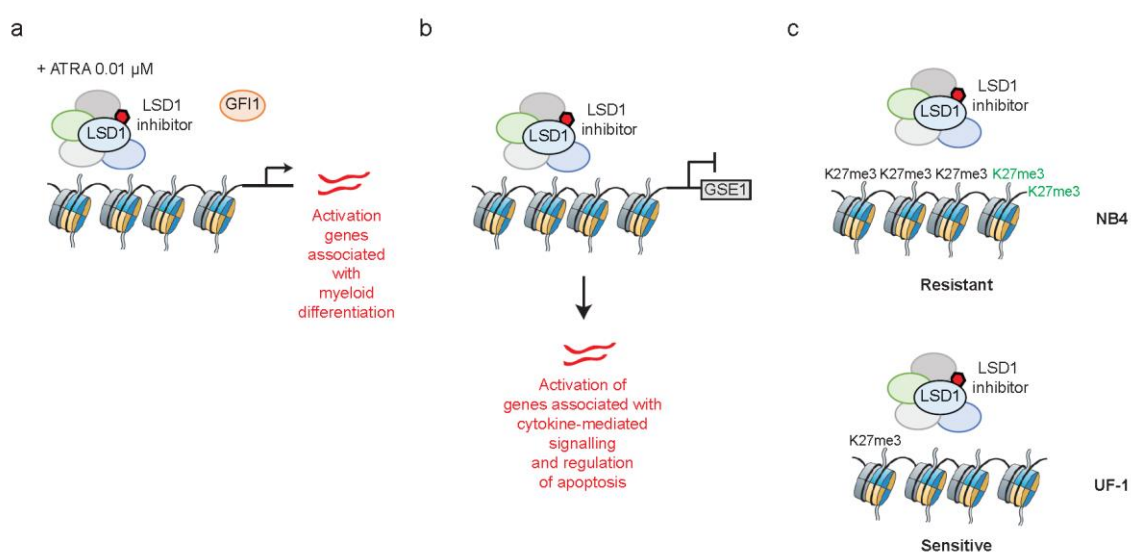


Figure 48: Novel possible mechanisms of action of the LSD1 inhibitors in APL. Quantitative MS-proteomics allowed identifying three possible molecular mechanisms underpinning the activity of LSD1 inhibitors in APL cells. First (a), the LSD1 inhibitor disrupts LSD1-GFI1 binding, thus sensitizing the cells to physiological doses of ATRA (ATRA 0.01 μM) and re-activating the myeloid differentiation process. Second (b), the drug reduces GSE1 transcription and thus induces the expression of genes associated with cytokine-signalling and regulation of apoptosis. Third (c), LSD1 inhibition leads to an increase of H3K27me3 levels (depicted in green) in NB4 and not in UF-1 cell, with bulk H3K27me3 levels being critical for sensitivity to LSD1 inhibition.

7. REFERENCES

1. Shi Y, Lan F, Matson C, et al. Histone demethylation mediated by the nuclear amine oxidase homolog LSD1. *Cell*. 2004. doi:10.1016/j.cell.2004.12.012
2. Humphrey GW, Wang Y, Russanova VR, et al. Stable Histone Deacetylase Complexes Distinguished by the Presence of SANT Domain Proteins CoREST/kiaa0071 and Mta-L1. *J Biol Chem*. 2001. doi:10.1074/jbc.M007372200
3. Shi Y, Sawada JI, Sui G, et al. Coordinated histone modifications mediated by a CtBP co-repressor complex. *Nature*. 2003. doi:10.1038/nature01550
4. Forneris F, Binda C, Vanoni MA, Mattevi A, Battaglioli E. Histone demethylation catalysed by LSD1 is a flavin-dependent oxidative process. *FEBS Lett*. 2005. doi:10.1016/j.febslet.2005.03.015
5. Chen Y, Yang Y, Wang F, et al. Crystal structure of human histone lysine-specific demethylase 1 (LSD1). *Proc Natl Acad Sci*. 2006. doi:10.1073/pnas.0606381103
6. Stavropoulos P, Blobel G, Hoelz A. Crystal structure and mechanism of human lysine-specific demethylase-1. *Nat Struct Mol Biol*. 2006. doi:10.1038/nsmb1113
7. Hosseini A, Minucci S. A comprehensive review of lysine-specific demethylase 1 and its roles in cancer. *Epigenomics*. 2017. doi:10.2217/epi-2017-0022
8. Hou H, Yu H. Structural insights into histone lysine demethylation. *Curr Opin Struct Biol*. 2010. doi:10.1016/j.sbi.2010.09.006
9. Shi YJ, Matson C, Lan F, Iwase S, Baba T, Shi Y. Regulation of LSD1 histone demethylase activity by its associated factors. *Mol Cell*. 2005. doi:10.1016/j.molcel.2005.08.027
10. Adamo A, Sesé B, Boue S, et al. LSD1 regulates the balance between self-renewal and differentiation in human embryonic stem cells. *Nat Cell Biol*. 2011. doi:10.1038/ncb2246
11. He Y, Zhao Y, Wang L, et al. LSD1 promotes S-phase entry and tumorigenesis via chromatin co-occupation with E2F1 and selective H3K9 demethylation. *Oncogene*. 2018. doi:10.1038/onc.2017.353
12. Sakamoto A, Hino S, Nagaoka K, et al. Lysine demethylase LSD1 coordinates glycolytic and mitochondrial metabolism in hepatocellular carcinoma cells. *Cancer Res*. 2015. doi:10.1158/0008-5472.CAN-14-1560
13. Hino S, Sakamoto A, Nagaoka K, et al. FAD-dependent lysine-specific demethylase-1 regulates cellular energy expenditure. *Nat Commun*. 2012. doi:10.1038/ncomms1755
14. Li Y, Wan X, Wei Y, et al. LSD1-mediated epigenetic modification contributes to ovarian cancer cell migration and invasion. *Oncol Rep*. 2016. doi:10.3892/or.2016.4729
15. Lin Y, Wu Y, Li J, et al. The SNAG domain of snail1 functions as a molecular hook

- for recruiting lysine-specific demethylase 1. *EMBO J.* 2010. doi:10.1038/emboj.2010.63
16. Boulding T, McCuaig RD, Tan A, et al. LSD1 activation promotes inducible EMT programs and modulates the tumour microenvironment in breast cancer. *Sci Rep.* 2018. doi:10.1038/s41598-017-17913-x
 17. Ambrosio S, Amente S, Saccà CD, et al. LSD1 mediates MYCN control of epithelial-mesenchymal transition through silencing of metastatic suppressor NDRG1 gene. *Oncotarget.* 2017. doi:10.18632/oncotarget.12924
 18. Amente S, Lania L, Majello B. The histone LSD1 demethylase in stemness and cancer transcription programs. *Biochim Biophys Acta - Gene Regul Mech.* 2013. doi:10.1016/j.bbagr.2013.05.002
 19. Wang Y, Zhang H, Chen Y, et al. LSD1 Is a Subunit of the NuRD Complex and Targets the Metastasis Programs in Breast Cancer. *Cell.* 2009. doi:10.1016/j.cell.2009.05.050
 20. Forneris F, Binda C, Dall'Aglio A, Fraaije MW, Battaglioli E, Mattevi A. A highly specific mechanism of histone H3-K4 recognition by histone demethylase LSD1. *J Biol Chem.* 2006. doi:10.1074/jbc.M607411200
 21. Lan F, Collins RE, De Cegli R, et al. Recognition of unmethylated histone H3 lysine 4 links BHC80 to LSD1-mediated gene repression. *Nature.* 2007. doi:10.1038/nature06034
 22. Magliulo D, Bernardi R, Messina S. Lysine-Specific Demethylase 1A as a Promising Target in Acute Myeloid Leukemia. *Front Oncol.* 2018. doi:10.3389/fonc.2018.00255
 23. Saleque S, Kim J, Rooke HM, Orkin SH. Epigenetic Regulation of Hematopoietic Differentiation by Gfi-1 and Gfi-1b Is Mediated by the Cofactors CoREST and LSD1. *Mol Cell.* 2007. doi:10.1016/j.molcel.2007.06.039
 24. Patel D, Shimomura A, Majumdar S, Holley MC, Hashino E. The histone demethylase LSD1 regulates inner ear progenitor differentiation through interactions with Pax2 and the NuRD repressor complex. *PLoS One.* 2018. doi:10.1371/journal.pone.0191689
 25. Ouyang H, Qin Y, Liu Y, Xie Y, Liu J. Prox1 Directly Interacts with LSD1 and Recruits the LSD1/NuRD Complex to Epigenetically Co-Repress CYP7A1 Transcription. *PLoS One.* 2013. doi:10.1371/journal.pone.0062192
 26. McClellan D, Casey MJ, Bareyan D, et al. Growth Factor Independence 1B-Mediated Transcriptional Repression and Lineage Allocation Require Lysine-Specific Demethylase 1-Dependent Recruitment of the BHC Complex. *Mol Cell Biol.* 2019. doi:10.1128/mcb.00020-19
 27. Yang Y, Huang W, Qiu R, et al. LSD1 coordinates with the SIN3A/HDAC complex and maintains sensitivity to chemotherapy in breast cancer. *J Mol Cell Biol.* 2018. doi:10.1093/jmcb/mjy021

28. Han X, Ranganathan P, Tzimas C, et al. Notch Represses Transcription by PRC2 Recruitment to the Ternary Complex. *Mol Cancer Res.* 2017. doi:10.1158/1541-7786.mcr-17-0241
29. Metzger E, Wissmann M, Yin N, et al. LSD1 demethylates repressive histone marks to promote androgen-receptor-dependent transcription. *Nature.* 2005. doi:10.1038/nature04020
30. Perillo B, Ombra MN, Bertoni A, et al. DNA oxidation as triggered by H3K9me2 demethylation drives estrogen-induced gene expression. *Science (80-).* 2008. doi:10.1126/science.1147674
31. Metzger E, Imhof A, Patel D, et al. Phosphorylation of histone H3T6 by PKCB i controls demethylation at histone H3K4. *Nature.* 2010. doi:10.1038/nature08839
32. Nair SS, Nair BC, Cortez V, et al. PELP1 is a reader of histone H3 methylation that facilitates oestrogen receptor- α target gene activation by regulating lysine demethylase 1 specificity. *EMBO Rep.* 2010. doi:10.1038/embor.2010.62
33. Chen C, Wang Y, Wang S, et al. LSD1 sustains estrogen-driven endometrial carcinoma cell proliferation through the PI3K/AKT pathway via di-demethylating H3K9 of cyclin D1. *Int J Oncol.* 2017. doi:10.3892/ijo.2017.3849
34. Scionti I, Hayashi S, Mouradian S, et al. LSD1 Controls Timely MyoD Expression via MyoD Core Enhancer Transcription. *Cell Rep.* 2017. doi:10.1016/j.celrep.2017.01.078
35. Musri MM, Carmona MC, Hanzu FA, Kaliman P, Gomis R, Párrizas M. Histone demethylase LSD1 regulates adipogenesis. *J Biol Chem.* 2010. doi:10.1074/jbc.M110.151209
36. Ahmed M, Streit A. Lsd1 interacts with cMyb to demethylate repressive histone marks and maintain inner ear progenitor identity. *Development.* 2018. doi:10.1242/dev.160325
37. Huang J, Sengupta R, Espejo AB, et al. p53 is regulated by the lysine demethylase LSD1. *Nature.* 2007. doi:10.1038/nature06092
38. Kontaki H, Talianidis I. Lysine Methylation Regulates E2F1-Induced Cell Death. *Mol Cell.* 2010. doi:10.1016/j.molcel.2010.06.006
39. Wang J, Hevi S, Kurash JK, et al. The lysine demethylase LSD1 (KDM1) is required for maintenance of global DNA methylation. *Nat Genet.* 2009. doi:10.1038/ng.268
40. Cho HS, Suzuki T, Dohmae N, et al. Demethylation of RB regulator MYPT1 by histone demethylase LSD1 promotes cell cycle progression in cancer cells. *Cancer Res.* 2011. doi:10.1158/0008-5472.CAN-10-2446
41. Kim Y, Nam HJ, Lee J, et al. Methylation-dependent regulation of HIF-1 α stability restricts retinal and tumour angiogenesis. *Nat Commun.* 2016. doi:10.1038/ncomms10347
42. Lee JY, Park JH, Choi HJ, et al. LSD1 demethylates HIF1 α to inhibit hydroxylation

- and ubiquitin-mediated degradation in tumor angiogenesis. *Oncogene*. 2017. doi:10.1038/onc.2017.158
43. Cuella-Martin R, Oliveira C, Lockstone HE, Snellenberg S, Grolmusova N, Chapman JR. 53BP1 Integrates DNA Repair and p53-Dependent Cell Fate Decisions via Distinct Mechanisms. *Mol Cell*. 2016. doi:10.1016/j.molcel.2016.08.002
 44. Schulte JH, Lim S, Schramm A, et al. Lysine-specific demethylase 1 is strongly expressed in poorly differentiated neuroblastoma: Implications for therapy. *Cancer Res*. 2009. doi:10.1158/0008-5472.CAN-08-1735
 45. Zhao ZK, Dong P, Gu J, et al. Overexpression of LSD1 in hepatocellular carcinoma: A latent target for the diagnosis and therapy of hepatoma. *Tumor Biol*. 2013. doi:10.1007/s13277-012-0525-x
 46. Hayami S, Kelly JD, Cho HS, et al. Overexpression of LSD1 contributes to human carcinogenesis through chromatin regulation in various cancers. *Int J Cancer*. 2011. doi:10.1002/ijc.25349
 47. Lim S, Janzer A, Becker A, et al. Lysine-specific demethylase 1 (LSD1) is highly expressed in ER-negative breast cancers and a biomarker predicting aggressive biology. *Carcinogenesis*. 2010. doi:10.1093/carcin/bgp324
 48. Serce N, Gnatzy A, Steiner S, Lorenzen H, Kirfel J, Buettner R. Elevated expression of LSD1 (Lysine-specific demethylase 1) during tumour progression from pre-invasive to invasive ductal carcinoma of the breast. *BMC Clin Pathol*. 2012. doi:10.1186/1472-6890-12-13
 49. Pollock JA, Larrea MD, Jasper JS, McDonnell DP, McCafferty DG. Lysine-specific histone demethylase 1 inhibitors control breast cancer proliferation in ER α -dependent and -independent manners. *ACS Chem Biol*. 2012. doi:10.1021/cb300108c
 50. Ferrari-Amorotti G, Chiodoni C, Shen F, et al. Suppression of Invasion and Metastasis of Triple-Negative Breast Cancer Lines by Pharmacological or Genetic Inhibition of Slug Activity. *Neoplasia*. 2014. doi:10.1016/j.neo.2014.10.006
 51. Bai JW, Chen MN, Wei XL, et al. The zinc-finger transcriptional factor Slug transcriptionally downregulates ER α by recruiting lysine-specific demethylase 1 in human breast cancer. *Oncogenesis*. 2017. doi:10.1038/oncsis.2017.38
 52. Ketscher A, Jilg CA, Willmann D, et al. LSD1 controls metastasis of androgen-independent prostate cancer cells through PXN and LPAR6. *Oncogenesis*. 2014. doi:10.1038/oncsis.2014.34
 53. Kashyap V, Ahmad S, Nilsson EM, et al. The lysine specific demethylase-1 (LSD1/KDM1A) regulates VEGF-A expression in prostate cancer. *Mol Oncol*. 2013. doi:10.1016/j.molonc.2013.01.003
 54. Hotte SJ, Saad F. Current management of castrate-resistant prostate cancer. *Curr Oncol*. 2010.
 55. Liang Y, Ahmed M, Guo H, et al. LSD1-mediated epigenetic reprogramming drives CENPE expression and prostate cancer progression. *Cancer Res*. 2017.

doi:10.1158/0008-5472.CAN-17-0496

56. Yang GJ, Lei PM, Wong SY, Ma DL, Leung CH. Pharmacological inhibition of LSD1 for cancer treatment. *Molecules*. 2018. doi:10.3390/molecules23123194
57. Cai C, He HH, Chen S, et al. Androgen Receptor Gene Expression in Prostate Cancer Is Directly Suppressed by the Androgen Receptor Through Recruitment of Lysine-Specific Demethylase 1. *Cancer Cell*. 2011. doi:10.1016/j.ccr.2011.09.001
58. Sehrawat A, Gao L, Wang Y, et al. LSD1 activates a lethal prostate cancer gene network independently of its demethylase function. *Proc Natl Acad Sci*. 2018. doi:10.1073/pnas.1719168115
59. Lv T, Yuan D, Miao X, et al. Over-expression of LSD1 promotes proliferation, migration and invasion in non-small cell lung cancer. *PLoS One*. 2012. doi:10.1371/journal.pone.0035065
60. Doll S, Kriegmair MC, Santos A, et al. Rapid proteomic analysis for solid tumors reveals LSD1 as a drug target in an end-stage cancer patient. *Mol Oncol*. 2018. doi:10.1002/1878-0261.12326
61. Huang Z, Li S, Song W, et al. Lysine-Specific Demethylase 1 (LSD1/KDM1A) Contributes to Colorectal Tumorigenesis via Activation of the Wnt/B-Catenin Pathway by Down-Regulating Dickkopf-1 (DKK1). *PLoS One*. 2013. doi:10.1371/journal.pone.0070077
62. Chen J, Ding J, Wang Z, Zhu J, Wang X, Du J. Identification of downstream metastasis-associated target genes regulated by LSD1 in colon cancer cells. *Oncotarget*. 2017. doi:10.18632/oncotarget.14778
63. Ding J, Zhang ZM, Xia Y, et al. LSD1-mediated epigenetic modification contributes to proliferation and metastasis of colon cancer. *Br J Cancer*. 2013. doi:10.1038/bjc.2013.364
64. Fernandez-Zapico ME, Kozub MM, Carr RM, Lomberk GL. LSD1, a double-edged sword, confers dynamic chromatin regulation but commonly promotes aberrant cell growth. *F1000Research*. 2017. doi:10.12688/f1000research.12169.1
65. De Kouchkovsky I, Abdul-Hay M. 'Acute myeloid leukemia: A comprehensive review and 2016 update.' *Blood Cancer J*. 2016. doi:10.1038/bcj.2016.50
66. Yamamoto JF, Goodman MT. Patterns of leukemia incidence in the United States by subtype and demographic characteristics, 1997-2002. *Cancer Causes Control*. 2008. doi:10.1007/s10552-007-9097-2
67. Bennett JM, Catovsky D, Daniel M -T, et al. Proposals for the Classification of the Acute Leukaemias French-American-British (FAB) Co-operative Group. *Br J Haematol*. 1976. doi:10.1111/j.1365-2141.1976.tb03563.x
68. Bloomfield CD, Brunning RD. FAB M7: Acute megakaryoblastic leukemia. Beyond morphology. *Ann Intern Med*. 1985. doi:10.7326/0003-4819-103-3-450
69. Lee EJ, Pollak A, Leavitt RD, Testa JR, Schiffer CA. Minimally differentiated acute

- nonlymphocytic leukemia: a distinct entity. *Blood*. 1987.
70. Degos L, Dombret H, Chomienne C, et al. All-trans-retinoic acid as a differentiating agent in the treatment of acute promyelocytic leukemia. *Blood*. 1995.
 71. De Thé H, Chen Z. Acute promyelocytic leukaemia: Novel insights into the mechanisms of cure. *Nat Rev Cancer*. 2010. doi:10.1038/nrc2943
 72. Chen G-Q, Zhu J, Shi X-G, et al. In vitro studies on cellular and molecular mechanisms of arsenic trioxide (As₂O₃) in the treatment of acute promyelocytic leukemia: As₂O₃ induces NB₄ cell apoptosis with downregulation of Bcl-2 expressi. *Blood*. 1996.
 73. Guillemain M-C, Raffoux E, Vitoux D, et al. In Vivo Activation of cAMP Signaling Induces Growth Arrest and Differentiation in Acute Promyelocytic Leukemia. *J Exp Med*. 2002. doi:10.1084/jem.20021129
 74. Zhu Q, Zhang JW, Zhu HQ, et al. Synergic effects of arsenic trioxide and cAMP during acute promyelocytic leukemia cell maturation subtends a novel signaling crosstalk. *Blood*. 2002. doi:10.1182/blood.V99.3.1014
 75. Muto A, Kizaki M, Kawamura C, et al. A novel differentiation-inducing therapy for acute promyelocytic leukemia with a combination of arsenic trioxide and GM-CSF. *Leukemia*. 2001. doi:10.1038/sj.leu.2402162
 76. Coombs CC, Tavakkoli M, Tallman MS. Acute promyelocytic Leukemia: Where did we start, where are we now, and the future. *Blood Cancer J*. 2015. doi:10.1038/bcj.2015.25
 77. dos Santos GA, Kats L, Pandolfi PP. Synergy against PML-RARα: targeting transcription, proteolysis, differentiation, and self-renewal in acute promyelocytic leukemia. *J Exp Med*. 2013. doi:10.1084/jem.20131121
 78. Insinga A, Monestiroli S, Ronzoni S, et al. Impairment of p53 acetylation, stability and function by an oncogenic transcription factor. *EMBO J*. 2004. doi:10.1038/sj.emboj.7600109
 79. Lallemand-Breitenbach V, de Thé H. PML nuclear bodies. *Cold Spring Harb Perspect Biol*. 2010. doi:10.1101/cshperspect.a000661
 80. Licht JD. Reconstructing a disease: What essential features of the retinoic acid receptor fusion oncoproteins generate acute promyelocytic leukemia? *Cancer Cell*. 2006. doi:10.1016/j.ccr.2006.01.024
 81. Di Croce L, Raker VA, Corsaro M, et al. Methyltransferase recruitment and DNA hypermethylation of target promoters by an oncogenic transcription factor. *Science* (80-). 2002. doi:10.1126/science.1065173
 82. Drumea K, Yang ZF, Rosmarin A. Retinoic acid signaling in myelopoiesis. *Curr Opin Hematol*. 2008. doi:10.1097/MOH.0b013e3282f20a9c
 83. Melnick A, Licht JD. Deconstructing a disease: RARα, its fusion partners, and their roles in the pathogenesis of acute promyelocytic leukemia. *Blood*. 1999.

84. Zhu J, Gianni M, Kopf E, et al. Retinoic acid induces proteasome-dependent degradation of retinoic acid receptor alpha (RARalpha) and oncogenic RARalpha fusion proteins. *Proc Natl Acad Sci*. 1999. doi:10.1073/pnas.96.26.14807
85. de Thé H, Pandolfi PP, Chen Z. Acute Promyelocytic Leukemia: A Paradigm for Oncoprotein-Targeted Cure. *Cancer Cell*. 2017. doi:10.1016/j.ccell.2017.10.002
86. Niebel D, Kirfel J, Janzen V, Höller T, Majores M, Gütgemann I. To the editor: Lysine-specific demethylase 1 (LSD1) in hematopoietic and lymphoid neoplasms. *Blood*. 2014. doi:10.1182/blood-2014-04-569525
87. Harris WJ, Huang X, Lynch JT, et al. The Histone Demethylase KDM1A Sustains the Oncogenic Potential of MLL-AF9 Leukemia Stem Cells. *Cancer Cell*. 2012. doi:10.1016/j.ccr.2012.03.014
88. Fang J, Ying H, Mao T, et al. Upregulation of CD11b and CD86 through LSD1 inhibition promotes myeloid differentiation and suppresses cell proliferation in human monocytic leukemia cells. *Oncotarget*. 2017. doi:10.18632/oncotarget.18564
89. Maiques-Diaz A, Spencer GJ, Lynch JT, et al. Enhancer Activation by Pharmacologic Displacement of LSD1 from GFI1 Induces Differentiation in Acute Myeloid Leukemia. *Cell Rep*. 2018. doi:10.1016/j.celrep.2018.03.012
90. Schenk T, Chen WC, Göllner S, et al. Inhibition of the LSD1 (KDM1A) demethylase reactivates the all-trans-retinoic acid differentiation pathway in acute myeloid leukemia. *Nat Med*. 2012. doi:10.1038/nm.2661
91. Lee MG, Wynder C, Schmidt DM, McCafferty DG, Shiekhattar R. Histone H3 Lysine 4 Demethylation Is a Target of Nonselective Antidepressive Medications. *Chem Biol*. 2006. doi:10.1016/j.chembiol.2006.05.004
92. Schmidt DMZ, McCafferty DG. trans-2-phenylcyclopropylamine is a mechanism-based inactivator of the histone demethylase LSD1. *Biochemistry*. 2007. doi:10.1021/bi0618621
93. Zheng YC, Ma J, Wang Z, et al. A Systematic Review of Histone Lysine-Specific Demethylase 1 and Its Inhibitors. *Med Res Rev*. 2015. doi:10.1002/med.21350
94. Maes T, Mascaró C, Tirapu I, et al. ORY-1001, a Potent and Selective Covalent KDM1A Inhibitor, for the Treatment of Acute Leukemia. *Cancer Cell*. 2018. doi:10.1016/j.ccell.2018.02.002
95. Lu Z, Guo Y, Zhang X, et al. ORY-1001 Suppresses Cell Growth and Induces Apoptosis in Lung Cancer Through Triggering HK2 Mediated Warburg Effect. *Front Pharmacol*. 2018. doi:10.3389/fphar.2018.01411
96. Augert A, Eastwood E, Ibrahim AH, et al. Targeting NOTCH activation in small cell lung cancer through LSD1 inhibition. *Sci Signal*. 2019. doi:10.1126/scisignal.aau2922
97. Maes T, Mascaró C, Ortega A, et al. KDM1 histone lysine demethylases as targets for treatments of oncological and neurodegenerative disease. *Epigenomics*. 2015. doi:10.2217/epi.15.9

98. Mohammad HP, Kruger RG. Antitumor activity of LSD1 inhibitors in lung cancer. *Mol Cell Oncol*. 2016. doi:10.1080/23723556.2015.1117700
99. Mohammad HP, Smitheman KN, Kamat CD, et al. A DNA Hypomethylation Signature Predicts Antitumor Activity of LSD1 Inhibitors in SCLC. *Cancer Cell*. 2015. doi:10.1016/j.ccell.2015.06.002
100. Huang M, Chen C, Geng J, et al. Targeting KDM1A attenuates Wnt/ β -catenin signaling pathway to eliminate sorafenib-resistant stem-like cells in hepatocellular carcinoma. *Cancer Lett*. 2017. doi:10.1016/j.canlet.2017.03.038
101. Lillico R, Lawrence CK, Lakowski TM. Selective DOT1L, LSD1, and HDAC Class I Inhibitors Reduce HOXA9 Expression in MLL-AF9 Rearranged Leukemia Cells, but Dysregulate the Expression of Many Histone-Modifying Enzymes. *J Proteome Res*. 2018. doi:10.1021/acs.jproteome.8b00118
102. Goossens S, Peirs S, Van Loocke W, et al. Oncogenic ZEB2 activation drives sensitivity toward KDM1A inhibition in T-cell acute lymphoblastic leukemia. *Blood*. 2017. doi:10.1182/blood-2016-06-721191
103. Smitheman KN, Severson TM, Rajapurkar SR, et al. Lysine specific demethylase 1 inactivation enhances differentiation and promotes cytotoxic response when combined with all-trans retinoic acid in acute myeloid leukemia across subtypes. *Haematologica*. 2019. doi:10.3324/haematol.2018.199190
104. Takagi S, Ishikawa Y, Mizutani A, et al. LSD1 inhibitor T-3775440 inhibits SCLC cell proliferation by disrupting LSD1 interactions with SNAG domain proteins INSM1 and GFI1B. *Cancer Res*. 2017. doi:10.1158/0008-5472.CAN-16-3502
105. Ishikawa Y, Gamo K, Yabuki M, et al. A Novel LSD1 Inhibitor T-3775440 Disrupts GFI1B-Containing Complex Leading to Transdifferentiation and Impaired Growth of AML Cells. *Mol Cancer Ther*. 2017. doi:10.1158/1535-7163.MCT-16-0471
106. McGrath JP, Williamson KE, Balasubramanian S, et al. Pharmacological inhibition of the histone lysine demethylase KDM1A suppresses the growth of multiple acute myeloid leukemia subtypes. *Cancer Res*. 2016. doi:10.1158/0008-5472.CAN-15-2333
107. Sugino N, Kawahara M, Tatsumi G, et al. A novel LSD1 inhibitor NCD38 ameliorates MDS-related leukemia with complex karyotype by attenuating leukemia programs via activating super-enhancers. *Leukemia*. 2017. doi:10.1038/leu.2017.59
108. Fiskus W, Sharma S, Shah B, et al. Highly effective combination of LSD1 (KDM1A) antagonist and pan-histone deacetylase inhibitor against human AML cells. *Leukemia*. 2014. doi:10.1038/leu.2014.119
109. Haydn T, Metzger E, Schuele R, Fulda S. Concomitant epigenetic targeting of LSD1 and HDAC synergistically induces mitochondrial apoptosis in rhabdomyosarcoma cells. *Cell Death Dis*. 2017. doi:10.1038/cddis.2017.239
110. Levis M. Quizartinib for the treatment of FLT3/ITD acute myeloid leukemia. *Futur Oncol*. 2014. doi:10.2217/fon.14.105

111. Daigle SR, Olhava EJ, Therkelsen CA, et al. Potent inhibition of DOT1L as treatment of MLL-fusion leukemia. *Blood*. 2013. doi:10.1182/blood-2013-04-497644
112. Dafflon C, Craig VJ, Méreau H, et al. Complementary activities of DOT1L and Menin inhibitors in MLL-rearranged leukemia. *Leukemia*. 2017. doi:10.1038/leu.2016.327
113. Maiques-Diaz A, Somervaille TCP. LSD1: Biologic roles and therapeutic targeting. *Epigenomics*. 2016. doi:10.2217/epi-2016-0009
114. Lokken AA, Zeleznik-Le NJ. Breaking the LSD1/KDM1A Addiction: Therapeutic Targeting of the Epigenetic Modifier in AML. *Cancer Cell*. 2012. doi:10.1016/j.ccr.2012.03.027
115. Barth J, Abou-El-Ardat K, Dalic D, et al. LSD1 inhibition by tranlycypromine derivatives interferes with GFI1-mediated repression of PU.1 target genes and induces differentiation in AML. *Leukemia*. 2019. doi:10.1038/s41375-018-0375-7
116. Binda C, Valente S, Romanenghi M, et al. Biochemical, structural, and biological evaluation of tranlycypromine derivatives as inhibitors of histone demethylases LSD1 and LSD2. *J Am Chem Soc*. 2010. doi:10.1021/ja101557k
117. Vianello P, Botrugno OA, Cappa A, et al. Discovery of a novel inhibitor of histone lysine-specific demethylase 1A (KDM1A/LSD1) as orally active antitumor agent. *J Med Chem*. 2016;59(4):1501-1517.
118. Kizaki M, Matsushita H, Takayama N, et al. Establishment and characterization of a novel acute promyelocytic leukemia cell line (UF-1) with retinoic acid-resistant features. *Blood*. 1996.
119. Hengesbach LM, Hoag KA. Physiological Concentrations of Retinoic Acid Favor Myeloid Dendritic Cell Development over Granulocyte Development in Cultures of Bone Marrow Cells from Mice. *J Nutr*. 2004. doi:10.1093/jn/134.10.2653
120. Idres N, Benoît G, Flexor MA, Lanotte M, Chabot GG. Granulocytic differentiation of human NB4 promyelocytic leukemia cells induced by all-trans retinoic acid metabolites. *Cancer Res*. 2001.
121. Dunham WH, Mullin M, Gingras AC. Affinity-purification coupled to mass spectrometry: Basic principles and strategies. *Proteomics*. 2012. doi:10.1002/pmic.201100523
122. Cox J, Mann M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat Biotechnol*. 2008. doi:10.1038/nbt.1511
123. Perkins DN, Pappin DJC, Creasy DM, Cottrell JS. Probability-based protein identification by searching sequence databases using mass spectrometry data. In: *Electrophoresis*. ; 1999. doi:10.1002/(SICI)1522-2683(19991201)20:18<3551::AID-ELPS3551>3.0.CO;2-2
124. Yates JR, Eng JK, McCormack AL, Schieltz D. Method to Correlate Tandem Mass Spectra of Modified Peptides to Amino Acid Sequences in the Protein Database. *Anal Chem*. 1995. doi:10.1021/ac00104a020

125. Hillenkamp F, Karas M. Mass spectrometry of peptides and proteins by matrix-assisted ultraviolet laser desorption/ionization. *Methods Enzymol.* 1990. doi:10.1016/0076-6879(90)93420-P
126. Hillenkamp F, Karas M, Beavis RC, Chait BT. Matrix-Assisted Laser Desorption/Ionization Mass Spectrometry of Biopolymers. *Anal Chem.* 1991. doi:10.1021/ac00024a002
127. Fenn JB, Mann M, Meng CK, Wong SF, Whitehouse CM. Electrospray ionization for mass spectrometry of large biomolecules. *Science (80-)*. 1989. doi:10.1126/science.2675315
128. Emmett MR, Caprioli RM. Micro-electrospray mass spectrometry: Ultra-high-sensitivity analysis of peptides and proteins. *J Am Soc Mass Spectrom.* 1994. doi:10.1016/1044-0305(94)85001-1
129. Griffin TJ, Goodlett DR, Aebersold R. Advances in proteome analysis by mass spectrometry. *Curr Opin Biotechnol.* 2001. doi:10.1016/S0958-1669(01)00268-3
130. Shukla AK, Futrell JH. Tandem mass spectrometry: Dissociation of ions by collisional activation. *J Mass Spectrom.* 2000. doi:10.1002/1096-9888(200009)35:9<1069::AID-JMS54>3.0.CO;2-C
131. Olsen J V., Macek B, Lange O, Makarov A, Horning S, Mann M. Higher-energy C-trap dissociation for peptide modification analysis. *Nat Methods.* 2007. doi:10.1038/nmeth1060
132. Nagaraj N, D'Souza RCJ, Cox J, Olsen J V., Mann M. Feasibility of large-scale phosphoproteomics with higher energy collisional dissociation fragmentation. *J Proteome Res.* 2010. doi:10.1021/pr100637q
133. Kelleher NL, Zubarev RA, Bush K, et al. Localization of labile posttranslational modifications by electron capture dissociation: The case of γ -carboxyglutamic acid. *Anal Chem.* 1999. doi:10.1021/ac990684x
134. Syka JEP, Coon JJ, Schroeder MJ, Shabanowitz J, Hunt DF. Peptide and protein sequence analysis by electron transfer dissociation mass spectrometry. *Proc Natl Acad Sci U S A.* 2004. doi:10.1073/pnas.0402700101
135. Zubarev RA, Horn DM, Fridriksson EK, et al. Electron capture dissociation for structural characterization of multiply charged protein cations. *Anal Chem.* 2000. doi:10.1021/ac990811p
136. Sandhu C, Hewel JA, Badis G, et al. Evaluation of data-dependent versus targeted shotgun proteomic approaches for monitoring transcription factor expression in breast cancer. *J Proteome Res.* 2008. doi:10.1021/pr700836q
137. Jazurek M, Ciesiolka A, Starega-Roslan J, Bilinska K, Krzyzosiak WJ. Identifying proteins that bind to specific RNAs - Focus on simple repeat expansion diseases. *Nucleic Acids Res.* 2016. doi:10.1093/nar/gkw803
138. Bantscheff M, Schirle M, Sweetman G, Rick J, Kuster B. Quantitative mass spectrometry in proteomics: A critical review. *Anal Bioanal Chem.* 2007.

139. Bantscheff M, Lemeer S, Savitski MM, Kuster B. Quantitative mass spectrometry in proteomics: Critical review update from 2007 to the present. *Anal Bioanal Chem*. 2012. doi:10.1007/s00216-012-6203-4
140. Ong SE, Blagoev B, Kratchmarova I, et al. Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Mol Cell Proteomics*. 2002. doi:10.1074/mcp.M200025-MCP200
141. Blagoev B, Kratchmarova I, Ong SE, Nielsen M, Foster LJ, Mann M. A proteomics strategy to elucidate functional protein-protein interactions applied to EGF signaling. *Nat Biotechnol*. 2003. doi:10.1038/nbt790
142. Blagoev B, Ong SE, Kratchmarova I, Mann M. Temporal analysis of phosphotyrosine-dependent signaling networks by quantitative proteomics. *Nat Biotechnol*. 2004. doi:10.1038/nbt1005
143. Henriksen P, Wagner SA, Weinert BT, et al. Proteome-wide analysis of lysine acetylation suggests its broad regulatory scope in *Saccharomyces cerevisiae*. *Mol Cell Proteomics*. 2012. doi:10.1074/mcp.M112.017251
144. de Godoy LMF, Olsen J V., de Souza GA, Li G, Mortensen P, Mann M. Status of complete proteome analysis by mass spectrometry: SILAC labeled yeast as a model system. *Genome Biol*. 2006. doi:10.1186/gb-2006-7-6-r50
145. Ross PL, Huang YN, Marchese JN, et al. Multiplexed protein quantitation in *Saccharomyces cerevisiae* using amine-reactive isobaric tagging reagents. *Mol Cell Proteomics*. 2004. doi:10.1074/mcp.M400129-MCP200
146. Thompson A, Schäfer J, Kuhn K, et al. Tandem mass tags: A novel quantification strategy for comparative analysis of complex protein mixtures by MS/MS. *Anal Chem*. 2003. doi:10.1021/ac0262560
147. Giambruno R, Grebien F, Stukalov A, et al. Affinity purification strategies for proteomic analysis of transcription factor complexes. *J Proteome Res*. 2013. doi:10.1021/pr4003323
148. Miteva Y V., Budayeva HG, Cristea IM. Proteomics-based methods for discovery, quantification, and validation of protein-protein interactions. *Anal Chem*. 2013. doi:10.1021/ac3033257
149. Gavin AC, Bösch M, Krause R, et al. Functional organization of the yeast proteome by systematic analysis of protein complexes. *Nature*. 2002. doi:10.1038/415141a
150. Ho Y, Gruhler A, Heilbut A, et al. Systematic identification of protein complexes in *Saccharomyces cerevisiae* by mass spectrometry. *Nature*. 2002. doi:10.1038/415180a
151. Budayeva HG, Cristea IM. A mass spectrometry view of stable and transient protein interactions. *Adv Exp Med Biol*. 2014. doi:10.1007/978-3-319-06068-2_11
152. Alanis-Lobato G. Mining protein interactomes to improve their reliability and support the advancement of network medicine. *Front Genet*. 2015.

doi:10.3389/fgene.2015.00296

153. Redon S, Reichenbach P, Lingner J. Protein-RNA and protein-protein interactions mediate association of human EST1A/SMG6 with telomerase. *Nucleic Acids Res.* 2007. doi:10.1093/nar/gkm724
154. Soldi M, Bonaldi T. The Proteomic Investigation of Chromatin Functional Domains Reveals Novel Synergisms among Distinct Heterochromatin Components. *Mol Cell Proteomics.* 2013. doi:10.1074/mcp.M112.024307
155. Van Eijl RAPM, Van Den Brand T, Nguyen LN, Mulder KW. Reactivity of human AGO2 monoclonal antibody 11A9 with the SWI/SNF complex: A case study for rigorously defining antibody selectivity. *Sci Rep.* 2017. doi:10.1038/s41598-017-07539-4
156. Selbach M, Mann M. Protein interaction screening by quantitative immunoprecipitation combined with knockdown (QUICK). *Nat Methods.* 2006. doi:10.1038/nmeth972
157. Lavallée-Adam M, Cloutier P, Coulombe B, Blanchette M. Modeling contaminants in AP-MS/MS experiments. *J Proteome Res.* 2011. doi:10.1021/pr100795z
158. Sowa ME, Bennett EJ, Gygi SP, Harper JW. Defining the Human Deubiquitinating Enzyme Interaction Landscape. *Cell.* 2009. doi:10.1016/j.cell.2009.04.042
159. Tackett AJ, DeGrasse JA, Sekedat MD, Oeffinger M, Rout MP, Chait BT. I-DIRT, a general method for distinguishing between specific and nonspecific protein interactions. *J Proteome Res.* 2005. doi:10.1021/pr050225e
160. Wang X, Huang L. Identifying dynamic interactors of protein complexes by quantitative mass spectrometry. *Mol Cell Proteomics.* 2008. doi:10.1074/mcp.M700261-MCP200
161. Marcilla M, Albar JP. Quantitative proteomics: A strategic ally to map protein interaction networks. *IUBMB Life.* 2013. doi:10.1002/iub.1081
162. Torrente MP, Zee BM, Young NL, et al. Proteomic interrogation of human chromatin. *PLoS One.* 2011. doi:10.1371/journal.pone.0024747
163. Noberini R, Sigismondo G, Bonaldi T. The contribution of mass spectrometry-based proteomics to understanding epigenetics. *Epigenomics.* 2016. doi:10.2217/epi.15.108
164. Shiio Y, Eisenman RN, Yi EC, Donohoe S, Goodlett DR, Aebersold R. Quantitative proteomic analysis of chromatin-associated factors. *J Am Soc Mass Spectrom.* 2003. doi:10.1016/S1044-0305(03)00204-6
165. Dignani JD, Lebovitz RM, Roeder RG. Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. *Nucleic Acids Res.* 1983. doi:10.1093/nar/11.5.1475
166. Wang X, Andreassen PR, D'Andrea AD. Functional Interaction of Monoubiquitinated FANCD2 and BRCA2/FANCD1 in Chromatin. *Mol Cell Biol.* 2004. doi:10.1128/mcb.24.13.5850-5862.2004

167. Lelong C, Chevallet M, Diemer H, Luche S, Van Dorsselaer A, Rabilloud T. Improved proteomic analysis of nuclear proteins, as exemplified by the comparison of two myeloid cell lines nuclear proteomes. *J Proteomics*. 2012. doi:10.1016/j.jprot.2012.09.034
168. Alajem A, Biran A, Harikumar A, et al. Differential association of chromatin proteins identifies BAF60a/SMARCD1 as a regulator of embryonic stem cell differentiation. *Cell Rep*. 2015. doi:10.1016/j.celrep.2015.02.064
169. Kustatscher G, Hégarat N, Wills KLH, et al. Proteomics of a fuzzy organelle: Interphase chromatin. *EMBO J*. 2014. doi:10.1002/emboj.201387614
170. Sutherland BW, Toews J, Kast J. Utility of formaldehyde cross-linking and mass spectrometry in the study of protein-protein interactions. *J Mass Spectrom*. 2008. doi:10.1002/jms.1415
171. Garcia BA, Mollah S, Ueberheide BM, et al. Chemical derivatization of histones for facilitated analysis by mass spectrometry. *Nat Protoc*. 2007. doi:10.1038/nprot.2007.106
172. Sidoli S, Yuan ZF, Lin S, et al. Drawbacks in the use of unconventional hydrophobic anhydrides for histone derivatization in bottom-up proteomics PTM analysis. *Proteomics*. 2015. doi:10.1002/pmic.201400483
173. Maile TM, Izrael-Tomasevic A, Cheung T, et al. Mass Spectrometric Quantification of Histone Post-translational Modifications by a Hybrid Chemical Labeling Method. *Mol Cell Proteomics*. 2015. doi:10.1074/mcp.O114.046573
174. Molden RC, Garcia BA. Middle-down and top-down mass spectrometric analysis of co-occurring histone modifications. *Curr Protoc Protein Sci*. 2014. doi:10.1002/0471140864.ps2307s77
175. Karch KR, Sidoli S, Garcia BA. Identification and Quantification of Histone PTMs Using High-Resolution Mass Spectrometry. In: *Methods in Enzymology*. ; 2016. doi:10.1016/bs.mie.2015.12.007
176. Moradian A, Kalli A, Sweredoski MJ, Hess S. The top-down, middle-down, and bottom-up mass spectrometry approaches for characterization of histone variants and their post-translational modifications. *Proteomics*. 2014. doi:10.1002/pmic.201300256
177. Nicosia L, Noberini R, Soldi M, et al. Mass Spectrometry and Epigenetics. In: *Handbook of Nutrition, Diet, and Epigenetics*. ; 2019. doi:10.1007/978-3-319-55530-0_115
178. Pesavento JJ, Mizzen CA, Kelleher NL. Quantitative analysis of modified proteins and their positional isomers by tandem mass spectrometry: Human histone H4. *Anal Chem*. 2006. doi:10.1021/ac0600050
179. Leroy G, Dimaggio PA, Chan EY, et al. A quantitative atlas of histone modification signatures from human cancer cells. *Epigenetics and Chromatin*. 2013. doi:10.1186/1756-8935-6-20

180. Henry RA, Mancuso P, Kuo YM, et al. Interaction with the DNA Repair Protein Thymine DNA Glycosylase Regulates Histone Acetylation by p300. *Biochemistry*. 2016. doi:10.1021/acs.biochem.6b00841
181. Noberini R, Bonaldi T. A Super-SILAC Strategy for the Accurate and Multiplexed Profiling of Histone Posttranslational Modifications. In: *Methods in Enzymology*. ; 2017. doi:10.1016/bs.mie.2016.09.036
182. Noberini R, Uggetti A, Pruneri G, Minucci S, Bonaldi T. Pathology Tissue-quantitative Mass Spectrometry Analysis to Profile Histone Post-translational Modification Patterns in Patient Samples. *Mol Cell Proteomics*. 2016. doi:10.1074/mcp.M115.054510
183. Nie L, Shuai L, Zhu M, et al. The landscape of histone modifications in a high-fat diet-induced obese (DIO) mouse model. *Mol Cell Proteomics*. 2017. doi:10.1074/mcp.M117.067553
184. Darwanto A, Curtis MP, Schrag M, et al. A modified “cross-talk” between histone H2B Lys-120 ubiquitination and H3 Lys-79 methylation. *J Biol Chem*. 2010. doi:10.1074/jbc.M110.126813
185. Gao J, Liao R, Yu Y, et al. Absolute quantification of histone PTM marks by MRM-based LC-MS/MS. *Anal Chem*. 2014. doi:10.1021/ac502333a
186. Roussel MJS, Lanotte M. Maturation sensitive and resistant t(15;17) NB4 cell lines as tools for APL physiopathology: Nomenclature of cells and repertory of their known genetic alterations and phenotypes. *Oncogene*. 2001. doi:10.1038/sj.onc.1204863
187. Lanotte M, Martin-Thouvenin V, Najman S, Balerini P, Valensi F, Berger R. NB4, a maturation inducible cell line with t(15;17) marker isolated from a human acute promyelocytic leukemia (M3). *Blood*. 1991.
188. Ong SE, Mann M. A practical recipe for stable isotope labeling by amino acids in cell culture (SILAC). *Nat Protoc*. 2006. doi:10.1038/nprot.2006.427
189. Konze KD, Ma A, Li F, et al. An orally bioavailable chemical probe of the lysine methyltransferases EZH2 and EZH1. *ACS Chem Biol*. 2013. doi:10.1021/cb400133j
190. Baghirova S, Hughes BG, Hendzel MJ, Schulz R. Sequential fractionation and isolation of subcellular proteins from tissue or cultured cells. *MethodsX*. 2015. doi:10.1016/j.mex.2015.11.001
191. Shevchenko A, Tomas H, Havliš J, Olsen J V., Mann M. In-gel digestion for mass spectrometric characterization of proteins and proteomes. *Nat Protoc*. 2007. doi:10.1038/nprot.2006.468
192. Rappsilber J, Mann M, Ishihama Y. Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. *Nat Protoc*. 2007. doi:10.1038/nprot.2007.261
193. Hu Q, Noll RJ, Li H, Makarov A, Hardman M, Cooks RG. The Orbitrap: A new mass spectrometer. *J Mass Spectrom*. 2005. doi:10.1002/jms.856

194. Michalski A, Damoc E, Hauschild JP, et al. Mass spectrometry-based proteomics using Q exactive, a high-performance benchtop quadrupole orbitrap mass spectrometer. *Mol Cell Proteomics*. 2011. doi:10.1074/mcp.M111.011015
195. Cox J, Neuhauser N, Michalski A, Scheltema RA, Olsen J V., Mann M. Andromeda: A peptide search engine integrated into the MaxQuant environment. *J Proteome Res*. 2011. doi:10.1021/pr101065j
196. Reimand J, Kull M, Peterson H, Hansen J, Vilo J. G:Profiler-a web-based toolset for functional profiling of gene lists from large-scale experiments. *Nucleic Acids Res*. 2007. doi:10.1093/nar/gkm226
197. Kingston RE, Chen CA, Rose JK. Calcium Phosphate Transfection. *Curr Protoc Mol Biol*. 2003. doi:10.1002/0471142727.mb0901s63
198. Finer MH, Dull TJ, Qin L, Parson D, Roberts MR. kat: A high-efficiency retroviral transduction system for primary human T lymphocytes. *Blood*. 1994.
199. Barozzi I, Termanini A, Minucci S, Natoli G. Fish the ChIPs: A pipeline for automated genomic annotation of ChIP-Seq data. *Biol Direct*. 2011. doi:10.1186/1745-6150-6-51
200. Langmead B, Trapnell C, Pop M, Salzberg SL. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol*. 2009. doi:10.1186/gb-2009-10-3-r25
201. Zhang Y, Liu T, Meyer CA, et al. Model-based analysis of ChIP-Seq (MACS). *Genome Biol*. 2008. doi:10.1186/gb-2008-9-9-r137
202. Cesaroni M, Cittaro D, Brozzi A, Pelicci PG, Luzi L. CARPET: A web-based package for the analysis of ChIP-chip and expression tiling data. *Bioinformatics*. 2008. doi:10.1093/bioinformatics/btn542
203. Ramírez F, Ryan DP, Grüning B, et al. deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res*. 2016. doi:10.1093/nar/gkw257
204. Dobin A, Davis CA, Schlesinger F, et al. STAR: Ultrafast universal RNA-seq aligner. *Bioinformatics*. 2013. doi:10.1093/bioinformatics/bts635
205. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*. 2014. doi:10.1186/s13059-014-0550-8
206. Chen EY, Tan CM, Kou Y, et al. Enrichr: Interactive and collaborative HTML5 gene list enrichment analysis tool. *BMC Bioinformatics*. 2013. doi:10.1186/1471-2105-14-128
207. Bonaldi T, Imhof A, Regula JT. A combination of different mass spectroscopic techniques for the analysis of dynamic changes of histone modifications. In: *Proteomics*. ; 2004. doi:10.1002/pmic.200300743
208. Cuomo A, Moretti S, Minucci S, Bonaldi T. SILAC-based proteomic analysis to dissect the “histone modification signature” of human breast cancer cells. *Amino*

209. Ruepp A, Waegele B, Lechner M, et al. CORUM: The comprehensive resource of mammalian protein complexes-2009. *Nucleic Acids Res.* 2009. doi:10.1093/nar/gkp914
210. Morey L, Brenner C, Fazi F, et al. MBD3, a Component of the NuRD Complex, Facilitates Chromatin Alteration and Deposition of Epigenetic Marks. *Mol Cell Biol.* 2008. doi:10.1128/MCB.00467-08
211. Senyuk V, Chakraborty S, Mikhail FM, Zhao R, Chi Y, Nucifora G. The leukemia-associated transcription repressor AML1/MDS1/EV11 requires CtBP to induce abnormal growth and differentiation of murine hematopoietic cells. *Oncogene.* 2002. doi:10.1038/sj/onc/1205436
212. Khan DH, Gonzalez C, Cooper C, et al. RNA-dependent dynamic histone acetylation regulates MCL1 alternative splicing. *Nucleic Acids Res.* 2014. doi:10.1093/nar/gkt1134
213. Hnilicová J, Hozeifi S, Dušková E, Icha J, Tománková T, Staněk D. Histone deacetylase activity modulates alternative splicing. *PLoS One.* 2011. doi:10.1371/journal.pone.0016727
214. Alló M, Buggiano V, Fededa JP, et al. Control of alternative splicing through siRNA-mediated transcriptional gene silencing. *Nat Struct Mol Biol.* 2009. doi:10.1038/nsmb.1620
215. Fenger-Grøn M, Fillman C, Norrild B, Lykke-Andersen J. Multiple processing body factors and the ARE binding protein TTP activate mRNA decapping. *Mol Cell.* 2005. doi:10.1016/j.molcel.2005.10.031
216. Chai P, Tian J, Zhao D, et al. GSE1 negative regulation by miR-489-5p promotes breast cancer cell proliferation and invasion. *Biochem Biophys Res Commun.* 2016. doi:10.1016/j.bbrc.2016.01.168
217. Ding K, Tan S, Huang X, et al. GSE1 predicts poor survival outcome in gastric cancer patients by SLC7A5 enhancement of tumor growth and metastasis. *J Biol Chem.* 2018. doi:10.1074/jbc.RA117.001103
218. Huang M, Tailor J, Zhen Q, et al. Engineering Genetic Predisposition in Human Neuroepithelial Stem Cells Recapitulates Medulloblastoma Tumorigenesis. *Cell Stem Cell.* 2019.
219. Van Der Meer LT, Jansen JH, Van Der Reijden BA. Gfi1 and Gfi1b: Key regulators of hematopoiesis. *Leukemia.* 2010. doi:10.1038/leu.2010.195
220. Möröy T, Vassen L, Wilkes B, Khandanpour C. From cytopenia to leukemia: The role of Gfi1 and Gfi1b in blood formation. *Blood.* 2015. doi:10.1182/blood-2015-06-655043
221. Hakimi MA, Dong Y, Lane WS, Speicher DW, Shiekhattar R. A candidate X-linked mental retardation gene is a component of a new family of histone deacetylase-containing complexes. *J Biol Chem.* 2003. doi:10.1074/jbc.M208992200

222. Yamamoto R, Kawahara M, Ito S, et al. Selective dissociation between LSD1 and GFI1B by a LSD1 inhibitor NCD38 induces the activation of ERG super-enhancer in erythroleukemia cells. *Oncotarget*. 2018. doi:10.18632/oncotarget.24774
223. Laurent B, Ruitu L, Murn J, et al. A specific LSD1/KDM1A isoform regulates neuronal differentiation through H3K9 demethylation. *Mol Cell*. 2015;57(6):957-970.
224. Velinder M, Singer J, Bareyan D, et al. GFI1 functions in transcriptional control and cell fate determination require SNAG domain methylation to recruit LSD1. *Biochem J*. 2016. doi:10.1042/bcj20160558
225. Salvesen GS. Caspases: Opening the boxes and interpreting the arrows. *Cell Death Differ*. 2002. doi:10.1038/sj.cdd.4400963
226. Kurotaki D, Nakabayashi J, Nishiyama A, et al. Transcription Factor IRF8 Governs Enhancer Landscape Dynamics in Mononuclear Phagocyte Progenitors. *Cell Rep*. 2018. doi:10.1016/j.celrep.2018.02.048
227. Solomon LA, Podder S, He J, et al. Coordination of Myeloid Differentiation with Reduced Cell Cycle Progression by PU.1 Induction of MicroRNAs Targeting Cell Cycle Regulators and Lipid Anabolism. *Mol Cell Biol*. 2017. doi:10.1128/mcb.00013-17
228. Bland JB, Peralta JR, Tse WT. Elevated EGR2 Expression Provides a Potential Link Between Cell Cycle Arrest and Induced Differentiation in Myeloid Progenitor Cells. *Blood*. 2014.
229. Xiong Y, Hannon GJ, Zhang H, Casso D, Kobayashi R, Beach D. P21 is a universal inhibitor of cyclin kinases. *Nature*. 1993. doi:10.1038/366701a0
230. Deng C, Zhang P, Wade Harper J, Elledge SJ, Leder P. Mice Lacking p21 CIP1/WAF1 undergo normal development, but are defective in G1 checkpoint control. *Cell*. 1995. doi:10.1016/0092-8674(95)90039-X
231. Gartel AL, Radhakrishnan SK. Lost in transcription: p21 repression, mechanisms, and consequences. *Cancer Res*. 2005. doi:10.1158/0008-5472.CAN-04-3995
232. Vaure C, Liu Y. A comparative review of toll-like receptor 4 expression and functionality in different animal species. *Front Immunol*. 2014. doi:10.3389/fimmu.2014.00316
233. Song JJ, Lee YJ. Differential cleavage of Mst1 by caspase-7/-3 is responsible for TRAIL-induced activation of the MAPK superfamily. *Cell Signal*. 2008. doi:10.1016/j.cellsig.2008.01.001
234. Leonard WJ, Lin JX. Cytokine receptor signaling pathways. *J Allergy Clin Immunol*. 2000. doi:10.1067/mai.2000.106899
235. Zee BM, Young NL, Garcia BA. Quantitative Proteomic Approaches to Studying Histone Modifications. *Curr Chem Genomics*. 2011. doi:10.2174/1875397301005010106
236. Sidoli S, Schwämmle V, Ruminowicz C, et al. Middle-down hybrid

- chromatography/tandem mass spectrometry workflow for characterization of combinatorial post-translational modifications in histones. *Proteomics*. 2014. doi:10.1002/pmic.201400084
237. Schmitges FW, Prusty AB, Faty M, et al. Histone Methylation by PRC2 Is Inhibited by Active Chromatin Marks. *Mol Cell*. 2011. doi:10.1016/j.molcel.2011.03.025
238. Musselman CA, Avvakumov N, Watanabe R, et al. Molecular basis for H3K36me3 recognition by the Tudor domain of PHF1. *Nat Struct Mol Biol*. 2012. doi:10.1038/nsmb.2435
239. Yuan W, Xu M, Huang C, Liu N, Chen S, Zhu B. H3K36 methylation antagonizes PRC2-mediated H3K27 methylation. *J Biol Chem*. 2011. doi:10.1074/jbc.M110.194027
240. Zheng Y, Sweet SMM, Popovic R, et al. Total kinetic analysis reveals how combinatorial methylation patterns are established on lysines 27 and 36 of histone H3. *Proc Natl Acad Sci*. 2012. doi:10.1073/pnas.1205707109
241. Shimazu T, Horinouchi S, Yoshida M. Multiple histone deacetylases and the CREB-binding protein regulate pre-mRNA 3'-end processing. *J Biol Chem*. 2007. doi:10.1074/jbc.M609745200
242. Ray SK, Li HJ, Metzger E, Schule R, Leiter AB. CtBP and Associated LSD1 Are Required for Transcriptional Activation by NeuroD1 in Gastrointestinal Endocrine Cells. *Mol Cell Biol*. 2014. doi:10.1128/MCB.01600-13
243. Baron R, Binda C, Tortorici M, McCammon JA, Mattevi A. Molecular mimicry and ligand recognition in binding and catalysis by the histone demethylase LSD1-CoREST complex. *Structure*. 2011. doi:10.1016/j.str.2011.01.001
244. Santos MA, Faryabi RB, Ergen A V., et al. DNA-damage-induced differentiation of leukaemic cells as an anti-cancer barrier. *Nature*. 2014. doi:10.1038/nature13483
245. Cazzalini O, Scovassi AI, Savio M, Stivala LA, Prosperi E. Multiple roles of the cell cycle inhibitor p21CDKN1A in the DNA damage response. *Mutat Res - Rev Mutat Res*. 2010. doi:10.1016/j.mrrev.2010.01.009
246. Jin YH, Yoo KJ, Lee YH, Lee SK. Caspase 3-mediated cleavage of p21(WAF1/CIP1) associated with the cyclin A-cyclin-dependent kinase 2 complex is a prerequisite for apoptosis in SK-HEP-1 cells. *J Biol Chem*. 2000. doi:10.1074/jbc.M001902200
247. Rosa A, Ballarino M, Sorrentino A, et al. The interplay between the master transcription factor PU.1 and miR-424 regulates human monocyte/macrophage differentiation. *Proc Natl Acad Sci U S A*. 2007. doi:10.1073/pnas.0706963104
248. Wei W, Liu C, Qin D, et al. Targeting peroxiredoxin i potentiates 1,25-dihydroxyvitamin D3-induced cell differentiation in leukemia cells. *Mol Med Rep*. 2016. doi:10.3892/mmr.2016.4787
249. Yáñez A, Ng MY, Hassanzadeh-Kiabi N, Goodridge HS. IRF8 acts in lineage-committed rather than oligopotent progenitors to control neutrophil vs monocyte production. *Blood*. 2015. doi:10.1182/blood-2014-09-600833

250. Sheng W, LaFleur MW, Nguyen TH, et al. LSD1 Ablation Stimulates Anti-tumor Immunity and Enables Checkpoint Blockade. *Cell*. 2018. doi:10.1016/j.cell.2018.05.052
251. Jin Y, Huo B, Fu X, et al. LSD1 collaborates with EZH2 to regulate expression of interferon-stimulated genes. *Biomed Pharmacother*. 2017. doi:10.1016/j.biopha.2017.01.055
252. Jin Y, Huo B, Fu X, et al. LSD1 knockdown reveals novel histone lysine methylation in human breast cancer MCF-7 cells. *Biomed Pharmacother*. 2017. doi:10.1016/j.biopha.2017.05.106
253. Tsai MC, Manor O, Wan Y, et al. Long noncoding RNA as modular scaffold of histone modification complexes. *Science (80-)*. 2010. doi:10.1126/science.1192002
254. Min GL, Villa R, Trojer P, et al. Demethylation of H3K27 regulates polycomb recruitment and H2A ubiquitination. *Science (80-)*. 2007. doi:10.1126/science.1149042
255. Rafiee MR, Girardot C, Sigismondo G, Krijgsveld J. Expanding the Circuitry of Pluripotency by Selective Isolation of Chromatin-Associated Proteins. *Mol Cell*. 2016. doi:10.1016/j.molcel.2016.09.019
256. Wen S, Wang J, Liu P, et al. Novel combination of histone methylation modulators with therapeutic synergy against acute myeloid leukemia in vitro and in vivo. *Cancer Lett*. 2018. doi:10.1016/j.canlet.2017.10.015

8. APPENDIX

Appendix 1: List of putative LSD1 interactors in NB4 cells

Gene names	Log2H/L IP- LSD1 DIR1	Log2L/H IP- LSD1 REV	Log2H/L IP- LSD1 DIR2	Known LSD1 Interactors
ADNP	4,1	4,1	4,1	
AQR	3,3	3,2	3,2	
ASCC3	2,7	3,1	3,0	
BUB3	2,6	2,7	2,9	
CAD	2,6	3,1	3,0	
CASP8AP2	4,0	3,4	4,4	
CDYL	4,0	3,7	3,9	
CHD4	3,9	3,8	4,0	YES
CIT	2,6	3,2	2,5	
CPSF1	5,4	5,3	5,2	
CPSF3	5,2	5,0	4,9	
CPSF7	4,2	3,9	3,0	
CRNKL1	3,4	3,1	3,5	
CSNK2A2	3,0	2,9	3,7	
CSTF1	3,1	2,8	2,6	
CSTF3	3,4	3,6	3,4	
CTBP1	4,0	3,8	4,2	
CTBP2	4,0	3,7	4,0	
CWC22	3,7	3,1	2,6	
DDX21	3,2	3,0	3,0	
DDX3X;DDX3Y	3,9	3,9	3,7	
DDX5	3,3	3,1	3,4	
DHX15	2,8	2,6	2,6	
DNTTIP1	2,9	3,6	2,9	
EDC4	3,4	4,4	3,7	
EFTUD2	3,3	3,2	3,3	
EHMT1	3,9	3,3	3,2	
EIF4A3	3,6	3,5	3,6	
ETV6	3,8	3,5	4,0	
FIP1L1	4,1	4,6	3,2	
GATAD2A	3,7	3,1	3,9	
GFI1	3,1	3,1	2,8	YES
GOPC	4,0	3,4	4,4	
GSE1	2,5	2,5	2,6	YES
GTF2I	3,8	3,8	3,1	YES
GTF3C1	3,4	3,5	3,7	
GTF3C2	3,6	3,6	3,8	
GTF3C3	3,6	3,4	3,4	
GTF3C4	3,1	3,2	3,5	
H2AFY	3,7	2,5	3,3	
H3F3B	3,7	4,0	3,4	

HADHA	2,1	2,7	2,5	
HDAC1	3,6	3,4	3,8	YES
HDAC2	3,5	3,5	3,9	YES
HIRA	3,8	4,2	3,8	
HIST1H2BN	2,9	3,9	3,6	
HIST1H4A	3,2	4,0	3,3	
HMG20A	2,6	2,6	2,7	YES
HMG20B	3,3	3,0	3,5	YES
HNRNPF	2,8	2,7	2,8	
HNRNPH1	2,3	2,5	2,9	
HNRNPM	2,6	2,3	2,7	
HNRNPU	3,3	3,4	3,0	
HSPA5	2,6	3,0	2,6	
HSPA8	2,8	2,9	2,9	
IKZF1	4,0	3,5	4,3	
INTS1	3,0	2,6	3,1	
INTS10	3,2	2,9	3,4	
KDM1A	3,7	3,5	3,8	BAIT
KEAP1	2,9	3,0	2,9	
KIAA1429	4,0	3,8	3,3	
MCM5	2,5	2,5	2,7	
MCM7	2,5	2,7	2,6	
MRPL15	4,1	4,8	5,6	
MRPL19	5,7	6,0	2,8	
MRPL24	4,1	4,5	4,2	
MRPL37	4,6	5,0	6,2	
MRPL4	4,2	3,6	3,8	
MRPL44	3,2	2,5	3,0	
MRPS2	2,4	4,2	2,8	
MRPS22	4,0	3,4	4,0	
MRPS30	3,7	4,1	5,0	
MTA2	3,3	3,2	3,5	
NPAT	3,7	2,9	3,1	
NPM1	3,6	3,2	3,8	
PELP1	3,6	4,0	3,3	
PGAM5	3,1	3,4	3,0	
PHF21A	3,5	3,1	3,7	YES
PLK1	2,8	3,0	3,0	
PLRG1	3,3	2,9	3,1	
PNN	3,6	3,6	3,3	
POLR2A	4,0	3,3	4,2	
POLR2B	3,2	2,6	3,4	
PRPF19	3,1	2,8	3,1	
PRPF6	3,4	3,3	3,4	
PRPF8	3,4	3,2	3,4	
RBBP4	3,4	3,1	3,4	
RBBP5	3,3	3,6	2,7	
RBBP7	2,0	2,7	3,8	
RBM15	4,0	3,5	3,3	

RCOR1	3,6	3,6	3,8	YES
RCOR2	2,4	3,7	2,7	YES
RCOR3	3,7	3,7	3,3	YES
RFC2	2,9	3,0	3,3	
RFC4	2,7	3,0	2,9	
RNPS1	3,5	3,6	2,7	
RPS3	3,0	4,1	2,7	
RREB1	4,0	3,8	4,3	
RUVBL1	2,8	2,9	3,0	
RUVBL2	2,3	2,8	3,2	
SCAF11	2,9	2,5	2,7	
SF3A1	3,0	2,7	3,1	
SF3A3	2,3	2,7	3,2	
SF3B1	3,1	3,0	3,2	
SF3B3	3,1	3,0	3,0	
SFMBT2	3,1	3,7	3,8	
SIN3A	3,2	3,0	3,4	
SKIV2L2	2,6	2,7	3,0	
SMARCC2	3,1	3,3	3,9	
SMU1	3,3	3,1	3,0	
SNRNP200	3,4	3,2	3,4	
SNRNP70	3,0	3,6	3,3	
SNRPB;SNRPN	3,2	3,4	2,5	
SRRM2	3,2	3,6	3,2	
SRRT	3,6	3,5	3,3	
SRSF1	4,4	4,5	4,1	
SRSF10	3,8	3,8	2,7	
SRSF2	2,9	3,9	2,9	
SRSF3	3,6	4,0	2,9	
SRSF5	3,9	2,5	3,8	
SRSF6	4,1	3,8	3,9	
SRSF7	4,1	3,9	3,7	
SRSF9	4,1	4,2	2,7	
SUPT6H	3,1	4,0	2,5	
SYMPK	3,7	3,1	3,4	
TFIP11	2,9	2,3	2,9	
TOP2B	2,4	2,8	2,7	
TRA2A	3,9	3,8	3,6	
TRA2B	4,0	4,0	3,8	
TUBB	2,8	3,3	3,1	
TUBB4B	2,5	2,5	3,0	
U2AF1	3,3	3,4	3,5	
U2AF2	3,4	3,7	3,7	
WDR33	4,1	5,2	3,8	
WTAP	3,4	3,2	3,5	
XAB2	3,3	3,0	3,3	
YBX3	4,3	4,7	3,8	YES
ZC3H18	3,3	3,5	3,1	
ZEB2	3,9	3,2	3,9	

ZMYM2	4,2	4,2	4,4	YES
ZMYM3	3,9	4,1	4,0	YES
ZMYM4	2,9	4,1	4,3	
ZMYND8	4,8	4,4	4,2	
ZNF217	3,9	3,5	4,2	YES
ZNF516	3,7	3,5	3,8	
ZNF592	5,0	4,3	3,7	
ZNF687	4,2	4,1	3,7	

Appendix 2: A Chromatin Proteomics (ChroP) approach dissects the histone post-translational modifications (PTMs) enriched at enhancers of mouse macrophages

BACKGROUND

Gene transcription is tightly regulated in time and space by *cis*-regulatory elements such as promoters and enhancers which are proximal and distal regulatory regions of coding genes, respectively. The first are located upstream of the transcription start sites (TSS), while the second activate transcription of their target genes at large distance (up to many hundred kilobases) and in an orientation-independent manner^{1,2}. Enhancers recruit transcriptional co-regulators that synergise to enforce specific transcriptional programs in various processes, like development and response to environmental signals. Enhancer activation usually involves the interaction of multiple transcription factors (TFs), including ubiquitous, lineage-determining (LD) and signal-dependent (SD) TFs^{3,4}. For instance, LD-TFs recognize and bind to specific DNA motifs, generating a nucleosome-free region that can be bound by other TFs and co-activators like histone modifiers or chromatin remodellers that ultimately regulate transcription^{5,6}.

Genome-wide studies demonstrated that enhancers are characterized by a specific histone PTM signature, including high levels of H3K4me1 and low levels of H3K4me3 (H3K4me1^{Hi}/K4me3^{Low})⁷. The fraction of enhancers known as active enhancers are also marked by high levels of H3K27ac⁸. Chromatin modifications exert an important function for enhancer regulation, by either acting as docking sites for SD-TFs or facilitating the access to other TFs. These mechanisms promote the assembly of transcriptional complexes and long-range chromatin interactions between regulatory elements⁹ that activate specific transcriptional responses^{4,10}.

Characterization of genomic features associated with enhancer regions has been enabled by genome-wide approaches like chromatin immuno-precipitation-sequencing (ChIP-seq) analysis. Nevertheless, a global and unbiased study of the complete repertoire of histone modifications specifically associated with these regions is still missing ¹¹. This information could allow identifying novel histone marks distinguishing distinct subpopulations of genomic enhancers and to better decode the molecular mechanisms underlying the different classes of enhancers.

Native chromatin proteomics (N-ChroP) is an approach that combines native ChIP with mass-spectrometry (MS) and allows determining the histone PTMs differentially associated with different genomic regions ¹². In this second project in which I was involved during my PhD, I employed the N-ChroP approach to study combination of histone modifications at different chromatin regions, namely enhancers, promoters and heterochromatin regions. Focusing on enhancers, I performed the same analysis on the subpopulation of active enhancers. The study has been carried out in RAW264.7 murine macrophage cells as model system, since they have been well-studied at both genomic and epigenomic levels and elicit a rapid and dynamic stimulation of hundreds of inflammatory genes involved in the innate and adaptive immune response. The perspective is to follow-up the results collected so far by analysing the changes of histone modifications upon lipopolysaccharide (LPS) treatment which activate inflammatory genes in RAW cells.

MATERIALS AND METHODS

Cell culture

RAW 264.7 mouse macrophage-like cells were cultured in high-glucose Dulbecco's modified Eagle's medium (DMEM) with 10% low endotoxin of fetal bovine serum (FBS),

1% glutamine and 1% Penicillin/Streptomycin (P/S). Cultures were maintained in a humidified tissue culture incubator at 37°C in 5% CO₂.

Native chromatin immunoprecipitation (N-ChIP)

100×10⁶ of RAW264.7 cells were re-suspended in lysis buffer (10% sucrose, 0.5 mM EGTA, 15 mM NaCl, 60 mM KCl, 15 mM HEPES, 0.5% Triton-X 100, 0.5 mM PMSF, 1 mM DTT, 5 mM sodium butyrate, 5 µg/ml aprotinin, 5 µg/ml leupeptin, 0.03 µg/µl spermine, 0.03 µg/µl spermidine), and nuclei were separated from the cytoplasm by centrifugation at 3750 rpm for 30 minutes on sucrose cushions at 4°C. Then, nuclei were washed twice with phosphate buffered saline (PBS) and re-suspended in micrococcal nuclease (MNase) digestion buffer (0.32 M sucrose, 50 mM Tris-HCl [pH 7.6], 4 mM MgCl₂, 1 mM CaCl₂ and 0.1 mM PMSF). MNase (Roche, 10107921001) was added to the solution at a final concentration of 0.01 U/µl and digestion was performed at 37°C for 90 minutes. The reaction was, then, stopped with 1 mM EDTA and cooling the samples on ice. At this point, nuclei were centrifuged at 10000 rpm for 10 minutes and the soluble fraction of chromatin (S1) including smaller fragments (mono-nucleosomes) were collected in the supernatant and separated from the bigger ones (contained in the pellets). Pellets were, then, re-suspended in dialysis buffer (10 mM Tris-HCl [pH 7.6], 1 mM EDTA, 0.5 mM PMSF, 5 mM sodium butyrate and 1X Roche protease inhibitors) and dialyzed overnight at 4°C in dialysis tubes (cut off 3.5 kDa). The second soluble fraction of chromatin (S2), containing large chromatin fragments (from di- to hepta- nucleosomes), was achieved after dialysis and gathered as supernatant after centrifugation at 10000 rpm for 10 minutes at 4°C. The entire S1 fraction was mixed with 5% of the S2 fraction and diluted in ChIP incubation buffer (50 mM Tris-HCl [pH 7.6], 50 mM NaCl and 5 mM EDTA) supplemented with 0.5 mM PMSF and 1X Roche protease inhibitors. A fraction (1/20) of Chromatin Input was collected before adding 30 µg of antibody to each sample. Meanwhile, 300 µl of Dynabeads protein G (Invitrogen

10004D) were blocked in 0.5% BSA in PBS overnight at 4°C. Blocked beads were added to the chromatin samples and incubated for 3 hours at 4°C on a rotating wheel. Beads were, then, washed three times with Buffer Washing A (50 mM Tris-HCl [pH 7.6], 10 mM EDTA, 75 mM NaCl), once with Buffer Washing B (50 mM Tris-HCl [pH 7.6], 10 mM EDTA, 125 mM NaCl) and once with Buffer Washing C (50 mM Tris-HCl [pH 7.6], 10 mM EDTA, 175 mM NaCl), all supplemented with protease inhibitors. Elution of the ChIP samples was performed by adding the LDS sample buffer (Invitrogen NP0007) supplemented with 100 mM DTT to the beads and incubated for 5 minutes at 95°C. Immuno-precipitated proteins including histones were separated on 4-12% SDS-PAGE gradient gel (Invitrogen, NP0335BOX) and stained with colloidal coomassie (Invitrogen, LC6025). The antibodies used for the N-ChroP approach were: H3K4me1 (AB8895, Abcam), H3K4me3 (39159, active motif), H3K9me3 (AB8898, Abcam) and H3K27ac (AB4729, Abcam).

In-gel protease digestion of H3 and H4 histones

Gel bands corresponding to histone H3 and H4 were cut in small pieces and de-stained with 50% acetonitrile (ACN) alternated to H₂O and 100% ACN. Histones were, then, chemically acylated by incubation with D₆-acetic anhydride (Sigma Aldrich, 175641) diluted 1:9 in 1 M NH₄HCO₃ and saturated sodium acetate (CH₃COONa) solution as catalyzer for 3 hours at 37°C in a thermomixer at 1400 rpm. Then, gel slices were washed with digestion buffer (50 mM NH₄HCO₃ in H₂O) alternated to increasing concentration of ACN (from 50% to 100%). Digestion was performed with 100 ng/μl of trypsin (Promega, V5113) in digestion buffer at 37°C overnight, in order to obtain an “ArgC-like” digestion. Indeed, in this case, trypsin cleaves at the C-terminus of only arginine residues since lysine residues have been chemically-derivatized and are not recognized by the enzyme. Finally, digested peptides were extracted with 5% formic acid (FA) alternated to ACN 100% and, then, desalted through a combination of micro-columns based on reverse-phase C18/carbon ‘sandwich’

system and ion-exchange (SCX) chromatography¹³. After loading on C18/carbon and SCX stage tips, peptides were eluted with Buffer B (80% ACN, 0.5% acetic acid) and SCX elution Buffer (5% NH₄OH, 30% methanol), respectively. Eluted peptides were lyophilized, re-suspended in 1% TFA and subjected to LC-MS/MS analysis.

For the immuno-precipitated histones from the H3K27ac N-ChroP, the protocol of sample preparation prior to the LC-MS/MS analysis was similar to that described above, with some modifications. In particular, lysines were chemically derivatized with propionic anhydride instead of D₆-acetic anhydride. In this case, saturated sodium propionate (C₂H₅COO) solution was used as catalyzer of the reaction. After digestion and extraction, histone peptides were lyophilized, re-suspended in aqueous solution containing 100 mM triethylammonium bicarbonate and 30 mM phenyl-isocyanate (PIC) and incubated for 60 minutes at 37°C. The reaction was stopped with 1% TFA and then peptides were desalted by using the C18 stage tip prior to elution and LC-MS/MS analysis.

LC-MS/MS analysis

Histone peptides were analyzed by online nano-flow liquid chromatography tandem mass spectrometry using an EASY-nLC™ 1200 (Thermo Fisher Scientific, Odense, Denmark) coupled to a hybrid quadrupole/Orbitrap mass spectrometer (Q Exactive HF, Thermo Fisher Scientific) through a nanoelectrospray ion source. The nano-LC system worked with one column set up with an EASY-Spray™ LC Columns (25-cm length, 75-µm inner diameter) packed with C18 reversed-phase resin (1.9 µm). Solvent A was 0.1% FA in ddH₂O and solvent B was 80% ACN with 0.1% FA. Peptides were injected at a flow rate of 500 nL/min and separated with a gradient of 10-45% solvent B over 100 minutes, followed by a gradient of 45-95% for 2 and 95% over 3 minutes at a flow rate of 250 nL/min. The Q Exactive HF instrument was operated in the data-dependent acquisition (DDA) mode to automatically switch between full scan MS and MS/MS acquisition. Survey full scan MS spectra (m/z 300-

1650) were acquired in the Orbitrap detector with a resolution of 60 000 at m/z 400. The 10 most intense peptide ions with charge states ≥ 2 were sequentially isolated to a target value for MS1 of 3×10^6 and fragmented by HCD with a normalized collision energy setting of 28%. The maximum permitted ion accumulation times were 20 ms for full scans and 80 ms for MS/MS and the target value for MS/MS was fixed to 1×10^6 . The dynamic exclusion time was set to 20 s, and the standard mass spectrometric conditions were: spray voltage of 1.8 kV, no sheath and auxiliary gas flow.

Histone PTM analysis from MS-data

MaxQuant software v.1.5.2.8 was used for the analysis of MS data, for both protein and peptides identification including as variable modifications mono-, di- and tri-methyl lysine, lysine acetylation, D₆-acylation in the case of peptides chemically modified with D₆-acetic anhydride and propionylation for peptides chemically modified with the propionic anhydride. The Uniprot MOUSE 1401 database was used for histone peptide identification. Enzyme specificity was set to Arg-C. A maximum of 3 missed cleavages were permitted, and the minimum peptide length was fixed at 6 amino acids. In the case of H3K27ac N-ChroP, PIC at the N-terminus of each peptide was set as a fixed modification. Quantitation of histone PTMs was performed by a label-free approach, in particular by calculating the extracted ion chromatogram (XIC) of each modified histone peptide and, then, extrapolating their percentage relative abundance (%RA). RA corresponds to the ratio between the XIC of a specific histone modified peptide over the sum of the XICs of all detected unmodified and modified forms of the same peptide¹². Relative Enrichment (RE) was calculated as the ratio of the %RA of each histone modification in the ChIP sample over that in the Input. Hierarchical clustering analysis was obtained with Perseus Software.

RESULTS AND DISCUSSION

MS-based analysis of histone PTMs co-enriched at different genomic regions in macrophages

To characterize the histone modifications specifically associated with different genomic regions in mouse macrophage-derived RAW264.7 cells, we employed the N-ChroP approach¹², using as bait the histone marks H3K4me1, H3K4me3 and H3K9me3, which are specifically enriched at enhancers, promoters and heterochromatic regions, respectively^{7,14,15}. Enrichment/depletion of histone modifications at the distinct chromatin fractions were determined on the basis of the ratio of their %RA in the ChIP over the bulk chromatin used as Input (Figure 1A). With this strategy, we assessed the co-association of 42 different histone modified peptides in all the native ChIP experiments. Unsupervised hierarchical clustering allowed distinguishing three different groups according to the bait used in each experiment (Figure 1B). All the three baits were strongly enriched in the respective ChIP experiments, which served as positive control of the efficacy of the immuno-precipitation. The enrichment of H3K4me1 at putative enhancers was associated with a slight depletion of H3K4me3, confirming the existence of the H3K4me1^{Hi}/K4me3^{Low} signature at these specific regulatory genomic regions⁷. Putative enhancers as well as promoters showed a global depletion of H3K9me2 and H3K9me3, which is in line with their association with inactive chromatin regions. Moreover, both enhancers and promoters displayed the enrichment of hyper-acetylated H3 and H4 peptides, in particular the peptides H3(9-17), H3(18-26) and H4(4-17) bearing acetylated H3K9 and H3K14, acetylated H3K18 and H3K23, tri-acetylated H4K16/K12/K8 and tetra-acetylated H4K16/K12/K8/K5 respectively. This result is consistent with the well-known functional link between acetylation levels and active regulatory regions. In addition to these marks, H3K79me1 and H3K79me2 were also significantly enriched at both enhancers and promoters (Figure 1B-C), in agreement with their known accumulation at actively transcribed genes¹⁶. Furthermore, the N-ChroP

experiment revealed a strong association of the H3K4me1 with H3K36me2 and H3K36me3 (Figure 1B-C). Interestingly, H3K36me2 and H3K36me3 were not enriched at the promoters, suggesting that these histone marks are specifically associated with enhancer regions. Unexpectedly, H3K27ac -which is the best-known marker of active enhancer⁸- was not enriched in the enhancers' nucleosomes: to explain this result we reasoned that with the N-ChroP we enriched the complete pool of all possible macrophage H3K4me1 enhancer regions while H3K27ac marks only a fraction of transcriptionally active enhancers; therefore, the lack of detection could be due to the low abundance of this fraction for reliable MS detection. In addition, the use of the D₆-acetic anhydride as chemical agent to derivatize unmodified and mono-methylated lysines prior to trypsin digestion does not help the detection of the H3(27-40) bearing acetylated K27 with a peak that is barely detectable. The marks H3K18me1 and H3K27me1 were found depleted at the enhancer regions, with H3K18me1 showing enrichment in H3K9me3-regions, in line with previous results suggesting its role in chromatin silencing (Figure 1B)¹⁷.

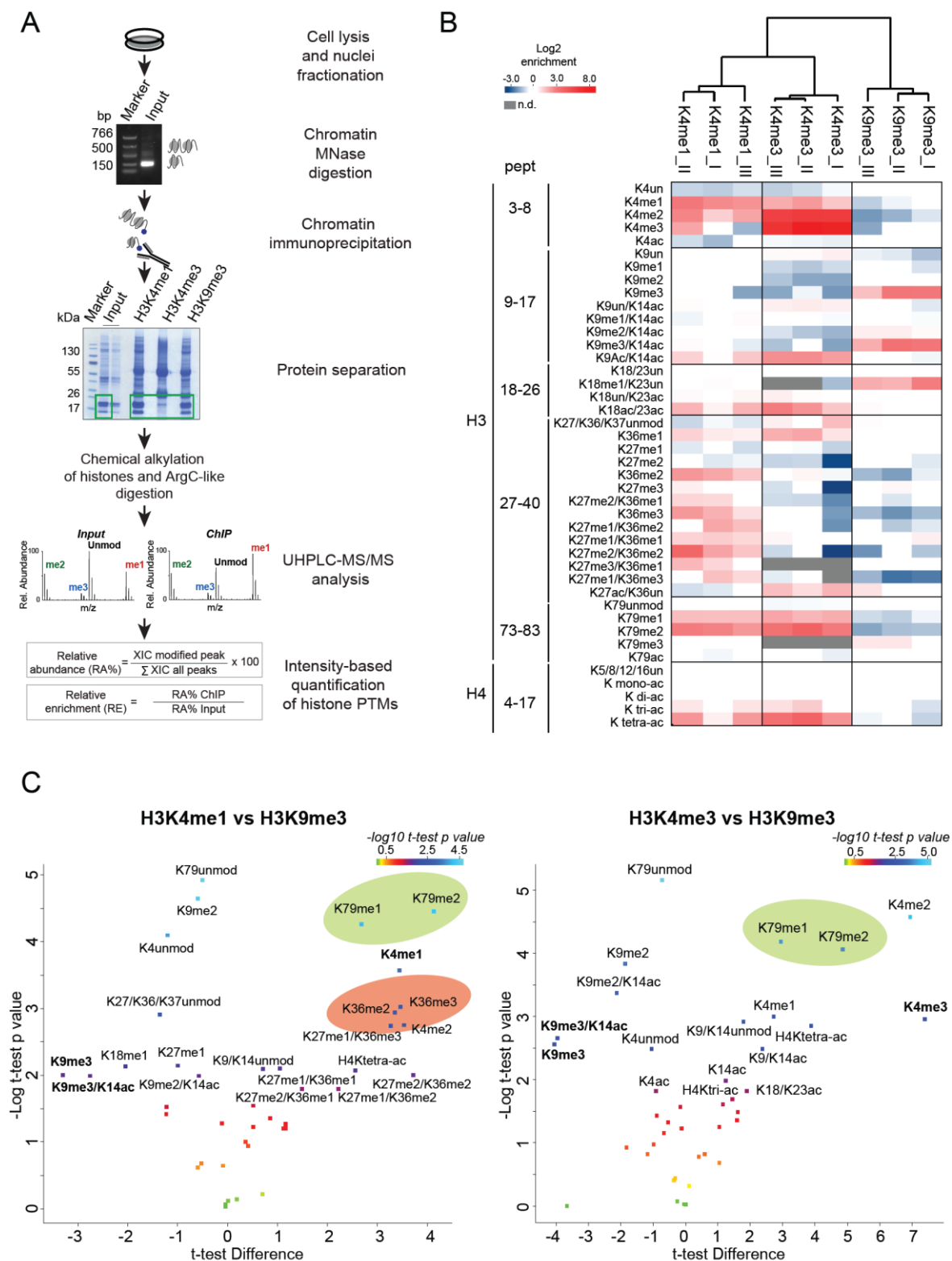


Figure 1: MS-based profiling of histone PTMs co-associated at different chromatin regions in RAW264.7 cells. A) Workflow of the N-ChroP approach. RAW264.7 cells are lysed and chromatin digested with MNase in order to enrich for mono-nucleosomes. The digested chromatin is, then, used as Input for the different chromatin immuno-precipitation. Immuno-purified H3 and H4 histones from the different N-ChroP are excised from the gel, chemically modified with D₆-acetic anhydride and digested with trypsin prior to LC-MS/MS analysis. Histone modifications are quantified by calculating their percentage relative

abundance (%RA) value and then the relative enrichment (RE) compared to its own Input, as explained in the paragraph “Histone PTM analysis from MS-data” of Materials and Methods of this section. B) Unsupervised hierarchical clustering displaying the relative enrichment of 42 different modified histone peptides in three biological replicates (n=3) of H3K4me1, H3K4me3 and H3K9me3 N-ChroPs. Results are plotted as log₂ enrichment of a specific histone modification in the N-ChroP compared to the Input. n.d. = not detected. C) Volcano plots displaying the comparison of H3K4me1 N-ChroP (on the left) and H3K4me3 (on the right) with that of the H3K9me3 N-ChroP. Histone modification with a p-value ≤ 0.05 were selected and defined as significantly enriched, using two-side t-test and Bonferroni correction.

Application of N-ChroP strategy to detect histone modification signatures that mark different subpopulations of macrophage enhancers

The results described in the previous paragraph were part of a paper published in 2017 in Nucleic Acids Research (which I was co-author) and presented the basis for a functional follow-up investigation of H3K36me2, H3K36me3 and H3K79me2, three histone modifications that resulted co-enriched within the H3K4me1 mono-nucleosomes. Through the intersection of the proteomics data with extensive ChIP-seq and RNA-seq analysis both at basal and LPS-stimulated condition, we could unravel the existence of sub-populations of enhancers that are marked by distinct combinations of histone modifications, in particular: H3K4me1/K36me2 marks transcribed enhancers, H3K4me1/K36me3 and H3K4me1/K79me2 combinations associate with different subsets of intronic enhancers. The detailed description of this results can be found in the paper ¹⁸.

MS-based profiling of histone PTMs co-enriched at the active enhancer regions

Once established the histone modifications that co-associated at bulk macrophage enhancers we focused on the specific class of active enhancers, by performing the N-ChroP, using H3K27ac as bait, the best-known marker of these regions ⁸. We followed essentially the previously described workflow (Figure 1A), with only one major modification in the histone

digestion step, where we substituted the D₆-acetic anhydride derivatization followed by trypsin digestion with the hybrid chemical derivatization “Pro-PIC” method. In this approach, unmodified and mono-methylated lysines were initially converted to their propionylated forms prior to trypsin digestion and labelling of new peptide N-termini with phenyl isocyanate (PIC). With this technique, the detection of short and hydrophilic peptides, such as H3(3-8) and some of low abundant modified histone peptides, among which also the one bearing the H3K27ac were highly improved ¹⁹.

We assessed the co-association of the H3K27ac with 42 different histone modified peptides belonging to both histones H3 and H4 (Figure 2A). As a positive control, we detected a strong enrichment of the H3 peptide bearing K27ac in the N-ChroP compared to the Input, which indicates the efficient immuno-purification of a significant pool of chromatin marked by this histone mark. In addition, we also detected the increase of both H3K4me1 and H3K4me3, further corroborating the presence of the H3K27ac at both enhancer and promoter regions ^{8,20}.

Interestingly, we observed a strong increase of all acetylated H3 and H4 peptides, with the hyper-acetylated isoform of each peptide being far more enriched than the single-acetylated version of the same peptide sequence. Some of these hyper-acetylated peptides, such as H3(9-17) bearing both K9ac/K14ac and the tetra-acetylated H4(4-17) were detected only in the N-ChroP and not in the Input (\log_2 Enrichment = Infinite). This result is perfectly in agreement with the model implicating that acetylation at multiple sites at both promoters and enhancers stimulates transcription of genes and enhancer RNAs (eRNAs), respectively ²¹. The acetylated marks H3K9ac, H3K14ac, H4K8ac and H4K16ac that we found co-enriched in H3K27ac positive regions were already reported in literature to co-associate at active enhancer regions ²¹⁻²⁴. More interestingly, also other less known acetylation sites -such as H3K64ac and H3K122ac- were found enriched in this specific subset of enhancers ²⁵.

Mirroring this strong increase of H3/H4 acetylation, we detected a strong depletion of several methylated sites, including H3K9me2/me3 and H3K27me2/me3, in agreement with their role as repressive marks ²⁶. We also detected the overall depletion of methylated H3K36 in H3K27ac enriched chromatin, more evident for H3K36me3, probably due to the localization of these PTMs at the gene body of active genes ^{14,27}. Unexpectedly, H3K36me2 that we previously found to be a novel marker of transcribed enhancers ¹⁸ was slightly depleted in H3K27ac positive regions (Figure 2A-B).

Overall, our data suggest that H3K27ac mainly co-associated with several acetylated marks of histones H3 and H4 while K- methylation, apart from H3K4me1 and H3K4me3, are all localized in different regions and are therefore excluded by H3K27ac-labelled nucleosomes.

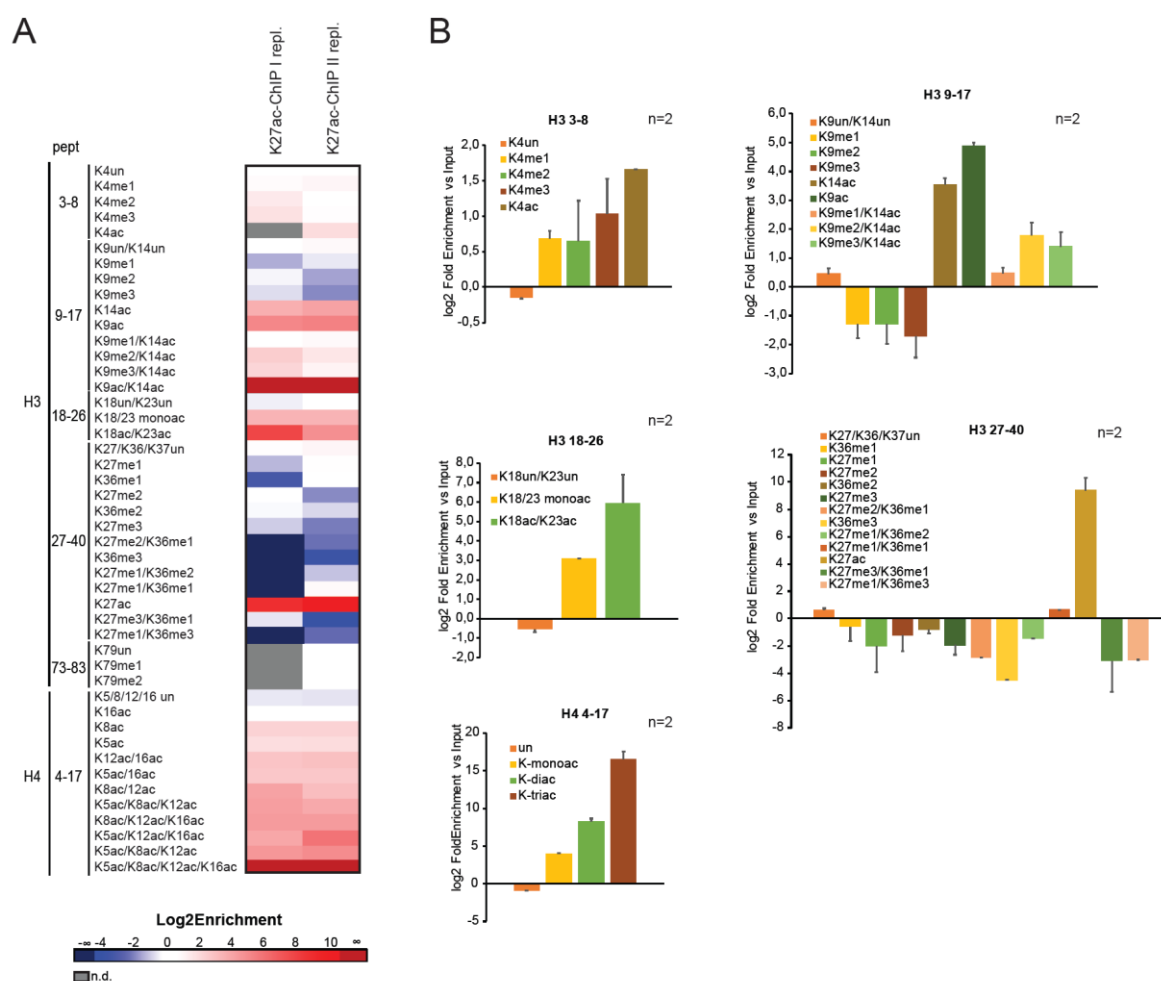


Figure 2: MS-based profiling of histone PTMs co-enriched at the H3K27ac positive regions. A) Heatmap displaying the enrichment of 42 different modified histone peptides in

two biological replicates of H3K27ac N-ChroPs. Results are plotted as \log_2 enrichment of a specific histone modification in the N-ChroP compared to the Input. n.d. = not detected. ∞ = Infinite, namely histone modified peptides identified in the N-ChroP and not in the Input, $-\infty$ = -Infinite, namely histone modified peptides identified in the Input and not in the N-ChroP. B) Bar graph displaying the \log_2 enrichment of a panel of histone modification in the H3K27ac N-ChroP compared to the Input. Chart represents mean \pm SEM (n=2 biological replicates).

Given that active enhancers are typically marked by both H3K4me1 and H3K27ac²⁸, we compared the results of the N-ChroP of these two histone marks in order to find a common signature of histone modifications that are enriched in both N-ChroPs and could be defined, as novel identifiers of active enhancer marks in macrophages. This comparative analysis demonstrated that several acetylation marks were enriched in both H3K4me1 and H3K27ac regions such as the H3K14ac and H3K23ac, alone or in combination with H3K9ac and H3K18ac respectively, as well as the combination of the tri- and tetra- acetylation of the H4 tail, including the K5ac, K8ac, K12ac and K16ac (Figure 3A). The enrichment level of these modifications was different among the two N-ChroPs, with H3K27ac N-ChroP displaying a much more significant increase than the H3K4me1 N-ChroP (Figure 3A-B). Since H3K27ac also binds to the promoters of the active genes²⁰, it could associate with these acetylation marks also in these genomic regions. In agreement with this, these acetylated sites were found increased in the H3K4me3 N-ChroP (Figure 1B) further suggesting that H3K27ac co-associated with these histone acetylation sites at both promoter and enhancer regions.

In contrast to the acetylation marks, several methylated peptides, including the H3(27-40) bearing methylated K27 and K36 or combination of them, were enriched in H3K4me1 positive regions and strongly depleted in H3K27ac-rich regions. This data hints towards a model whereby H3K4me1 covers different sub-populations of enhancers that are both functionally distinct and physically separated, and that are characterized by specific chromatin states: poised/inactive enhancers marked by both H3K4me1 and methylated

H3K27²⁹ and active ones characterized by H3K4me1, together with H3K27ac and other acetylation marks.

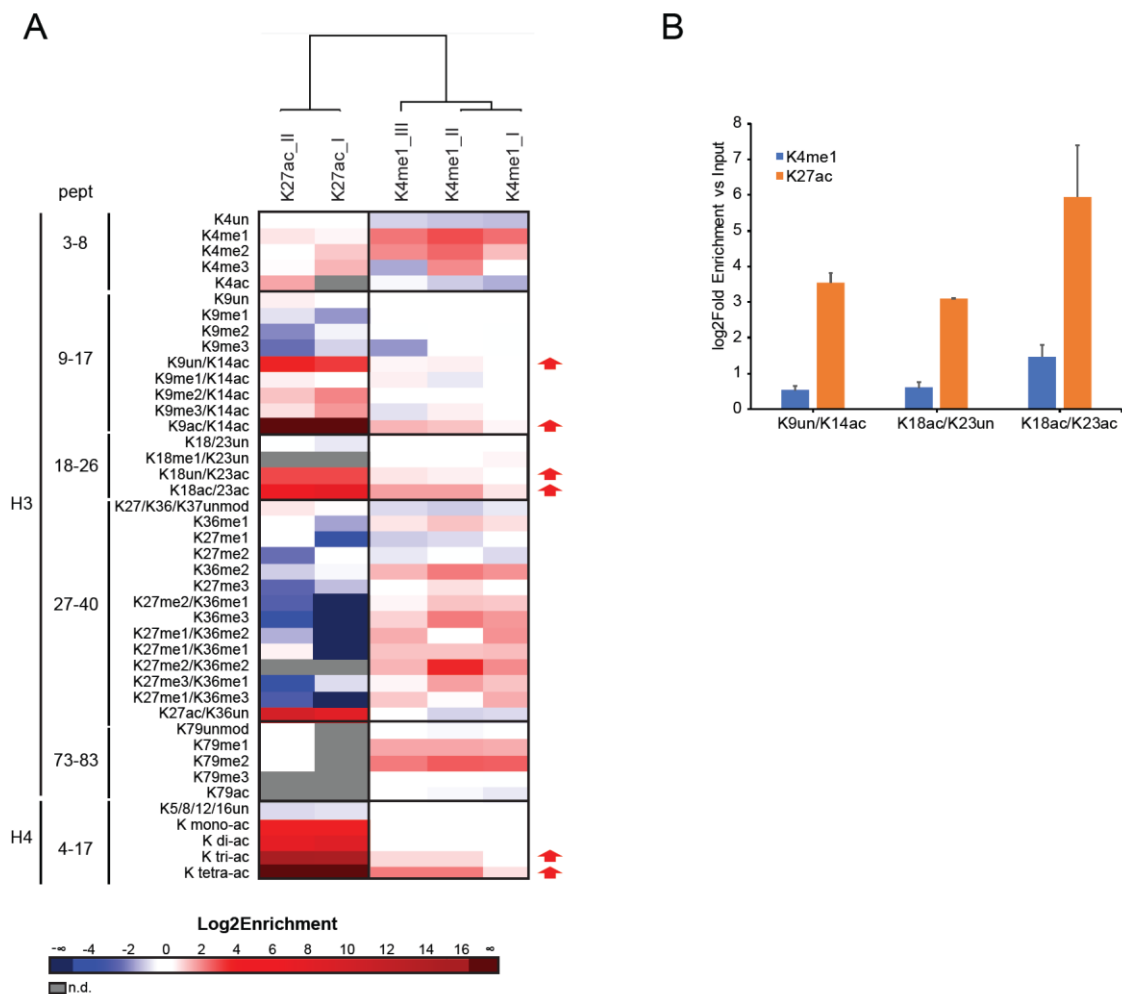


Figure 3: Acetylated marks associate at active enhancer regions. A) Unsupervised hierarchical clustering showing the enrichment of 42 different modified histone peptides in both H3K4me1 and H3K27ac N-ChroPs. Results are plotted as log₂ enrichment of a specific histone modification in the N-ChroP compared to the Input. Arrows point out modifications that show enrichment in both H3K27ac and H3K4me1 regions. n.d. = not detected. ∞ = Infinite, namely histone modified peptides identified in the N-ChroP and not in the Input, -∞ = -Infinite, namely histone modified peptides identified in the Input and not in the N-ChroP B) Bar graph displaying the log₂enrichment of three different histone modifications in both H3K27ac and H3K4me1 N-ChroPs compared to the Input. Chart represents mean ± SEM (from 2 to 3 biological replicates).

Ongoing experiments and future perspectives

Currently, we are re-analysing the MS-data from the ChroP experiments in search of other less characterised histone PTMs like ubiquitination, butyrylation, crotonylation, to assess if they specifically co-enrich at active enhancers, with a specific focus on crotonylation because it has already been described to mark active genomic regions (i.e. promoters and potential enhancers) ³⁰. In parallel, by using the same ChroP approach we are also investigating the interactors specifically enriched with H3K4me1- and H3K27ac- chromatin. Following a description of the basal enhancer interactome and histone PTM pattern, we plan to assess their changes upon the strong transcriptional burst elicited by LPS treatment, which triggers the inflammatory response of macrophages.

REFERENCES

1. Bulger M, Groudine M. Functional and mechanistic diversity of distal transcription enhancers. *Cell*. 2011. doi:10.1016/j.cell.2011.01.024
2. Ong CT, Corces VG. Enhancer function: New insights into the regulation of tissue-specific gene expression. *Nat Rev Genet*. 2011. doi:10.1038/nrg2957
3. Heinz S, Romanoski CE, Benner C, Glass CK. The selection and function of cell type-specific enhancers. *Nat Rev Mol Cell Biol*. 2015. doi:10.1038/nrm3949
4. Heinz S, Benner C, Spann N, et al. Simple Combinations of Lineage-Determining Transcription Factors Prime cis-Regulatory Elements Required for Macrophage and B Cell Identities. *Mol Cell*. 2010. doi:10.1016/j.molcel.2010.05.004
5. Roeder RG. Transcriptional regulation and the role of diverse coactivators in animal cells. In: *FEBS Letters*. ; 2005. doi:10.1016/j.febslet.2004.12.007
6. Weake VM, Workman JL. Inducible gene expression: Diverse regulatory mechanisms. *Nat Rev Genet*. 2010. doi:10.1038/nrg2781
7. Heintzman ND, Stuart RK, Hon G, et al. Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. *Nat Genet*. 2007. doi:10.1038/ng1966
8. Creighton MP, Cheng AW, Welstead GG, et al. Histone H3K27ac separates active

- from poised enhancers and predicts developmental state. *Proc Natl Acad Sci*. 2010. doi:10.1073/pnas.1016071107
9. Carter D, Chakalova L, Osborne CS, Dai Y feng, Fraser P. Long-range chromatin regulatory interactions in vivo. *Nat Genet*. 2002. doi:10.1038/ng1051
 10. Calo E, Wysocka J. Modification of Enhancer Chromatin: What, How, and Why? *Mol Cell*. 2013. doi:10.1016/j.molcel.2013.01.038
 11. Park PJ. ChIP-seq: Advantages and challenges of a maturing technology. *Nat Rev Genet*. 2009. doi:10.1038/nrg2641
 12. Soldi M, Bonaldi T. The Proteomic Investigation of Chromatin Functional Domains Reveals Novel Synergisms among Distinct Heterochromatin Components. *Mol Cell Proteomics*. 2013. doi:10.1074/mcp.M112.024307
 13. Rappsilber J, Mann M, Ishihama Y. Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. *Nat Protoc*. 2007. doi:10.1038/nprot.2007.261
 14. Barski A, Cuddapah S, Cui K, et al. High-Resolution Profiling of Histone Methylations in the Human Genome. *Cell*. 2007. doi:10.1016/j.cell.2007.05.009
 15. Visel A, Rubin EM, Pennacchio LA. Genomic views of distant-acting enhancers. *Nature*. 2009. doi:10.1038/nature08451
 16. Sawado T, Halow J, Im H, et al. H3 K79 dimethylation marks developmental activation of the 2-globin gene but is reduced Upon lcr-mediated high-level transcription. *Blood*. 2008. doi:10.1182/blood-2007-12-128983
 17. Zee BM, Levin RS, Xu B, LeRoy G, Wingreen NS, Garcia BA. In vivo residue-specific histone methylation dynamics. *J Biol Chem*. 2010. doi:10.1074/jbc.M109.063784
 18. Soldi M, Mari T, Nicosia L, et al. Chromatin proteomics reveals novel combinatorial histone modification signatures that mark distinct subpopulations of macrophage enhancers. *Nucleic Acids Res*. 2017. doi:10.1093/nar/gkx821
 19. Maile TM, Izrael-Tomasevic A, Cheung T, et al. Mass Spectrometric Quantification of Histone Post-translational Modifications by a Hybrid Chemical Labeling Method. *Mol Cell Proteomics*. 2015. doi:10.1074/mcp.O114.046573
 20. Wang Z, Zang C, Rosenfeld JA, et al. Combinatorial patterns of histone acetylations and methylations in the human genome. *Nat Genet*. 2008. doi:10.1038/ng.154
 21. Pradeepa MM. Causal role of histone acetylations in enhancer function.

22. Karmodiya K, Krebs AR, Oulad-Abdelghani M, Kimura H, Tora L. H3K9 and H3K14 acetylation co-occur at many gene regulatory elements, while H3K14ac marks a subset of inactive inducible promoters in mouse embryonic stem cells. *BMC Genomics*. 2012. doi:10.1186/1471-2164-13-424
23. Li QL, Wang DY, Ju LG, et al. The hyper-activation of transcriptional enhancers in breast cancer. *Clin Epigenetics*. 2019. doi:10.1186/s13148-019-0645-x
24. Taylor GCA, Eskeland R, Hekimoglu-Balkan B, Pradeepa MM, Bickmore WA. H4K16 acetylation marks active genes and enhancers of embryonic stem cells, but does not alter chromatin compaction. *Genome Res*. 2013. doi:10.1101/gr.155028.113
25. Pradeepa MM, Grimes GR, Kumar Y, et al. Histone H3 globular domain acetylation identifies a new class of enhancers. *Nat Genet*. 2016. doi:10.1038/ng.3550
26. Kouzarides T. Chromatin Modifications and Their Function. *Cell*. 2007. doi:10.1016/j.cell.2007.02.005
27. Schwartz S, Meshorer E, Ast G. Chromatin organization marks exon-intron structure. *Nat Struct Mol Biol*. 2009. doi:10.1038/nsmb.1659
28. Natoli G, Andrau J-C. Noncoding Transcription at Enhancers: General Principles and Functional Models. *Annu Rev Genet*. 2012. doi:10.1146/annurev-genet-110711-155459
29. Zentner GE, Tesar PJ, Scacheri PC. Epigenetic signatures distinguish multiple classes of enhancers with distinct cellular functions. *Genome Res*. 2011. doi:10.1101/gr.122382.111
30. Tan M, Luo H, Lee S, et al. Identification of 67 histone marks and histone lysine crotonylation as a new type of histone modification. *Cell*. 2011. doi:10.1016/j.cell.2011.08.008

9. List of publications achieved during PhD

Published papers:

1. Soldi M., Mari T., **Nicosia L.**, Musiani D., Sigismondo G., Cuomo A., Pavesi G., Bonaldi T. (2017). Chromatin proteomics reveals novel combinatorial histone modification signatures that mark distinct sub-populations of macrophage enhancers. *Nucleic Acids Res.*, 45(21):12195-12213. *Impact factor: 11.56.*
2. **Nicosia L.**, Noberini R., Soldi M., Cuomo A., Musiani D., Spadotto V., Bonaldi T. (2017). Mass spectrometry and epigenetics. *Handbook of Nutrition, Diet and Epigenetics*, https://doi.org/10.1007/978-3-319-31143-2_115-1.

Papers under revision:

1. Ravasio R., Ceccacci E. *, **Nicosia L.** *, Hosseini A., Rossi PL., Barozzi I., Fornasari L., Dal Zuffo R., Valente S., Fioravanti R., Mercurio C., Varasi M., Mattevi A., Mai A., Bonaldi T. and Minucci S. (2019). Targeting the scaffolding role of LSD1 (KDM1A) poises acute myeloid leukemia cells for Retinoic Acid induced differentiation. *Science Advances*, under revision. *Impact factor: 11.51.* * equal contribution. This paper contains some of the data presented in this thesis: the basal and dynamic LSD1 interactome and the follow-up analysis on the effects of the LSD1-GFI1 interaction upon disruption of this binding.
2. Musiani D., Giambruno R., Massignani E., Ippolito M.R., Maniaci M., Jammula S., Manganaro D., Cuomo A., **Nicosia L.**, Pasini D. and Bonaldi T. (2019). PRMT1 controls the Senescence-Associated Secretory Phenotype Transcriptional Program during replication stress response in cancer cells. *Cell Report*, under revision. *Impact factor: 8.032.*

10. Acknowledgments

I would like to thank my PhD supervisor, Dr. Tiziana Bonaldi, for giving me the opportunity to work in her laboratory and for her guidance, suggestions and support in the research projects during the years of the PhD study.

I would like to thank my Internal Advisor Prof. Saverio Minucci for his important contribution to the continuation of my main project and suggestions/ideas on experiments and models to use. Thanks also to some members of his group, in particular Roberto Ravasio, Elena Ceccacci and Amir Hosseini with whom I collaborated and shared part of the project.

I would like to thank my External Supervisor Prof. Axel Imhof for his kind suggestions and comments during the years of the PhD study.

I would like to thank all members of the TB group, in particular Roberto and Roberta for their assistance, precious advices on design/interpreting/analysing experiments. A special thanks also to Angelo that helped me in performing experiments and Fabio for his important contribution in the bioinformatics analysis of the transcriptomics data.

I would like to thank friends and colleagues at the Campus IFOM/IEO, with whom I shared positive and negative moments of this study, in particular Debora, Cecilia, Laura, Silvia, Zivojin, Paolo.

Last but not least, a special thanks to Giulia with whom I fully shared these years. Thanks because you have always supported me especially during difficult moments. Thanks also to my family and my best friends Gabriele and Federica for their fundamental encouragement during these years.