

erriari ta zabal zazu



Universidad  
del País Vasco

Euskal Herriko  
Unibertsitatea

# Proteostasis and Rare Diseases: Congenital Erythropoietic Porphyria and Townes-Brocks Syndrome

**Immacolata Giordano**

July 2020



# Proteostasis and Rare Diseases: Congenital Erythropoietic Porphyrria and Townes-Brocks Syndrome

**Immacolata Giordano**

July 2020

**Supervisor: Dr. María Rosa Barrio Olano**

Tutor: Prof. Miguel Angel Trueba Conde



*A mio padre,  
che mi ha dato l'esempio per non arrendermi.*

*A mia sorella Mary,  
che mi ha dato l'esempio per tornare in piedi.*

*A Alfredo, per esserci stato sempre.*



*None but those who have experienced them can conceive of the enticements of science. In other studies, you go as far as others have gone before you, and there is nothing more to know; but in a scientific pursuit, there is continual food for discovery and wonder...*

*...If the study to which you apply yourself has a tendency to weaken your affections and to destroy your taste for those simple pleasures in which no alloy can possibly mix, then that study is certainly unlawful, that is to say, not befitting the human mind.*

***Frankenstein; or, The Modern Prometheus***

**Mary Wollstonecraft Shelley**





<b>ABBREVIATIONS.....</b>	<b>5</b>
<b>Figures and Tables index .....</b>	<b>9</b>
<b>Summary .....</b>	<b>17</b>
<b>Resumen .....</b>	<b>19</b>
<b>1. INTRODUCTION .....</b>	<b>23</b>
<b>1.1. Proteostasis: the maintenance of proteome equilibrium.....</b>	<b>25</b>
1.1.1 Ubiquitin and Ubiquitin-like pathways .....	27
1.1.2 Degradation of misfolded proteins.....	31
1.1.3 Proteostasis dysfunction and therapeutic strategies .....	33
<b>1.2 Rare diseases .....</b>	<b>34</b>
<b>1.3 Congenital Erythropoietic Porphyria: a metabolic rare disease .....</b>	<b>37</b>
1.3.1 Porphyrins and alterations in the heme biosynthesis pathway.....	37
1.3.2 Congenital Erythropoietic Porphyria: a rare disease based on protein stability alterations .....	40
1.3.3 Diagnosis and treatment of Congenital Erythropoietic Porphyria.....	42
1.3.4 Uroporphyrinogen III Synthase mutations and stability .....	44
1.3.5 Genome editing technology: CRISPR/Cas9 .....	46
1.3.6 FLEX-Switch System .....	48
<b>1.4 Townes-Brocks Syndrome: a developmental rare disease.....</b>	<b>50</b>
1.4.1 The SALL family of transcription factors .....	52
1.4.2 Truncations of SALL1 cause Townes-Brocks Syndrome .....	53
1.4.3 SALL1 SUMOylation .....	54
1.4.4 SALL1 Interactors .....	57
1.4.5 CBX4 is a member of the Polycomb group .....	60
<b>2. MATERIALS AND METHODS .....</b>	<b>67</b>
Cell culture .....	69
Cell Transfection .....	69
CRISPR/Cas9 genome editing .....	69
Cycloheximide assay .....	70
Fluorescence-activated cell sorting.....	70
Generation of vectors .....	71
Genotyping and PCR primers .....	72
GFP-Trap Pulldown .....	74
GST-SUMO proteins purification.....	75

GST-Pulldown .....	75
High-performance liquid chromatograph analysis .....	75
Immunoprecipitation .....	76
Immunostaining .....	76
Lentiviral transduction .....	77
Porphyryns extraction.....	77
Proliferation assay .....	78
Quantitative reverse transcriptional PCR analysis .....	78
Ubiquitination assay .....	79
Western blot .....	80
Statistical analysis .....	81
<b>3. CONGENITAL ERYTHROPOIETIC PORPHYRIA .....</b>	<b>83</b>
<b>3.1 Hypothesis.....</b>	<b>¡Error! Marcador no definido.</b>
<b>3.2 Objectives.....</b>	<b>¡Error! Marcador no definido.</b>
<b>3.3 Results.....</b>	<b>¡Error! Marcador no definido.</b>
3.3.1 Design and generation of DNA constructs to use as donor for the knock-in genome editing .....	<b>¡Error!</b>
<b>Marcador no definido.</b>	
Components .....	<b>¡Error! Marcador no definido.</b>
Assembly of the constructs .....	<b>¡Error! Marcador no definido.</b>
3.3.2 Generation of humanized cells using CRISPR/Cas9 editing	<b>¡Error! Marcador no definido.</b>
Cell line selection and antibody validation.....	<b>¡Error! Marcador no definido.</b>
sgRNA validation .....	<b>¡Error! Marcador no definido.</b>
Generation of CRISPR_UROS cellular clones .....	<b>¡Error! Marcador no definido.</b>
3.3.3 Technical validation of the FLEX-Switch system in cells ....	<b>¡Error! Marcador no definido.</b>
Validation of CRSPR_UROS cells at protein level .....	<b>¡Error! Marcador no definido.</b>
Validation of CRSPR_UROS cells at DNA level .....	<b>¡Error! Marcador no definido.</b>
Validation of UROS-C73R cells viability .....	<b>¡Error! Marcador no definido.</b>
3.3.4 Functional validation of the cell model .....	<b>¡Error! Marcador no definido.</b>
FACS analysis .....	<b>¡Error! Marcador no definido.</b>
HPLC analysis.....	<b>¡Error! Marcador no definido.</b>
3.3.5 Generation of an inducible UROS-C73R mutant mouse model .....	<b>¡Error! Marcador no definido.</b>
<b>defenido.</b>	
3.3.6 Ciclopirox treatment.....	<b>¡Error! Marcador no definido.</b>
<b>3.4 Discussion.....</b>	<b>¡Error! Marcador no definido.</b>

Targeting Uros locus in mouse cells .....	<b>¡Error! Marcador no definido.</b>
Purification of DNA for mouse transgenesis .....	<b>¡Error! Marcador no definido.</b>
Generation of the inducible humanized UROS-C73R mouse model .....	<b>¡Error! Marcador no definido.</b>
Ciclopirox as a promising but still not definitive CEP treatment .	<b>¡Error! Marcador no definido.</b>
<b>4. TOWNES-BROCKS SYNDROME .....</b>	<b>89</b>
<b>4.1 Hypothesis.....</b>	<b>91</b>
<b>4.2 Objectives.....</b>	<b>92</b>
<b>4.3 Results.....</b>	<b>93</b>
4.3.1 Validation of the interaction between SALL1 and CBX4 .....	93
SALL1 and CBX4 interact with each other at the protein level .....	93
Truncated SALL1-826C>T interacts with CBX4 .....	94
4.3.2 Analysis of the role of SALL1 SUMOylation in the interaction with CBX4 .....	96
SUMOylation of SALL1 in not required for CBX4 interaction .....	96
SALL1 SIM domains are not required for CBX4 interaction.....	97
SUMO-related mutant forms of SALL1 do not localize to Pc bodies .....	100
Analysis of the SUMO binding capacity of SALL1 predicted SIMs .....	101
4.3.3 Analysis of the functional effects of SALL1 – CBX4 interaction .....	104
SALL1 increases endogenous CBX4 protein levels .....	104
SALL1 increases CBX4 levels at post-transcriptional level.....	105
SALL1 influences size and number of Pc bodies .....	107
CBX4 targets are downregulated in presence of SALL1 .....	108
SALL1 stabilizes CBX4 avoiding its degradation via the UPS .....	110
CBX4 is degraded by the Ubiquitin-Proteasome System .....	112
CBX4 ubiquitination is reduced in presence of SALL1 .....	113
<b>4.4 Discussion.....</b>	<b>115</b>
SALL1-CBX4 interaction .....	116
Involvement of SALL1 SUMOylation in Pc bodies regulation and CBX4 stability .....	117
<b>5. GENERAL DISCUSSION.....</b>	<b>123</b>
<b>6. CONCLUSIONS .....</b>	<b>129</b>
<b>7. BIBLIOGRAPHY .....</b>	<b>135</b>
<b>CONTRIBUTED PUBLICATION .....</b>	<b>151</b>
<b>CURRICULUM VITAE.....</b>	<b>157</b>







## ABBREVIATIONS

<b>ALA</b>	$\delta$ -aminolevulinic acid
<b>ALAS</b>	$\delta$ -aminolevulinic synthase
<b>ANOVA</b>	Analysis of Variance
<b>APC</b>	allophycocyanin
<b>R</b>	Arginine
<b>BCA</b>	Bicinchoninic acid assay
<b>BSA</b>	Bovine serum albumin
<b>Cas9</b>	CRISPR associated protein 9
<b>CBX</b>	Chromobox
<b>cDNA</b>	Complementary DNA
<b>CEP</b>	Congenital Erythropoietic Porphyria
<b>CHX</b>	Cycloheximide
<b>CMA</b>	Chaperone-mediated autophagy
<b>CMV</b>	Cytomegalovirus
<b>CoA</b>	Coenzyme A
<b>COPRO</b>	Coproporphyrinogen
<b>Cpf1</b>	CRISPR-associated endonuclease in <i>Prevotella</i> and <i>Francisella</i> 1
<b>cPRC1</b>	Canonical PRC1
<b>CPX</b>	Ciclopirox
<b>CRISPR</b>	Clustered Regularly Interspaced Short Palindromic Repeats
<b>crRNA</b>	Crispr RNA
<b>CtBP</b>	Carboxyl-terminus binding protein
<b>Cys</b>	Cysteine
<b>DAPI</b>	4',6-diamino-2-phenylindol
<b>DMEM</b>	Dulbecco's modified Eagle medium
<b>DMSO</b>	Dimethyl sulfoxide
<b>dox</b>	Doxycycline
<b>DSB</b>	Double-strand break

<b>DTT</b>	Dithiothreitol
<b>DUB</b>	Deubiquitinating enzymes
<b>EDTA</b>	Ethylenediaminetetraacetic acid
<b>ER</b>	Endoplasmic reticulum
<b>ERT2</b>	Tamoxifen-inducible estrogen receptor
<b>EV</b>	Empty vector
<b>FACS</b>	Fluorescence-activated cell sorter
<b>FBS</b>	Fetal bovine serum
<b>FLEX Switch</b>	Flip-excision switch
<b>GAPDH</b>	Glyceraldehyde-3-Phosphate Dehydrogenase
<b>GATA</b>	GATA binding protein
<b>GFP</b>	Green fluorescent protein
<b>GFS</b>	GFP-Flag-Streptavidin
<b>GG</b>	Dyglycin motif
<b>G</b>	Glycine
<b>GO</b>	Gene Ontology
<b>gRNA / sgRNA</b>	guide RNA / single guide RNA
<b>GST</b>	Glutathione S-transferase
<b>h</b>	Hours
<b>HA</b>	Human influenza hemagglutinin
<b>HDR</b>	Homology-directed repair
<b>HEK 239FT</b>	Human embryonic kidney fibroblasts
<b>HMB</b>	Hydroxymethylbilane
<b>Hox</b>	Homeobox
<b>HPLC</b>	High-performance liquid chromatography
<b>HRP</b>	Horseradish Peroxidase
<b>kb</b>	Kilobase
<b>kDa</b>	Kilodaltons
<b>KO</b>	Knock-out
<b>lox</b>	Locus of crossover
<b>K</b>	Lysine



<b>M</b>	Molar
<b>me3</b>	Trimethylation
<b>MS</b>	Mass spectrometry
<b>NHEJ</b>	Non-homologous end joining
<b>NuRD</b>	Nucleosome remodeling deacetylase
<b>°C</b>	Celsius degree
<b>PAM</b>	Protospacer Adjacent Motif
<b>PBS</b>	Phosphate buffered saline
<b>Pc</b>	Polycomb
<b>PcG</b>	Polycomb grup
<b>PCGF</b>	Polycomb group ring finger
<b>PCR</b>	Polymerase Chain Reaction
<b>PFA</b>	Paraformaldehyde
<b>PFMS</b>	Phenylmethylsulfonyl fluoride
<b>PHC</b>	Polyhomeotic-like
<b>PLA</b>	Proximity Ligation Assay
<b>PolyA</b>	Polyadenylation tail
<b>PolyQ</b>	Polyglutamine
<b>PRs</b>	Proteostasis regulators
<b>PRC</b>	Polycomb Repressive Complex
<b>puro</b>	puromycin
<b>qPCR</b>	Quantitative Polymerase Chain Reaction
<b>RING1</b>	Ring finger protein 1
<b>RIPA</b>	Radioimmunoprecipitation assay buffer
<b>RNF4/11</b>	RING finger protein 4/11
<b>ROS</b>	Reactive oxygen species
<b>RS1</b>	RAD51-stimulatory compound 1
<b>RT</b>	Room temperature
<b>SALL</b>	Spalt-like protein
<b>SDS</b>	Sodium dodecyl sulfate
<b>SDS-PAGE</b>	Sodium dodecyl sulfate polyacrylamide gel electrophoresis

<b>SENP</b>	SUMO specific peptidase
<b>SIM</b>	Sumo Interaction Motif
<b>StUbL</b>	SUMO-targeted ubiquitin ligases
<b>SUMO</b>	Small Ub-like Modifier
<b>Tam</b>	4-OH Tamoxifen
<b>TBS</b>	Townes-Brocks syndrome
<b>tracrRNA</b>	Trans-activating crispr RNA
<b>Ub</b>	Ubiquitin
<b>Ub1</b>	Monoubiquitylation
<b>UBC9</b>	Ubiquitin conjugating enzyme
<b>Ubl</b>	Ubiquitin-like modifier
<b>UPR</b>	Unfolded protein response
<b>UPS</b>	Ubiquitin-proteasome system
<b>URO</b>	Uroporphyrinogen
<b>UROD</b>	Uroporphyrinogen decarboxylase
<b>UROS</b>	Uroporphyrinogen III synthase
<b>UV</b>	Ultraviolet
<b>WB</b>	Washing buffer
<b>WT</b>	Wild type
<b>YFP</b>	Yellow fluorescence protein
<b>ZF</b>	Zinc finger domain
<b>β-GAL</b>	β-Galactosidase

## Figures and Tables index

Figure 1. Protein homeostasis processes.....	26
Figure 2. Ubiquitination cycle .....	28
Figure 3. The Ubiquitin code.....	29
Figure 4. Unfolded protein response .....	33
Figure 5. Distribution of different types of rare diseases.....	35
Figure 6. Schematic representation of heme pathway-associated porphyrias.....	37
Figure 7. Schematic representation of the heme biosynthesis pathway .....	40
Figure 8. Uroporphyrinogen formation .....	41
Figure 9. Symptoms of Congenital Erythropoietic Porphyria .....	43
Figure 10. Schematic representation of UROS gene .....	45
Figure 11. CRISPR/Cas9 system.....	48
Figure 12. The directionality of the Lox sites determines the outcome of the Cre mediated recombination .....	49
Figure 13. FLEX-Switch System .....	50
Figure 14. Classical symptoms of Townes-Brocks Syndrome .....	51
Figure 15. Schematic representation of the main conserved domains present in SALL proteins.....	53
Figure 16. Schematic representation of SALL1 protein .....	54
Figure 17. SALL1 is modified by SUMO1 and SUMO2/3 in mammalian cells.....	56
Figure 18. Schematic representation of identified and mutated SUMO consensus sites in SALL1.....	57
Figure 19. Search of SALL1 interactors by proximity proteomics.....	58
Figure 20. Analysis of SALL1 interactors by Mass Spectrometry.....	59
Figure 21. Chromatin silencing marks by PRC1 and PRC2 .....	61

Figure 22. Pc bodies result from several engaged silenced-chromatin regions.....62

Figure 23. Clustering of Hox genes in mammals.....63

Figure 24. Schematic representation of CBX4 important motifs and domains.....64

Figure 25. Structural similarity of human and mouse UROS ..... **¡Error! Marcador no definido.**

Figure 26. DNA donor constructs for CRSPR/Cas9 knock-in editing....**¡Error! Marcador no definido.**

Figure 27. Assembly of the constructs.....**¡Error! Marcador no definido.**

Figure 28. Cell line selection and sgRNAs validation .....**¡Error! Marcador no definido.**

Figure 29. CRISPR\_UROS clones screening by Western blot ..... **¡Error! Marcador no definido.**

Figure 30. Schematic representation of the FLEX-Switch strategy **¡Error! Marcador no definido.**

Figure 31. Analysis of CRISPR\_UROS cells upon tamoxifen treatment**¡Error! Marcador no definido.**

Figure 32. DNA analysis of CRISPR\_UROS clones.....**¡Error! Marcador no definido.**

Figure 33. Growth assay of CRISPR\_UROS cells.....**¡Error! Marcador no definido.**

Figure 34. FACS analysis of CRISPR\_UROS WT and C73R cells ..... **¡Error! Marcador no definido.**

Figure 35. HPLC analysis of URO I levels in CRISPR\_UROS cells..... **¡Error! Marcador no definido.**

Figure 36. DNA digestion and analysis for the preparation and purification of the mouse-transgenesis construct.....**¡Error! Marcador no definido.**

Figure 37: Crystal violet assay of combined tamoxifen and ciclopirox treatment in CRISPR\_UROS and parental 3T3 cells .....**¡Error! Marcador no definido.**

Figure 38. SALL1 and CBX4 interact with each other at the protein level.....94

Figure 39. Schematic representation of SALL1 WT and the SALL1-826C>T mutant....95

Figure 40. SALL1-826 interacts with CBX4 .....96

Figure 41. Both SALL1 WT and SALL1ΔSUMO interact with CBX4.....97

Figure 42. Identification of SIMs and SUMOylation sites in human SALL1 protein.....98

Figure 43. SUMO-related SALL1 mutants interact with CBX4 at protein level.....100

Figure 44. SALL1 WT or SUMO-related mutants do not co-localize with CBX4 at the nuclear bodies.....101

Figure 45. Western blot analysis of recombinant GST-SUMO proteins .....102

Figure 46. Analysis of SALL1 SIMS through GST-SUMO beads pulldown .....103

Figure 47. SALL1 increases CBX4 protein levels.....104

Figure 48. Analysis of SALL1 and CBX4 levels in HEK 293FT cell model.....106

Figure 49. Analysis of Pc bodies.....108

Figure 50. SALL1 expression downregulates CBX4 targets.....110

Figure 51. SALL1 stabilizes CBX4 protein .....111

Figure 52. CBX4 is ubiquitinated and degraded by the proteasome.....113

Figure 53. Analysis of CBX4 ubiquitination by

Figure 1. Protein homeostasis processes.....26

Figure 2. Ubiquitination cycle .....28

Figure 3. The Ubiquitin code.....29

Figure 4. Unfolded protein response .....33

Figure 5. Distribution of different types of rare diseases .....35

Figure 6. Schematic representation of heme pathway-associated porphyrias.....37

Figure 7. Schematic representation of the heme biosynthesis pathway .....40

Figure 8. Uroporphyrinogen formation .....41

Figure 9. Symptoms of Congenital Erythropoietic Porphyria .....43

Figure 10. Schematic representation of UROS gene .....45

Figure 11. CRISPR/Cas9 system.....48

Figure 12. The directionality of the Lox sites determines the outcome of the Cre mediated recombination .....49

Figure 13. FLEX-Switch System .....50

Figure 14. Classical symptoms of Townes-Brocks Syndrome .....51

Figure 15. Schematic representation of the main conserved domains present in SALL proteins .....53

Figure 16. Schematic representation of SALL1 protein .....54

Figure 17. SALL1 is modified by SUMO1 and SUMO2/3 in mammalian cells.....56

Figure 18. Schematic representation of identified and mutated SUMO consensus sites in SALL1.....57

Figure 19. Search of SALL1 interactors by proximity proteomics.....58

Figure 20. Analysis of SALL1 interactors by Mass Spectrometry.....59

Figure 21. Chromatin silencing marks by PRC1 and PRC2 .....61

Figure 22. Pc bodies result from several engaged silenced-chromatin regions.....62

Figure 23. Clustering of Hox genes in mammals.....63

Figure 24. Schematic representation of CBX4 important motifs and domains.....64

Figure 25. Structural similarity of human and mouse UROS ..... **¡Error! Marcador no definido.**

Figure 26. DNA donor constructs for CRISPR/Cas9 knock-in editing....**¡Error! Marcador no definido.**

Figure 27. Assembly of the constructs.....**¡Error! Marcador no definido.**

Figure 28. Cell line selection and sgRNAs validation .....**¡Error! Marcador no definido.**

Figure 29. CRISPR\_UROS clones screening by Western blot ..... **¡Error! Marcador no definido.**

Figure 30. Schematic representation of the FLEX-Switch strategy **¡Error! Marcador no definido.**

Figure 31. Analysis of CRISPR\_UROS cells upon tamoxifen treatment **¡Error! Marcador no definido.**

Figure 32. DNA analysis of CRISPR\_UROS clones.....**¡Error! Marcador no definido.**

Figure 33. Growth assay of CRISPR\_UROS cells.....**¡Error! Marcador no definido.**

Figure 34. FACS analysis of CRISPR\_UROS WT and C73R cells ..... **¡Error! Marcador no definido.**

Figure 35. HPLC analysis of URO I levels in CRISPR\_UROS cells..... **¡Error! Marcador no definido.**

Figure 36. DNA digestion and analysis for the preparation and purification of the mouse-transgenesis construct.....**¡Error! Marcador no definido.**

Figure 37: Crystal violet assay of combined tamoxifen and ciclopirox treatment in CRISPR\_UROS and parental 3T3 cells .....**¡Error! Marcador no definido.**

Figure 38. SALL1 and CBX4 interact with each other at the protein level.....94

Figure 39. Schematic representation of SALL1 WT and the SALL1-826C>T mutant....95

Figure 40. SALL1-826 interacts with CBX4 .....96

Figure 41. Both SALL1 WT and SALL1ΔSUMO interact with CBX4.....97

Figure 42. Identification of SIMs and SUMOylation sites in human SALL1 protein.....98

Figure 43. SUMO-related SALL1 mutants interact with CBX4 at protein level.....100

Figure 44. SALL1 WT or SUMO-related mutants do not co-localize with CBX4 at the nuclear bodies.....101

Figure 45. Western blot analysis of recombinant GST-SUMO proteins .....102

Figure 46. Analysis of SALL1 SIMS through GST-SUMO beads pulldown .....103

Figure 47. SALL1 increases CBX4 protein levels.....104

Figure 48. Analysis of SALL1 and CBX4 levels in HEK 293FT cell model.....106

Figure 49. Analysis of Pc bodies ..... 108

Figure 50. SALL1 expression downregulates CBX4 targets ..... 110

Figure 51. SALL1 stabilizes CBX4 protein ..... 111

Figure 52. CBX4 is ubiquitinated and degraded by the proteasome ..... 113

Figure 53. Analysis of CBX4 ubiquitination by bioUb assay ..... 114

Figure 54. Schematic model of possible CBX4 regulation by SALL1 ..... 120

bioUb assay ..... 114

Figure 54. Schematic model of possible CBX4 regulation by SALL1 ..... 120

Table 1. Relevant sequences ..... 70

Table 2. Vectors used in the study ..... 71

Table 3. Genotyping and PCR oligonucleotide sequences ..... 72

Table 4. Oligonucleotide sequences used for qPCR. .... 78







## Summary

Proteostasis refers to all those biological processes that control the biogenesis, folding, trafficking and degradation of proteins present within and outside the cell. The alteration of the proteostasis regulatory pathways can seriously affect cell function and cause the onset of diseases, including rare diseases such as Congenital Erythropoietic Porphyria (CEP) and Townes-Brocks Syndrome (TBS).

CEP is caused by a deficiency in the fourth enzyme of the heme pathway, UROS. The most common pathogenic allele, C73R, generates an unstable protein with <1% of normal activity, leading to severe phenotypes. Although mouse models exist for this disease, the classical C73R mouse model has poor viability and chronic health issues, posing challenges for experimental setups. In order to generate an inducible humanized mouse model for CEP, we designed a novel single-insertion genetic switch to replace the mouse *Uros* allele with a conditionally controlled human mutant version. Human UROS replaces and rescues mouse UROS function and, upon tamoxifen induction of Cre recombinase, the allele is converted to express the mutant C73R version. To verify functionality of the system, we tested it in murine fibroblasts. CRISPR/Cas9 was used to introduce the construct into cells. The correct functioning of the system was validated at both DNA and protein levels. Moreover, FACS and HPLC analyses were performed to measure porphyrin levels and test the UROS activity. This cellular model, as well as the planned *knock-in* mouse model, will be novel tools to allow deeper studies into CEP and testing of potential therapies.

TBS is caused by mutations in the zinc-finger transcriptional repressor SALL1 and is characterized by a spectrum of malformations in the digits, ears, and kidneys. In order to elucidate SALL1 functions and regulation in the context of post-translational modifications, proximity proteomics BioID methodology in combination with mass spectrometry made possible the identification of putative interactors of human SALL1. Among them, we found CBX4, an E3 SUMO ligase and a key member of the Polycomb Repressive Complex 1 involved in chromatin remodeling and gene silencing. We confirmed that SALL1 and CBX4 interact with each other at the protein level and that they do co-localize in the nucleoplasm. By reducing its ubiquitination

and subsequent degradation via the ubiquitin proteasome system, SALL1 stabilizes and increases CBX4 protein levels. This affects the Pc nuclear bodies formation, which appear more abundant and bigger than in wild type cells. In addition, high levels of SALL1 increased the transcriptional repression capacity of CBX4 on some of its target genes. Interestingly, the SUMOylation of SALL1 seems to be necessary to enhance the transcriptional repression capacity of CBX4. Our data could contribute to understand the relationship underlying the interaction between SALL1 and CBX4, which could help to better understand the role of SALL1 during normal development and the pathogenesis of TBS.

Altogether, the results obtained throughout this work could help to better understand the pathogenesis of both these rare diseases and could advance the development of new potential therapies.

## Resumen

El término proteostasis se refiere a todos aquellos procesos biológicos críticos para el buen funcionamiento de la célula. La alteración de las vías reguladoras de la proteostasis puede afectar seriamente la función celular y causar la aparición de enfermedades, incluidas enfermedades raras como el Síndrome de Townes-Brocks (TBS) y la Porfiria Eritropoyética Congénita (CEP).

La CEP es causada por una deficiencia en la cuarta enzima de la vía de producción del grupo hemo, UROS. El alelo patógeno más común, C73R, genera una proteína inestable con menos del 1% de la actividad normal, lo que conduce a la acumulación de porfirinas y a fenotipos graves. El modelo clásico de ratón C73R tiene poca viabilidad y presenta problemas de salud crónicos, lo que plantea desafíos para su uso experimental. Para generar un modelo de ratón humanizado, inducible, mutante para CEP, diseñamos un nuevo interruptor genético de inserción única para reemplazar el alelo *Uros* del ratón con una versión controlada condicionalmente. El UROS humano reemplaza y rescata la función UROS del ratón y, tras la inducción con tamoxifeno de la recombinasa Cre, el alelo expresa la versión mutante C73R. Para verificar su funcionalidad, se testó el sistema en fibroblastos murinos. Se usó CRISPR/Cas9 para introducir la construcción en las células. El correcto funcionamiento del sistema fue validado tanto a nivel de ADN como de proteína. Se realizaron análisis FACS and HPLC para medir los niveles de porfirinas e investigar la actividad de UROS. Este modelo celular, así como el modelo de ratón *knock-in* planificado, serán herramientas novedosas para permitir estudios más profundos en CEP y probar posibles terapias.

El TBS está causado por mutaciones en el represor transcripcional de dedos de zinc SALL1 y se caracteriza por un espectro de malformaciones en dígitos, oídos y riñones. Para dilucidar las funciones y la regulación de SALL1 en el contexto de las modificaciones postraduccionales, se utilizó un ensayo de proteómica de proximidad BiID en combinación con espectrometría de masas para identificar posibles interactores de SALL1 humano. Entre éstos, encontramos CBX4, una E3 SUMO ligasa y miembro clave del Complejo Polycomb 1 de Represión involucrado en la

remodelación de la cromatina y el silenciamiento génico. Confirmamos que SALL1 y CBX4 interactúan entre sí a nivel de proteína y que colocalizan en el nucleoplasma. Al modificar su ubiquitinación y su posterior degradación a través del proteasoma, SALL1 estabiliza y aumenta los niveles de proteína de CBX4, afectando así la formación de cuerpos nucleares Pc que parecen más abundantes y más grandes. Además, altos niveles de SALL1 aumentan la capacidad de represión transcripcional de CBX4 en algunos de sus genes diana. Curiosamente, la SUMOilación de SALL1 parece ser necesaria para mejorar el efecto de represión transcripcional de CBX4. Elucidar la relación subyacente a la interacción entre SALL1 y CBX4, podría ayudar a comprender mejor el papel de SALL1 durante el desarrollo normal y en la patogénesis de TBS.

En conjunto, los resultados obtenidos a lo largo de este trabajo podrían ayudar a comprender mejor la patogénesis de estas dos enfermedades raras y podrían ayudar en el avance del desarrollo de potenciales terapias.







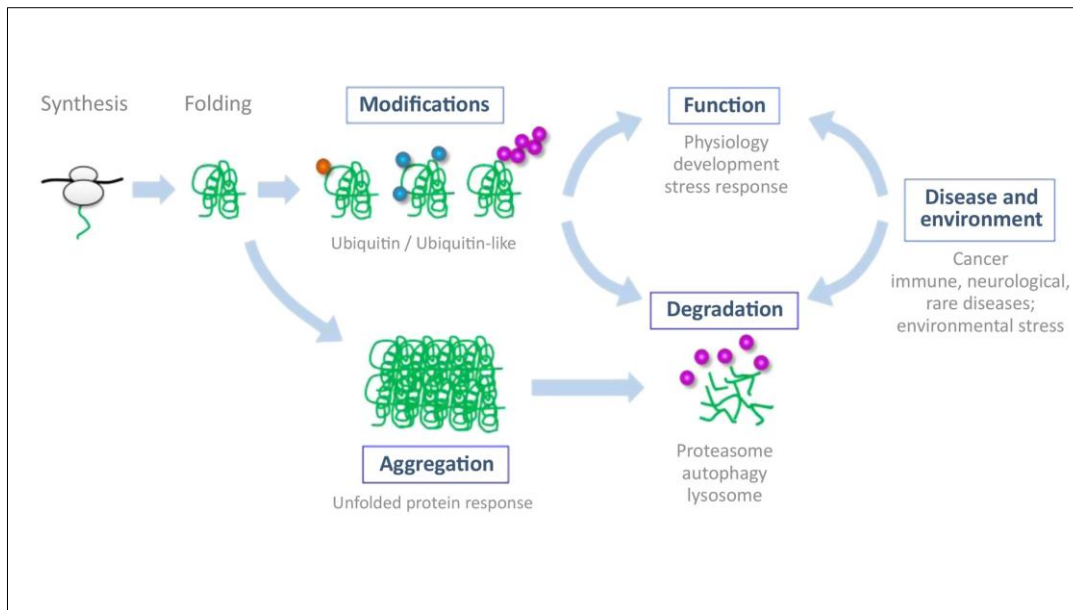
## **1. INTRODUCTION**



## **1.1. Proteostasis: the maintenance of proteome equilibrium**

Proteostasis, is a blend word formed by 'protein' and 'homeostasis' and referred to all those biological pathways that control the biogenesis, folding, trafficking and degradation of proteins present within and outside the cell. Proteostasis, enables healthy cell, organismal development and aging, and protects against disease (Balch et al. 2008). To preserve proteostasis cells must ensure correct proteins folding and assembling, as well as appropriated abundance and compartmentalization (Jayaraj, Hipp, and Hartl 2020). Highly sophisticated mechanisms provide a tight regulation of those processes on multiple levels (Hartl, Bracher, and Hayer-Hartl 2011).

Proteostasis network is a set of numerous interacting activities that keep proteins in a balanced equilibrium by maintaining their folding and regulating their expression and degradation (Figure 1). These pathways are critical processes for the proper functioning of the cell and comprise a wide range of general and specialized chaperones, folding enzymes, post-translational modifications and degradation components, as well as trafficking components, which can influence proteins compartmentalization (Hipp, Kasturi, and Hartl 2019; Powers et al. 2009).



**Figure 1. Protein homeostasis processes.** Proteostasis depends on the biogenesis, folding, post-translational modifications, trafficking and degradation of proteins present within and outside the cell to keep its proper functioning. Adapted from (Dissmeyer et al. 2019).

The generation of a new protein is regulated by many factors involving transcription, translocation of mRNA out of the nucleolus and translation. When the new polypeptide chain emerges from the ribosome, it has to be properly folded. Molecular chaperones promote folding and maintenance within the cell, largely by minimizing misfolding and aggregation. Protein folding, unfolding, and refolding are highly dynamic processes constantly occurring throughout the lifetime of proteins. In fact, chaperones can participate in the folding of neo-synthesized proteins or interact with pre-existing proteins that may suffer denaturation upon acute stress. Specialized chaperones for various compartments of the cell such as the endoplasmic reticulum (ER) and mitochondria have also been reported (Brehme et al. 2014).

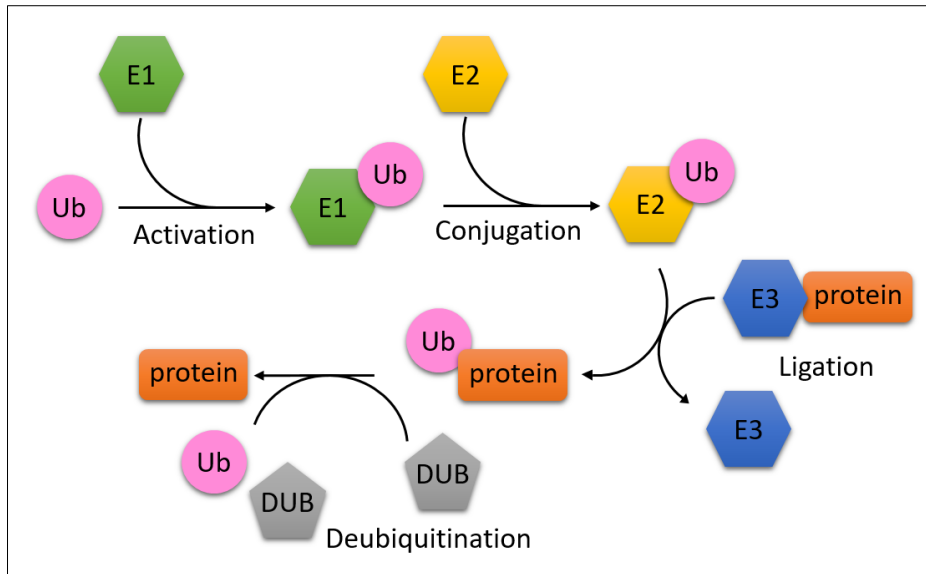
Once proteins are folded, they can undergo numerous post-translational modifications that determine their function, localization, interaction with other proteins and half-life. These modifications could involve small chemical molecules as it occurs in phosphorylation, acetylation and methylation, or proteins could covalently modify other proteins as in the case of ubiquitin (Ub) (Hochstrasser 2009). Ub is a small protein that can be transiently attached to thousands of different proteins.

Proteins similar to Ub, either in sequence or in their three-dimensional structure, form the family of Ub-like modifiers (UbL). Nearly 20 UbL proteins have been described in mammalian cells (Pirone et al. 2017). The main members of the UbL family include SUMO1-4 (Small Ub-related Modifier 1-4), NEDD8 (Neural precursor cell expressed developmentally down-regulated 8), ISG15 (Interferon-stimulated gene 15), FAT10 (HLA-F-adjacent transcript 10), UFM1 (Ub fold modifier 1), Atg-8 and Atg-12 (autophagy-related Ub-like modifier 8 and 12) and URM1 (Ub-related modifier 1) (Cappadocia and Lima 2018).

### **1.1.1 Ubiquitin and Ubiquitin-like pathways**

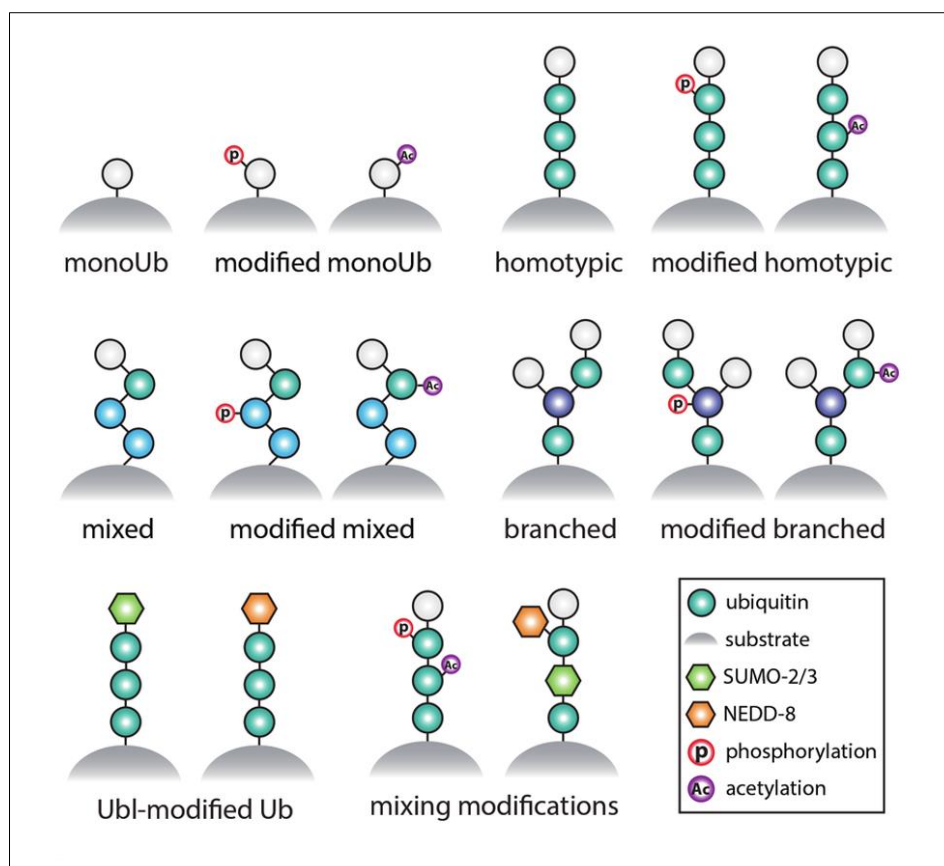
The Ub and UbL pathways represent essential branches of the proteostasis network and play a central role in the management of misfolded and aggregated proteins. Ub and UbLs are conjugated to the substrates as result of sequential reactions catalysed by several enzymes (Figure 2). First, an E1 activating enzyme forms a transient thioester bond with the Ub C-terminal glycine residue, exposing the characteristic di-glycine motif present in most UbLs. Afterwards, the activated Ub is transferred to a cysteine residue located in the active site of an E2 conjugating enzyme. Ubiquitination is finalized by an E3 ligase, which catalyses conjugation of the UbL to the substrate, typically to a lysine residue (Bett 2016).

Protein ubiquitination is a reversible process regulated by deubiquitinating enzymes (DUBs). The role of DUBs is the specific removal of Ub from substrates, which contribute to the activation and/or deactivation, recycling and localization of many regulatory proteins in different cellular processes (Farshi et al. 2015). Although modification by all the different UbLs follow the same reversible cycle, the enzymes implicated are specific for each of the UbL moieties.



**Figure 2. Ubiquitination cycle.** E1 activating enzyme forms a transient thioester bond with the ubiquitin C-terminal glycine residue. Afterwards, the activated ubiquitin is transferred to the active-site cysteine residue of an E2 conjugating enzyme. Ubiquitination is finalized by an E3 ligase, which catalyses conjugation of ubiquitin to the substrate, typically to a lysine residue. Eventually, deubiquitinating enzymes (DUBs) can remove ubiquitin from substrates. All UbLs undergo this modification/demodification cycle by means of enzymes specific for each UbL.

Substrate proteins can be ubiquitinated by a single ubiquitin moiety or multi-monoubiquitinated at different lysine residues. Ubiquitin may also form homotypic or heterotypic chains by modifying itself in any of the lysine residues in its sequence. Ubiquitin chains are homotypic if the same residue is modified during elongation, for instance K11-, K48-, or K63-linked chains. Chains are heterotypic if different linkages alternate at succeeding positions of the chain. Each moiety of the chain can be further ubiquitinated to form a branched structure. Moreover, ubiquitin moieties could undergo additional modifications as acetylation, phosphorylation or UbL molecules conjugation, among others, generating a great range of signals responsible for specific cellular outcomes. This intricate configuration system is known as the “ubiquitin code” (Figure 3) (Swatek and Komander 2016).



**Figure 3. The Ubiquitin code.** Conceptual representation of some of the possible ubiquitin and Ubl (NEDD8, SUMO2/3, etc.) chains, and chemical modifications of ubiquitin. Edited from (Swatek and Komander 2016).

Besides ubiquitination, SUMOylation represents the best-studied Ubl modification and can compete with ubiquitination for modification of the same lysine residues. SUMO is also covalently attached to an acceptor lysine of target proteins. In general, many SUMOylation sites follow a consensus motif  $\psi$ -K-X-E/D ( $\psi$  is a hydrophobic amino acid, K is the target lysine, X is any amino acid and D/E is aspartic or glutamic acid), but in addition to this, different non-canonical consensus sequences have been described (Matic et al. 2010; Hietakangas et al. 2006).

There are four SUMO isoforms in humans, designated SUMO1, SUMO2, SUMO3 and SUMO4. While SUMO1 is distinct from the other family members (47% sequence identity between SUMO1 and SUMO2/3), SUMO2 and SUMO3 share a

sequence identity of 97% and are mostly referred to as SUMO2/3. SUMO4 shares 87% sequence identity with SUMO2 however, little is known about its expression and function (Bohren et al. 2004). SUMO1 and SUMO2/3 are ubiquitously expressed. The main target of SUMO1 is RANGAP1 (Ran GTPase-activating protein 1), while SUMO2/3 is normally present in the cells unconjugated and is rapidly attached to substrates under cellular stress conditions (Eifler and Vertegaal 2015). Indeed, SUMOylation levels are dynamically regulated by various stresses such as heat shock, hypoxia or nutrient depletion. While SUMO2/3 can form polymeric chains, SUMO1 is generally considered as a chain terminator when incorporated in SUMO2/3 chains (Enserink 2015).

Before to be attached on targets, a SUMO precursor protein is processed by SUMO proteases (SENPs) to expose the carboxyl-terminal di-glycine motif. Next, similarly to ubiquitin, the mature form of SUMO is first activated by the heterodimeric SUMO activating enzyme (E1) composed of SAE1 and SAE2 (also known as UBA2) and then transferred to the single SUMO E2 conjugation enzyme UBC9 (also known as UBE2I). Finally, the activity of SUMO E3 ligases is necessary for efficient conjugation of SUMO to its substrates. Since SUMOylation is a reversible process, SUMO proteases are also responsible of SUMO cleavage from its substrate.

SUMO proteins could also interact with other proteins in a non-covalent way through specific recognition sequences called SUMO Interaction Motifs (SIMs), which can contribute to the mechanism and consequences of SUMOylation. SIMs are the predominant motifs that bind SUMO proteins and generally are composed of four hydrophobic residues that are flanked by acidic residues or residues that can be phosphorylated to generate negative charge (Merrill et al. 2010). SIMs are found in the SUMO activating enzyme UBA2, in all known SUMO E3s, in some SUMO substrates and receptors, as well as in some ubiquitin E3s called StUbls (SUMO targeted Ubiquitin Ligases). This is a conserved subfamily of ubiquitin E3 ligases, which can specifically target and bind SUMOylated proteins facilitating their ubiquitination. Besides containing a RING domain required for their ubiquitination activity, members of this family possess multiple SIMs to target SUMOylated substrates for ubiquitination (Ohkuni et al. 2018). Two StUbls have been identified in humans, RNF4 and RNF111.



SUMO-ubiquitin chains on a protein either can act as a recruitment signal or can target proteins to the proteasome.

Besides degradation, SUMOylation might affect the target protein in many different ways, including changes in localization, protein-protein interactions, protein activity and stability. While ubiquitination is generally linked to degradation, this is not necessarily the case for SUMOylation (Celen and Sahin 2020). For example, SUMOylation of Oct4 leads to significantly increased Oct4 stability and increased DNA binding (Wei, Schöler, and Atchison 2007). Regulatory factors responsible for chromatin remodelling and modification are also subject to SUMOylation, which alters their activities. For instance, SUMOylation enhances DNMT1 (DNA methyltransferase 1) activity promoting DNA methylation (Lee and Muller 2009). Moreover, SUMOylation of HDAC1 (histone deacetylase 1) is required for its full transcriptional repression activities at defined promoters (David, Neptune, and DePinho 2002). SUMOylation is therefore involved in many biological processes, including regulation of transcription, DNA repair, immunity and development, among others (Talamillo et al. 2020; Garvin 2019).

The imbalance of SUMOylation and deSUMOylation has been associated with the occurrence and progression of various diseases. Cancer development can be attributed to abnormal SUMOylation as consequence of the enhanced expression of the enzymes involved in the SUMO modification pathway (Seeler and Dejean 2017). Moreover, SUMOylation contributes to different neurological disorders such as Parkinson's disease, Alzheimer's disease and Huntington's disease in which SUMO has been found to either enhance or protect against the aggregation of neuronal proteins (Yang et al. 2017).

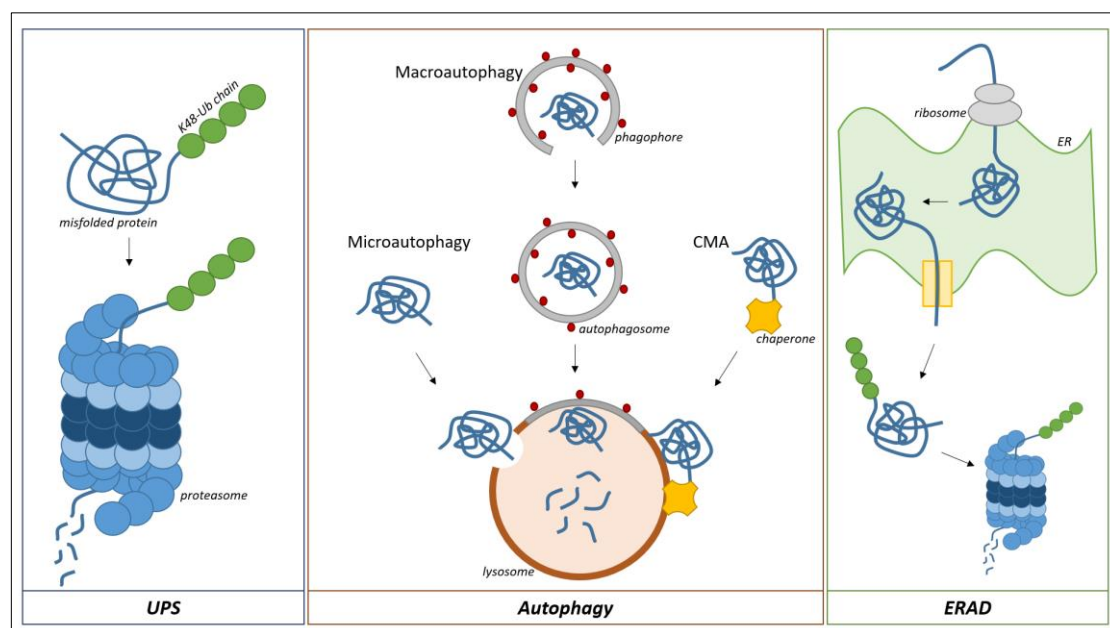
### **1.1.2 Degradation of misfolded proteins**

Protein folding is an intrinsically error-prone process that frequently results in toxic protein-protein interactions and that must be efficiently managed by the cellular proteostasis network. Several different pathways exist for carrying out these degradation processes. When proteins are unfolded or misfolded, they are typically degraded via the ubiquitin-proteasome system (UPS) or the endoplasmic-reticulum-

associated protein degradation (ERAD). Autophagy can also be used as a proteostatic degradation mechanism (Figure 4).

The UPS represent the primary pathway the cells use to recognize and clear soluble misfolded proteins. When misfolded proteins need to be degraded, they are typically tagged with a K48-linked polyubiquitin chain to mediate binding by ubiquitin receptors on the 26S proteasome (Rpn10/S5a and Rpn13) required for substrate-mediated activation of the proteasome's unfolding ability (Cundiff et al. 2019). The unfolded protein response (UPR) in the endoplasmic reticulum (ER) is activated by imbalances of misfolded proteins inside the ER. When incorrectly folded, proteins are retro-translocated to the cytosol and ubiquitinated for proteasomal degradation.

Aggregated proteins that cannot be unfolded for proteasomal degradation may be removed by autophagy and lysosomal/vacuolar degradation. In mammalian cells, three types of autophagy have been described: macroautophagy, microautophagy, and chaperone-mediated autophagy (CMA). Macroautophagy requires sequestration of substrates in a double membrane vesicle, which then fuses with lysosomes; microautophagy involves direct engulfment of cytosolic components by lysosomes. In CMA, substrates are recognized by chaperones and targeted to lysosomes for degradation.



**Figure 4. Unfolded protein response.** Misfolded proteins are typically marked with K48-linked Ub chain and degraded by the proteasome (left panel). Three types of autophagy may remove proteins that cannot be unfolded for proteasomal degradation. Macroautophagy requires sequestration of substrates in a double membrane vesicle, which then fuses with lysosomes; microautophagy involves direct engulfment of cytosolic components by lysosomes. In chaperone-mediated autophagy (CMA), substrates are recognized by chaperones and targeted to lysosomes for degradation (middle panel). Misfolded proteins inside the ER are retro-translocated to the cytosol, and ubiquitinated for proteasomal degradation (left panel).

### 1.1.3 Proteostasis dysfunction and therapeutic strategies

Mutations in genes involved in protein synthesis, folding, aggregation, autophagy, mitophagy, ER stress or the UPS can result in proteostasis dysregulation. These mutations could be inherited and differ in phenotypic severity from having no noticeable effect to embryonic lethality. Disease develops when these mutations significantly affect proteins making them more prone to misfolding, aggregation and degradation. If this effect only alters the mutated protein, the negative consequences will only be local loss of function. However, mutations can also cause dominant negative effects interfering with the function of wild type (WT) proteins. Differently, if mutations occur in a chaperone or a protein that interacts with many other proteins, dramatic global alterations in the proteostasis boundary will occur (Powers et al. 2009; Hipp, Park, and Hartl 2014).

Two main strategies have been used for therapeutic development targeting the proteostasis network: pharmacologic chaperones and proteostasis regulators. Pharmacological chaperones are cell-permeant small molecules that bind to and stabilize target proteins. They are very specific and have effects on particular proteins. Most of the pharmacological chaperones are substrates or active site-directed competitive inhibitors and are very efficient at lower concentrations. However, a pharmacological chaperone cannot bind to all the variants present in a disorder (Haneef and Doss 2016). Target proteins, with the help of chaperones, fold correctly

and hence pass through the quality control system of the ER and traffic safely from there to the designated location.

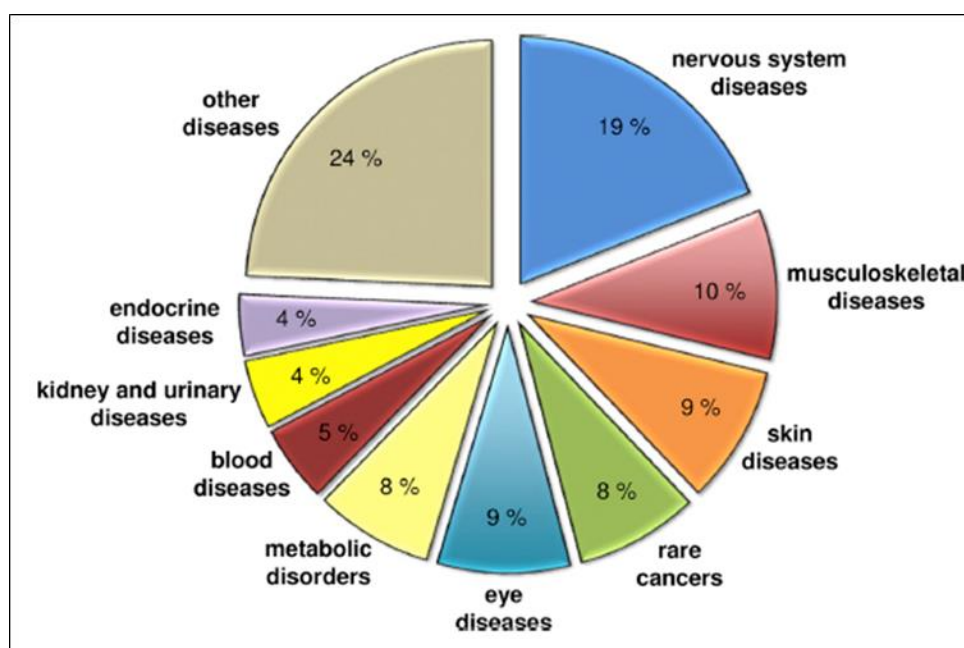
Instead, proteostasis regulators (PRs) work by manipulating signalling pathways, including the UPR, resulting in transcriptional and translational upregulation of proteostasis network components. PRs manipulate the system through different mechanisms. Some PRs negatively affect protein production to inhibit global protein generation. Other regulators are used to either increase the production of molecular chaperones or modify their function. Moreover, some PRs function by directly manipulating the proteasomal system either by increasing its activity when cells are under stress or through inhibiting the system to prevent premature degradation (Mohamed et al. 2017).

Understanding how proteins function is essential to understanding how cells work. Model systems, including laboratory animals and cell-based systems, continue to have an important role in the identification of functional mechanisms that safeguard proteostasis and are essential to discover and develop novel and better drugs for the treatment of human diseases. Model systems of diverse misfolding-prone disease proteins have so far revealed numerous chaperone and co-chaperone modifiers involved in proteostasis maintenance (Brehme and Voisine 2016). Furthermore, protein function is usually accomplished by interactions with other proteins. To unveil proteins behaviour and regulation can allow better predicting, preventing, diagnosing, and treating disease.

## 1.2 Rare diseases

Maintenance of cellular proteostasis is essential for cell functioning and survival. Alteration of proteostasis regulatory processes might severely affect cellular functions and might directly or indirectly be associated to the aetiology of many diseases, including those with very low prevalence (Osinalde et al. 2019). Rare diseases are those that affect a small number of people compared to the general population and specific issues are raised in relation to their rarity. In Europe, a disease is considered rare when it affects less than 1 in 2000 people. An individual rare disease

may affect only one person in a million but, all together, rare disease patients comprise 6% to 8% of the EU population (Eurodis). Distribution of different types of rare diseases is illustrated in Figure 5. Rare diseases are often severe, chronic and progressive disorders. Although their signs are usually observed at birth or in childhood, over 50% of rare diseases, appear during adulthood (Orphanet). Because the number of people affected with any one specific rare disease is relatively small, the development of effective drugs and medical devices to prevent, diagnose, treat, or cure these conditions becomes extremely complicated.



**Figure 5. Distribution of different types of rare diseases.** The pie chart represents rare disease categories and their frequencies, given as the percentage of the total number of rare diseases. Categories with a frequency < 4% were grouped into the category termed “Other diseases”. Categories and frequencies shown in the graph were extracted from information provided by the Genetic and Rare Diseases Information Center (GARD, [rarediseases.info.nih.gov](http://rarediseases.info.nih.gov)). Edited from (Osinalde et al. 2019).

Rare diseases can have multiple causes. They are mostly inherited but may also arise as result of de novo mutations. Besides, some diseases may be acquired because of a toxic exposure, infection or radiation. Many, if not most, are caused by defects in a single gene, as in the case of the alpha 1 antitrypsin deficiency (AATD), in which mutation in *SERPINA1* gene may cause serious lung or liver disease; or Friedreich's

ataxia, a neurological disorder caused by mutation in *FXN* gene that may also be accompanied by cardiac and other problems. Multiple different mutations in that single gene may result in disease with varying features or severity while, in other cases, mutations in a single gene can produce pleiotropic effects. For example, fragile X syndrome is caused upon transcriptional silencing of the *FMR1* gene. Its product, the FMRP protein, is responsible for binding and repressing the translation of hundreds of mRNAs so potentially affecting multiple pathways (Osinalde et al. 2019).

In some rare conditions, multiple genes may contribute collectively to the manifestations of the disorder. For instance, in the Williams-Beuren syndrome the gene responsible for the production of the elastin protein (ELN) is clearly identified as producing the cardiovascular problems, (Pober 2010). Further, some diseases may present variable expressivity and incomplete penetrance as in the case of the Neurofibromatosis type 1 and Marfan syndrome (Cimino and Gutmann 2018; Cañadas et al. 2010).

Due to their wide aetiology, the diagnosis and treatment of rare diseases is not straightforward. The complex set of factors that influence particular genetic changes can make the diagnosis extremely difficult. Moreover, for many rare diseases, there are no available therapies and lack of adequate models makes even more difficult the study of new possible treatments.

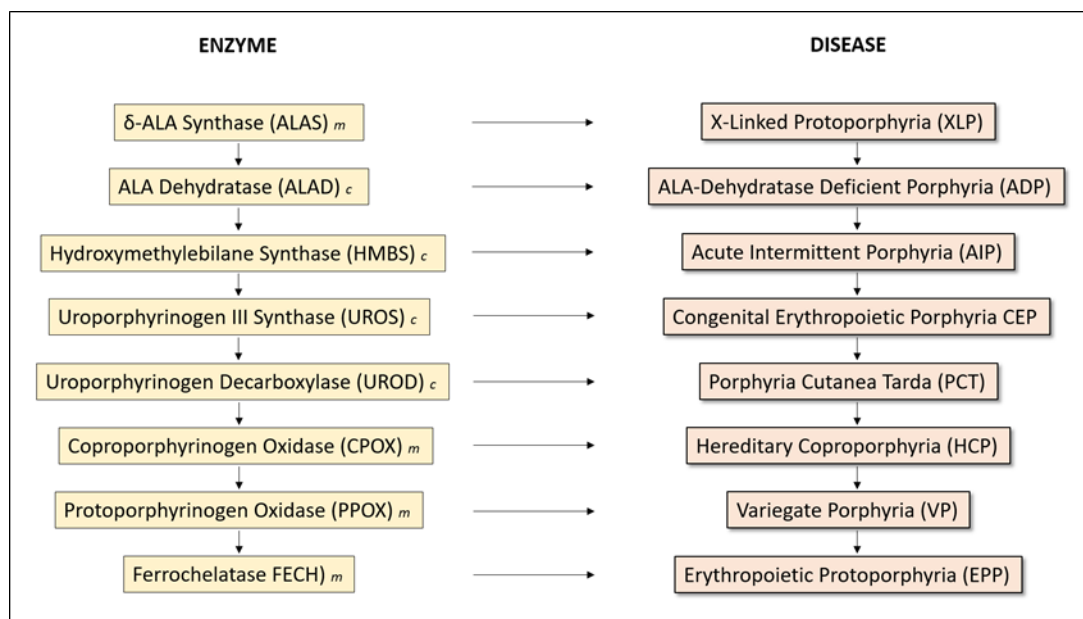
This work is focused on two different rare diseases: Congenital Erythropoietic Porphyria (CEP) and Townes-Brock Syndrome (TBS). Both diseases are extremely rare conditions caused by proteostasis alterations. However, the mechanisms underlying each condition are different. While CEP is caused by mutations that alter the stability of an enzyme involved in the synthesis of the heme group, TBS is originated by the

premature truncation of the *SALL1* gene, which triggers changes in the stability of partner proteins.

## 1.3 Congenital Erythropoietic Porphyria: a metabolic rare disease

### 1.3.1 Porphyrins and alterations in the heme biosynthesis pathway

Porphyrins belong to a group of rare diseases characterized by a deficiency (inherited or acquired) of the enzymes involved in the heme biosynthesis, resulting in an abnormal accumulation of precursors molecules called porphyrins. Heme biosynthesis is a chemical process, constituted by a series of sequential reactions carried out by eight different intracellular enzymes localized in the mitochondria or the cytosol. Deficiency of each one of those enzymes cause a different type of porphyria (Figure 6).



**Figure 6. Schematic representation of heme pathway-associated porphyrias.** Deficiencies of each one of the eight enzymes involved in heme biosynthetic pathway cause a different type of porphyria. Subcellular localization per each enzyme is indicated as m: mitochondrial; c: cytoplasmic.

Porphyrias are classified as hepatic or acute and erythropoietic/cutaneous or chronic. This classification is based on the nature of the clinical manifestation and whether the primarily site of the precursor and/or porphyrins deposition is the liver or the erythron. The hepatic porphyrias include Acute Intermittent Porphyria (AIP), Hereditary Coproporphyria (HCP), Variegate Porphyria (VP) and ALA-Dehydratase Deficient Porphyria (ADP). Erythropoietic porphyrias include Congenital Erythropoietic Porphyria (CEP), Erythropoietic Protoporphyria (EPP) and X-Linked Protoporphyria (XLP). The Porphyria Cutanea Tarda (PCT) is considered to be a hepatocutaneous porphyria, as it presents with cutaneous lesions but the primarily site of porphyrin accumulation is the liver. PCT is the most common porphyria, with a prevalence of 1 in 10.000. The most common acute porphyria, AIP, has a prevalence of approximately 1 in 20.000, and the prevalence of the most common erythropoietic porphyria, EPP, is estimated in 1 among 50.000 to 75.000 persons. CEP is extremely rare, with an estimated prevalence of 1 in 1.000.000 or less, existing approximately 280 reported cases worldwide. ADP is also extremely rare with only six confirmed cases reported worldwide (Ramanujam and Anderson 2015).

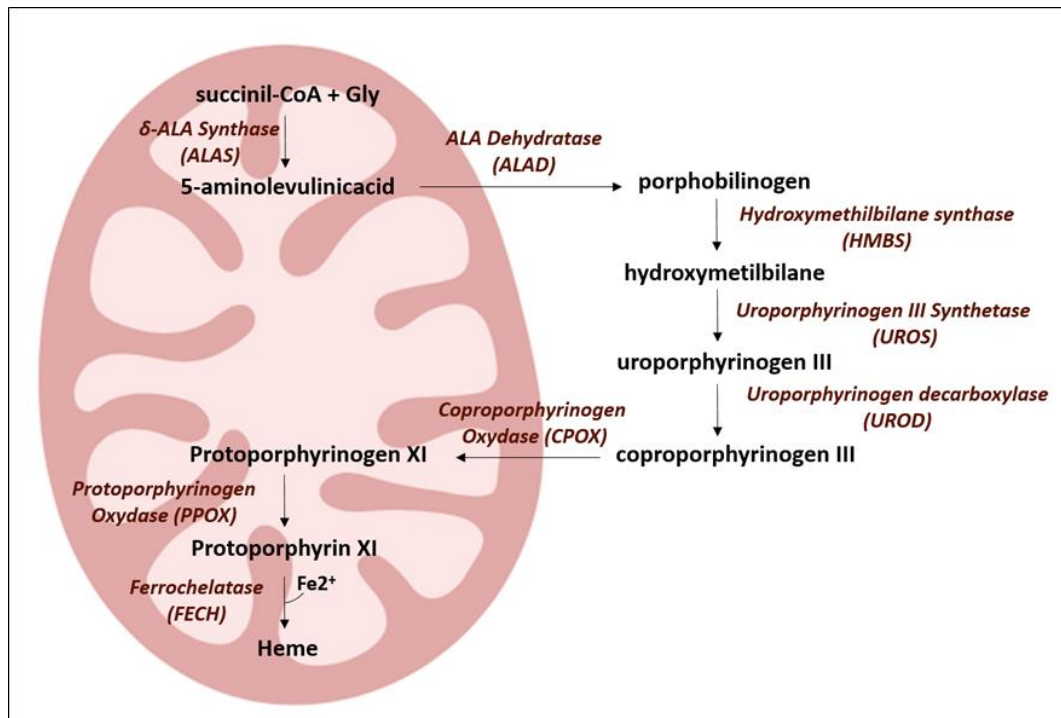
Heme is an important cofactor that plays an important role in the reactions of oxidation-reduction and in oxygen transport. It is necessary for the activity of a variety of hemoproteins, such as hemoglobin, myoglobin, respiratory cytochromes, and cytochrome P450 enzymes. The first enzyme in heme synthesis,  $\delta$ -aminolevulinic synthase (ALAS), as well as the last three enzymes, coproporphyrinogen oxidase (CPOX), protoporphyrinogen oxidase (PPOX) and ferrochelatase (FECH) are mitochondrial, whereas  $\delta$ -aminolevulinic acid dehydratase (ALAD), hydroxymethylbilane synthase (HMBS), uroporphyrinogen III synthase (UROS), and uroporphyrinogen decarboxylase (UROD) are localized in the cytosol.

The synthesis of heme begins with the formation of  $\delta$ -aminolevulinic acid (ALA) by ALAS from the amino acid glycine and succinyl-CoA from the citric acid cycle. This step represents the first and rate-limiting reaction in the pathway. In mammals, two different ALAS isoforms have been described, ALAS1 and ALAS2. Whereas ALAS1 is ubiquitously expressed in all tissues providing the basic needs of heme in the non-erythropoietic cells, ALAS2 is expressed exclusively in the erythropoietic cells sustaining high levels of heme needed during the late stages of erythropoietic



differentiation. In non-erythrocyte cells, *ALAS1* is regulated by negative feedback of the free heme pool produced. In the erythrocytes, the expression of *ALAS2* is regulated by erythroid-specific transcription factors, like GATA1, and, at the posttranscriptional level, by iron through a 5' iron responsive element (IRE) that interacts with iron responsive proteins (IRPs). Under conditions of iron deficiency, the translation of *ALAS2* mRNA is inhibited by IRPs binding to the 5' IRE. However, when intracellular iron level increases, IRPs are degraded allowing the translation of *ALAS2* mRNA (Chiabrando, Mercurio, and Tolosano 2014).

In the heme synthesis process, schematized in Figure 7, the *neo*-formed ALA exits the mitochondria into the cytosol where two molecules of ALA condense to produce the pyrrole ring compound, porphobilinogen (PBG). The next step of the pathway involves condensation of four molecules of PBG, aligned to form the linear hydroxymethylbilane (HMB). HMB is then closed to form an asymmetric pyrrole ring D called uroporphyrinogen III (URO III). Next, this is modified to produce coproporphyrinogen III (COPRO III). Following its synthesis, COPRO III is transported into mitochondria and decarboxylated to form the protoporphyrinogen IX. Finally, protoporphyrinogen IX is converted to protoporphyrin IX. The final reaction involves the insertion of ferrous iron ( $\text{Fe}^{2+}$ ) into protoporphyrin IX leading to the formation of heme.



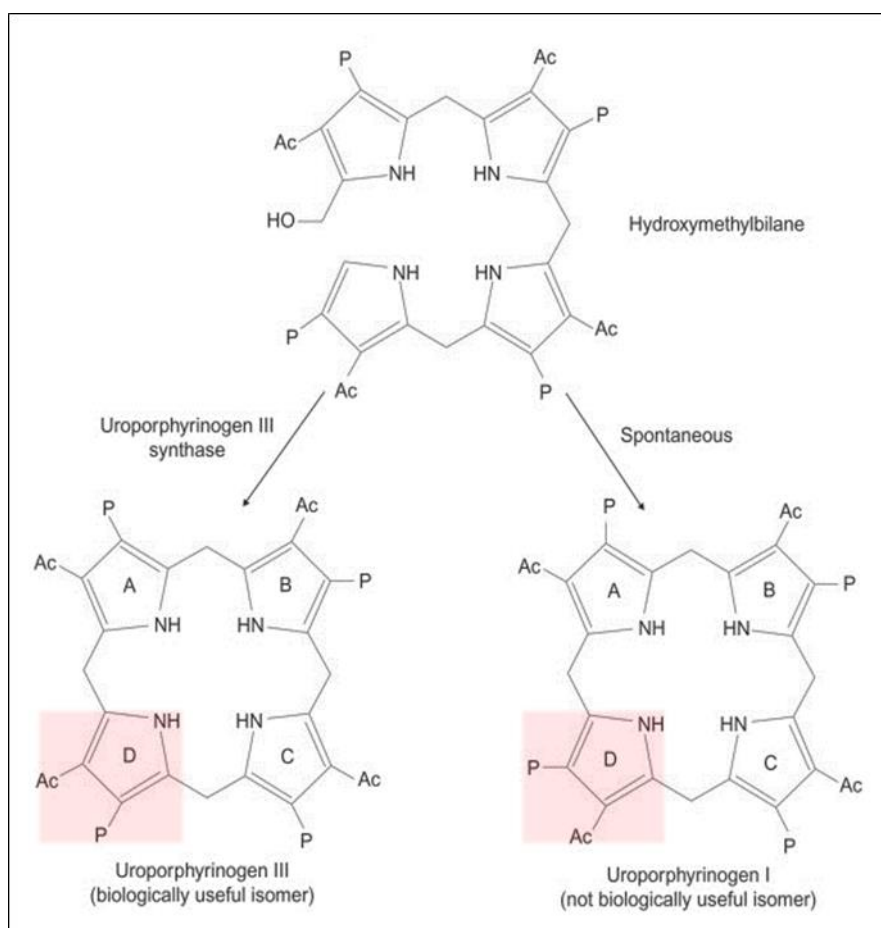
**Figure 7. Schematic representation of the heme biosynthesis pathway.** Enzymes are indicated in dark red, while the respective substrates are reported in black. ALAS, CPOX, PPOX and FECH are mitochondrial, whereas ALAD, HMBS, UROS and UROD are localized in the cytosol.

### 1.3.2 Congenital Erythropoietic Porphyria: a rare disease based on protein stability alterations

CEP (ICD-10 #E80.0; MIM #263700), also known as Gunther's disease, is an extremely rare disease inherited as an autosomal recessive trait. It is caused by an inborn error in the heme biosynthesis, mostly derived from the deficient activity, although not complete absence (<1 to ~10% of normal activity), of the fourth cytosolic enzyme of the pathway, UROS (EC 4.2.1.75).

As shown in Figure 8, UROS catalyses the rapid cyclization of most of the linear HMB, inverting the configuration of one of the four aromatic rings and leading to the formation of URO III. In normal conditions, a little amount of HMB suffers a spontaneous closure resulting into uroporphyrinogen I (URO I). URO I and III can then follow into the biosynthesis chain as substrates for the next enzyme of the pathway, UROD, which catalyses the transformation of URO I and URO III in COPRO I

(coproporphyrinogen I) and COPRO III, respectively. While URO III and COPRO III are proper substrates for the subsequent reactions of the pathway, URO I and COPRO I are accumulated in the tissues and excreted in urine and faeces. In CEP, the markedly reduction of the UROS activity leads to the specific and massive accumulation of type I porphyrins, URO I and COPRO I, which result toxic and are responsible for the symptoms.



**Figure 8. Uroporphyrinogen formation.** The closure of the HMB by UROS leads to the formation of uroporphyrinogen III isomer, while the spontaneous closure produces the isomer I. The red squares highlight the different configuration of the ring D. Adapted from (Bhagavan N. V., Chung-Eun Ha 2015).

Patients are either homozygous or composed heterozygous and the clinical severity is mostly mutation-dependent. Nevertheless, cases of variable expressivity have been reported (Desnick et al. 1998; Ged et al. 2009). Photosensitivity and

hemolytic anemia with associated splenomegaly are the main symptoms of CEP. In the bone marrow, it has been observed that developing erythropoietic cells contain large amounts of porphyrins. Under fluorescent microscope, most normoblasts show localized red fluorescence heme-containing inclusion bodies into the nucleolus. The overproduction of porphyrin isomers I, accumulated mainly in erythroid cells, is associated with dyserythropoiesis, poikilocytosis, and bone marrow erythroid-hyperplasia (Merino, To-Figueras, and Herrero 2006; Katugampola, Badminton, et al. 2012).

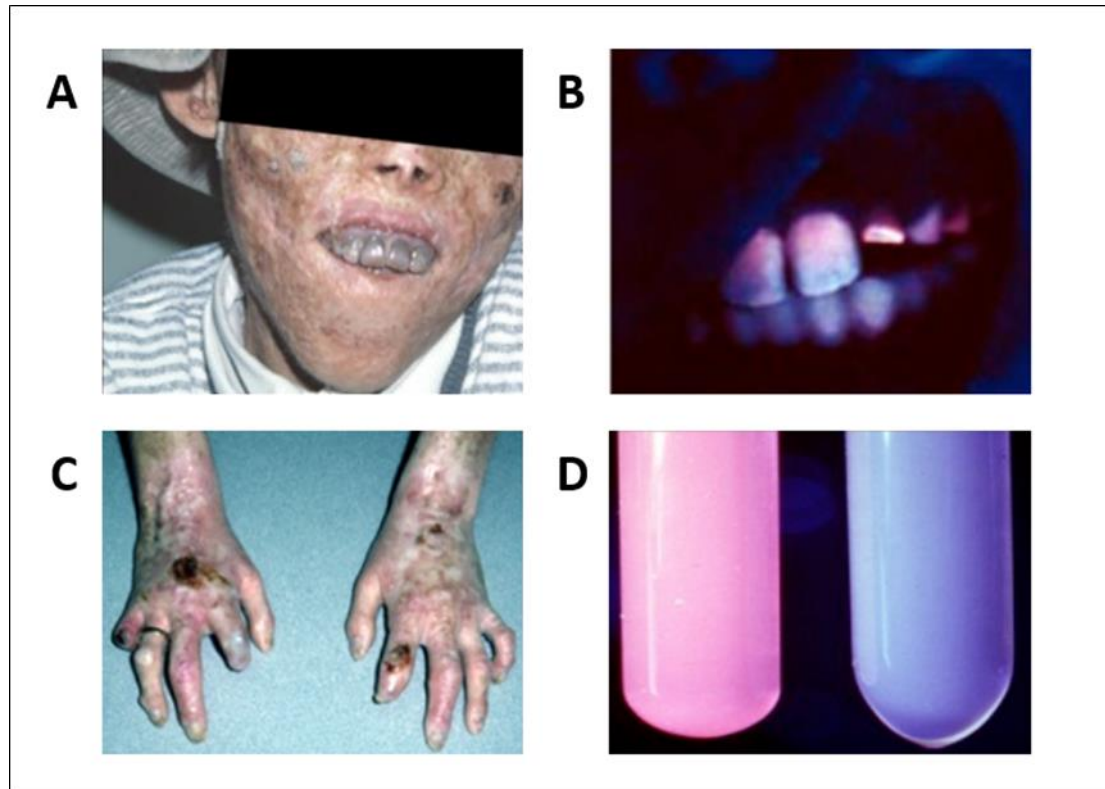
The electronic configuration of porphyrins endows them of photochemical properties that explain porphyrin phototoxicity. In fact, porphyrins are excited by visible and ultraviolet (UV) light (excitation 490 nm, emission 650 nm). Hemolysis causes porphyrins release into the plasma with subsequent deposits in the tissues. Porphyrins accumulation induces photodynamic reactions that stimulate the production of reactive oxygen species (ROS), which directly damage tissues and indirectly promote a proinflammatory response. The severity of cutaneous reaction depends on porphyrin amount in the tissue and the degree of light exposure (Di Pierro, Brancaleoni, and Granata 2016).

CEP photosensitivity starts early in life and causes a series of cutaneous wounds such as vesicular or bullous eruptions in the skin exposed to light (Figure 9). Ultimately, continuous injuries and secondary skin infections with scars and hyperpigmentation can lead to severe mutilation of fingers, hands, ears, lips and nose. Other characteristic symptoms are porphyrin accumulation in bone marrow, plasma and urine, hypertrichosis or hirsutism (abnormal amount of hair growth over the body and excessive body hair on parts of the body where hair is normally absent or minimal, respectively) and erythrodontia (red discoloration of teeth) (Erwin et al. 1993).

### **1.3.3 Diagnosis and treatment of Congenital Erythropoietic Porphyria**

CEP diagnosis usually occurs in childhood from the early onset of severe skin photosensitivity, pink to dark brown fluorescent urine under Wood's light (long-wave UV light lamp) and erythrodontia (Figure 9). The definitive diagnosis requires

biochemical analysis of porphyrins from biological samples as urine, stool, plasma, or red blood cells. Determination of the residual activity of UROS completes the diagnosis. Enzymatic methods are implied to analyse isomers I and III formation, which are preferentially detected by high-performance liquid chromatography (HPLC).



**Figure 9. Symptoms of Congenital Erythropoietic Porphyria.** A-C) A severely affected CEP patient who had multiple sun-induced skin lesions. The cutaneous bullae and vesicles burst and became secondarily infected, leading to bone involvement and resultant loss of facial features and digits. B) Note the brownish discolored teeth, which fluoresce (erythrodontia) when exposed to ultraviolet light. The erythrodontia is the result of the accumulation of uroporphyrin I and coproporphyrin I in his teeth. D) Urine from a CEP patient that fluoresces red under ultraviolet light (left) and from a healthy person (right). Adapted from (Balwani and Desnick 2012).

Currently, there are no treatments available for CEP. The only curative therapies are bone marrow or hematopoietic stem cell transplantation (HSCT). HSCT is the most effective treatment to CEP; however, it entails many operative and post-operative risks and complications for the patients. Oral administration of beta-carotene, photoprotection and sunlight avoidance are recommended for CEP patients

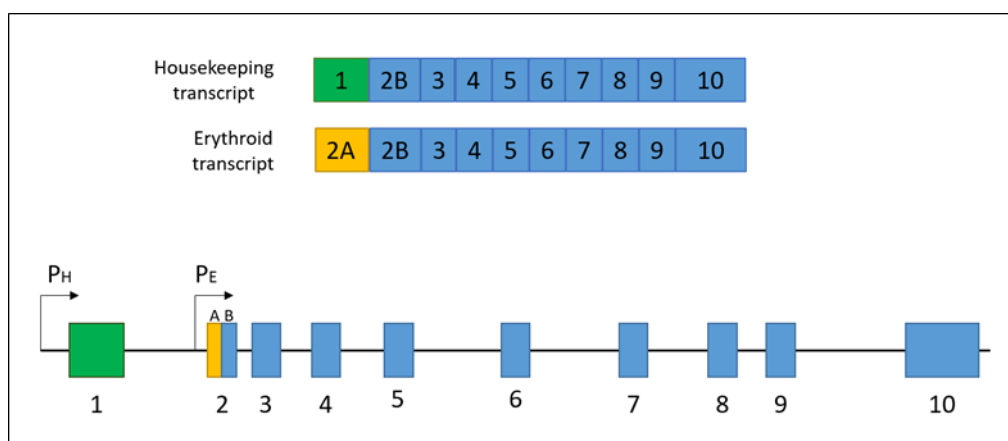
in order to prevent skin injuries (Katugampola, Anstey, et al. 2012; Kauffman et al. 1991; Hogeling et al. 2011; Dupuis-Girod et al. 2005).

Suppression of erythropoiesis, short or long-term blood transfusions, splenectomy in severe cases, and use of agents that help excretion of porphyrins can be employed as palliative therapies to alleviate symptomatology. However, these treatments show high variability among patients and produce side effects as iron overload, thrombocytopenia or infections, among others (Fritsch et al. 1997; Balwani and Desnick 2012).

In view of the current situation, the search for a definitive treatment is urgently needed for these patients. In order to better approach therapeutic development, a comprehensive analysis of the genetic alterations and their consequences at the protein level is required.

#### **1.3.4 Uroporphyrinogen III Synthase mutations and stability**

The UROS human protein is a monomer of 29.5 kDa encoded by the *UROS* gene (Figure 10). This is located on chromosome 10q25.2-q26.3 and consists of 10 exons and 10 introns. *UROS* gene has two different promoters that generate housekeeping and erythroid-specific transcripts (Figure 10). The two transcripts share nine common coding exons (2B to 10) but have unique 5'-untranslated sequences (exons 1 and 2A). The housekeeping transcript is expressed in all tissues, while the erythroid transcript is expressed only in erythropoietic tissues and its activity is increased during hemin-induced erythroid differentiation. An alternative erythroid-specific promoter is present also in the first introns of the genes that encode the second and third enzymes in the heme biosynthetic pathway. This suggests a common regulatory mechanism for erythropoiesis and erythroid differentiation, which may have evolved from an ancestral sequence that could have duplicated and diverged into separate enzymes in the pathway (Aizencang et al. 2000).



**Figure 10. Schematic representation of UROS gene.** Blue boxes indicate common exons, green and yellow boxes indicate alternative housekeeping and erythroid 5' UTRs, respectively. Arrows indicate the housekeeping (PH) and erythroid (PE) promoter regions. Edited from (Aizencang et al. 2000).

To date, over 49 UROS disease-causing mutations are reported, according with the Human Gene Mutation Database ([www.hgmd.org](http://www.hgmd.org)), including missense/nonsense mutations (29), splicing defects (4), insertions/deletions (10) and mutations in regulatory regions (6). Moreover, a gain of function mutation in the *ALAS2* gene and trans-acting mutation in *GATA1* gene (R216W) have also been described as modulator or causative of CEP, respectively (Phillips et al. 2007).

The C73R (c.217 T>C) missense mutation in the exon 4 is the most common pathologic allele found in about 40% of the patients and is associated with a severe phenotype (Frank et al. 1998). Homozygous patients for this mutation have less than 1% UROS activity as consequence of a misfolded and unstable protein, which is prematurely degraded via the proteasome (Bishop et al. 2006; 2011). Fortian and co-workers demonstrated that the cysteine in position 73 is not essential for the catalytic activity of the enzyme, but its mutation to arginine speeds up the process of irreversible unfolding and aggregation. The enzyme itself retains about 50% of normal activity but its quick degradation via the UPS dramatically reduces the amount of intracellular available enzyme. Moreover, mutant protein levels can be restored upon cell treatment with the proteasome inhibitor MG132 (Fortian et al. 2009). Recently,

Urquiza and co-worker showed that the off-patent synthetic antimicrobial ciclopirox (CPX) acts as an UROS pharmacological chaperone in vitro and in vivo. CPX acted as an allosteric chemical stabilizer of UROS and did not affect the enzyme's catalytic role (Urquiza et al. 2018).

Mouse models of human porphyrias have proven useful to investigate the pathogenesis of the diseases and to facilitate the development of new therapeutic approaches. To date, mouse models have been generated for all major porphyrias, including few different mouse models for CEP. They recapitulate many of the clinical and biochemical features of this severe human disease. Moreover, they have been particularly useful to study the mechanisms underlying the disease and for preclinical evaluation of novel therapeutic approaches. In spite of this, homozygous C73R mouse model has poor viability and chronic health issues, posing numerous practical and ethical challenges for experimental setups (Bishop et al. 2006). Therefore, it is urgent to find new strategies that ensure the generation of viable models. Since presently the most promising therapeutic approach for CEP is based on the use of pharmacological chaperons, a humanized model could improve drug discovery providing a better system for designing and testing highly specific compounds. Additionally, an inducible model could partially solve viability and health problems, facilitating experimental procedures.

### **1.3.5 Genome editing technology: CRISPR/Cas9**

Nowadays, precise targeted gene modification is considered the standard procedure to analyse gene function and to generate animal models providing the potential for therapeutical applications. During the last few decades, simple and economic methods for gene-targeted modification have been engineered, including zinc finger nucleases, transcription activator-like effector nucleases and CRISPR (Clustered Regularly Interspaced Short Palindromic Repeats)/Cas9 system. These nucleases generate a DNA double-strand break (DSB) at the targeted genome locus. The break induces the repair response through an error-prone non-homologous end joining (NHEJ) or homology-directed repair (HDR). NHEJ is activated in the absence of a template and results in insertions and/or deletions (indels) that disrupt the targeted

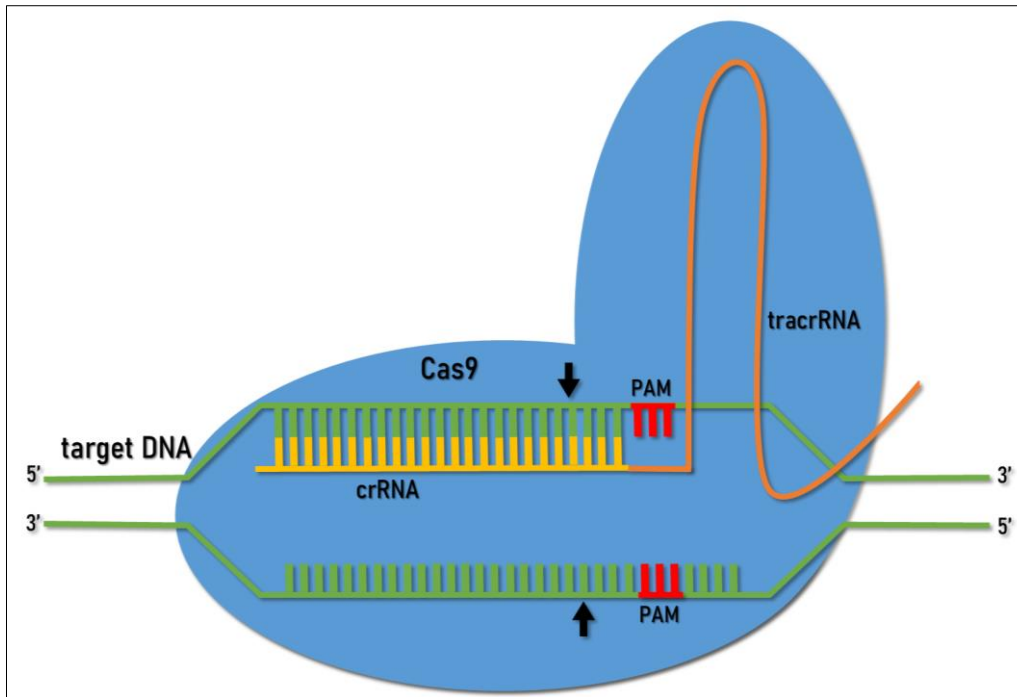


locus. In presence of a donor template with homology to the targeted locus, the HDR pathway is activated, allowing precise repair or the generation of specific mutations (Ma, Zhang, and Huang 2014).

The CRISPR/Cas9 system was first described as an adaptive immune system in prokaryotes, which use it as a powerful defensive strategy against viral invaders. Now, it has been engineered for genome editing. CRISPR genome editing system is based on a non-specific endonuclease (Cas9 or the closely related Cpf1) to cut the genome and on two RNAs: the 20 nucleotides crRNA (crispr RNA) that defines the genomic target for Cas9, and the tracrRNA (trans-activating crispr RNA), which acts as a scaffold linking the crRNA to Cas9 (Figure 11). In most of experimental procedures, these two small RNAs are condensed into one RNA sequence known as guide RNA (gRNA) or single guide RNA (sgRNA).

Cas9 and its variants have two endonuclease domains: The N-terminal RuvC-like nuclease domain and the HNH-like nuclease domain near the centre of the protein. Once bound to the target, Cas9 undergoes a conformational change that positions the nuclease domains to cleave opposite strands of the target DNA. A "nickase" Cas9 mutant, which cuts only one strand of DNA, is commonly used with paired gRNAs to lower off-target cuts frequency. gRNA length has also been optimized. Truncated gRNAs with <20 base homology shown less off-target activity (Fu et al. 2014).

In order to be cut by Cas9, the DNA target must contain a 3-base pair sequence, known as the Protospacer Adjacent Motif or PAM, immediately downstream (3') of the site targeted by the guide RNA. Thus, the result of Cas9-mediated DNA cut is a DSB within the target DNA ~3-4 nucleotides upstream of the PAM sequence. In the absence of either the gRNA or a PAM sequence, Cas9 will neither bind, nor cut the target.

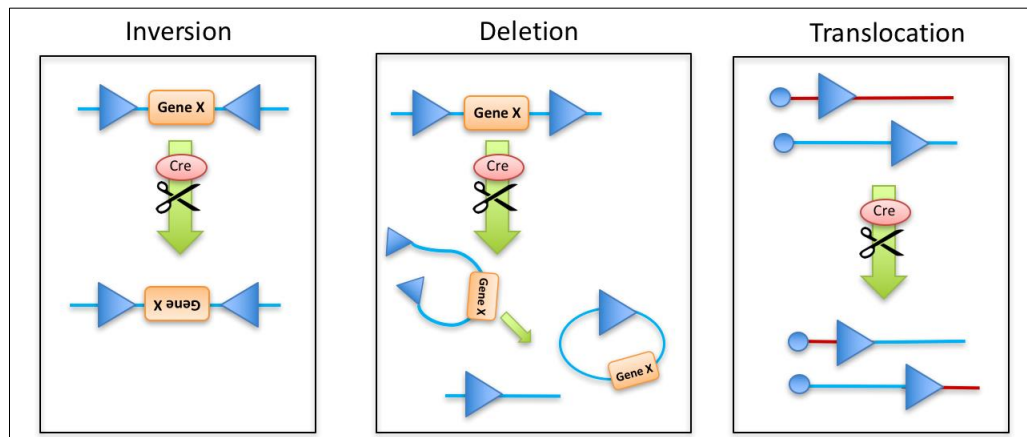


**Figure 11. CRISPR/Cas9 system.** tracrRNA acts as a scaffold for the crRNA, which guides the Cas9 on the target sequence. Cas9 produces DSB ~3-4 nucleotides upstream of PAM sequence. Arrows indicate the cutting site.

### 1.3.6 FLEX-Switch System

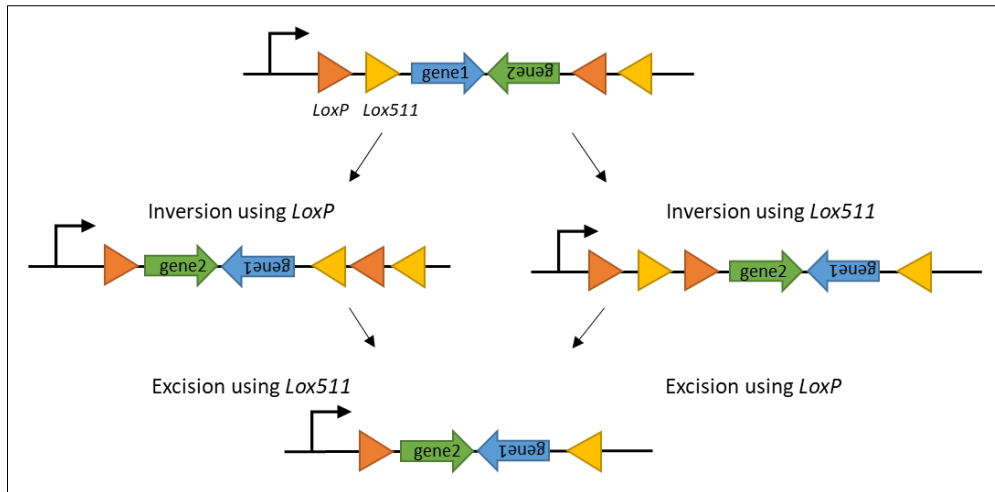
The Cre-LoxP system is a common site-specific genetic manipulation tool. This system, through a site-specific recombination, allows modulating the expression of targeted genes in cell lines and animal models, at a certain developmental stage or in specific tissues. Cre recombinase is a tyrosine recombinase enzyme derived from the P1 bacteriophage, and is a member of the  $\lambda$  integrase superfamily of SSRs (subtilisin-like serine proteases). Cre mediates DNA recombination, involving strand cleavage, exchange and ligation, through the recognition of DNA target sites, the 34 bp *LoxP* sites (locus of crossover (x) in P1). Each *LoxP* site includes two sets of 13 bp recognition sequences separated by 8 bp spacer sequence. The 13 bp sequences are palindromic while the 8 bp spacer is responsible for the directionality of the *Lox* site. Different Cre recombination outcomes may be generated based on the location and orientation of the *LoxP* sites. When the two *LoxP* sites are in opposite orientations, the sequence in

between is reversibly inverted (Figure 12, left panel). However, if two *LoxP* sites are in the same orientation, the sequence in between is excised (Figure 12, middle panel). This process is essentially irreversible. A reversible translocation event is generated at the *LoxP* sites when they are located on separate DNA molecules (Figure 12, right panel) (Friedel et al. 2011).



**Figure 12. The directionality of the *LoxP* sites determines the outcome of the Cre mediated recombination.** When the two *LoxP* sites are in the opposite orientations, the sequence in between is inverted (left panel). However, if two *LoxP* sites are in the same orientation, the sequence in between is excised (middle panel). A translocation event is generated at the *LoxP* sites when they are located on separate DNA molecules (right panel). Edited from Addgene.

Cre-dependent genetic switch (FLEX-Switch) system (Schnütgen et al. 2003) is a genetic tool based on the use of two pairs of heterotypic *Lox* sites, WT (*LoxP*) and mutant (e.g. *Lox511*). Cre recognizes both *Lox* variants, but only identical pairs of *Lox* sites can recombine with each other and not with any other variant. By combining the ability of Cre recombinase to invert or excise a DNA fragment and the availability of both WT and mutant *Lox* sites, the FLEX-Switch system allows a gene to be turned off, while the expression of another one is simultaneously turned on (Figure 13). This site-specific genetic manipulation can be combined with the CRISPR/Cas9 technology as a novel genetic tool for site-specific genetic manipulation *in vivo*.



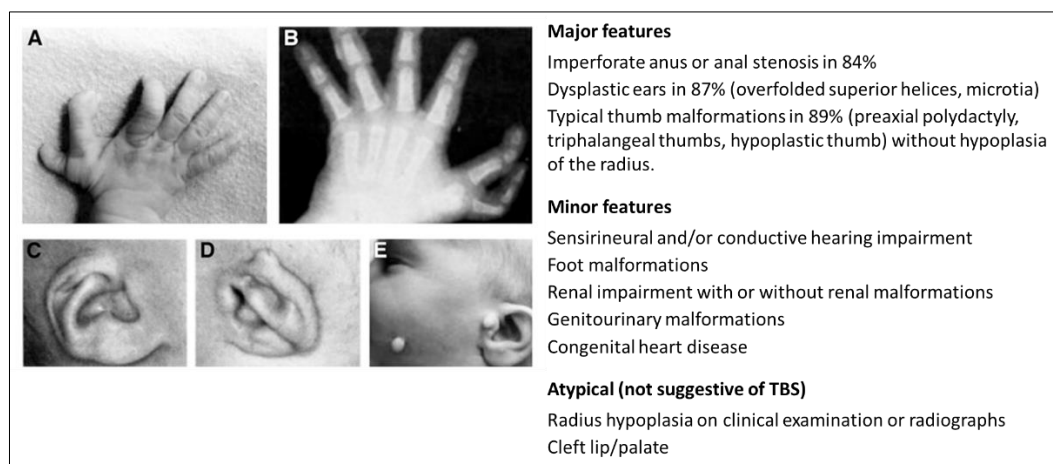
**Figure 13. FLEX-Switch System.** Incompatible pairs of *Lox* sites can be used to induce the expression of a gene of interest, simultaneously to the silencing of another gene, in two steps of recombination. First, using *LoxP* (orange arrowheads) or *Lox511* (yellow arrowheads) sites, Cre can flip the middle sequence inverting the strand sense of the genes. Next, the sequence between two *Lox* sites faced in the same direction is irreversibly excised.

Considering the advantages of this technology and the need for a new genome editing strategy, the FLEX-Switch system could represent the ideal approach for the generation of a new viable humanized inducible CEP C73R mouse model.

## 1.4 Townes-Brocks Syndrome: a developmental rare disease

TBS is a rare disease with an estimated prevalence of 1 per 350.000 live births. It is an autosomal dominant genetic disorder characterized by three major features: imperforate anus, dysplastic ears and thumb malformations such as triphalangeal thumbs, duplication of the thumb (preaxial polydactyly) and thumb hypoplasia (Powell and Michaelis 1999). TBS is also frequently associated with other minor features such as sensorineural and/or conductive hearing impairment, foot malformations, renal impairment including end-stage renal disease or polycystic kidneys, together with genitourinary malformations and congenital heart disease.

TBS features are schematized in Figure 14. The diagnosis of TBS is established by the presence of the three major features. If only two major features are present, the minor features and the absence of atypical features further support the diagnosis.



**Figure 14. Classical symptoms of Townes-Brocks Syndrome.** A-B) TBS patientes exhibit polydactyly in hands. D) Dysplastic ears are a common TBS feature, compared to normal ear in C). E) TBS individuals might also exhibit preauricular tags. F) Major, minor and atypical features in TBS. (Powell and Michaelis 1999; J. Kohlhase et al. 1998).

TBS is caused by mutations in the gene encoding the transcriptional repressor *SALL1* and is associated with the presence of a truncated protein that localizes to the cytoplasm in contrast to the WT protein, which resides primarily in the nucleus (Sato et al. 2004). Recently, Bozal-Basterra and co-workers proposed that TBS might be considered a ciliopathy, which is a group of diseases caused by malfunction of the primary cilia. In agreement with this classification, TBS individuals-derived primary fibroblasts show a higher rate of ciliogenesis, abnormally elongated cilia, aberrant cilia disassembly, and SHH (Sonic-Hedgehog) signaling defects (Bozal-Basterra et al. 2018).

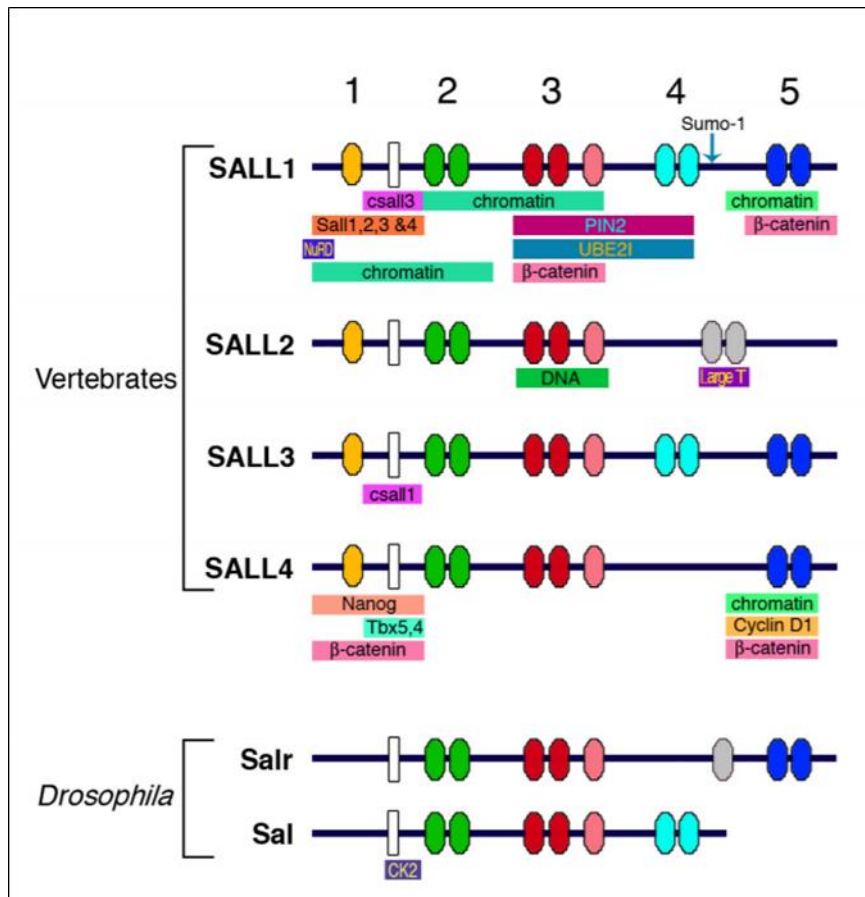
TBS is also referred as TBS1 since a TBS2 disease has been described. Unlike TBS1 patients, TB2 patients do not have mutations in the *SALL1* gene, but in the *DACT1* gene (Webb et al. 2017). Hear defects, as well as imperforate anus and renal defects are some of the symptoms that overlap with TBS1.

### 1.4.1 The SALL family of transcription factors

SALL (Spalt-like) proteins are zinc finger transcription factors characterized by the presence of stereotypical pairs of zinc finger domains along the protein, which are thought to mediate interactions with DNA via an AT-rich sequence (Netzer et al. 2006). The first zinc finger domain (ZF1) corresponds to a single zinc finger C2HC type, conserved only in vertebrates; the rest of the domains (ZF2-5) are organized in doublets or triplets of C2H2 type zinc fingers, connected by sequences conserved throughout evolution. In vertebrates, the N-terminal region of SALL proteins presents a conserved 12 amino acid sequence involved in transcriptional repression (Lauberth and Rauchman 2006). Another important N-terminal motif is the conserved polyglutamine (PolyQ) domain involved in protein dimerization with itself, with other SALL family members or with other factors (Sweetman et al. 2003). Finally, the ZF1 motif, which may mediate protein-protein interactions rather than DNA binding (Laity, Lee, and Wright 2001).

In *Drosophila*, two paralogs, Salm and Salr, are implicated in the formation of the wing via control of cell differentiation in imaginal discs, the trachea, the sensory organs and the nervous system (Cantera et al. 2002; J. F. de Celis and Barrio 2000; J. F. de Celis and Barrio 2009). In vertebrates, there are four members of the SALL family, defined as SALL1-4. SALL family is very important in different aspects of human health, since they are associated to hereditary syndromes and are involved in stem cells maintenance and cancer as tumor-suppressor factors (Misawa et al. 2018). Figure 15 shows a schematic representation of the main conserved domains present in SALL proteins.

While mutations in *SALL1* cause TBS, mutations in *SALL4* cause Okihiro Syndrome (Kohlhase et al. 2002), which is characterized by hearing dysfunction, external ear malformation, renal abnormalities, atrial septal defects and facial asymmetry. Mutations in *SALL2* have been identified as responsible for ocular coloboma in human and in mice, a congenital defect resulting from failure in the optic fissure normal closure (Kelberman et al. 2014).

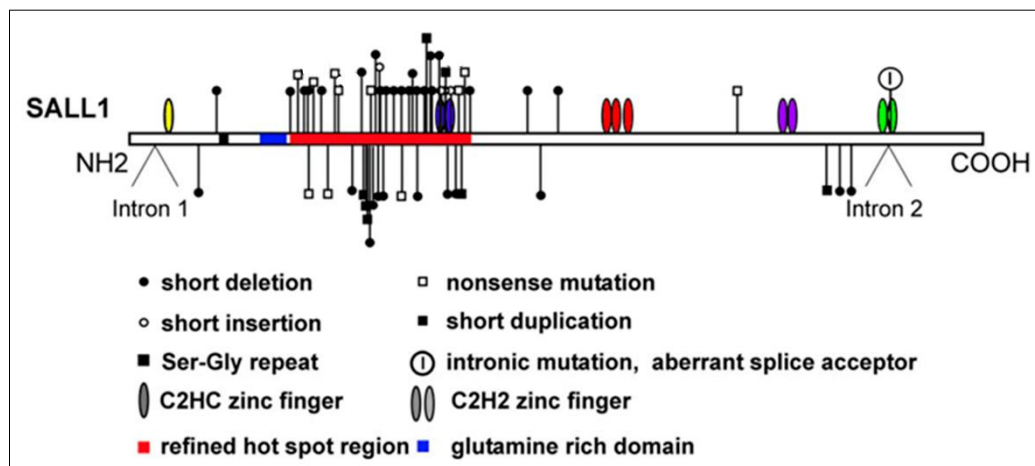


**Figure 15. Schematic representation of the main conserved domains present in SALL proteins.** Coloured ovals numbered 1 to 5 represent the zinc finger domains from vertebrate and *Drosophila* Sall homologues. White rectangles represent the polyQ regions. The arrow in SALL1 indicates the SUMOylation site described for this protein. Coloured horizontal bars below each protein indicate the SALL-interaction domains with other proteins. Adapted from (J. F. de Celis and Barrio 2009).

### 1.4.2 Truncations of SALL1 cause Townes-Brocks Syndrome

SALL1 (MIM: 602218) is one of the four members of the SALL family in vertebrates. SALL1 is a transcriptional repressor. The N-terminal region of SALL1 mediates transcriptional repression via the interaction with the nucleosome remodelling deacetylase (NuRD) complex through a conserved 12 amino acid sequence (Kiefer et al. 2002; Lauberth and Rauchman 2006). Moreover, SALL1 is linked to chromatin-mediated repression (J. F. de Celis and Barrio 2009).

More than 50 point-mutations in *SALL1* were described that cause premature stop codons by frame shifts, short insertions or deletions, mainly in a hot-spot region located between the N-terminal part of the protein and the ZF1, as shown in Figure 16 (Botzenhart et al. 2007; 2005; J. Kohlhase et al. 1998; Borozdin et al. 2006; Nakajo et al. 1990).



**Figure 16. Schematic representation of SALL1 protein.** Schematic representation of the SALL1 protein and localization of the mutations identified. Zinc fingers are indicated as coloured ovals. The red horizontal bar marks the refined hot-spot region and the blue bar assigns the glutamine rich domain. Adapted from (Botzenhart et al. 2007).

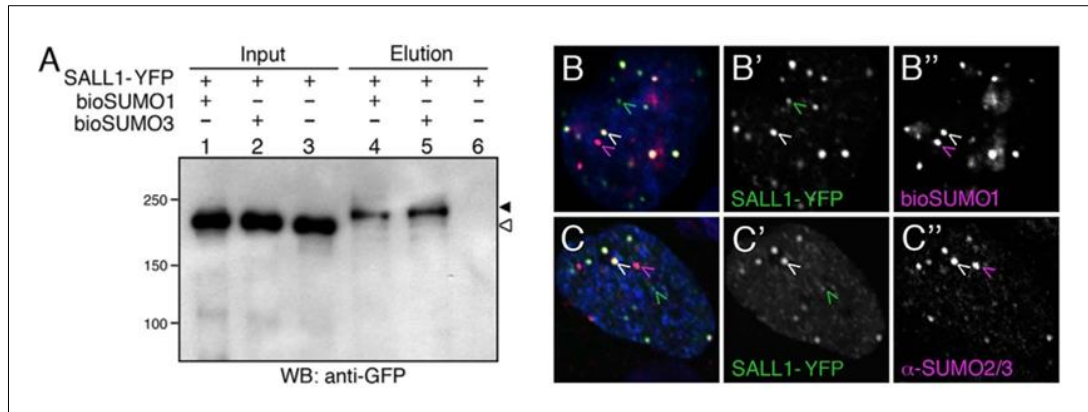
Many *SALL1* mutations reported in TBS result in truncated proteins that lack most of the zinc finger pairs, but retain the N-terminal domain, the ZF1 and the PolyQ region. In fact, truncated proteins causing TBS are still able to form multimers with themselves and with full-length SALL proteins (Lauberth et al. 2007; Kiefer et al. 2003). The c.826C>T (R267X) mutation in the exon 2 is the most common allele found in TBS patients.

### 1.4.3 SALL1 SUMOylation

Post-translation modifications can modulate transcription factors activity altering their regulation and localization. In *Drosophila*, SALL proteins can be SUMOylated. SUMOylation alters their nuclear localization and influence the vein pattern formation in the wing, as well as their transcriptional repressor activity

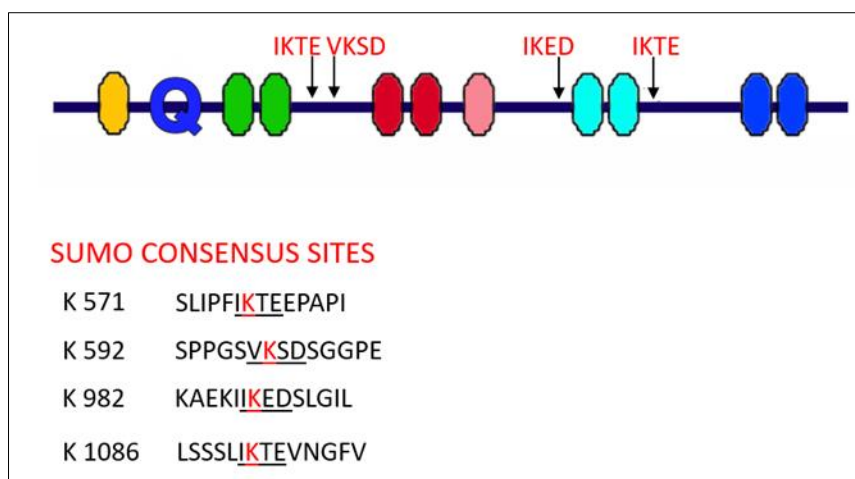


(Sánchez et al. 2010; 2011). Through human library-based yeast-two-hybrid screen, Netzer and co-worker identified UBC9 and SUMO1 as interactors of SALL1 and reported the SUMOylation of K1086 *in vitro* (Netzer et al. 2002). Pirone et al. confirmed SUMOylation of SALL1 in cells using the bioSUMO system, a comprehensive method for studying SUMOylated proteins. As shown in Figure 17, they showed that SALL1 can be modified by both SUMO1 and SUMO3 and that it partially colocalized in the nucleus with bioSUMO1, similarly than with the endogenous SUMO2/3 (Pirone et al. 2017).



**Figure 17. SALL1 is modified by SUMO1 and SUMO2/3 in mammalian cells.** A) Western blot of HEK 293FT cells showing that the transcription factor SALL1 fused to YFP was SUMOylated in presence (+) of bioSUMO1 (lane 4) or bioSUMO3 (lane 5; bioSUMO1-BirA or bioSUMO3-BirA, respectively). Black arrowhead indicates the modified SALL1-YFP in the elution panel (lanes 4, 5), which is shifted in comparison to the non-modified SALL1-YFP in the input panel (white arrowhead, lanes 1–3). Molecular weight markers are shown to the left. B-C) Partial colocalization between SALL1-YFP (B, green) and bioSUMO1 (purple) (bioSUMO1-BirA-UBC9) in U2OS cells (B) or with endogenous SUMO2/3 (C, purple). White arrowheads indicate colocalization. Nuclei were stained with DAPI (blue). B'-C') Green and purple channels are shown independently in black and white. Edited from (Pirone et al. 2017).

For a deeper analysis of SALL1 SUMOylation, our laboratory generated a SALL1 mutant SUMO unable to undergo SUMOylation (Pirone 2016). By analyzing the full-length amino acidic sequence of SALL1 by the SUMOplot program (<http://www.abgent.com/sumoplot>), 7 possible SUMO consensus sites were found with a high probability score. Four high score sequences were selected and mutated by substituting each lysine (K571, K592, K982 and K1086) with an arginine (SALL1-4KR) (Figure 18). The SALL1-4KR mutant lost the capacity to be SUMOylated in cells, therefore, SALL1-4KR was considered a SALL1 $\Delta$ SUMO mutant.

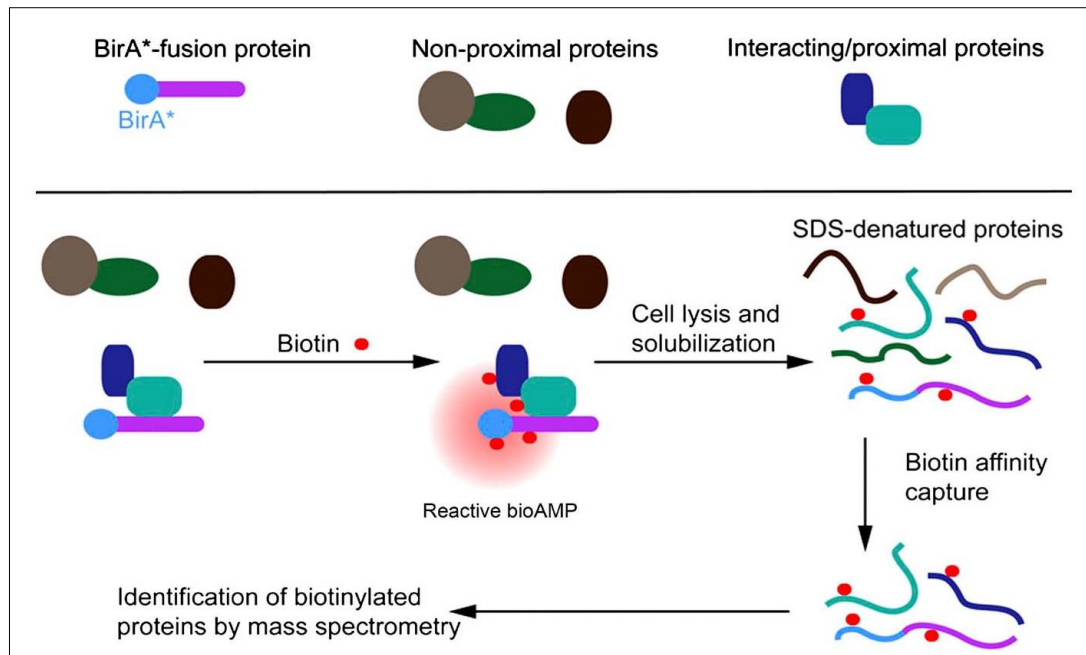


**Figure 18. Schematic representation of identified and mutated SUMO consensus sites in SALL1.** Ovals represent the zinc fingers distributed along the protein. Q represents the poly-glutamine domain. In red, SUMO consensus sites mutated in SALL1 $\Delta$ SUMO.

Interestingly, recent mass spectrometry experiments to find SUMOylated proteins in different cellular conditions, revealed that all the SALL human homologues can be SUMOylated (Hendriks and Vertegaal 2016a; 2016b). In the case of SALL1, various SUMOylation sites have been detected by mass spectrometry, corresponding to the consensus sites identified previously, the most prominent site being the K1086 previously identified (Netzer et al. 2002).

#### 1.4.4 SALL1 Interactors

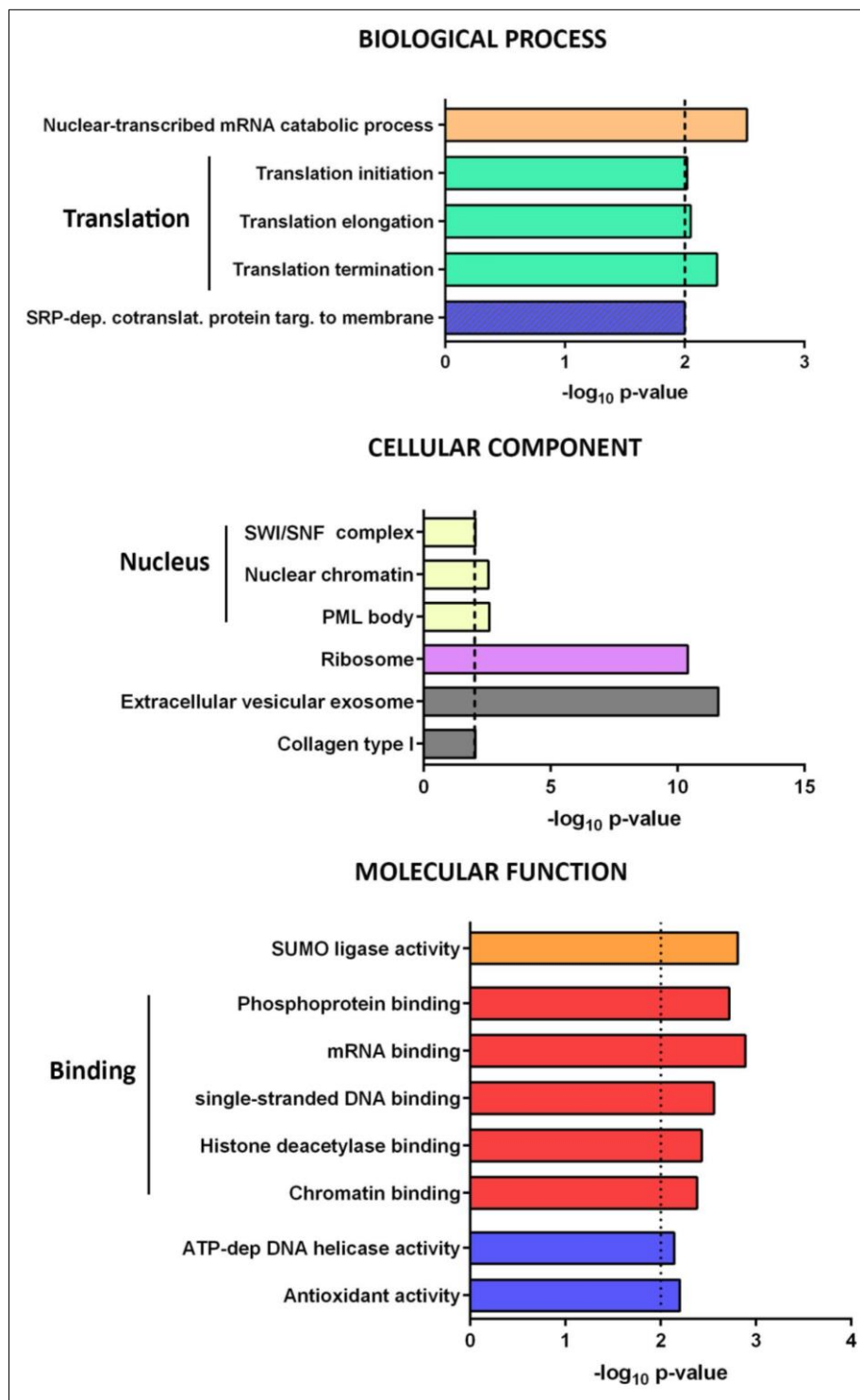
In order to gain insights on SALL1 function in the context of TBS, “proximity proteomics” BiID methodology in combination with Mass Spectrometry was used to identify possible interactors of human SALL1 (schematized in Figure 19) (Bozal-Basterra et al. 2018).



**Figure 19. Search of SALL1 interactors by proximity proteomics.** Schematic representation of the BioID methodology in combination with Mass Spectrometry (Bozal-Basterra et al. 2018). Adapted from (Roux et al. 2012).

Gene Ontology (GO) analysis revealed that many of the putative interactors of SALL1 were nuclear proteins, most of them factors involved in gene expression regulation. Interestingly, SALL1 interacted with many components of the UbL pathway, including E1, E2 and E3 enzymes. Among other categories, the SUMO ligase term was well represented in the list of interacting factors (Figure 20). The known mammalian SUMO E3 ligases CBX4, PIAS1, PIAS2 and PIAS4 were identified as putative SALL1 interactors, opening the possibility that one of those ligases was involved in the SUMOylation of SALL1. Interestingly, the genetic interaction between *Drosophila sall* genes and Polycomb (Pc), the *Drosophila* homolog of CBX4, was previously reported since mutations in *sall* enhanced the phenotype of Pc group mutations during embryogenesis (Casanova 1989; Landecker, Sinclair, and Brock 1994). However, the mechanisms underlying the interaction between CBX4 and SALL1 in human were still unknown. The opportunity to explore the relationship between human SALL1 and

CBX4 opened an opportunity of research with the aim to better understand the implications of SALL1 SUMOylation in gene regulation.

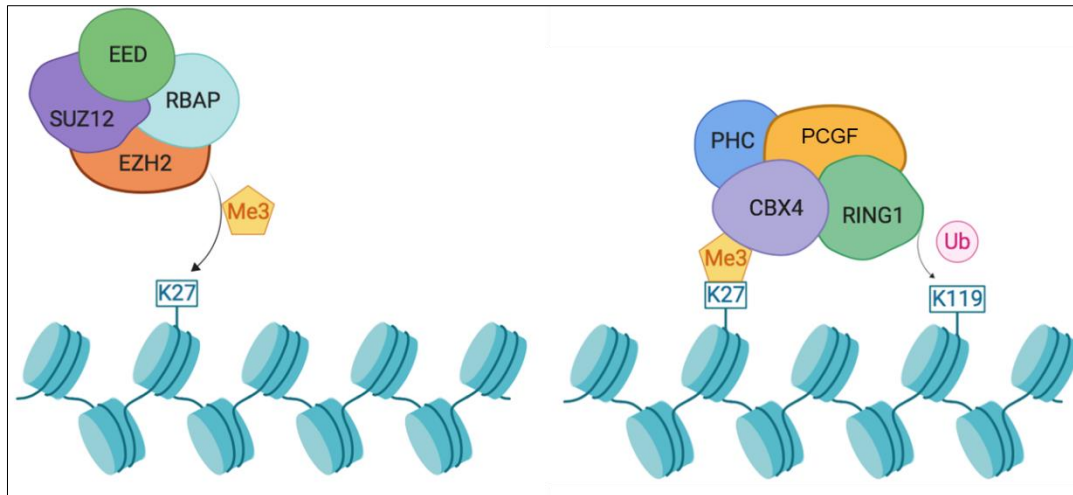


**Figure 20. Analysis of SALL1 interactors by Mass Spectrometry.** GO analysis for biological process, cellular component and molecular function performed using Innate DB. (Bozal-Basterra et al. 2018).

### 1.4.5 CBX4 is a member of the Polycomb group

The Polycomb group of proteins (PcG) is a conserved family of transcriptional repressors involved in epigenetic silencing and in the maintenance of cell identity during development. They regulate the expression of numerous genes that participate in different cellular processes including cell fate choices and cell cycle control.

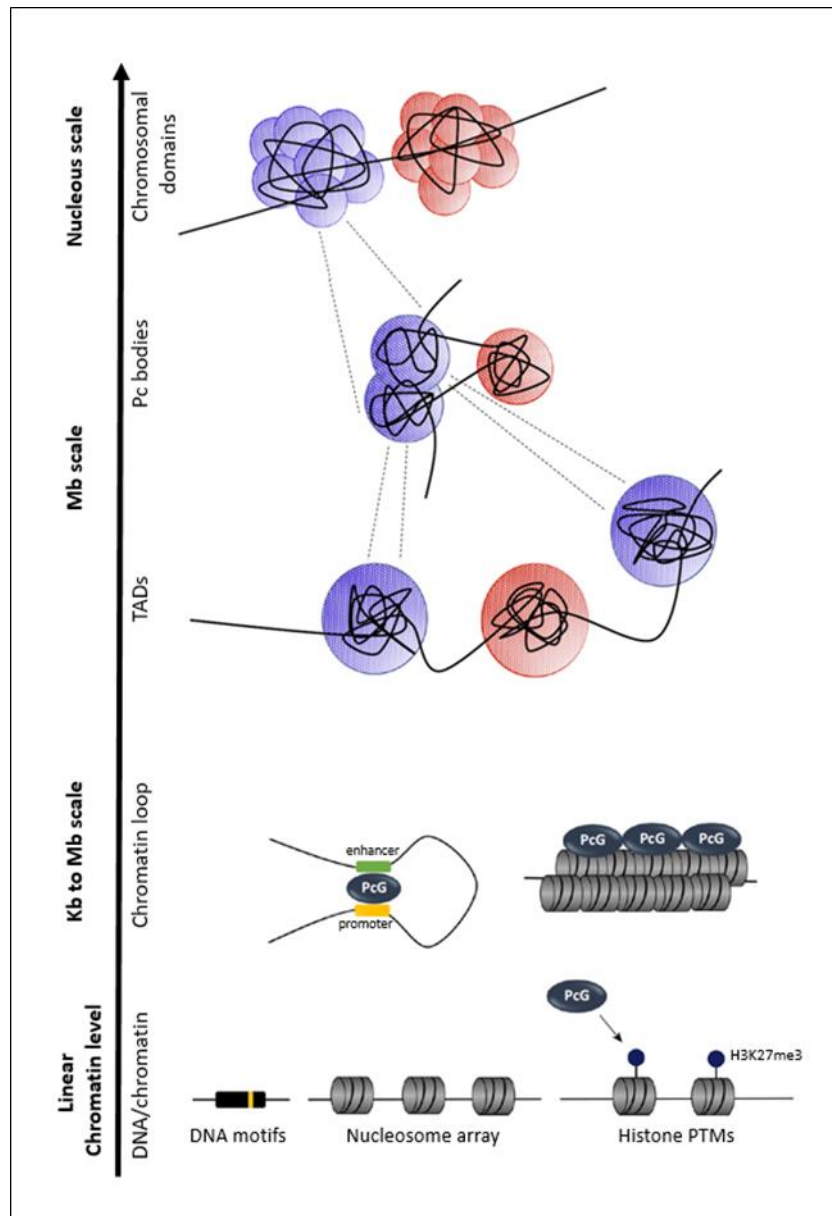
PcG proteins assemble into chromatin-associated multiprotein complexes known as Polycomb Repressive Complex 1 (PRC1) and Polycomb Repressive Complex 2 (PRC2), containing enzymatic subunits to modify histones (Bracken and Helin 2009; Simon and Kingston 2009; Entrevan, Schuettengruber, and Cavalli 2016). PRC2 complex is composed by the core components EED (Embryonic ectoderm development), SUZ12 (Suppressor of zeste 12) and EZH2 (Enhancer of zeste 2), which mediate the trimethylation of histone H3 at lysine 27 (H3K27me<sub>3</sub>), a landmark for gene transcriptional repression. The canonical PRC1 complex (cPRC1) is composed of CBX (Chromobox), PCGF (Polycomb group ring finger), RING (Ring finger protein), and PHC (Polyhomeotic-like) proteins. CBX proteins in PRC1 (CBX2,4,6,8) recognize and bind the Histone H3 trimethylated at lysine K27 (H3K27me<sub>3</sub>) facilitating the recruitment of cPRC1 to PRC2-target genes. RING1A/B proteins in the cPRC1 complex, together with PCGF2/4 proteins, can then catalyse the monoubiquitylation of histone H2A at lysine 119 (H2AK119ub<sub>1</sub>), which contribute to chromatin compaction and repression of lineage-specific genes (Figure 21) (Eskeland et al. 2010; Aranda, Mas, and Di Croce 2015).



**Figure 21. Chromatin silencing marks by PRC1 and PRC2.** PRC2, through the catalytic subunit EZH2, promotes the trimethylation (me3) of lysine 27 (K27) of the histone H3, which is then recognized by the CBX protein of the cPRC1. Next, the RING1 subunit of PRC1 catalyses the ubiquitination (Ub) of the lysine 119 (K119) of the histone H2A.

However, this hierarchical model of PRC2-dependent H2AK119 ubiquitination has been challenged in recent years by the identification of non-canonical PRC1 (nc-PRC1) complexes, which do not require PRC2 activity to mediate H2AK119ub1 and in which the CBX component is replaced by RYBP (RING1 and YY1 binding protein) or YAF2 (YY1 associated factor) (Gao et al. 2012; Morey et al. 2013; Zhao et al. 2020).

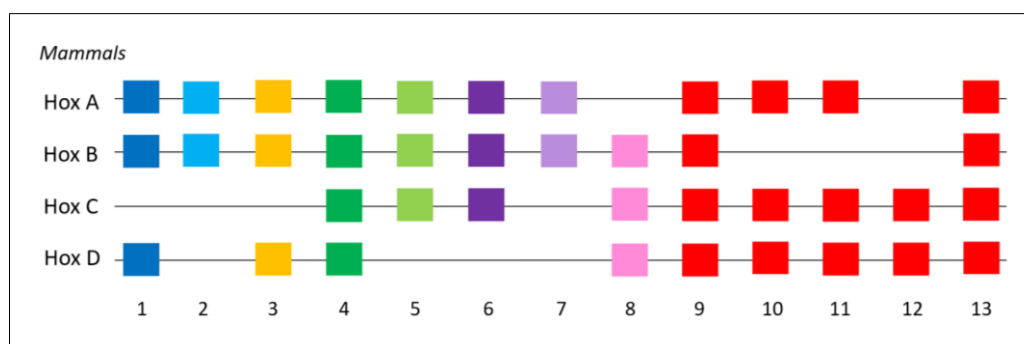
PcG proteins regulate chromatin structure at multiple levels, from the linear genome by modifying histones by binding Pc Regulatory Elements (PRE) in target genes, to organizing topologically associating domains (TADs) in the nuclear space. TADs represent linear units of chromatin that fold as 3D structures and may establish dynamic long-range interactions with each other, mostly due to the dispersed genomic distribution of the target sites, to form Pc nuclear bodies (Figure 22) (Entrevan, Schuettengruber, and Cavalli 2016).



**Figure 22. Pc bodies result from several engaged silenced-chromatin regions.** Binding of PcG proteins to chromatin can induce DNA compaction and PcG proteins are involved in mediating looping interactions between cis-regulatory elements like enhancers and promoters. Linear chromatin regions can fold into specific 3D structures to form TADs. In addition, Polycomb-repressed TADs show long-range interactions with each other mediated by PcG proteins. PcG proteins accumulate in the nucleus to form PcG foci, which are the nuclear counterparts of genomic domains, silenced by PcG proteins and may contain individual PcG TADs or multiple TADs engaged in long-range interactions. Adapted from (Entrevaan, Schuettengruber, and Cavalli 2016).



PcG genes were originally identified as factor mediating repression of homeotic genes. *Hox* (homeobox) genes are grouped in genomic clusters and encode transcription factors defining cellular identities along the major and secondary body axes. In human there are 39 *Hox* genes organized in four clusters on four different chromosomes: *HoxA*, *HoxB*, *HoxC* and *HoxD* (Figure 23). During development, collinear regulation of *Hox* genes in space and time is critical for patterning the body plans of many animals with bilateral symmetry. Dynamic patterns of histone marks and higher-order chromatin structure are important determinants of *Hox* gene regulation (Deschamps and Duboule 2017; Soshnikova 2014).



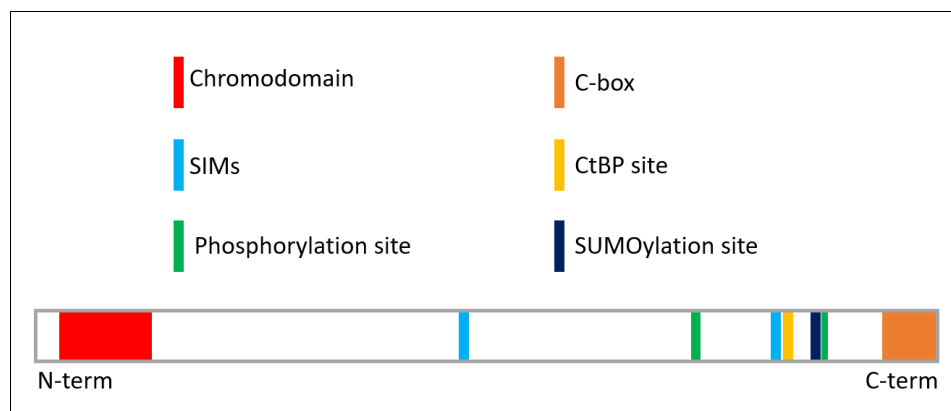
**Figure 23. Clustering of Hox genes in mammals.** In mammals there are 39 *Hox* genes organized in four clusters *HoxA*, *HoxB*, *HoxC* and *HoxD*. *Hox* genes are indicated by coloured boxes. Different colours indicate gene that share a common ancestor.

In addition to *Hox* genes, PcG complexes also occupy promoters of members of other transcription factor families, such as GATA family, that have key roles in a variety of developmental processes. Members of the *GATA4*, 5 and 6 subfamily are expressed in various mesoderm- and endoderm-derived tissues, where they play critical roles in regulating tissue-specific gene expression controlling the development and function of different cell types [(Aronson, Stapleton, and Krasinski 2014; Katsumura 2017; Boyer et al. 2006).

Different CBX proteins share the N-terminal chromodomain and a small hydrophobic region at the extreme C-terminal called C-box. However, among CBX proteins, CBX4 is the only one known to have an enzymatic activity. Kagey and

colleagues demonstrated that CBX4 is a SUMO E3 ligase for the transcriptional corepressor CtBP (carboxyl-terminus binding protein) and is itself SUMOylated. Several CBX4 SUMOylation targets have been described (Kagey, Melhuish, and Wotton 2003; Chen et al. 2018; J. Li et al. 2014; Ismail et al. 2012; B. Li et al. 2007).

Two SIMs have been identified in CBX4, which are required for the SUMO E3 ligase activity, together with the chromodomain and the C-box. Schematized representation of some important CBX4 domains is shown in Figure 24. The chromodomain of CBX4 is also necessary for the efficient binding to H3K27me<sub>3</sub>, as well as its SUMOylation. In fact, Kang and colleagues demonstrated that CBX4 is a target of the SUMO protease SENP2 and that SUMOylation is essential for CBX4-mediated PRC1 recruitment to H3K27me<sub>3</sub>. In *SENP2 null* mutant embryos, SUMOylated CBX4 accumulates and CBX4 occupancy on the promoters of PcG target genes is markedly increased with consequent increment of targets repression (Wotton and Merrill 2007; Merrill et al. 2010; Kang et al. 2010). Interestingly, a critical role for SUMOylation in the regulation of PcG target gene expression was previously reported in *Drosophila*. There, Pc SUMOylation was shown to regulate PcG-mediated silencing by modulating the kinetics of Pc binding to chromatin, as well as its ability to form Polycomb bodies (Gonzalez et al. 2014).



**Figure 24. Schematic representation of CBX4 important motifs and domains.** Some important domains are represented. Colours are used according to the legend. Edited from (Wotton and Merrill 2007).

Transcriptional repression and SUMO E3 ligase activities endows CBX4 with a key role in several essential pathways. CBX4 has been described to facilitate the differentiation of hematopoietic stem cells (Klauke et al. 2013), counteracting cellular senescence (Ren et al. 2019) and maintaining the epithelial lineage identity via repression of non-epidermal lineage and cell cycle inhibitor genes (Mardaryev et al. 2016). Moreover, CBX4 is known to be recruited rapidly to DNA damage (Ismail et al. 2012) and has emerged as a critical component of the DNA end resection machinery (Soria-Bretones et al. 2017). Finally, CBX4 plays an important role in the occurrence and development of tumours. In fact, dysregulation of CBX4 contributes to the progression of human cancers in which it can acts as both oncogene and tumour suppressor, depending on the cellular context (Wang et al. 2016).

Given the importance of SALL1 and CBX4 in gene regulation during development and their possible interaction in the cells, to understand the relationship between these two transcription factors might help to clarify their role in cellular functioning and regulation.



# **2. MATERIALS AND METHODS**



## Cell culture

Mouse NIH3T3-Shh-LIGHT2-Cas9-blast (Taipale et al. 2000), human U2OS (ATCC HTB-96) and HEK 293FT (Invitrogen) cells, as well as derived cell lines, were cultured at 37°C with 5% of CO<sub>2</sub> in DMEM (Dulbecco's modified Eagle's medium; Gibco) supplemented with 10% FBS and 1% penicillin/streptomycin (GIBCO).

## Cell Transfection

HEK 293FT cells were transiently transfected using calcium phosphate in 10 cm dishes with 10 µg of DNA using different sets of plasmids according to each experiment.

U2OS cells were transiently transfected using Effectene® (QIAGEN) according to the manufacturers datasheet.

## CRISPR/Cas9 genome editing

The mouse *Uros* locus was targeted in NIH3T3-based Shh-LIGHT2 fibroblasts [(Taipale et al. 2000); kind gift from A. McGee, Imperial College]. Cas9 was introduced into Shh-LIGHT2 cells by lentiviral transduction (Lenti-Cas9-blast; Addgene #52962; kind gift from F. Zhang, MIT) and selected with blasticidin (5 µg/ml). Two high-scoring sgRNAs were selected (<http://crispor.tefor.net/>) to target near the second exon of *Uros* gene (Table 1). To express both sgRNAs and additional Cas9 with puromycin selection, these sequences were cloned into px459 2.0 (Addgene #62988; kind gift of F. Zhang, MIT).

Transfections were performed in NIH3T3-based Shh-LIGHT2\_Cas9-blast cells with Lipofectamine 3000 (Thermo). 2 µg of *CondMmUROSv6\_puro\_Cre-ERT2* plasmid were transfected into the cells together with 0.67 µg of each sgRNA. Cells were treated with 10 µM of RS1 (Sigma) 16 hours after the transfection. 24 hours after transfection, transient puromycin selection (1 µg/ml) was applied for 48 hours to enrich for transfected cells. Cells were plated at clonal density, and well-isolated clones were picked and propagated individually. Western blotting was used to identify clones expressing human UROS.

**Table 1. Relevant sequences.**

Name	Sequence
T2A	(GSG) E G R G S L L T C G D V E E N P G P
P2A	(GSG) A T N F S L L K Q A G D V E E N P G P
hUROS <sup>opt</sup> (*)	GAGTTGGGGCTTTATGGGTTGGAGGCTACTCTGATCCCCGTGTTATCATTCGAATTCTTGAGC CTTCCATCATTCTCCGAAAAGCTGAGTCACCCCGAGGACTACGGGGGCCTCATCTTCACATCC CCACGGGCCGTGGAGGCCGAGAAGTGTGTCTGGAGCAGAACAACAAGACAGAAGTGTGG GAGCGCAGCCTGAAGGAGAAATGGAACGAAAATCAGTTTACGTGGTCCGGGAACGCCACCG CCAGTTTAGTGAGCAAAATCGGCTTAGACACCGAGGGCGAGACCTGTGGAAACGCAGAGAA GCTGGCTGAATACATCTGCTCCCGGAATCCTCCGCCCTTCTCTCTCTTTCCCTGTGGCAAC TTGAAGAGAGAGATCCTcCCAAAAGCTCTGAAGGACAAGGGCATTGCCATGGAAAGTATCAC TGTGTACCAAACCGTGGCCACCCCGGCATCCAAGGCAACCTCAACTCATATTACAGCCAGCA GGGCGTGCCTGCCTCTATCACCTTTTTCTCTCCCTCCGGCCTGACTTATAGCCTCAAACACATT CAGGAGCTTAGTGGGACAACATTGACCAGATCAAGTTTGCCGCAATCGGACCCACAACCTGC ACGGGCCTTGGCTGCCAAGGACTGCCTGTCTCATGCACAGCCGAGAGCCCCACTCCTCAAG CTTTGGCAACTGGGATCAGAAAAGCCTTGCAGCCCCACGGCTGCTGC
sgRNA1	CTTACTAAAAGACGCCA_AGG
sgRNA2	CACAATCGCAGCTCCTG_CAA

(\*) hUROS<sup>opt</sup>: cDNA sequence (exons 3-10) of human UROS optimized for mouse codon usage.

## Cycloheximide assay

3 x 10<sup>5</sup> HEK 293FT cells per well were plated in 6-well plates. Four hours later, cells were transfected with 2 µg of *CMV-SALL1-YFP*, *CMV-SALL1ΔSUMO-YFP* or *CMV-GFP-β-Galactosidase* plasmid per well using the calcium phosphate method. Cells were treated with 50 µg of cycloheximide (CHX) in combination or not with 10 µM of MG132 at different time points (0h, 4h, 8h or 16h). Cells were lysed in RIPA buffer [150 mM NaCl, 1.0% NP-40, 0.5% sodium deoxycholate, 0.1% SDS, 50 mM Tris, pH 8.0, protease inhibitors (Roche)] and analyzed by Western blot.

## Fluorescence-activated cell sorting

Live cells were collected in 1x PBS (phosphate buffered saline) and flow cytometry analyses were performed using the BD FACSCanto™ II system (BD Biosciences). Data were analyzed using the FlowJo software. Porphyrin fluorescence was detected by APC laser (Ex-Max 650 nm/Em-Max 660 nm).



## Generation of vectors

The following vectors were used in this study (Table 2). All constructs for this work were done by Dr. J. D. Sutherland and the candidate, unless otherwise specified. DNA fragments were amplified from the indicated plasmids by high-fidelity PCR Platinum SuperFi (Thermo) and, when required, digested using the indicated restriction enzymes (Fermentas; NEB). PCR products were purified using mi-Gel Extraction kit (Metabion) and afterwards ligation was performed using Gibson Assembly Master Mix (BioLabs). All vectors were checked by sequencing.

**Table 2. Vectors used in the study**

Name of the vectors	Reference	Cloning sites/notes
pCAG-ERT2CreERT2	Addgene #13777	Mlu I - Spe I
17ABAVMP_CondMmUROSv6_pMA	This work/GeneArt	Mlu I - Spe I
A (CondMmUROSv6_ERT2-Cre-ERT2)	This work	Mlu I - Spe I
B (CondMmUROSv6_Cre-ERT2)	This work	Mlu I - Spe I ; BsiW I - Asc I
C (CondMmUROSv6_puro_Cre-ERT2)	This work	Mlu I - Spe I
CMV-CBX4-YFP	J.D.Sutherland, unpublished	-
CMV-SALL1-YFP	(Pirone et al. 2017)	-
CMV-SALL1 $\Delta$ SUMO-YFP	J.D.Sutherland, unpublished	-
CMV-SALL1 $\Delta$ SIM-YFP	J.D.Sutherland, unpublished	-
CMV-YFP	(Pirone et al. 2017)	-
CMV-SALL1-2xHA	J.D.Sutherland, unpublished	-
CMV-SALL1-826-2xHA	J.D.Sutherland, unpublished	-
CB6-HA	(Pirone et al. 2017)	-
CB6-SALL1-HA	J.D.Sutherland, unpublished	-
CB6-SALL1 $\Delta$ SUMO-HA	J.D.Sutherland, unpublished	-
CB6-CBX4-HA	J.D.Sutherland, unpublished	-
CMV-GFP- $\beta$ -Galactosidase	J.D.Sutherland, unpublished	-
CMV-BirA-2A-puro	J.D.Sutherland, unpublished	-
CMV-BirA-2A-bioUb-puro	J.D.Sutherland, unpublished	-
px459 2.0	Addgene #62988	-
Lenti-Cas9-blast vector	Addgene #52962	-
pTRIPZ	Horizon Discovery	-

## Genotyping and PCR primers

The following primers were used in this study (Table 3). Primers were used for either Sanger sequencing or PCR amplification of the vectors.

**Table 3. Genotyping and PCR oligonucleotide sequences.**

Oligo Name	Oligo Sequence
ZfTia1l.sgRNA.for	5'-caccGGTATGTCGGGAACCTCTCC-3'
ZfTia1l.sgRNA.rev	5'-aacGGAGAGGTTCCCGACATACC-3'
pX330.seq.for	5'-CAGTTTTAAAATTATGTTTTAAAATGGAC-3'
MmUROS.sg1.for	5'-CACCGCTTACTAAAAGACGCCAAGG-3'
MmUROS.sg1.rev	5'-AAACCCTTGGCGTCTTTTAGTAAGC-3'
MmUROS.sg2.for	5'-CACCGTTGCAGGAGCTGCGATTGTG-3'
MmUROS.sg2.rev	5'-AAACCACAATCGCAGCTCCTGCAAC-3'
T7.MmUROS.sg1.for	5'-TAATACGACTCACTATAGGGCTTACTAAAAGACGCCAAGG-3'
T7.MmUROS.sg2.for	5'-TAATACGACTCACTATAGGGATAGGCCCTGTGCACTTCTC-3'
T7.sgRNA.rev	5'-AAAAGCACCGACTCGGTGCC-3'
R1.HsUROS.for	5'-GATCGAATTCGCCACCATGAAGGTTCTTTACTGAAGGATGCG AAGG-3'
Sal1.HsUROS.rev	5'-GATCGTCGACCCGCAGCAGCCATGGGGCTGGAG-3'
R1.MmUROS.for	5'-GATCGAATTCGCCACCATGAAGGTTCTTACTAAAAGACGCC AAGG-3'
Sal1.MmUROS.rev	5'-GATCGTCGACCCGCAACAGTGGTTTGGCTTCAGCAC-3'
HsUROS.C73R.for	5'-GAGCAGTGAAGCAGCAGAGTTACGTTTGGAGCAAAACAAT AAAAGTGAAGTCTGG-3'
HsUROS.C73R.rev	5'-CCAGACTTCAGTTTTATTGTTTGGCTCCAAACGTAACCTCTGCTG CTTCCACTGCTC-3'
MmUROS.KO.genotype.for	5'-TCTTGCCTCTCTCCTTCAA-3'
MmUROS.KO.genotype.rev	5'-CAGCACAGGAATCAGTGTGG-3'
UROS.ERT2.for	5'-GATGTCGAAGAGAATCCTGGACCGACGCGTGCCACCATGGCT GGAGACATG-3'
RGlobpA.UROS.rev	5'-CATTATACGAAGTTATGATCGATCGATACTAGTCTGCAGGTCG AGGGATCTTCATAAGAGAAG-3'
UROS.CRE.for	5'-GATGTCGAAGAGAATCCTGGACCGACGCGTATGAATTTACTG ACCGTACACCAAAATTTGCCTGCA-3'
T2A.CRE.for	5'-TGCGGAGATGTCGAAGAGAACCCTGGCCCTATGAATTTACTG ACCGTACACCAAAATTTGCCT-3'

CRE.T2A.rev	5'-AGGCAAATTTTGGTGTACGGTCAGTAAATTCATAGGGCCAGG GTTCTCTTCGACATCTCCGCA-3'
UROS.puro.for	5'-GATGTCGAAGAGAATCCTGGACCGACGCGTATGACCGAGTAC AAGCCACGGTG-3'
UROS.P2A.seq.for	5'-CGGCTGCTGCGGATCCGGCGCAACAA-3'
Lox511.UROS.seq.rev	5'-GGGGATGCGGTGGGCTCTATGGCGATC-3'
CRE50.seq.rev	5'-CATCAGTTTCTTGCGAACCTCATCACTC-3'
CRE700.seq.for	5'-CCGTCTCTGGTGTAGCTGATGATCCG-3'
ERT430.seq.for	5'-GGTGGAGATCTTCGACATGCTGCTGG-3'
RbglobinA.seq.for	5'-CCTCTCCTGACTACTCCAGTCATAGCTG-3'
UROS.3'HA.seq.rev	5'-CTTTGTAGACCAGGCTGGCCTCGAAC-3'
HSS1.2xHA.5'gen.for	5'-GATCGGTATGTCGGGAACCTCTCCAGGAGGGCATGCTAATTT GGGTCGGC-3'
HSS1.2xHA.5'gen.rev	5'-TGCATAGTCTGGTACGTCGTAAGGATAAGATCCTGCG TAATCAGGGACATCATATGGGTACATGCTGGCTCAAACA TCAGCTGGGG-3'
HSS1.2xHA.3'gen.for	5'-TCTTATCCTTACGACGTACCAGACTATGCAGGCAGCATGTCCG GGAGGAAGCAAGCGAAGCCTCAACATTTTCAATCCGACCCCGAAGTGGCC- 3'
HSS1.2xHA.3'gen.rev	5'-CAGCCCAGCGCCCGAAGTAACTCCT-3'
Cas9.seq.for	5'-TGGACGCCACCCTGATCCACCAGAGCA-3'
MmUROS.5'exterior.for	5'-GGCGGATTTCTGAGTTCAAG-3'
MmUROS.5'exteriorv2.for	5'-GCCAGCCTGGTCTACAGAGT-3'
MmUROS.3'exterior.rev	5'-TGAGGGACTAAGGGTGATGC-3'
MmUROS.3'exteriorv2.rev	5'-AACCTGTGGCTCAGAACACC-3'
HsUROS.degen.rev	5'-AACCCATAAAGCCCCAACTC-3'
HsUROS.degenv2.rev	5'-TTTCGGAGAATGATGGAAGG-3'
UROS.postrecomb.rev	5'-AGAGATGGCTCAGCGGATAA-3'
UROS_HAtag.rev	5'-AGCGTAATCTGGAACATCGTATGGGTAG-3'
UROSintron.5'loxP.for	5'-GGAAGAGCCAGTCTGTGCTCTTAA-3'
huUROS_post-flip 1.rev	5'-GCAAGCTTTTCTGCATTTCC-3'
huUROS_post-flip 2.for	5'-TGCCATGGAAAGCATAACTG-3'
Lox511_3'exterior_post-flip.rev	5'-GGCGTGCGATAACTTCGTAT-3'
CRE_start.3796.rev	5'-TCCATCGCTCGACCAGTTTA-3'
BGH_start.5897.for	5'-CCAGGGTCAAGGAAGGCA-3'
ERT2_start.4566.rev	5'-TTTTCCCTGGTTCCTGTCCA-3'
Intron_start125.for	5'-TTATTCCGTGACCCCTTGCT-3'

UROS_start.2068.rev	5'-CCTTCAGAGCTTTTGGGAGG-3'
UROS_start.1856_post-flip.for	5'-ACTGAAGTCTGGGAAAGGTCT-3'
HsUROS.geno5.rev	5'-CCTTCGCATCCTTCAGTAAAGAACCTTCAT-3'
HsUROS.geno5v2.rev	5'-CCTGATATACGGATCCTGGCCACAGTCATC-3'
MmUROS.geno5.for	5'-GCCTTTAATCCCAGCACTTG-3'
MmUROS.geno5v2.for	5'-TGCTCATAACACTTTTAGATAATAAAGAA-3'
BshTI_HuUROS.ATG.for	5'-ACGGGTTTGCCGCCAGAACACAGGACCGGTGCCACCATG AAGGTTCTTTTACTGAAGGATGCGAAG-3'
Clover3.T2A.puro.rev	5'-ACTTCTCTGCCCTCTCCGCTTCCGGATCCCTTGACAGCTC GTCCATGCCATG-3'
Hu.UROS.mAID.for	5'-CTCCAGCCCCATGGCTGCTGCGGGCGCGCCGGAAGCAAGG AGAAGAGTGCTTGTCTAAAGATC-3'
Hu.UROS.mAID.rev	5'-GATCTTTAGGACAAGCACTTCTCCTTGCTTCCGGCGCGCC CGCAGCAGCCATGGGGCTGGAG-3'

## GFP-Trap Pulldown

All steps were performed at 4°C. HEK 293FT cells were transfected using the calcium phosphate method with different combinations of plasmids and collected after 48 hours. Cells were washed 3 times with 1x PBS and lysed in 1 ml of lysis buffer [25 mM Tris-HCl pH 7.5, 150 mM NaCl, 1 mM EDTA, 1% NP-40, 0.5% Triton X-100, 5% glycerol, protease inhibitors (Roche)]. Lysates were sonicated and spun down at 25000 x g for 20 minutes. After saving 40 µl of supernatant (input), the rest of the lysate was incubated overnight with 30 µl of pre-washed GFP-Trap resin (Chromotek) in a rotating wheel. Beads were washed 5 times for 5 minutes each with washing buffer (WB) (25 mM Tris-HCl pH 7.5, 300 mM NaCl, 1 mM EDTA, 1% NP-40, 0.5% Triton X-100, 5% glycerol). Beads were centrifuged at 2000 x g for 2 minutes after each wash. For elution, samples were boiled for 5 minutes at 95°C in 2x Laemmli buffer.

Transfected plasmids: *CMV-CBX4-YFP*, *CMV-SALL1-YFP*, *CMV-SALL1ΔSUMO-YFP*, *CMV-SALL1ΔSIM-YFP*, *CMV-YFP*, *CMV-SALL1-2xHA*, *CMV-SALL1-826-2xHA*, *CB6-HA*, *CB6-SALL1-HA*, *CB6-SALL1ΔSUMO-HA*, *CB6-CBX4-HA*. 5 µg of each plasmid were used for a 100 mm dish transfection.

## **GST-SUMO proteins purification**

GST fusion proteins were produced from 200 ml of *E. coli* C41DE3PLYS cells grown in a rich medium. A standard induction protocol was used, shifting log-phase cultures (A600= about 0.6) from 37°C RT (room temperature); saved non-induced cells and added IPTG to a final concentration of 1 mM overnight at 20°C with constant vigorous shaking. Bacteria were recovered and centrifuged at top speed at 4°C for 15 minutes. Cells were lysed in lysis buffer [1mM PFMS in EtOH, lysozyme 1mg/ml, 1x protease inhibitors (Roche) in 1x PBS], sonicated and centrifuged at 4°C 20000 x g, 20 minutes. Supernatants were collected and passed through GST columns. GST-fused proteins were purified by dialysis. The amount of purified proteins was quantified by BCA kit (Thermo Fisher Scientific), checked by PAGE, and stained by Coomassie.

## **GST-Pulldown**

All steps were performed at 4°C. 100 µl of GST-fused SUMO proteins were incubated at 4°C overnight with 300 µl of glutathione-sepharose beads pre-equilibrated in lysis buffer [25 mM Tris-HCl pH 7.5, 150 mM NaCl, 1 mM EDTA, 1% NP-40, 0.5% Triton X-100, 5% glycerol, protease inhibitors (Roche)]. Next, beads were centrifuged 2 minutes at 2000 x g and washed four times with WB (25 mM Tris-HCl pH 7.5, 300 mM NaCl, 1 mM EDTA, 1% NP-40, 0.5% Triton X-100, 5% glycerol).

HEK 293FT cells were transfected using the calcium phosphate method with different combination of plasmids and collected after 48 hours. Cells were washed 3 times with 1x PBS and lysed in 1 ml of lysis buffer. Lysates were sonicated and spun down at 25000 x g for 20 minutes. After saving 40 µl of supernatant (input), the rest of the lysate was incubated overnight with 30 µl of GST-SUMO beads in a rotating wheel. Beads were washed 5 times for 5 minutes each with WB (25 mM Tris-HCl pH 7.5, 300 mM NaCl, 1 mM EDTA, 1% NP-40, 0.5% Triton X-100, 5% glycerol). Beads were centrifuged at 2000 x g for 2 minutes after each wash. For elution, samples were boiled for 5 minutes at 95°C in 2x Laemmli buffer.

## **High-performance liquid chromatograph analysis**

A high-performance liquid chromatograph (Shimadzu) with an autosampler was used to separate and quantify porphyrins. The analytical column was BDS Hypersil

C18 (250 x 3 mm; 5  $\mu\text{m}$  particle size) purchased from Thermo Scientific. The method to separate the porphyrins was facilitated by 60 minutes gradient elution and a two-component mobile phase consisting of ammonium acetate (1 M, pH 5.16) filtered through 0.1  $\mu\text{m}$  as solvent A and 100% acetonitrile as solvent B. Gradient elution commenced upon injection at 0% B, increased to 65% B for 30 minutes, remained for 5 minutes, returned to 0% B in 15 minutes and remained for 10 minutes in order to re-equilibrate the column at 0% B before the next injection. The flow rate was 1 ml  $\text{min}^{-1}$  and the sample injection volume was 100  $\mu\text{l}$ . All analyses were performed at 20°C and the spectra were taken at an excitation wavelength of 405 nm and an emission wavelength of 610 nm. The concentrations of porphyrins were calculated with five-point calibration curve ranging from 0.0 to 100 pM. This protocol was performed in collaboration with Dr. Pedro Urquiza and Itxaso San Juan Quintana from the Precision Medicine and Metabolism Lab at CIC bioGUNE.

### **Immunoprecipitation**

All steps were performed at 4°C. Cells were collected 24 hours after seeding, washed 3 times with 1x PBS and lysed in 1 ml of lysis buffer [20 mM Tris-HCl pH 7.5, 137 mM NaCl, 2 mM EDTA, 1% NP-40, 5% glycerol, protease inhibitor mixture (Roche)]. Lysates were incubated overnight with 1  $\mu\text{g}$  of anti-Cre antibody (Cell Signalling) and for additional 4 hours with 40  $\mu\text{l}$  of pre-washed Protein G Sepharose 4 Fast Flow beads (GE Healthcare) in a rotating wheel. Beads were washed 5 times for 5 minutes each with WB (10 mM Tris-HCl pH 7.5, 137 mM NaCl, 1 mM EDTA, 1% Triton X-100). Beads were centrifuged at 2000 x g for 2 minutes after each wash. For elution, samples were boiled for 5 minutes at 95°C in 2x Laemmli buffer.

### **Immunostaining**

HEK 293FT\_TripZ-SALL1-2xHA and U2OS cells were seeded on 11 mm coverslips (50000 cells per well in a 24 well-plate). After washing 3 times with cold 1x PBS, cells were fixed with 4% PFA (paraformaldehyde) supplemented with 0.1% Triton X-100 in PBS for 20 minutes at RT. Then, coverslips were washed 3 times with 1x PBS. Blocking was performed for 20 minutes at RT in blocking buffer (1% BSA, 0.3% Triton X-100, 25 mM NaCl in 1x PBS). Primary antibodies were incubated 1 hour at 37°C and

cells were washed with 1x PBS 3 times. To label Pc bodies we used the primary human antibody against CBX4 (Proteintech, 1:100). Donkey anti-rabbit secondary antibodies (Jackson ImmunoResearch) conjugated to Alexa 568 (1:200) was incubated for 1 hour at 37°C. Endogenous SALL1 or SALL1-2xHA in HEK 293FT\_TripZ-SALL1-2xHA cells were recognized by a primary antibody against SALL1 (R&D, 1:100) and a donkey anti-mouse secondary antibodies (Jackson ImmunoResearch) conjugated to Alexa 488 (1:200). Nuclei were stained with DAPI (10 minutes, 300 ng/ml in 1x PBS; Sigma). Fluorescence imaging was performed using a confocal microscope (Leica SP8) with 63x objective, zoom 4x. Pc bodies were analyzed using the Fiji software.

### **Lentiviral transduction**

Lentiviral expression constructs were packaged using psPAX2 and pVSVG (Addgene) in HEK 293FT cells, and lentiviral supernatants were used to transduce HEK 293FT cells to generate the TripZ-SALL1-2xHA-puro line (using the TRIPZ Inducible Lentiviral shRNA [Horizon Discovery]) and GFS-SALL1-blast line (using the Lenti-Cas9-blast vector [Addgene #52962]). Stable-expressing populations were selected using puromycin (1 µg/ml) or blasticidin (5 µg/ml). Lentiviral supernatants were concentrated 100 fold before use (Lenti-X concentrator, Clontech).

### **Porphyryns extraction**

Cells were lysed in 300 µl of 6 M HCl, sonicated 3 cycles of 25 seconds each and incubated at 37°C for 30 minutes in a thermoblock. The samples were then centrifuged for 10 minutes at 10000 g. Afterwards, pellets were removed and the supernatants were transferred to a fresh filter cellulose acetate membrane centrifuge tube, pore size 0.22 µm (Corning® Costar® Spin-X®) and centrifuged for 10 minutes at 4000 g. Samples were then analyzed by HPLC. The porphyryns standards were obtained from Frontier Scientific Europe (Carnforth, UK), including a chromatographic market kit containing the number I isomers of 8, 7, 6, 5, 4 carboxylate porphyryns and mesoporphyrin IX. All porphyrin standards were reconstituted in 3 M HCl.

## Proliferation assay

12 × 10<sup>3</sup> parental NIH3T3-Shh-LIGHT2-Cas9-blast or CRISPR\_UROS cells were plated in triplicate in 12-well dishes. 24 hours later, the cells were considered day 0 and were fixed in formalin 10% for 15 minutes. The same procedure was performed after 2 and 4 days. Cell proliferation was measured by staining with crystal violet (0.1% in 20% methanol) for 45 minutes at RT. After washing 3 times with water, all samples were air dried. The precipitate was solubilized in 10% acetic acid for 20 minutes at RT and the absorbance was measured at 595 nm. For each time point, 3 biological replicates were measured.

## Quantitative reverse transcriptional PCR analysis

HEK 293FT cells transfected with 5 µg of *CMV-SALL1-YFP*, *CMV-SALL1ΔSUMO-YFP* or *CMV-GFP-β-Galactosidase* plasmids, or HEK 293-TripZ-SALL1-2xHA\_puro cells induced with different concentrations of doxycycline, were used for qPCR analysis. 48 hours after transfection or 72 hours after induction, total RNA was obtained by using EZNA Total RNA Kit (Omega) and quantified by Nanodrop spectrophotometer. cDNAs were prepared using the SuperScript III First-Strand Synthesis System (Invitrogen) in 20 µl volume per reaction. qPCR was done using PerfeCTa SYBR Green SuperMix Low Rox (Quantabio). Reactions were performed in 20 µl, adding 5 µl of cDNA and 0.5 µl of each primer (10 µM), in a CFX96 thermocycler (BioRad) using the following protocol: 95°C for 5 minutes and 40 cycles of 95°C for 15 seconds, 56°C or 62°C for 30 seconds and 72°C 20 seconds. Melting curve analysis was performed for each pair of primers between 65°C and 95°C, with 0.5°C temperature increments every 5 seconds. Relative gene expression data were analyzed using the  $\Delta\Delta C_t$  method. Reactions were done in duplicates and results were derived from at least three independent experiments normalized to GAPDH and presented as relative expression levels. Primer sequences are reported in Table 4.

**Table 4. Oligonucleotide sequences used for qPCR.**

Name	Sequence
hHoxa11_for	5'-AACGGGAGTTCTTCTCAGCGTCT-3'



hHoxa11_rev	5'-ACTTGACGATCAGTGAGGTTGAGC-3'
hHoxb4_for	5'-AGGTCTTGGAGCTGGAGAAGGAAT-3'
hHoxb4_rev	5'-GGTGTGGGCAACTTGTGGTCTTT-3'
hHoxb7_for	5'-AGACCCTGGAGCTGGAGAAAGAAT-3'
hHoxb7_rev	5'-ATGCGCCGTTCTGAAACCAAATC-3'
hHoxb13_for	5'-TACGCTGATGCCTGCTGTCAACTA-3'
hHoxb13_rev	5'-AGTACCCGCCTCCAAAGTAACCAT-3'
hHoxc6_for	5'-AGGACCAGAAAGCCAGTATCCAGA-3'
hHoxc6_rev	5'-ATTCCTTCTCCAGTTCAGGGTCT-3'
hHoxc10_for	5'-TGAAATCAAGACGGAGCAGAGCCT-3'
hHoxc10_rev	5'-TTGCTGTCAGCCAATTTCTGTGG-3'
hHoxc12_for	5'-AGGGAActCTCAGACCGCTTGAAT-3'
hHoxc12_rev	5'-AGAGCTTGCTCCCTCAACAGAAGT-3'
hHoxd13_for	5'-ATGTGGCTCTAAATCAGCCGGACA-3'
hHoxd13_rev	5'-AGATAGGTTCTAGCAGCCGAGAT-3'
hGata4_for	5'-TCTCAGAAGGCAGAGAGTGTGTCA-3'
hGata4_rev	5'-GGTTGATGCCGTTTCATCTTGTGGT-3'
hGAPDH_for	5'-CATGTTTCGTCATGGGTGTGAACCA-3'
hGAPDH_rev	5'-AGTGATGGCATGGACTGTGGTCAT-3'
hSALL1_for	5'-GCTTGCACTATTTGTGGAAGAGC-3'
hSALL1_rev	5'-GAACTTGACGGGATTGCCTCCT-3'
hCBX4_for	5'-CATCGAGAAGAAGCGGATCCGCAAG-3'
hCBX4_rev	5'-CTGTTCTGGAAGGCGATCAGCAGCC-3'

## Ubiquitination assay

3 x 10<sup>6</sup> HEK 293FT cells were plated per condition to be analyzed. Four hours later, cells were transfected with specific plasmid combinations using the calcium phosphate method. 5 µg of each plasmid were used for a 100 mm dish transfection. Transfected plasmids: *CMV-SALL1-YFP*, *CMV-SALL1ΔSUMO-YFP*, *CMV-GFP-β-Galactosidase*, *CMV-BirA-2A-puro*, *CMV-BirA-2A-bioUb-puro*, *CB6-CBX4-HA*. 16 hours after transfection, medium was supplemented with biotin at 50 µM, and 24 hours after transfection some plates were treated with 10 µM of MG132 overnight. Cells were collected 48 hours after seeded, washed 3 times with 1x PBS and scraped in lysis buffer [8 M urea, 1% SDS, 1x protease inhibitor cocktail (Roche), 1x PBS; 0.5 ml per

100 mm dish]. Samples were sonicated and cleared by centrifugation at room temperature (RT). Cell lysates were incubated overnight with 30  $\mu$ l of equilibrated NeutrAvidin-agarose beads (Thermo Scientific). Beads were subjected to stringent washes using the following washing buffers all prepared in 1x PBS: WB1 (8 M urea, 0.25% SDS); WB2 (6 M Guanidine-HCl); WB3 (6.4 M urea, 1 M NaCl, 0.2% SDS), WB4 (4 M urea, 1 M NaCl, 10% isopropanol, 10% ethanol and 0.2% SDS); WB5 (8 M urea, 1% SDS); and WB6 (2% SDS). For elution of biotinylated proteins, beads were heated at 99°C in 50  $\mu$ l of Elution Buffer (4x Laemmli buffer, 100 mM DTT). Beads were separated by centrifugation (18000 x g, 5 minutes). Samples were used for Western blot analysis.

## Western blot

Cells were lysed in cold RIPA buffer [150 mM NaCl, 1.0% NP-40, 0.5% sodium deoxycholate, 0.1% SDS, 50 mM Tris, pH 8.0, protease inhibitors (Roche)] or assay specific lysis buffers. RIPA lysates were kept on ice for 30 minutes vortexing every 5 minutes and then cleared by centrifugation (25000 x g, 20 minutes, 4°C). Supernatants were collected and mixed with Laemmli buffer 5X.

After SDS-PAGE and transfer to nitrocellulose, membranes were blocked in 1x PBS with 5% milk and 0.1% Tween-20. In general, primary antibodies were incubated overnight at 4°C and secondary antibodies for 1 hour at RT. Antibodies used: UROS A-8 (Santa Cruz, 1:1000) detecting mouse protein, UROS G-9 (Santa Cruz, 1:1000) detecting human protein, HA (Sigma, 1:1000/1:500), Cre (Cell Signaling, 1:1000),  $\beta$ -Actin (Sigma, 1:1000),  $\alpha$ -Actin (Proteintech 1:100), GFP (Roche, 1:1000), GAPDH (Proteintech, 1:1000), SALL1 (R&D, 1:1000), CBX4 (Proteintech, 1:1000), Avitag (GeneScript, 1:1000), BirA (Sino Biological Inc., 1:1000), GST (Sigma, 1:1000), Myc (Cell Signaling, 1:1000), Vinculin (Cell Signaling, 1:1000).

HRP-conjugated secondary anti-mouse or anti-rabbit antibodies (Jackson ImmunoResearch, 1:5000) were used. Proteins were detected using Clarity ECL (BioRad) or Super Signal West Femto (Pierce). Quantification of bands was performed using Fiji software and normalized against Actin, GAPDH or Vinculin levels, unless otherwise indicated. At least three independent blots were quantified per experiment.

## Statistical analysis

Statistical analysis was performed using GraphPad 7.0 software. Data were analyzed by Shapiro-Wilk normality test and Levene's test of variance. We used Mann Whitney-U test or Unpaired T-test for comparing two groups, One-way ANOVA for more than two groups and Two-way ANOVA for comparing more than one variable in more than two groups. P values were represented by asterisks as follows: (\*) P-value < 0.05; (\*\*) P-value < 0.01; (\*\*\*) P-value < 0.001; (\*\*\*\*) P-value < 0.0001. Differences were considered significant when  $P < 0.05$ .



**3. CONGENITAL  
ERYTHROPOIETIC  
PORPHYRIA**

**Confidential**













# **4. TOWNES-BROCKS SYNDROME**



## 4.1 Hypothesis

A previous search for human SALL1 interactors by proximity proteomics followed by MS and GO analysis, revealed the E3 SUMO ligase CBX4 as a possible interactor of SALL1. Given the nature of both factors, we hypothesized that SALL1 interaction with CBX4 could have an important role in cellular function and gene expression regulation.

## 4.2 Objectives

Our **general objective** was to **elucidate the mechanisms behind the relationship between SALL1 and the SUMO E3 ligase CBX4 and its biological consequences.**

We defined the following specific objectives:

1. To validate and characterize the interaction between SALL1 and CBX4.
2. To analyze the role of SALL1 SUMOylation in the interaction with CBX4.
3. To analyze the functional effects of SALL1 – CBX4 interaction.

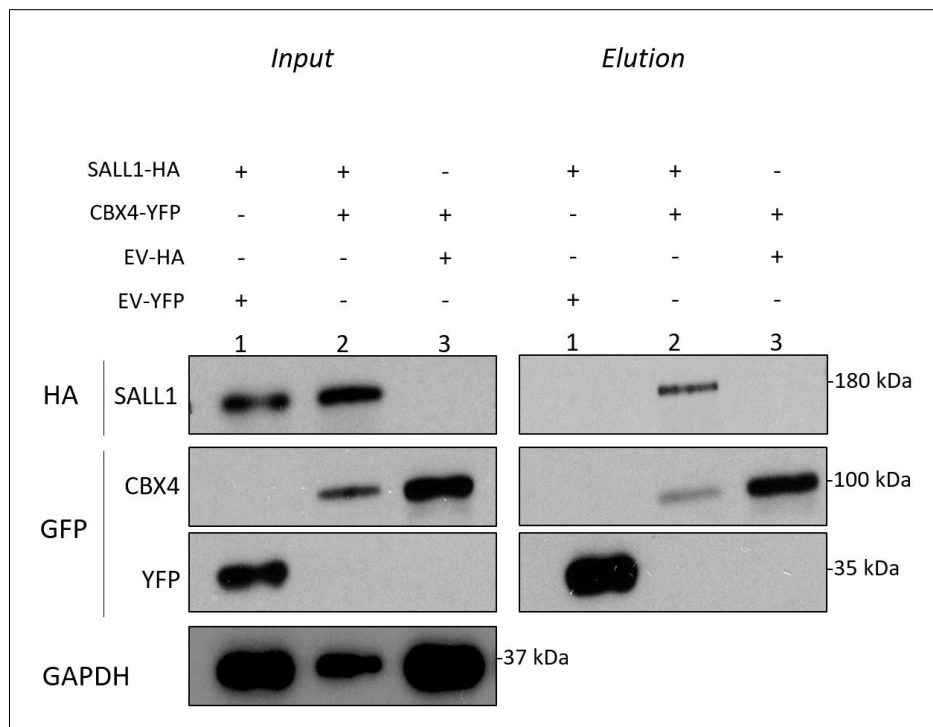
## 4.3 Results

Post-translation modifications can modulate the activity of transcription factors by altering their regulation, localization and interaction with other proteins. As mentioned in the introduction, a previous search for human SALL1 interactors by proximity proteomics followed by MS and GO analysis, revealed that SALL1 interacted with several SUMO E3 ligases, including PIAS1, PIAS2, PIAS4, RANBP2 and CBX4 (Bozal-Basterra et al. 2018). The E3 SUMO ligase CBX4 was chosen for further investigation, since the genetic interaction between the *Drosophila sall* genes and *Pc*, the CBX4 homolog, has been previously described to be important during embryogenesis (Casanova 1989; Landecker, Sinclair, and Brock 1994).

### 4.3.1 Validation of the interaction between SALL1 and CBX4

#### ***SALL1 and CBX4 interact with each other at the protein level***

In order to confirm the interaction between SALL1 and CBX4, we performed a pulldown experiment using GFP-Trap agarose beads (Chromotek). GFP-Trap is an affinity resin for immunoprecipitation of GFP-fusion proteins and consists of GFP nanobodies coupled to agarose beads. *CB6-SALL1-HA* and *CMV-CBX4-YFP* plasmids were transiently overexpressed in HEK 293FT cells and pulled down using GFP-Trap beads. The results were analyzed by Western blot. As shown in Figure 38, SALL1 is detectable in the elution when co-transfected with *CMV-CBX4-YFP* (lane 2) but not when co-transfected with the empty vector *CMV-YFP* (lane 1). These results confirmed the interaction between SALL1 and CBX4 at the protein level.

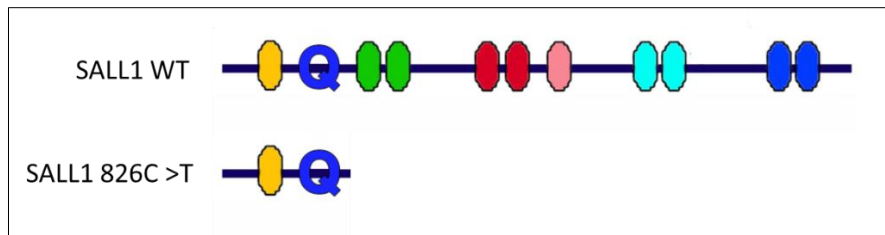


**Figure 38. SALL1 and CBX4 interact with each other at the protein level.** Western blot showing the expression of *CB6-SALL1-HA* transfected in HEK 293FT cells together with *CMV-CBX4-YFP* or *CMV-YFP* alone. *CMV-HA* empty vector was used as control. Both proteins were pulled down using GFP beads and analyzed by Western blot. As shown in the second lane of the pulldown (Elution), CBX4 interacts with SALL1. Molecular weight markers are shown to the right. Antibodies were used as indicated to the left. EV: empty vector.

### ***Truncated SALL1-826C>T interacts with CBX4***

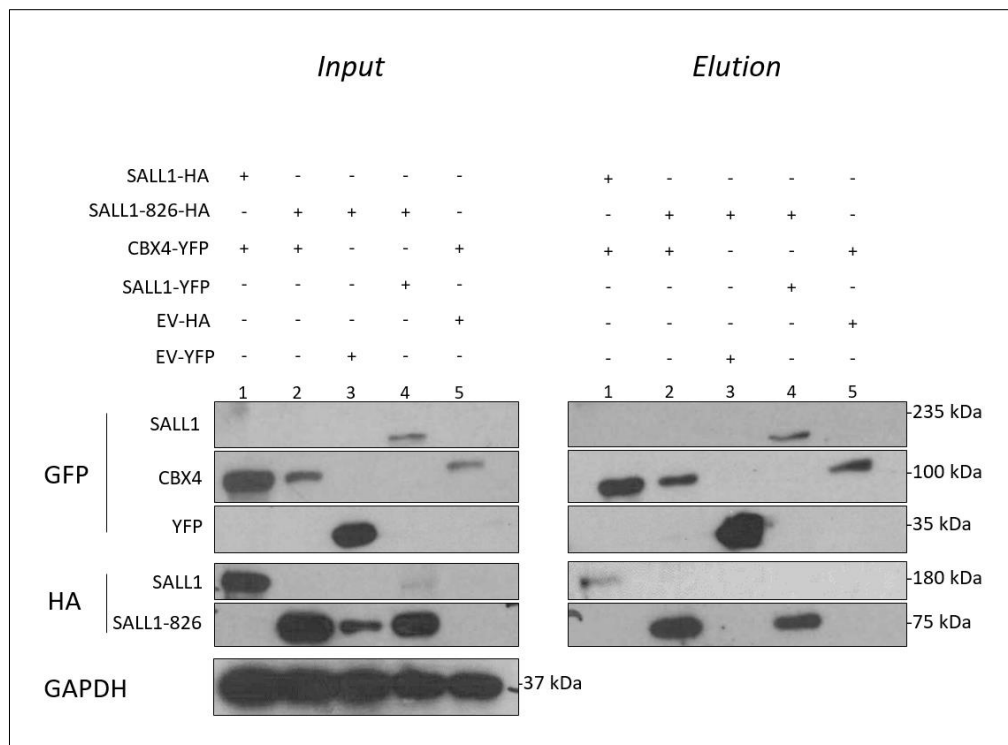
In most of the reported TBS cases, SALL1 presents a mutation in a hot spot region, which results in the expression of a truncated protein. The 826C>T mutation (henceforth referred to as SALL1-826) is the most common allele found in TBS patients (Botzenhart et al. 2007). This mutation generates a truncated protein, which lacks most of the zinc finger pairs, but retains the N-terminal domain, including the ZF1, and the PolyQ region (Figure 39). In fact, truncated proteins causing TBS are still able to form dimers with themselves, with full-length SALL1 or with other SALL proteins. Truncated forms of SALL1 could have lost the capacity to bind CBX4, which could eventually alter the cellular proteostasis.





**Figure 39. Schematic representation of SALL1 WT and the SALL1-826C>T mutant.** This mutation generates a truncated protein that lacks most of the zinc finger pairs, but retains the N-terminal domain of SALL1, including the ZF1 and the PolyQ regions. Ovals represent the zinc fingers (ZF) distributed along the protein. Q represents the poly-glutamine domain.

In order to analyze whether the truncated SALL1-826 mutant preserved the capacity of interaction with CBX4 we performed a GFP-Trap pulldown. *CMV-SALL1-2xHA* or the mutant *CMV-SALL1-826-2xHA* were transiently overexpressed in HEK 293FT cells in combination with *CMV-CBX4-YFP* or *CMV-YFP* alone. *CMV-SALL1-YFP* was used as a positive control, since it was known to bind to the truncated mutant. GAPDH was used as loading control. Our results showed that SALL1-826 mutant was able to interact with CBX4. This result suggested that the domain(s) of interaction between SALL1 and CBX4 could be located at the N-terminal part of the protein (Figure 40).



**Figure 40. SALL1-826 interacts with CBX4.** Western blot analysis of SALL1 full-length and truncated SALL1-826 fused with HA-tag, transfected in HEK 293FT cells together with CBX4 fused with YFP and pulled down using GFP beads. As shown in the lane 2 of the elution panel, SALL1-826 is able to bind CBX4. SALL1-826 dimerization with full-length SALL1 was used as positive control. Molecular weight markers are shown to the right. Antibodies were used as indicated to the left.

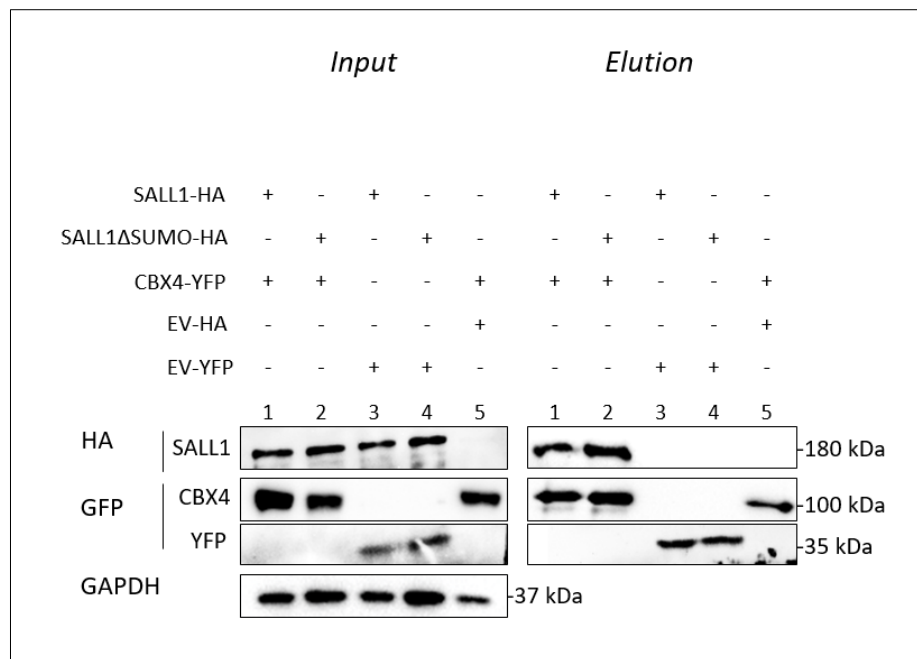
### 4.3.2 Analysis of the role of SALL1 SUMOylation in the interaction with CBX4

#### *SUMOylation of SALL1 is not required for CBX4 interaction*

SALL1 post-translational modifications could affect its interaction with other proteins, for instance those that contain SIM domains as in the case of CBX4. As mentioned in the introduction, for a deeper analysis of SALL1 SUMOylation, our laboratory generated a SALL1 $\Delta$ SUMO mutant in which four lysines were replaced by arginines (K571R, K592R, K982R and K1086R) losing therefore the ability to be SUMOylated in cells and *in vitro*.

In order to test whether SUMOylation could have a role on SALL1 binding to CBX4, we analyzed the SALL1 $\Delta$ SUMO capability to interact with CBX4 (Figure 41). *CMV-*

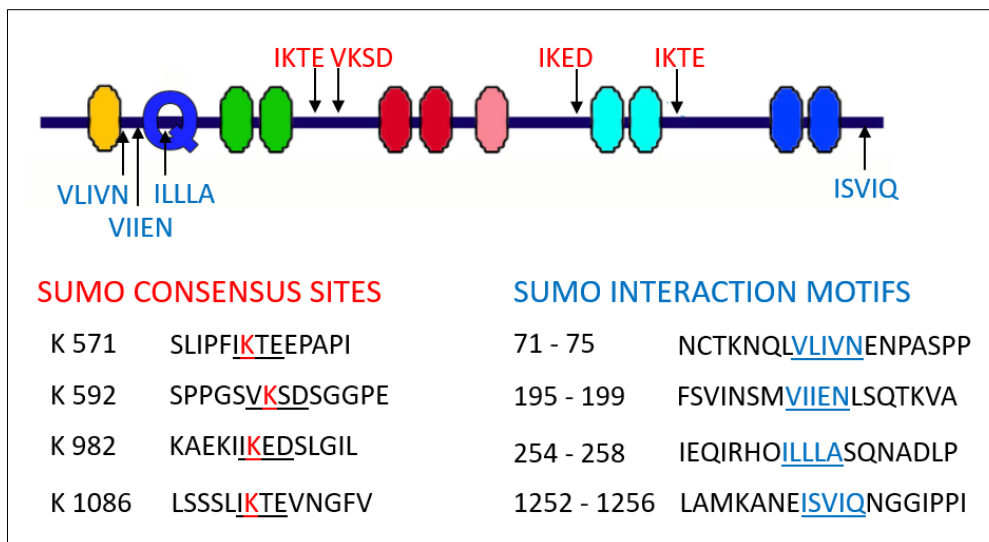
*SALL1-2xHA* and *CMV-SALL1ΔSUMO-2xHA* were transiently transfected in HEK 293FT cells together with *CMV-CBX4-YFP* (lanes 1 and 2 respectively) or with the empty vector *CMV-YFP* (lanes 3 and 4). A GFP-Trap pulldown was performed and analyzed by Western blot. Our results show that the SUMOylation-deficient mutant for SALL1 was still able to interact with CBX4. No visible differences were appreciated between WT SALL1 and SALL1ΔSUMO in the interaction with CBX4.



**Figure 41. Both SALL1 WT and SALL1ΔSUMO interact with CBX4.** *CMV-SALL1-2xHA* and *CMV-SALL1ΔSUMO-2xHA* were transfected in HEK 293FT cells together with *CMV-CBX4-YFP* or *CMV-YFP* alone. Proteins were pulled down using GFP beads and then analyzed by Western blot. As shown in the lanes 1 and 2 of the Elution panel, both WT and mutant versions of SALL1 interact with CBX4. Molecular weight markers are shown to the right. Antibodies were used as indicated to the left.

### ***SALL1 SIM domains are not required for CBX4 interaction***

The results of the pulldowns previously shown demonstrated that the SALL1-CBX4 interactions does not depend on the SUMOylation of SALL1 and that truncated mutant SALL1-826 maintains this interaction, suggesting that at least some interaction domain(s) could be located at the N-terminal part of the protein. By analyzing the full-length amino acid sequence of SALL1, we found the presence of four high-scored SIMs.

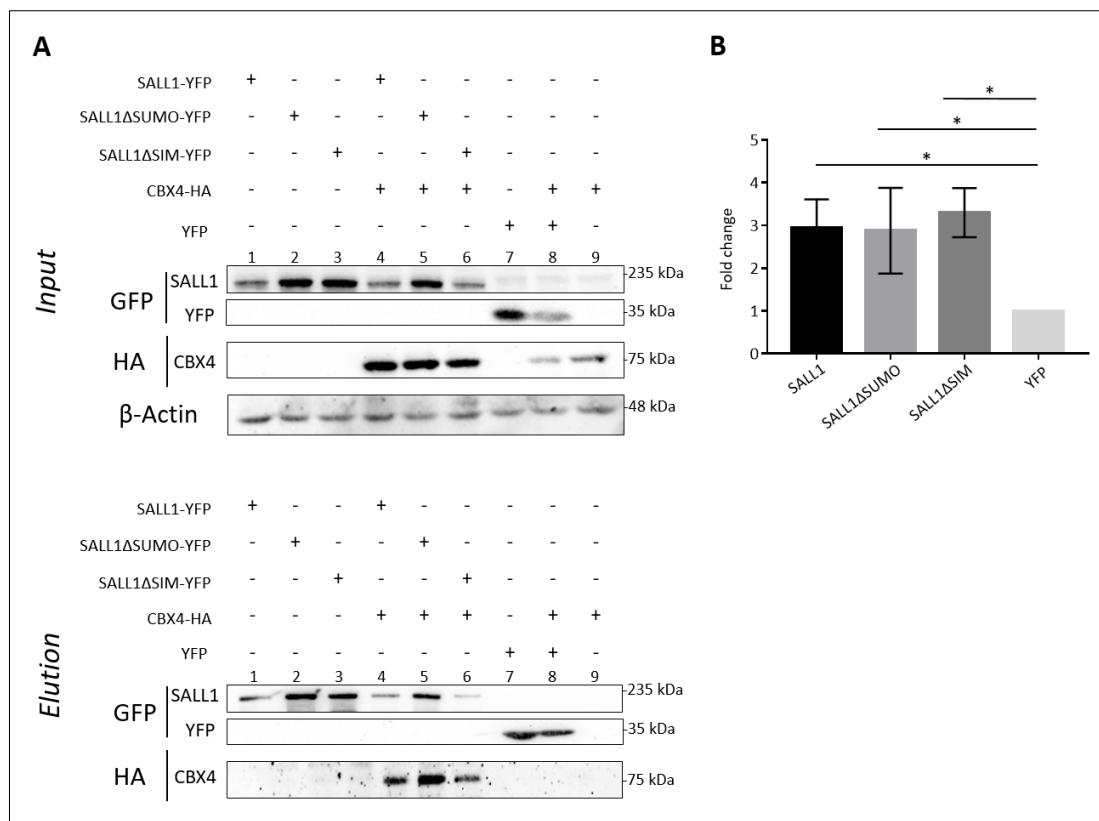


**Figure 42. Identification of SIMs and SUMOylation sites in human SALL1 protein.** Ovals represent the zinc fingers distributed along the protein. Q represents the poly-glutamine domain. In red, SUMO consensus sites mutated in SALL1 $\Delta$ SUMO and, in blue, the predicted SIMs of SALL1.

As shown in Figure 42, three out of the four predicted SIMs of SALL1, are located in the N-terminal part of the protein. Therefore, since CBX4 is known to be SUMOylated *in vitro*, we hypothesized a role of SALL1 SIMs in the interaction with CBX4. To investigate this possibility, we generated a SALL1 $\Delta$ SIM version in which the four predicted SIMs were mutated. At first, the isoleucine (I) residues of the C-terminal SIM of SALL1 were mutated to lysine in order to destroy the hydrophobic core of the domain (ISVIQN > KSVKQN). Afterward, the hydrophobic and acidic residues of the three N-terminal SIMs were mutated to alanine (VLIVN > AAAAN; VIIEN > AAAAN; ILLA > AAAAA). Subsequently, to analyze whether the binding between SALL1 and CBX4 requires SALL1 SIMs, we performed a pulldown experiment comparing the capability of SALL1 WT and both SUMO-related mutants to interact with CBX4 (Figure 43A). *CMV-SALL1-YFP*, *CMV-SALL1 $\Delta$ SUMO-YFP* or *SALL1 $\Delta$ SIM-YFP* were transiently overexpressed in HEK 293FT cells together with *CB6-CBX4-HA* or *CB6-HA* alone.  $\beta$ -Actin was used as loading control. Despite our predictions, as shown in the lanes 4, 5 and 6 of the Elution panel, SALL1 WT, SALL1 $\Delta$ SUMO and SALL1 $\Delta$ SIM showed similar capacity to bind CBX4. Nevertheless, slight differences in the intensity of CBX4 signals between SALL1 WT, SALL1 $\Delta$ SUMO and SALL1 $\Delta$ SIM can be observed. These differences were mostly due to

the initial levels of the YFP-fused proteins. Of note, SALL1 $\Delta$ SUMO levels were higher compared to SALL1 WT, which is reflected in higher levels of CBX4-HA in the pulldown. Our experiments suggested that nor SALL1 SUMOylation or predicted SALL1 SIM sites were necessary for the interaction with CBX4.

As shown in Figure 43A, the levels of CBX4 were higher when cells were transfected with SALL1 proteins. In order to test this, the levels of CBX4 in SALL1 WT, SALL1 $\Delta$ SUMO, SALL1 $\Delta$ SIM and YFP alone expressing cells were quantified in three independent experiment, normalized to the  $\beta$ -Actin levels, and plotted in a graph (Figure 43B). In presence of SALL1 or any of its mutant forms, the levels of CBX4 were significantly higher than in YFP-transfected control cells. No significant differences in CBX4 levels were observed between SALL1 WT and SALL1 $\Delta$ SUMO or SALL1 $\Delta$ SIM.

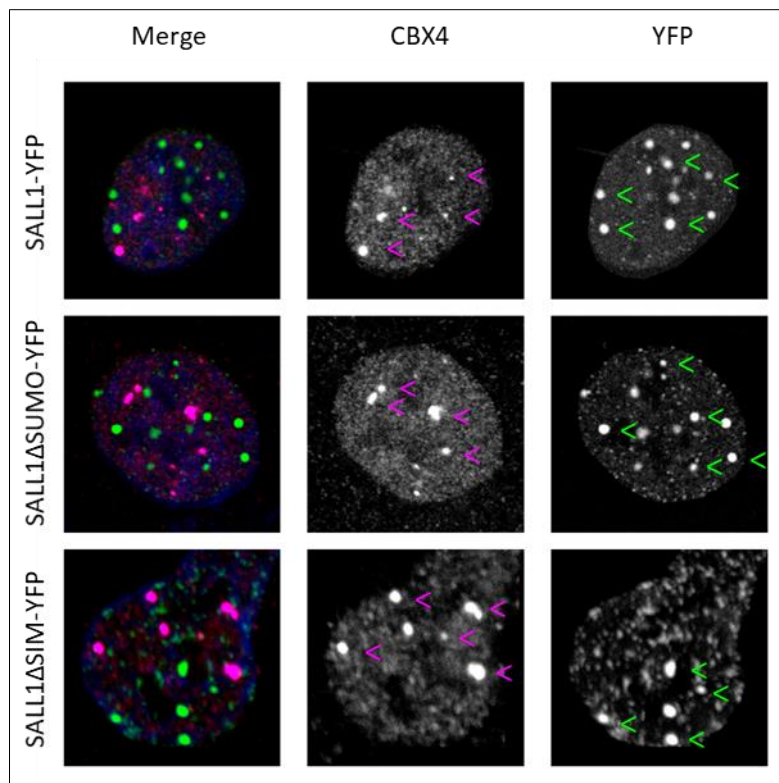


**Figure 43. SUMO-related SALL1 mutants interact with CBX4 at the protein level.** A) Western blot analysis of HEK 293FT cells transfected with *CB6-CBX4-HA* together with *CMV-SALL1-YFP*, *CMV-SALL1 $\Delta$ SUMO-YFP*, *CMV-SALL1 $\Delta$ SIM-YFP* or *CMV-YFP* as control. Pulldowns were performed using the GFP-Trap technique. Interaction between CBX4 and WT SALL1 or SALL1 mutants was detected in the Elution. Molecular weight markers are shown to the right. Antibodies were used as indicated to the left. B) CBX4 levels in SALL1 WT, SALL1 $\Delta$ SUMO, SALL1 $\Delta$ SIM or YFP alone expressing cells were quantified, normalized to  $\beta$ -Actin and reported as relative fold change. P-values were calculated on n= 3 using Mann Whitney test. \*p< 0,05. Graph represents mean plus SEM.

### ***SUMO-related mutant forms of SALL1 do not localize to Pc bodies***

The mammalian cell nucleus contains several substructures, well defined despite the absence of a delineating membrane. Many types of nuclear “bodies” have been identified in the nucleus such as PML bodies and Pc bodies. Pc bodies are nuclear foci that show high concentration of PcG proteins and have been defined as centers of chromatin regulation for transcriptional repression of target genes. CBX4 localizes to the Pc bodies. SALL1 has been also reported to localize at nuclear bodies in cultured cells and *in vivo*, but the nature and function of these bodies are still unknown. In contrast to the results of SALL1-CBX4 interaction, previous experiments showed that SALL1 does not localize to Pc bodies. However, Proximity Ligation Assay (PLA) results obtained in the laboratory, showed the co-localization of SALL1 and CBX4 in the nucleoplasm (Pirone 2016). Since the interaction between WT and mutant forms of SALL1 and CBX4 was demonstrated, we wondered whether mutations in SALL1 could influence the subcellular localization of the protein and, consequently, alter its localization respect to CBX4.

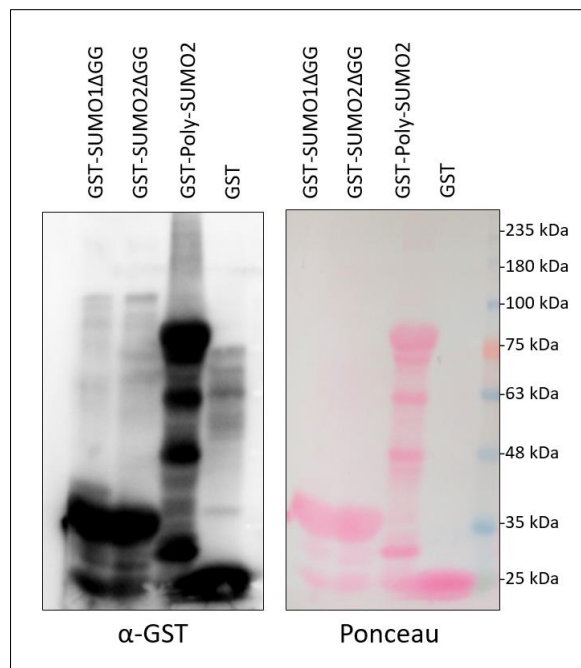
SALL1 or its SUMO-related mutants fused to YFP were transfected into U2OS cells, where endogenous CBX4 was visualized by immunofluorescence using a specific anti-CBX4 antibody. Surprisingly, as shown in Figure 44, our confocal pictures showed that, despite the interaction between the two proteins, there was not co-localization of SALL1 and CBX4 in the nuclear bodies, neither the WT nor the SALL1 $\Delta$ SUMO or SALL1 $\Delta$ SIM mutant forms. Overall, these results indicated that SALL1 co-localizes with CBX4 in the nucleoplasm, but not on chromatin.



**Figure 44. SALL1 WT or SUMO-related mutants do not co-localize with CBX4 at the nuclear bodies.** Immunofluorescence confocal pictures of SALL1-YFP, SALL1 $\Delta$ SUMO-YFP or SALL1 $\Delta$ SIM-YFP expressed in U2OS cells (green), and endogenous CBX4 (magenta) using a specific antibody. Nuclei were stained with DAPI. Black and white pictures show single green or magenta channels. Arrowheads indicate SALL1 (green arrowheads) or Pc bodies (magenta arrowheads).

#### ***Analysis of the SUMO binding capacity of SALL1 predicted SIMs***

The results obtained with the SALL1 $\Delta$ SIM mutant form indicated that the identified SIMs were not involved in the binding to CBX4. We decided then to examine their capacity to bind SUMO in a no-covalent way. Hence, we generated constructs to produce three recombinant SUMO proteins in *E. coli*, fused to a GST-tag (Glutathione S-transferase): *GST-SUMO1- $\Delta$ GG*, *GST-SUMO2- $\Delta$ GG* and *GST-PolySUMO2*. The C-terminal di-glycine motif of SUMO was removed to prevent covalent modifications. GST alone was used as a control. Protocol details are reported in Material and Methods section. Purified proteins were verified by Western blot as shown in Figure 45. The nitrocellulose membrane was stained with ponceau and then GST-proteins were recognized using an anti-GST antibody.



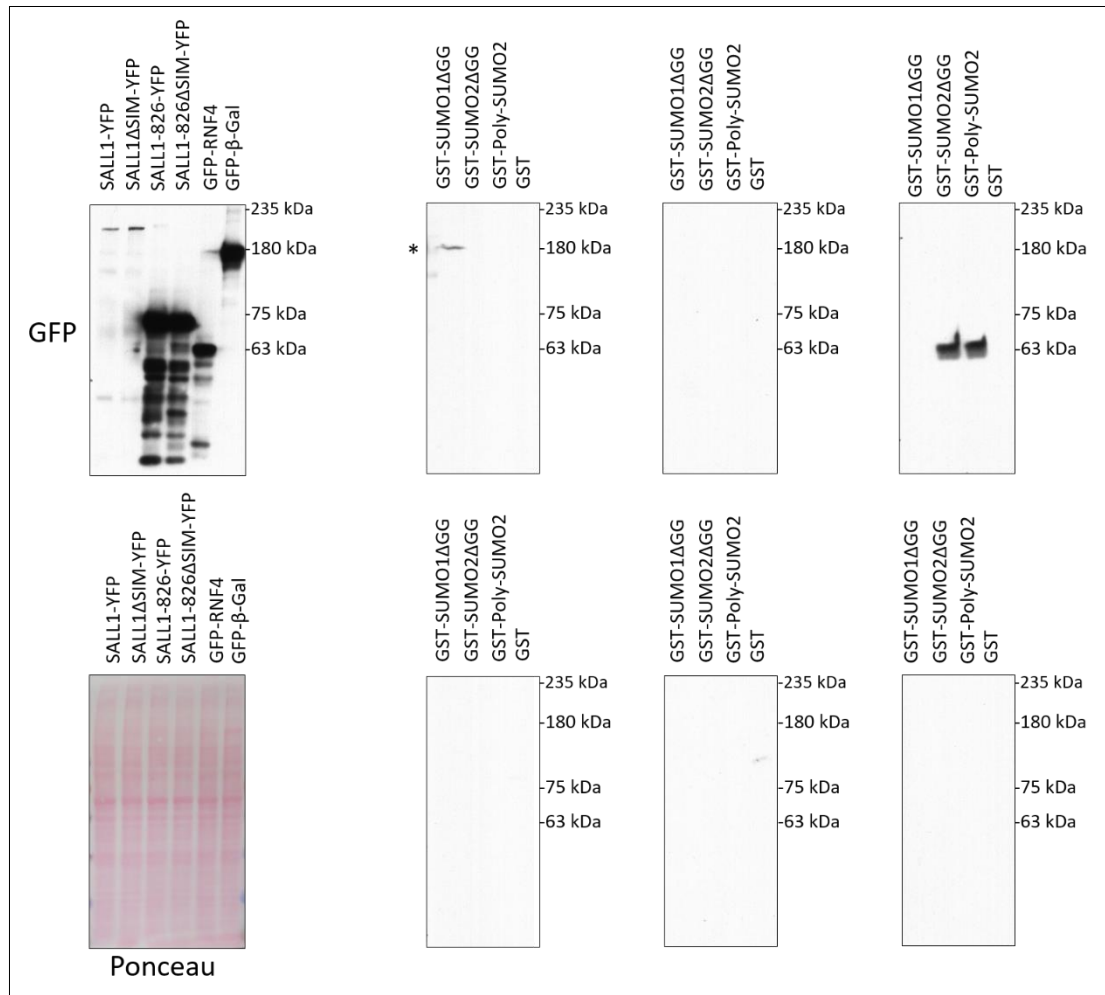
**Figure 45. Western blot analysis of recombinant GST-SUMO proteins.** *GST-SUMO1-ΔGG*, *GST-SUMO2-ΔGG*, *GST-PolySUMO2* and GST-alone proteins were purified and analyzed by Western blot. Nitrocellulose membrane was stained with ponceau (right) and then proteins were recognized using an anti-GST antibody. Molecular weight markers are shown to the right.

GST-fused proteins or GST alone were incubated with glutathione-beads. Next, charged beads were used for pulldown experiments. *CMV-SALL1-YFP* and *CMV-SALL1ΔSIM-YFP*, as well as *CMV-SALL1-826-YFP* and *CMV-SALL1-826 ΔSIM-YFP* were transiently overexpressed in HEK 293FT cells and cell extracts were pulled down using the GST-SUMO beads. *CMV-GFP-RNF4* was used as positive control since this protein is known to contain four tandem SIM repeats that preferentially interact with poly-SUMO chains *versus* SUMO moieties (Tatham et al. 2008). Conversely, *CMV-GFP-β-Galactosidase* ( $\beta$ -Gal) was used as negative control. Results were analyzed by Western blot (Figure 46).

We could not detect the interaction of SALL1 WT, neither of SALL1-826 WT with the SUMO moieties, nor the polySUMO2 chains, in the Elution (panels 1 and 4, respectively), while the positive control RNF4 was detected in both SUMO2 and Poly-SUMO 2 pulldowns (panel 3 of the Elution). However, no differences could be appreciated between RNF4 binding affinity for SUMO2 or PolySUMO2 in our conditions.



These experiments revealed that the predicted and selected SALL1 SIMs might not be functional, or alternatively, better tools should be used for their validation (see Discussion). In view of these results, the SIM mutant form of SALL1 was not taken into consideration for the following experiments.



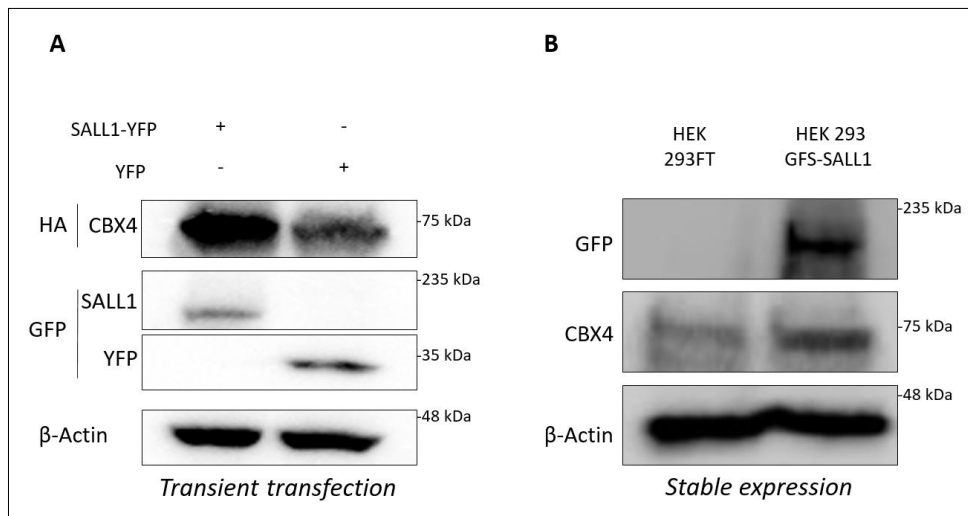
**Figure 46. Analysis of SALL1 SIMs through GST-SUMO beads pulldown.** Western blot analysis of SALL1-YFP or SALL1 $\Delta$ SIM-YFP and SALL1-826 or SALL1-826  $\Delta$ SIM overexpressed in HEK 293FT cells and pulled down with GST-SUMO beads. GFP-RNF4 was used as positive control while GFP- $\beta$ -Galactosidase was used as negative control. Molecular weight markers are shown to the right. Antibodies were used as indicated to the left. Asterisk indicate unspecific band.

### 4.3.3 Analysis of the functional effects of SALL1 – CBX4 interaction

#### *SALL1 increases endogenous CBX4 protein levels*

Previous results showed that, when co-transfected with SALL1, CBX4 protein levels were increased (Figure 43). In order to confirm those results, we transiently transfected HEK 293FT cells with *CB6-CBX4-HA* together with *CMV-SALL1-YFP* or with *CMV-YFP* alone. Western blot analysis revealed higher levels of CBX4-HA in presence of SALL1-YFP than in presence of YFP alone (Figure 47A).

In order to discard any artefactual effect due to the overexpression conditions, we generated a HEK 293FT cell line expressing constitutively the GFS tagged (GFP-Flag-Strep tag; see Materials and Methods section) version of SALL1. This cell line allowed us to obtain constant SALL1 expression levels moderately increased over the WT HEK 293FT cells. We analyzed the endogenous levels of CBX4 in HEK 293FT WT cells compared with GFS-SALL1 expressing cells. Western blot analysis showed increased levels of CBX4 in HEK 293FT\_GFS-SALL1 cells compared with HEK 293FT WT cells (Figure 47B).



**Figure 47. SALL1 increases CBX4 protein levels.** A) SALL1-YFP was transfected in HEK 293FT cells and CBX4-HA levels were detected by Western blot using anti-HA antibodies. B) Endogenous CBX4 levels in stable HEK 293FT GFS-SALL1 cells were higher than in HEK 293FT WT cells. Molecular weight markers are shown to the right. Antibodies were used as indicated to the left.

In the interest to modulate the expression of SALL1 in cells, and so provide a more robust evidence of its effect on CBX4 levels, we decided to generate a stable HEK

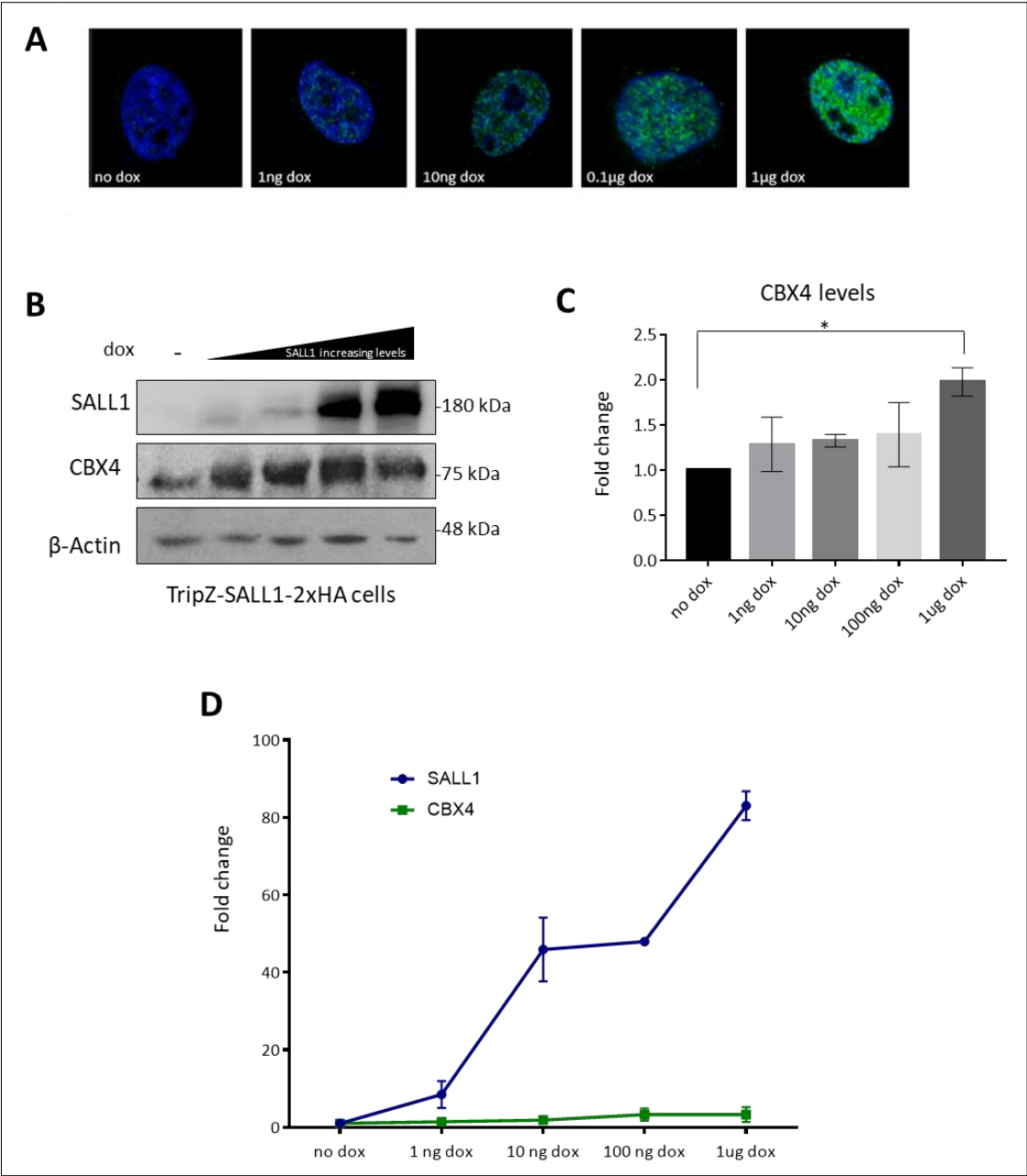
293FT cell line using the inducible lentiviral TripZ vector (see Materials and Methods section). The HEK 293FT\_TripZ-SALL1-2xHA cell line was based on the Tet-On system. Hence, this model allowed us to induce SALL1-2xHA expression in cells in a doxycycline dependent way, while the endogenous SALL1 expression was preserved.

To test the effectiveness of this new cell model, we tested SALL1 expression at increasing concentrations of doxycycline. HEK 293FT\_TripZ-SALL1-2xHA cells were treated with 1 ng, 10 ng, 0.1  $\mu$ g or 1  $\mu$ g per ml of doxycycline and the progressive increment of the SALL1 expression was verified by immunofluorescence (Figure 48A). SALL1-2xHA and endogenous SALL1 were detected using anti-SALL1 primary antibody and anti-mouse Alexa 488 secondary antibody. Confocal microscopy pictures showed how SALL1 signal, in green, increased according to doxycycline dose. The same doses of doxycycline were used to perform a Western blot experiment in which CBX4 protein levels were also analyzed (Figure 48B). The quantification of three independent experiments showed that CBX4 levels were significantly increased when the cells were treated with 1  $\mu$ g of doxycycline compared to untreated cells (Figure 48C).

#### ***SALL1 increases CBX4 levels at post-transcriptional level***

The experiments described above showed that CBX4 levels increased in presence of SALL1. As SALL1 is a transcription factor, we wondered whether this effect could be due to the transcriptional activation of CBX4 expression in presence of SALL1, probably in an indirect way being SALL1 a transcriptional repressor. To address this possibility, we took advantage from the inducible HEK 293FT\_TripZ-SALL1-2xHA cell model.

Cells were treated with the above-mentioned concentration of doxycycline and SALL1 and CBX4 mRNA expression was analyzed by quantitative reverse transcription PCR (qPCR). As shown in Figure 48D, and consistent with previous results showing an increase in SALL1 protein levels, SALL1 mRNA expression increased in a doxycycline dependent manner. However, CBX4 mRNA expression levels did not vary significantly. Altogether these results demonstrated that high levels of SALL1 induce an increment of CBX4 protein levels and, importantly, that this effect occurs at post-transcriptional level.



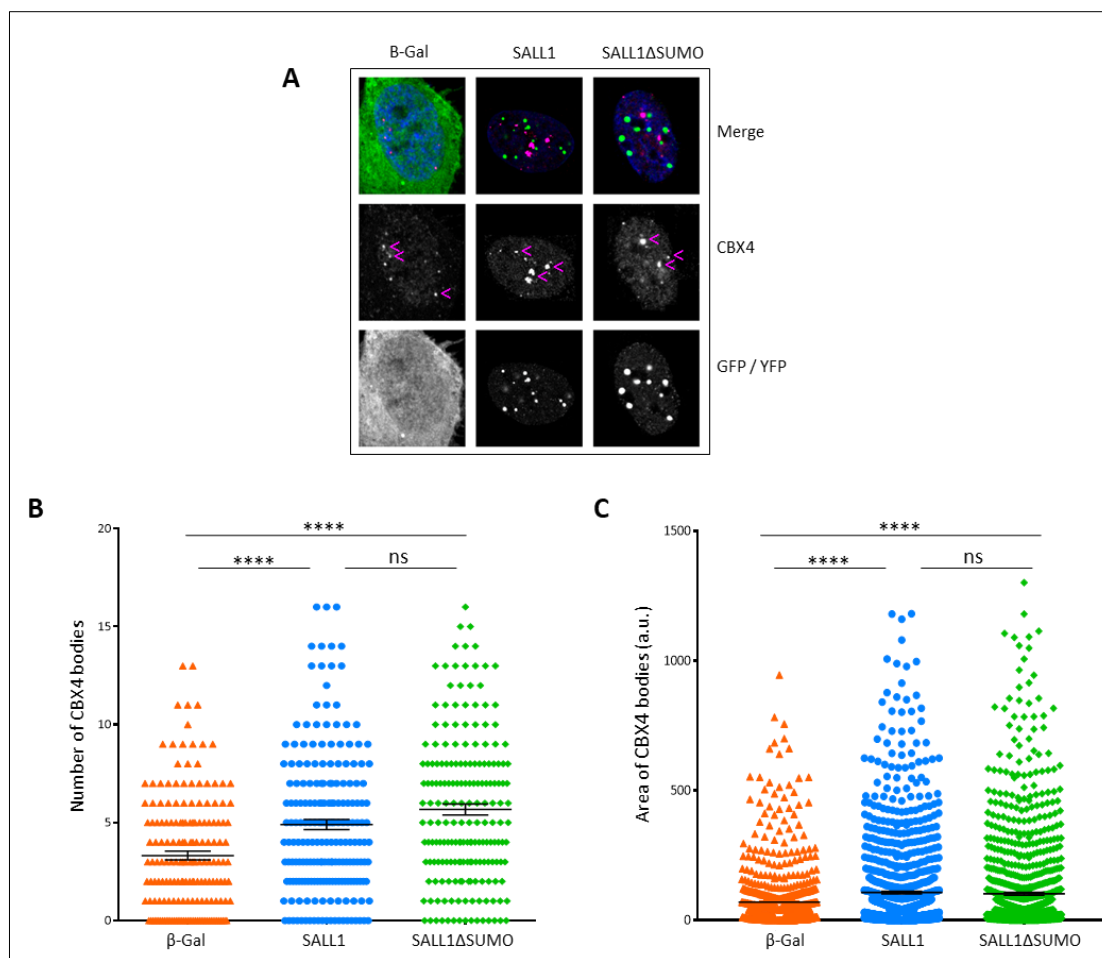
**Figure 48. Analysis of SALL1 and CBX4 levels in HEK 293FT cell model.** A) Immunofluorescence confocal microscopy pictures of HEK 293FT\_TripZ-SALL1-2xHA cells. Cells were treated with different concentration of doxycycline (dox) to induce SALL1 expression. SALL1-2xHA was detected using anti-SALL1 primary antibody (green). Cell nuclei were stained with DAPI (blue). B) Western blot analysis showing CBX4 endogenous levels in 293FT\_TripZ-SALL1-2xHA cells at different concentration of doxycycline. Molecular weight markers are shown to the right. Antibodies were used as indicated to the left. C) Quantification of 3 independent experiments. CBX4 values were normalized to β-Actin. P-values were calculated using Two-way ANOVA test. \* P-value < 0,05. D) qPCR of *SALL1* and *CBX4* mRNA expression at different concentrations of doxycycline in HEK 293FT\_TripZ-SALL1-2xHA cells. *SALL1* and *CBX4* expression is shown as fold change and normalized using *GAPDH*. Graphs in C) and D) represent mean with SEM.

### ***SALL1 influences size and number of Pc bodies***

One of the most important functions of CBX4 is its role as a member of the PRC1 complex required for chromatin remodelling. As previously mentioned in the introduction, PcG protein can form nuclear bodies as a result of the concentration of several chromatin repression points. The previous experiment demonstrated that SALL1 does not co-localize with CBX4 in the Pc bodies, neither in its WT nor SUMOylation or SIM mutant forms. However, in light of the latest results, we decided to further analyze the relationship between these two proteins at the level of the Pc bodies via immunofluorescence.

We transiently transfected *CMV-SALL1-YFP* or its mutant *CMV-SALL1ΔSUMO-YFP* in U2OS cells. *CMV-GFP-β-Galactosidase* ( $\beta$ -Gal) was transfected as control. While SALL1 or SALL1ΔSUMO localized to nuclear bodies,  $\beta$ -Gal localized both in the nucleus and in the cytoplasm. To visualize CBX4 bodies, transfected cells were stained with specific CBX4 primary antibody (magenta).

By analyzing confocal microscopy pictures, we observed that cells transfected with SALL1 or SALL1ΔSUMO showed a higher number of Pc bodies compared with the cells transfected with  $\beta$ -Gal (Figure 49A). Number of bodies, as well as the bodies area were examined in more than 100 cells per condition using Fiji software (Figure 49B-C). Our results showed that Pc bodies are significantly larger and more abundant in cells expressing SALL1 or SALL1ΔSUMO than in cells expressing  $\beta$ -Gal. However, no significant differences were observed between cells expressing SALL1 and SALL1ΔSUMO, neither in the number nor in the area of the Pc bodies. These results revealed that Pc bodies are modified in the presence of SALL1, independently of its SUMOylation status.



**Figure 49. Analysis of Pc bodies.** A) Confocal microscopy images of U2OS cells expressing SALL1-YFP, SALL1ΔSUMO-YFP or β-Gal-GFP, visible in green. CBX4 bodies, in magenta, were stained with CBX4 specific antibody. DAPI was used to stain cell nuclei. B) CBX4 bodies were counted and quantified as well as the CBX4 bodies areas using Fiji software and processed data were represented. Graph shows number of bodies per cell in the three different conditions. C) Graph shows the mean area of the bodies per cell per condition. P-values were calculated using One-way ANOVA test. ns: no significant; \*\*\*\*  $p < 0.0001$ . Error bars indicate mean plus SEM. Graphs represent mean with SEM.

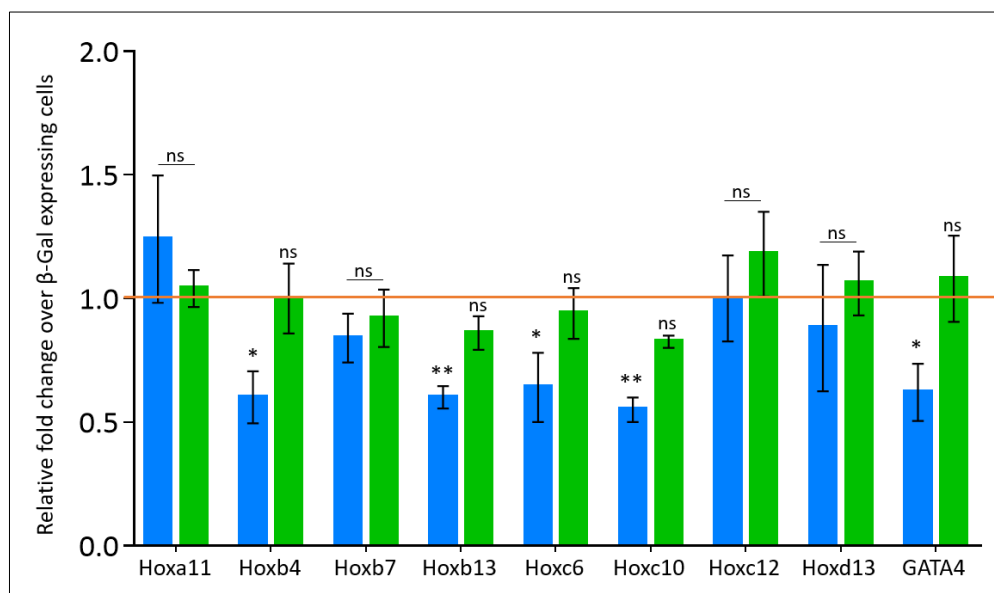
### ***CBX4 targets are downregulated in presence of SALL1***

PcG multiprotein PRC1 complex plays a central role in the transcriptional silencing of target genes. Enhanced Pc bodies formation may lead to increased transcriptional repression of several target genes, including *Hox* genes, of which PcG proteins are well known regulators (Gonzalez et al. 2014; Cheutin and Cavalli 2018; Soshnikova 2014). As shown previously, CBX4 protein levels, as well as the size and

number of CBX4 bodies, increased as consequence of SALL1 overexpression. Therefore, we hypothesized that this could produce an increased transcriptional repression of the CBX4 target genes.

Among direct CBX4 targets, we selected genes belonging to *Hox* clusters and *GATA* family for their expression analysis. HEK 293FT cells were transiently transfected with *CMV-SALL1-YFP*, *CMV-SALL1 $\Delta$ SUMO-YFP* or *CMV-GFP- $\beta$ -Galactosidase* as control, and *Hoxa11*, *Hoxb4*, *Hoxb7*, *Hoxb13*, *Hoxc6*, *Hoxc10*, *Hoxc12*, *Hoxd13* and *GATA4* expression levels were analyzed by qPCR. Our analysis showed significant differences in the expression levels of *Hoxb4*, *Hoxb13*, *Hoxc6*, *Hoxc10* and *GATA4* between SALL1 and  $\beta$ -GAL expressing cells (Figure 50). However, no significant differences were observed between SALL1 $\Delta$ SUMO and control cells.

Taken together, these results indicated that high SALL1 levels increases the transcriptional repression capacity of CBX4 on some of its target genes, probably by increasing CBX4 presence on chromatin. Interestingly, SUMOylation of SALL1 seemed to be necessary for this transcriptional effect.



**Figure 50. SALL1 expression downregulates CBX4 targets.** Graphical representation of qPCR analysis of CBX4 target genes expression. HEK 293FT cells were transfected with *CMV-SALL1-YFP* (blue columns), *CMV-SALL1 $\Delta$ SUMO-YFP* (green columns) or *CMV-GFP- $\beta$ -GAL* as control (orange line). Afterwards, the expression levels of the indicated CBX4 target genes were analyzed by qPCR. Relative gene expression data were normalized on *GAPDH* and relative fold change over  $\beta$ -Gal expressing cells was represented. P-values were calculated on n= 5 using One-way ANOVA test. ns: no significant; \*p< 0,05; \*\*p<0,01. Graph represents mean plus SEM.

### ***SALL1 stabilizes CBX4 avoiding its degradation via the UPS***

Increasing the levels of a given protein can be linked to multiple cellular processes such as alteration in protein stability, changes in subcellular localization or protein degradation. The results obtained from the analysis of Pc bodies revealed that SALL1 is involved in the regulation of Pc bodies number and size. Moreover, SALL1 overexpression increased the transcriptional repression capacity of CBX4 over its target genes. These findings led us to hypothesize that SALL1 could influence the formation of Pc bodies by stabilizing CBX4 and thus increasing its presence on the chromatin.

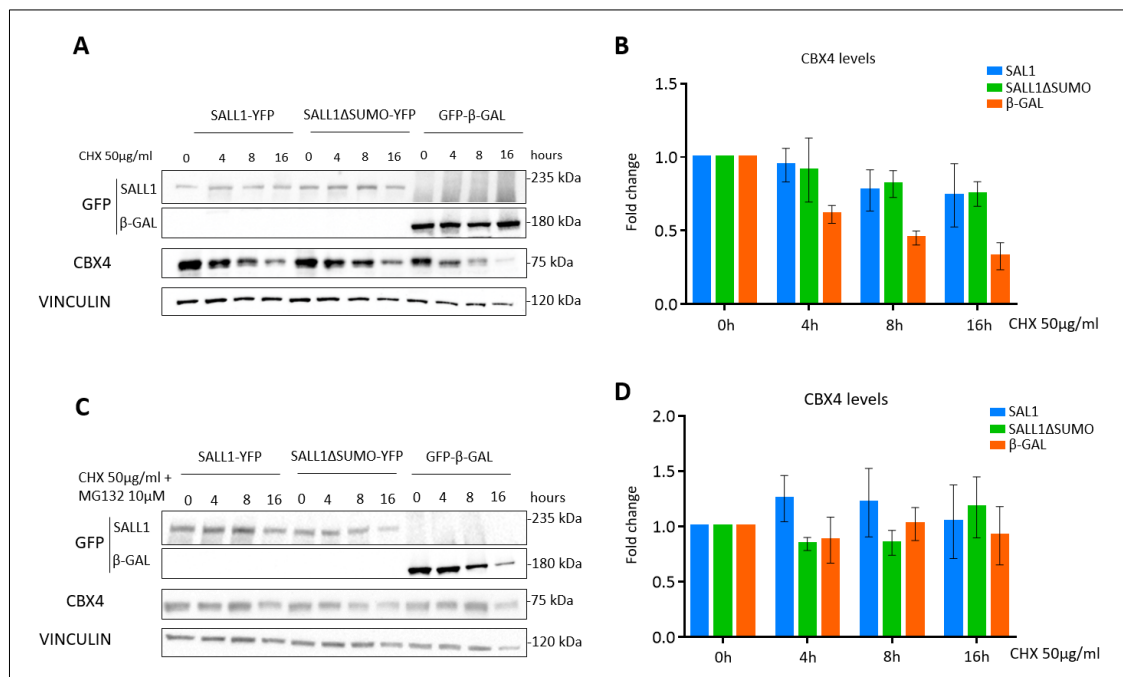
In order to test this hypothesis, we analyzed the half-life of CBX4 by using cycloheximide (CHX). CHX treatment blocks the translational elongation of protein synthesis allowing the observation of the half-life of a protein of interest without confounding contributions from *de novo* translation. Therefore, in order to investigate whether SALL1 stabilizes CBX4 at the protein level, HEK 293FT cells were transfected with *CMV-SALL1-YFP*, *CMV-SALL1 $\Delta$ SUMO-YFP* or *CMV-GFP- $\beta$ -Galactosidase* and treated with 50  $\mu$ g/ml of CHX in presence or absence of 10  $\mu$ M of the proteasome inhibitor MG132. Cells were collected at different time points (0, 4, 8 and 16 hours after initiation of treatment) and CBX4 levels were analyzed by Western blot. Vinculin was used as loading control.

This time-course experiment revealed that after 4 h of CHX treatment, the levels of CBX4 start to decrease. However, in SALL1 WT or SALL1 $\Delta$ SUMO transfected cells the reduction in CBX4 levels was slower than in control cells (Figure 51A). Quantification of four independent experiments is shown in Figure 51B. When cells were co-treated with CHX and MG132 (Figure 51C), proteasome degradation was inhibited and CBX4 levels did not decline at 4 h. Consequently, as shown in the Western blot quantification, no



clear differences in the CBX4 levels were observed between cells transfected with SALL1, SALL1 $\Delta$ SUMO or control (Figure 51D). Overall, these results suggested that in presence of SALL1 or SALL1 $\Delta$ SUMO the levels of CBX4 decreased slower than in the control. Moreover, the accumulation of CBX4 levels in the cells treated with the MG132 indicated that the CBX4 protein degradation occurs through the proteasome.

Therefore, we concluded that SALL1 stabilizes CBX4 protein slowing down its degradation via the proteasome and that the SUMOylation of SALL1 seemed not essential for CBX4 stabilization.



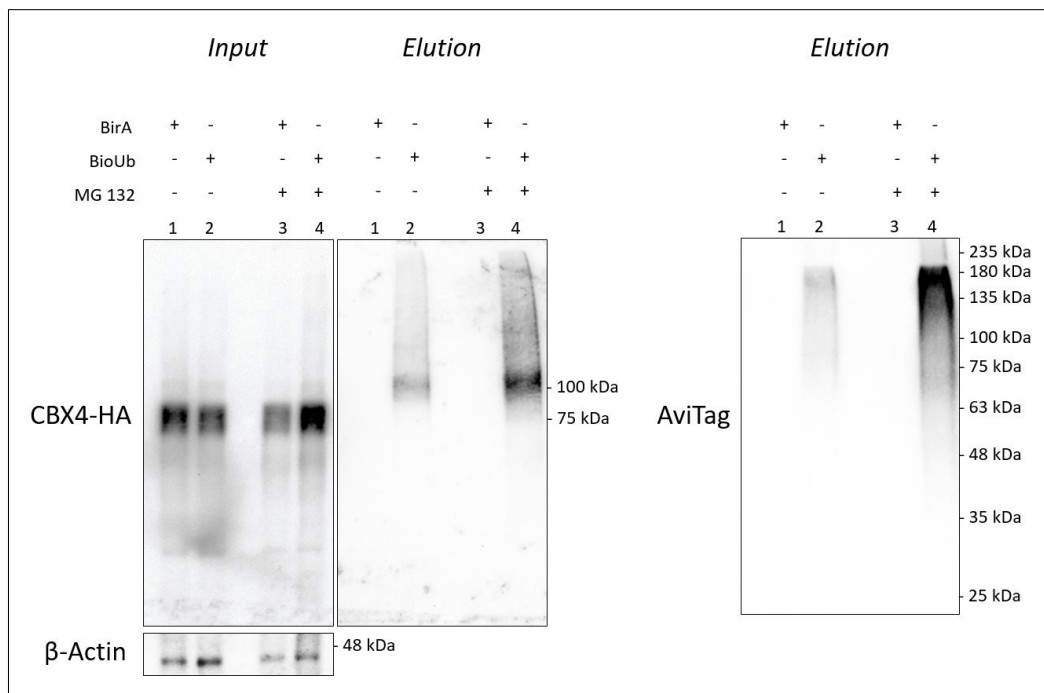
**Figure 51. SALL1 stabilizes CBX4 protein.** Cycloheximide chase experiments were performed in HEK 293FT cells transfected with *CMV-SALL1-YFP*, *CMV-SALL1 $\Delta$ SUMO-YFP* or *CMV-GFP- $\beta$ -Galactosidase*. Cells were treated with 50  $\mu$ g/ml of cycloheximide (CHX) in presence (C) or absence (A) of 10  $\mu$ M of the proteasome inhibitor MG132. Cells were collected at different time points (0, 4, 8 and 16 hours after initiation of treatment) and endogenous CBX4 levels were analyzed by Western blot. Vinculin was used as loading control. Molecular weight markers are shown to the right. Antibodies were used as indicated to the left. CBX4 levels were quantified after CHX treatment alone or in combination with MG132, normalized to Vinculin, and data from three different independent experiments were pooled together in the graphs B and D, respectively. The graph in B shows a trend of SALL1 WT or SALL1 $\Delta$ SUMO in stabilizing CBX4 levels, although no significant differences were obtained by Two-way ANOVA test. Graphs represent mean with SEM.

### ***CBX4 is degraded by the Ubiquitin-Proteasome System***

The degradation of proteins represents an important aspect of cell regulation. For example, protein degradation is an essential process for the rapid turnover of regulatory molecules, such as enzymes or transcription factors. In addition, as mentioned before, misfolded or damaged proteins are quickly degraded to prevent proteostasis alterations. In general, most of the proteins destined for degradation in the proteasome are previously marked with K48-linked polyubiquitin chains on specific lysine residues.

Evidences reported in the literature have shown that CBX4 is covalently modified by ubiquitin before to be sent for degradation through the proteasome (Ning et al. 2017). The bioUb system is an efficient strategy for the isolation of ubiquitin conjugates in cells (Pirone et al. 2017). This method is based on multicistronic expression from a single vector containing the *E. coli* biotin protein ligase BirA and an Avi-tag-fused to Ub. Avi-tag is a unique 15-amino acid peptide, which is specifically biotinylated by the BirA enzyme. Avi-tagged proteins can be purified using streptavidin beads. Denaturing conditions and stringent washes are used to inactivate deconjugating enzymes and remove Ub interactors, as well as non-specific background.

We used this approach to investigate CBX4 modification in relation to SALL1 expression. Therefore, we first tested the efficiency of this system to detect CBX4 ubiquitinated fraction. We transiently transfected HEK 293FT cells with *CB6-CBX4-HA* together with *CMV-BirA-2A-bioUb* or *CMV-BirA* as control. Cells were treated with biotin in presence or absence of the protease inhibitor MG132. Protein lysates were processed for bioUb assay and results were analyzed by Western blot (Figure 52). Ubiquitinated CBX4 is shown in the Elution panel. A prominent band of 135 KDa and a polyubiquitination smear is shown, consistent with a poly-ubiquitinated form of CBX4. As expected, levels of ubiquitinated CBX4 increased in presence of the proteasome inhibitor. Anti-Avi-tag antibodies also showed an increase in the general ubiquitination levels in presence of MG132, as shown in the Elution panel. These results confirmed the modification of CBX4 by Ub and its degradation via UPS. In addition, they proved the usefulness of bioUb assay to study modifications in the CBX4 ubiquitination status.



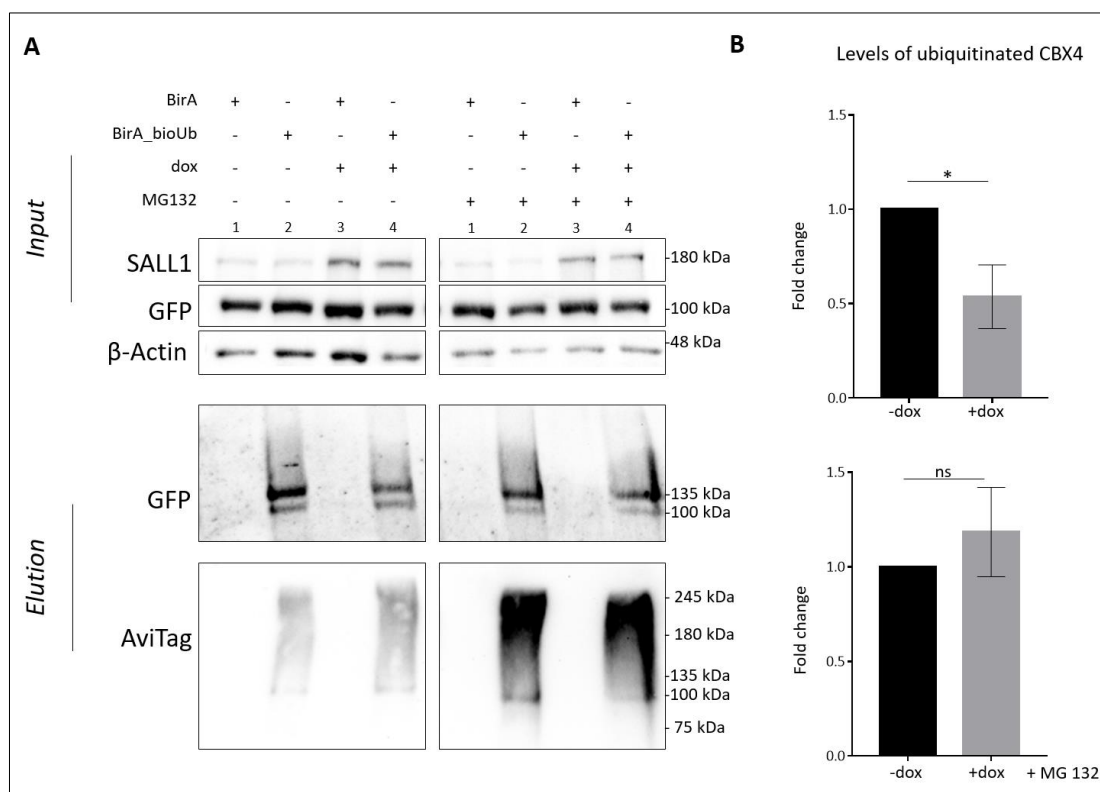
**Figure 52. CBX4 is ubiquitinated and degraded by the proteasome.** Western blot of HEK 293FT cells transfected with *CB6-CBX4-HA* together with *CMV-BirA-2A-bioUb* or *CMV-BirA* as control. Cells were treated with 50  $\mu$ M of biotin in presence or absence of 10  $\mu$ M MG132. Protein lysates were incubated with streptavidin beads and were analyzed by Western blot. Molecular weight markers are shown to the right. Antibodies were used as indicated to the left.

### ***CBX4 ubiquitination is reduced in presence of SALL1***

In light of the obtained results, we speculated that SALL1 could increase CBX4 stability, and consequently enhance its repressive function, by impairing its ubiquitination and subsequent proteasomal degradation. Initially, we planned to compare the ubiquitination levels of CBX4 after overexpression of SALL1, SALL1 $\Delta$ SUMO or  $\beta$ -Gal. However, when more than three plasmids were transfected at the same time, we encountered difficulties to finely control the expression levels of the different overexpressed proteins. Thus, we decided to study CBX4 ubiquitination in the inducible TripZ-SALL1-2xHA cells. For this reason, we could not analyze the effect of SALL1 $\Delta$ SUMO on CBX4 ubiquitination.

TripZ-SALL1-2xHA cells were transiently transfected with *CMV-CBX4-YFP* together with *CMV-BirA-2A-bioUb* or *CMV-BirA* as control. The cells were treated or not

with 1  $\mu\text{g}/\text{ml}$  of doxycycline, in presence or absence of 10  $\mu\text{M}$  of MG132. Protein lysates were passed through streptavidin beads to isolate bio-ubiquitin conjugated proteins and results were analyzed by Western blot (Figure 53A). According to our hypothesis, a statistically significant reduction of CBX4 ubiquitination was observed in presence of high levels of SALL1 (Figure 53B, upper panel). However, in presence of MG132, no significant differences were appreciated between induced and not induced cells (Figure 53B, lower panel). These results indicated that SALL1 is able to stabilize CBX4 protein by reducing its ubiquitination and subsequent degradation via the proteasome.



**Figure 53. Analysis of the effect of SALL1 on CBX4 ubiquitination by bioUb assay.** A) Western blot of HEK TripZ-SALL1-2xHA cells transiently transfected with *CMV-CBX4-YFP* together with *CMV-BirA-2A-bioUb* or *CMV-BirA* as control. The cells were treated or not with 1  $\mu\text{g}/\text{ml}$  of dox, in presence or absence of 10  $\mu\text{M}$  of MG132. Protein lysates were passed through streptavidin beads to isolate bioUb conjugated proteins and results were analyzed by Western blot.  $\beta$ -Actin was used as loading control. B) The levels of ubiquitinated CBX4 in doxycycline induced and not induced cells, in presence (bottom panel) or absence (upper panel) of MG132, were quantified and normalized to the CBX4 levels in the input. P-values were calculated on  $n=4$  using Mann Whitney test. ns: no significant;  $*p < 0,05$ . Graphs represent mean with SEM.

## 4.4 Discussion

Townes-Brocks syndrome is a rare disease caused by mutations in the zinc-finger transcriptional repressor *SALL1* and characterized by a spectrum of malformations in digits, ears and kidneys. Mutations in *SALL1* gene are associated with the presence of a truncated form of the protein that, in contrast to the WT that resides primarily in the nucleus (Sato et al. 2004), localizes to the cytoplasm and impedes primary cilia function (Bozal-Basterra et al. 2018).

Understanding the function and regulation of proteins is essential to understand how cells work and to advance in the treatment of diseases. The study of post-translational modifications can reveal much about proteins regulation. PTMs, in fact, can influence numerous properties of proteins including protein interactions, localization, and half-life. In *Drosophila*, SUMOylation of Sall proteins alter their nuclear localization, as well as their transcriptional repressor activity (Sánchez et al. 2010; 2011). Previous results obtained in our laboratory had confirmed *SALL1* SUMOylation in human cells. Furthermore, a proximity proteomics approach combined with MS analysis was used to identify possible interactors of human *SALL1* (Bozal-Basterra et al. 2018). Among them, GO analysis revealed that the SUMO ligase term was well represented in the list of interacting factors. Since the genetic interaction between *sall* genes and *Pc* had been previously described to be important during *Drosophila* embryogenesis, the E3 SUMO ligase CBX4 was chosen for further investigation and, at first, was hypothesized as possible E3 SUMO ligase involved in *SALL1* SUMOylation. Many efforts have been made in the past to address this possibility. *In vitro* results were compatible with a role of CBX4 promoting the SUMOylation of *SALL1* (unpublished results). However, the results obtained in cells were inconclusive mostly due to technical difficulties experienced during the performance of the experiments (Pirone 2016). Whether the interaction between CBX4 and *SALL1* could have an important role in cellular function and regulation beyond SUMOylation, and the mechanisms underlying this interaction, had not been elucidated.

### ***SALL1-CBX4 interaction***

In this work, we confirmed that SALL1 and CBX4 interact with each other at the protein level. Moreover, using the truncated SALL1-826 mutant we demonstrated that the domain(s) of interaction between SALL1 and CBX4 could be located at the N-terminal part of the protein.

The N-terminal part of SALL1 contains important features that determine TBS etiology. This region contains a conserved PolyQ domain involved in dimerization with other SALL family members (Sweetman et al. 2003), a zinc finger (ZF1) conserved in mammals and a conserved 12 amino acids sequence through which it interacts with the NuRD complex to mediate transcriptional repression (Kiefer et al. 2002; Lauberth and Rauchman 2006). Is any of these motifs involved in CBX4 binding? In order to define the precise site of SALL1 binding to CBX4, different SALL1 truncated mutants could be generated to analyze the minimal residues needed for the interaction. This strategy would also help to study whether the ZF1, a C2HC-type zinc finger that might mediate protein-protein interactions rather than DNA binding (Laity, Lee, and Wright 2001), could have a role in the interaction with CBX4, as well as with other proteins.

By analyzing the full-length amino acid sequence, three SIMs of SALL1 were predicted and, interestingly, two of them were located at the N-terminal part of SALL1. The results obtained with the SALL1 $\Delta$ SIM mutant form or the truncated version SALL1-826 $\Delta$ SIM indicated that the identified SIMs were not necessary for the binding of SALL1 to CBX4. However, it should be noted that these experiments were performed on cells that maintained the endogenous version of SALL1. Therefore, it cannot be excluded that binding to CBX4 was still carried out by the WT moiety through the dimerization of endogenous and overexpressed SALL1 forms. *In vitro* experiments or the use of a SALL1-KO cell line might help to clarify the role of those moieties in the interaction with CBX4. On the other hand, it should be considered that we were not able to prove the binding capacity of these SIM sequences to SUMO. A deeper investigation of those predicted SIMs by different tools, as well as the selection and test of other predicted SALL1 SIMs would be necessary to clarify the properties of these SALL1 domains.

CBX4 also contains two functional SIMs, both of which contribute to its SUMO E3 activity in mammalian cells, being also required for SUMOylation of CBX4 itself (Merrill

et al. 2010). Therefore, we also hypothesized that CBX4 SIMs would be involved in its binding with SALL1 when SUMOylated. However, our results demonstrated that SALL1 $\Delta$ SUMO mutant was still able to interact with CBX4 suggesting that SUMOylation of SALL1 is not required for CBX4 interaction. Nevertheless, the possible contribution of the endogenous SALL1 in the interaction should be considered.

### ***Involvement of SALL1 SUMOylation in Pc bodies regulation and CBX4 stability***

In cells, CBX4 localizes to the Pc bodies, which have been defined as centers of chromatin regulation for transcriptional repression of target genes (Entrevan, Schuettengruber, and Cavalli 2016). Previous experiments showed the co-localization of SALL1 and CBX4 in the nucleoplasm by PLA (Pirone 2016). However, in spite of their interaction, neither SALL1 WT nor the SALL1 $\Delta$ SUMO or SALL1 $\Delta$ SIM mutant forms showed co-localization with CBX4 in the nuclear bodies, indicating that the SALL1-CBX4 interaction do not occurs on chromatin.

Although it does not co-localize to Pc bodies, we demonstrated that SALL1, as well as its SUMOylation-deficient mutant form, affect the Pc bodies formation, which appear more abundant and bigger. The non-localization of SALL1 at Pc bodies could be explained considering its binding with CBX4 as a dynamic and transitory interaction. SALL1 could interact with CBX4 in the nucleoplasm altering its levels but it would not be involved in the formation of PC bodies. Indeed, in this work we demonstrated that SALL1 stabilizes and increases CBX4 protein levels in a post-translational way, reducing its ubiquitination with subsequent reduction of its degradation via the proteasome. Interestingly, a not significant effect was previously observed on Pc bodies in presence of the truncated SALL1-826 form (Pirone 2016).

Different scenarios could explain our results. As a transcriptional repressor, SALL1 could be inhibiting the transcription of ubiquitin E3 ligase(s) involved in CBX4 modification, such is the CHIP (STUB1) ubiquitin E3 ligase, and so impairing its degradation (Wang et al. 2020). Alternatively, SALL1 binding could reduce CBX4 ubiquitination by decreasing its recognition by the ubiquitin E3 ligase. For instance, at least in the case of CHIP, CBX4 phosphorylation is required for its ubiquitin-mediated degradation. The binding with SALL1 could be masking the phosphorylation site and so

impeding CBX4 ubiquitination by CHIP. Also, SALL1 could be facilitating the binding and/or the recognition of CBX4 by DUBs such as USP26, a known deubiquitinase for CBX4 (Ning et al. 2017). Further experiments are required to investigate these possibilities. The identification of the E3 ligase involved in CBX4 ubiquitination in this context will open the way to ulterior experiments. Moreover, DUBs inhibitors could be used to test whether these enzymes could have a role in this regulatory process.

Furthermore, in this work we also demonstrated that high SALL1 levels increase the transcriptional repression capacity of CBX4 on some of its target genes. Although it seems not necessary for CBX4 protein levels regulation, the results obtained by qPCR analysis indicated that SUMOylation of SALL1 is important to modulate the transcriptional repression activity of CBX4 on some gene targets. Therefore, the binding to SALL1 might be sufficient for CBX4 stabilization; however, only when SALL1 is SUMOylated, the recruitment of CBX4 on the chromatin results in a functional effect. One possible explanation of these results could be the involvement of a third component. For instance, SUMOylation of SALL1 could facilitate the contemporary interaction with other members of the PRC1, such as RING1 or PHC1. Interestingly, those factors were also found as possible SALL1 interactors in the MS analysis of the BioID (Bozal-Basterra et al. 2018). Otherwise, SUMOylation of SALL1 could facilitate the interaction with co-factor required for chromatin silencing. For example, the corepressor KAP1 binds PRC1 enhancing its binding at the promoters of differentiation-inducible genes and their transcriptional repression (Cheng, Ren, and Kerppola 2014).

This interpretation could be translated into the following model. SALL1 would interact with CBX4, either in its SUMOylated form or as an unmodified protein. This interaction would result in less ubiquitination of CBX4 with its consequent stabilization. Thus, CBX4 would be recruited on chromatin, either via PREs or by the recognition of the H3K27me3 mark, where it would act as a transcriptional repressor of its target genes directly or promoting the chromatin remodeling via PRC1. However, the mere presence of CBX4 on chromatin would not be sufficient for its function (Figure 54A). When SALL1 is SUMOylated, in addition to interacting with CBX4, it would also be able to interact with repression cofactors or other components of PRC1, which will be recruited on chromatin along with CBX4. The recruitment of transcriptional cofactor(s), or various

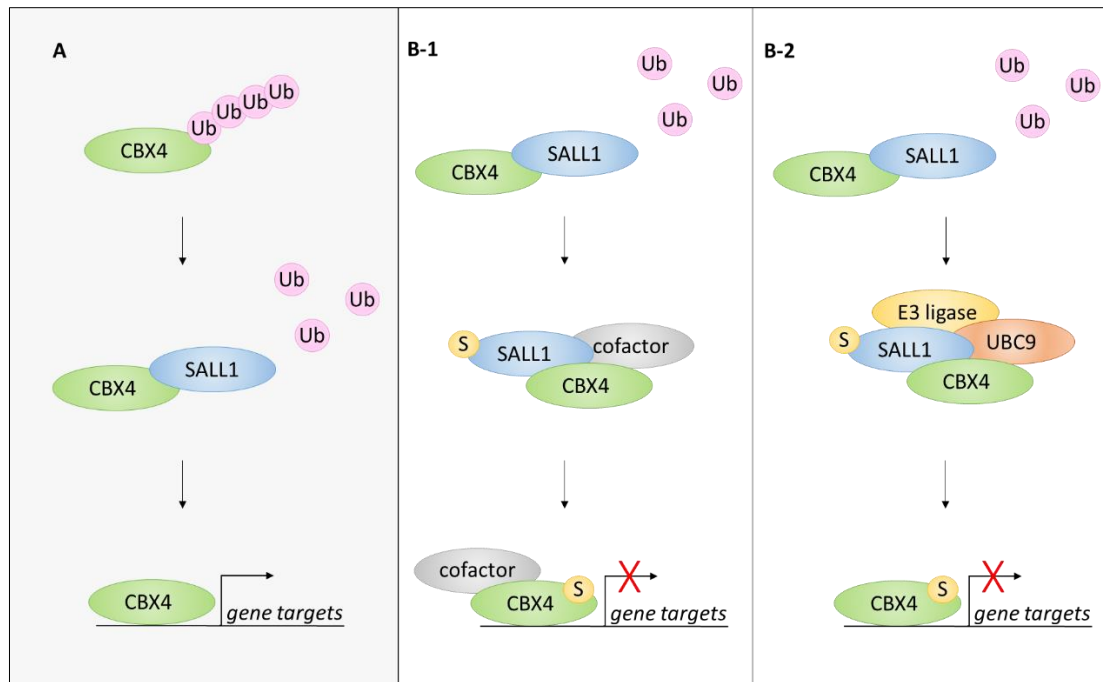


components of PRC1, would result in the activation of the multiprotein complex with consequent repression of the target genes (Figure 54B-1).

Alternatively, SUMOylated SALL1 could enhance CBX4 repression capacity by facilitating its SUMOylation. The SUMOylation of CBX4 is known to be necessary for its repression activity on the chromatin (Kang et al. 2010). Previous results obtained in our laboratory by bioSUMO assay, showed that in presence of high levels of SALL1, the SUMOylation of CBX4 increased (Pirone 2016). However, this result was not easy to interpret due to the increased total levels of the protein. In addition, SALL1 was demonstrated to interact with UBC9 and SUMO1 in a yeast two-hybrid system (Netzer et al. 2002). Interestingly, some members of the SUMOylation pathway were also found in in the MS analysis of the BioID (Bozal-Basterra et al. 2018). In this alternative scenario, once it promoted CBX4 stabilization impairing its ubiquitination, SUMOylated SALL1 would be able also to promote CBX4 SUMOylation by recruiting an E3 SUMO ligase or other components of the SUMOylation machinery (Figure 54B-2). Unlike in the previous hypothesized model in which it would be SALL1-independent, in this case CBX4 SUMOylation would be dependent of SALL1.

Interestingly, a prediction analysis of CBX4 ubiquitination sites by UbPred (<http://www.ubpred.org/>), revealed that the K224 involved in CBX4 SUMOylation, could also works as ubiquitination site. Additionally, the adjacent K209 and K247 were predicted as high score ubiquitination sites. In light of this, we could also speculate that PTM of CBX4 by ubiquitin and SUMO would be mutually exclusive events and that SALL1 might be involved in their regulation. Further experiments are necessary to corroborate the feasibility of both these hypothesized scenarios, which, actually, are not mutually exclusive.

The validation of the interaction of SALL1 WT and SALL1 $\Delta$ SUMO with PRC1 proteins and SUMOylation machinery components by pulldown assay would give us a starting point for the design of the following key experiments. Moreover, a chromatin immunoprecipitation assay could be performed to verify the presence of CBX4 on the promoter of repressed genes, as a more robust proof of its recruitment on chromatin.



**Figure 54. Schematic model of possible regulatory scenarios between SALL1 and CBX4.** A) SALL1 binds and stabilizes CBX4 impeding its ubiquitination and its consequent degradation by the proteasome. CBX4 stabilization entails an increment of its protein levels and its accumulation at the Pc bodies but not an increase of its transcriptional repression capacity. However, when SUMOylated SALL1 binds CBX4 this interaction translates into transcriptional repression of its target genes. B-1) This effect could be due to the contemporary recruitment of other essential transcription factors. B-2) Alternatively, although non-exclusively, SUMOylated SALL1 could increase CBX4 transcriptional repression facilitating its SUMOylation via the recruitment of SUMOylation machinery components.

Altogether, these results suggest that SALL1 plays an important role in the control of the expression of key developmental genes through post-transcriptional regulation of CBX4. The loss of SALL1 SUMOylation, originated by the truncation of SALL1, could alter the regulation of important genes during development, which could eventually contribute to the TBS onset. The elucidation of the SALL1-CBX4 interaction could help to better understand the role of SALL1 during normal development and clarify more aspects of the TBS pathogenesis.





# 5. GENERAL DISCUSSION



The concept of proteostasis, and specially its maintenance, is essential to understand the cause of diseases associated with alterations in proteins misfolding, degradation, trafficking or aggregation. Proteostasis preservation is ensured by a set of interacting pathways that are critical for the proper functioning of the cell. Alteration of those regulatory processes can seriously affect cellular functions and cause the onset of diseases, including rare diseases such as TBS and CEP.

Rare diseases affect a small number of people compared to the general population (less than 1 in 2000 people) and are usually characterized by specific and uncommon issues. Because of their unusual nature, rare diseases often risk being neglected by the pharmacological industry and research policies. However, even though a single rare disease can affect only one person in a million, overall rare disease patients comprise about 7% of the EU population. Besides, these patients often have serious clinical manifestations and fewer available options for treatment.

Investigating the underlying mechanism of rare diseases can lead to interesting scientific discoveries. The research into rare diseases may help us to understand fundamental mechanisms of human biology and disease, often applicable to disorders that are more prevalent. For example, research on Liddle syndrome (a rare inherited kidney disorder associated with early and severe hypertension) has contributed to increase the knowledge about the pathology of hypertension, and studies of Fanconi anemia have illuminated disease mechanisms of bone marrow failure, cancer, and resistance to chemotherapy (Lifton, Gharavi, and Geller 2001; D'Andrea 2010). Examining how and why a gene is not working properly in a particular disease can also help to discover more about its function under normal conditions, which in turn can shine light on the function of other genes that interact with it, as well as help to our understanding of related common diseases. *Vice versa*, the analysis of the physiological regulation of a disease-causing gene may help us to understand or unveil the pathological consequences of its alteration.

This thesis work was focused on two different rare diseases: Townes-Brock Syndrome and Congenital Erythropoietic Porphyria. Both diseases are extremely rare conditions caused by proteostasis alterations. However, the mechanisms underlying

each of these conditions are different, as well as the premises that drove our research interest on them.

Although numerous advances have been made in understanding the molecular alterations of the disease, CEP remains a metabolic disorder for which no effective treatments are available yet. In addition, the lack of adequate models makes the study of new possible therapies even more difficult. A new humanized and inducible CEP C73R mouse could improve drug discovery by providing a more useful model that could partially solve viability and health problems experienced with previous models, facilitating experimental procedures.

TBS presents developmental defects that are known to be caused by mutations in *SALL1*. Recently, TBS has been proposed to be a ciliopathy, as the truncated *SALL1* might interfere with primary cilia regulation (Bozal-Basterra et al. 2018). However, very little is known about the function of this transcription factor, which makes difficult to understand the molecular basis of the disease. The discovery of a role for *SALL1* in the regulation of chromatin remodelling and modulation of developmental gene expression through the regulation of *CBX4* could shed light on its role in human development and help to better understand molecular mechanism underlying TBS.

In summary, the results obtained throughout this work constitute a step forward in the understanding of *SALL1* function, and would help to advance in the development of new CEP therapies, as well as to improve our knowledge about the pathogenesis of both rare diseases.







## **6. CONCLUSIONS**



The results obtained along this thesis on the analysis of the UROS-C73R humanized and inducible mouse cell model for CEP lead to the following conclusions:

- I. Upon FLEX-Switch construct insertion by CRISPR/Cas9, human UROS correctly replaces and rescues mouse UROS function.
- II. Upon tamoxifen induction of Cre recombinase, UROS WT allele is correctly converted to express the mutant C73R version.
- III. FACS and HPLC analyses confirmed increased porphyrin levels after induction by tamoxifen.
- IV. MG132 and ciclopirox treatments cause a reduction in porphyrin levels.
- V. This cellular model represent a valid CEP model.

The results obtained along this thesis on the analysis of the relation between SALL1 and CBX4 lead to the following conclusions:

- I. SALL1 interacts with CBX4 at the protein level, probably through its N-terminal part domain(s).
- II. SALL1-826 truncated mutant interact with CBX4 at protein level.
- III. SUMOylation of SALL1 is not required for its interaction with CBX4.
- IV. The analyzed SALL1 SIMs domains are not required for its interaction with CBX4.
- V. SALL1, as well as its SUMO-related mutants, do not co-localize with CBX4 at Pc nuclear bodies but they do co-localize in the nucleoplasm.

- VI. At high levels of SALL1, as well as of its SUMO-related mutants, endogenous CBX4 protein levels increase.
- VII. SALL1 increases CBX4 levels at post-transcriptional level.
- VIII. SALL1, as well as its SALL1 $\Delta$ SUMO mutant form, influences size and number of Pc bodies, which appeared more abundant and larger.
- IX. CBX4 targets are downregulated in presence of high levels of SALL1, however not in presence of its SALL1 $\Delta$ SUMO mutant form.
- X. SALL1 stabilizes CBX4 impairing its ubiquitination and consequent degradation via the UPS.







## **7. BIBLIOGRAPHY**



- Aizencang, G., C. Solis, D. F. Bishop, C. Warner, y R. J. Desnick. 2000. «Human Uroporphyrinogen-III Synthase: Genomic Organization, Alternative Promoters, and Erythroid-Specific Expression». *Genomics* 70 (2): 223-31. <https://doi.org/10.1006/geno.2000.6373>.
- Aranda, Sergi, Gloria Mas, y Luciano Di Croce. 2015. «Regulation of Gene Transcription by Polycomb Proteins». *Science Advances* 1 (11): e1500737. <https://doi.org/10.1126/sciadv.1500737>.
- Arnold, Konstantin, Lorenza Bordoli, Jürgen Kopp, y Torsten Schwede. 2006. «The SWISS-MODEL Workspace: A Web-Based Environment for Protein Structure Homology Modelling». *Bioinformatics (Oxford, England)* 22 (2): 195-201. <https://doi.org/10.1093/bioinformatics/bti770>.
- Aronson, Boaz E., Kelly A. Stapleton, y Stephen D. Krasinski. 2014. «Role of GATA Factors in Development, Differentiation, and Homeostasis of the Small Intestinal Epithelium». *American Journal of Physiology. Gastrointestinal and Liver Physiology* 306 (6): G474-490. <https://doi.org/10.1152/ajpgi.00119.2013>.
- Balch, William E., Richard I. Morimoto, Andrew Dillin, y Jeffery W. Kelly. 2008. «Adapting Proteostasis for Disease Intervention». *Science (New York, N.Y.)* 319 (5865): 916-19. <https://doi.org/10.1126/science.1141448>.
- Balwani, Manisha, y Robert J. Desnick. 2012. «The Porphyrrias: Advances in Diagnosis and Treatment». *Blood* 120 (23): 4496-4504. <https://doi.org/10.1182/blood-2012-05-423186>.
- Benkert, Pascal, Marco Biasini, y Torsten Schwede. 2011. «Toward the Estimation of the Absolute Quality of Individual Protein Structure Models». *Bioinformatics (Oxford, England)* 27 (3): 343-50. <https://doi.org/10.1093/bioinformatics/btq662>.
- Bett, John S. 2016. «Proteostasis Regulation by the Ubiquitin System». *Essays in Biochemistry* 60 (2): 143-51. <https://doi.org/10.1042/EBC20160001>.
- Bhagavan N. V., Chung-Eun Ha. 2015. *Essentials of Medical Biochemistry With Clinical Cases*. Second edition. Vol. Chapter 27. Elsevier.
- Biasini, Marco, Stefan Bienert, Andrew Waterhouse, Konstantin Arnold, Gabriel Studer, Tobias Schmidt, Florian Kiefer, et al. 2014. «SWISS-MODEL: Modelling Protein Tertiary and Quaternary Structure Using Evolutionary Information». *Nucleic Acids Research* 42 (Web Server issue): W252-258. <https://doi.org/10.1093/nar/gku340>.
- Bishop, David F., Sonia Clavero, Narla Mohandas, y Robert J. Desnick. 2011. «Congenital Erythropoietic Porphyria: Characterization of Murine Models of the Severe Common (C73R/C73R) and Later-Onset Genotypes». *Molecular Medicine (Cambridge, Mass.)* 17 (7-8): 748-56. <https://doi.org/10.2119/molmed.2010.00258>.
- Bishop, David F., Annika Johansson, Robert Phelps, Amr A. Shady, Maria C. M. Ramirez, Makiko Yasuda, Andres Caro, y Robert J. Desnick. 2006. «Uroporphyrinogen III Synthase Knock-in Mice Have the Human Congenital Erythropoietic Porphyria Phenotype, Including the Characteristic Light-Induced Cutaneous Lesions». *American Journal of Human Genetics* 78 (4): 645-58. <https://doi.org/10.1086/502667>.
- Blouin, Jean-Marc, Yann Duchartre, Pierre Costet, Magalie Lalanne, Cécile Ged, Ana Lain, Oscar Millet, Hubert de Verneuil, y Emmanuel Richard. 2013. «Therapeutic

- Potential of Proteasome Inhibitors in Congenital Erythropoietic Porphyria». *Proceedings of the National Academy of Sciences of the United States of America* 110 (45): 18238-43. <https://doi.org/10.1073/pnas.1314177110>.
- Bohren, Kurt M., Varsha Nadkarni, Jian H. Song, Kenneth H. Gabbay, y David Owerbach. 2004. «A M55V Polymorphism in a Novel SUMO Gene (SUMO-4) Differentially Activates Heat Shock Transcription Factors and Is Associated with Susceptibility to Type I Diabetes Mellitus». *The Journal of Biological Chemistry* 279 (26): 27233-38. <https://doi.org/10.1074/jbc.M402273200>.
- Borozdin, Wiktor, Katharina Steinmann, Beate Albrecht, Armand Bottani, Koenraad Devriendt, Michael Leipoldt, y Jürgen Kohlhase. 2006. «Detection of Heterozygous SALL1 Deletions by Quantitative Real Time PCR Proves the Contribution of a SALL1 Dosage Effect in the Pathogenesis of Townes-Brocks Syndrome». *Human Mutation* 27 (2): 211-12. <https://doi.org/10.1002/humu.9396>.
- Botzenhart, Elke M., Gabriella Bartalini, Edward Blair, Angela F. Brady, Frances Elmslie, Karen L. Chong, Katie Christy, et al. 2007. «Townes-Brocks Syndrome: Twenty Novel SALL1 Mutations in Sporadic and Familial Cases and Refinement of the SALL1 Hot Spot Region». *Human Mutation* 28 (2): 204-5. <https://doi.org/10.1002/humu.9476>.
- Botzenhart, Elke M., Andrew Green, Helena Ilyina, Rainer König, R. Brian Lowry, Ivan F. M. Lo, Mordechai Shohat, et al. 2005. «SALL1 Mutation Analysis in Townes-Brocks Syndrome: Twelve Novel Mutations and Expansion of the Phenotype». *Human Mutation* 26 (3): 282. <https://doi.org/10.1002/humu.9362>.
- Boyer, Laurie A., Kathrin Plath, Julia Zeitlinger, Tobias Brambrink, Lea A. Medeiros, Tong Ihn Lee, Stuart S. Levine, et al. 2006. «Polycomb Complexes Repress Developmental Regulators in Murine Embryonic Stem Cells». *Nature* 441 (7091): 349-53. <https://doi.org/10.1038/nature04733>.
- Bozal-Basterra, Laura, Itziar Martín-Ruíz, Lucia Pirone, Yinwen Liang, Jón Otti Sigurðsson, Maria Gonzalez-Santamarta, Immacolata Giordano, et al. 2018. «Truncated SALL1 Impedes Primary Cilia Function in Townes-Brocks Syndrome». *American Journal of Human Genetics* 102 (2): 249-65. <https://doi.org/10.1016/j.ajhg.2017.12.017>.
- Bracken, Adrian P., y Kristian Helin. 2009. «Polycomb Group Proteins: Navigators of Lineage Pathways Led Astray in Cancer». *Nature Reviews. Cancer* 9 (11): 773-84. <https://doi.org/10.1038/nrc2736>.
- Brehme, Marc, y Cindy Voisine. 2016. «Model Systems of Protein-Misfolding Diseases Reveal Chaperone Modifiers of Proteotoxicity». *Disease Models & Mechanisms* 9 (8): 823-38. <https://doi.org/10.1242/dmm.024703>.
- Brehme, Marc, Cindy Voisine, Thomas Rolland, Shinichiro Wachi, James H. Soper, Yitan Zhu, Kai Orton, et al. 2014. «A Chaperome Subnetwork Safeguards Proteostasis in Aging and Neurodegenerative Disease». *Cell Reports* 9 (3): 1135-50. <https://doi.org/10.1016/j.celrep.2014.09.042>.
- Cantera, Rafael, Karin Lürer, Tor Erik Rusten, Rosa Barrio, Fotis C. Kafatos, y Gerhard M. Technau. 2002. «Mutations in Spalt Cause a Severe but Reversible Neurodegenerative Phenotype in the Embryonic Central Nervous System of *Drosophila Melanogaster*». *Development (Cambridge, England)* 129 (24): 5577-86. <https://doi.org/10.1242/dev.00158>.

- Cañadas, Victoria, Isidre Vilacosta, Isidoro Bruna, y Valentin Fuster. 2010. «Marfan Syndrome. Part 1: Pathophysiology and Diagnosis». *Nature Reviews. Cardiology* 7 (5): 256-65. <https://doi.org/10.1038/nrcardio.2010.30>.
- Cappadocia, Laurent, y Christopher D. Lima. 2018. «Ubiquitin-like Protein Conjugation: Structures, Chemistry, and Mechanism». *Chemical Reviews* 118 (3): 889-918. <https://doi.org/10.1021/acs.chemrev.6b00737>.
- Casanova, Jordi. 1989. «Mutations in Thespalt Gene OfDrosophila Cause Ectopic Expression OfUltrabithorax AndSex Combs Reduced». *Roux's Archives of Developmental Biology: The Official Organ of the EDBO* 198 (3): 137-40. <https://doi.org/10.1007/BF02438938>.
- Celen, Arda B., y Umut Sahin. 2020. «Sumoylation on Its 25th Anniversary: Mechanisms, Pathology, and Emerging Concepts». *The FEBS Journal*, abril. <https://doi.org/10.1111/febs.15319>.
- Celis, J. F. de, y R. Barrio. 2000. «Function of the Spalt/Spalt-Related Gene Complex in Positioning the Veins in the Drosophila Wing». *Mechanisms of Development* 91 (1-2): 31-41. [https://doi.org/10.1016/s0925-4773\(99\)00261-0](https://doi.org/10.1016/s0925-4773(99)00261-0).
- Celis, Jose F. de, y Rosa Barrio. 2009. «Regulation and Function of Spalt Proteins during Animal Development». *The International Journal of Developmental Biology* 53 (8-10): 1385-98. <https://doi.org/10.1387/ijdb.072408jd>.
- Chen, Qingbo, Lei Huang, Dongning Pan, Lihua J. Zhu, y Yong-Xu Wang. 2018. «Cbx4 Sumoylates Prdm16 to Regulate Adipose Tissue Thermogenesis». *Cell Reports* 22 (11): 2860-72. <https://doi.org/10.1016/j.celrep.2018.02.057>.
- Cheng, Bo, Xiaojun Ren, and Tom K. Kerppola. 2014. «KAP1 Represses Differentiation-Inducible Genes in Embryonic Stem Cells through Cooperative Binding with PRC1 and Derepresses Pluripotency-Associated Genes». *Molecular and Cellular Biology* 34 (11): 2075-91. <https://doi.org/10.1128/MCB.01729-13>.
- Cheutin, Thierry, and Giacomo Cavalli. 2018. «Loss of PRC1 Induces Higher-Order Opening of Hox Loci Independently of Transcription during Drosophila Embryogenesis». *Nature Communications* 9 (1): 3898. <https://doi.org/10.1038/s41467-018-05945-4>.
- Chiabrando, Deborah, Sonia Mercurio, and Emanuela Tolosano. 2014. «Heme and Erythropoiesis: More than a Structural Role». *Haematologica* 99 (6): 973-83. <https://doi.org/10.3324/haematol.2013.091991>.
- Cimino, Patrick J., and David H. Gutmann. 2018. «Neurofibromatosis Type 1». *Handbook of Clinical Neurology* 148: 799-811. <https://doi.org/10.1016/B978-0-444-64076-5.00051-X>.
- Cundiff, Mary D., Christina M. Hurley, Jeremy D. Wong, Joseph A. Boscia, Aarti Bashyal, Jake Rosenberg, Eden L. Reichard, Nicholas D. Nassif, Jennifer S. Brodbelt, and Daniel A. Kraut. 2019. «Ubiquitin Receptors Are Required for Substrate-Mediated Activation of the Proteasome's Unfolding Ability». *Scientific Reports* 9 (1): 14506. <https://doi.org/10.1038/s41598-019-50857-and>.
- D'Andrea, Alan D. 2010. «Susceptibility Pathways in Fanconi's Anemia and Breast Cancer». *The New England Journal of Medicine* 362 (20): 1909-19. <https://doi.org/10.1056/NEJMra0809889>.
- David, Gregory, Mychell A. Neptune, and Ronald A. DePinho. 2002. «SUMO-1 Modification of Histone Deacetylase 1 (HDAC1) Modulates Its Biological

- Activities». *The Journal of Biological Chemistry* 277 (26): 23658-63. <https://doi.org/10.1074/jbc.M203690200>.
- Deschamps, Jacqueline, and Denis Duboule. 2017. «Embryonic Timing, Axial Stem Cells, Chromatin Dynamics, and the Hox Clock». *Genes & Development* 31 (14): 1406-16. <https://doi.org/10.1101/gad.303123.117>.
- Desnick, R. J., I. A. Glass, W. Xu, C. Solis, and K. H. Astrin. 1998. «Molecular Genetics of Congenital Erythropoietic Porphyria». *Seminars in Liver Disease* 18 (1): 77-84. <https://doi.org/10.1055/s-2007-1007143>.
- Di Pierro, Elena, Valentina Brancaloni, and Francesca Granata. 2016. «Advances in Understanding the Pathogenesis of Congenital Erythropoietic Porphyria». *British Journal of Haematology* 173 (3): 365-79. <https://doi.org/10.1111/bjh.13978>.
- Dissmeyer, Nico, Olivier Coux, Manuel S. Rodriguez, Rosa Barrio, and Core Group Members of PROTEOSTASIS. 2019. «PROTEOSTASIS: A European Network to Break Barriers and Integrate Science on Protein Homeostasis». *Trends in Biochemical Sciences* 44 (5): 383-87. <https://doi.org/10.1016/j.tibs.2019.01.007>.
- Dupuis-Girod, Sophie, Véronique Akkari, Cécile Ged, Claire Galambrun, Kamila Kebaïli, Jean-Charles Deybach, Alain Claudy, et al. 2005. «Successful Match-Unrelated Donor Bone Marrow Transplantation for Congenital Erythropoietic Porphyria (Günther Disease)». *European Journal of Pediatrics* 164 (2): 104-7. <https://doi.org/10.1007/s00431-004-1575-x>.
- Eifler, Karolin, and Alfred C. O. Vertegaal. 2015. «Mapping the SUMOylated Landscape». *The FEBS Journal* 282 (19): 3669-80. <https://doi.org/10.1111/febs.13378>.
- Enserink, Jorrit M. 2015. «Sumo and the Cellular Stress Response». *Cell Division* 10: 4. <https://doi.org/10.1186/s13008-015-0010-1>.
- Entrevan, Marianne, Bernd Schuettengruber, and Giacomo Cavalli. 2016. «Regulation of Genome Architecture and Function by Polycomb Proteins». *Trends in Cell Biology* 26 (7): 511-25. <https://doi.org/10.1016/j.tcb.2016.04.009>.
- Erwin, Angelika, Manisha Balwani, Robert J. Desnick, and Porphyrias Consortium of the NIH-Sponsored Rare Diseases Clinical Research Network. 1993. «Congenital Erythropoietic Porphyria». En *GeneReviews*<sup>®</sup>, editado por Margaret P. Adam, Holly H. Ardinger, Roberta A. Pagon, Stephanie E. Wallace, Lora JH Bean, Karen Stephens, and Anne Amemiya. Seattle (WA): University of Washington, Seattle. <http://www.ncbi.nlm.nih.gov/books/NBK154652/>.
- Eskeland, R., E. Freyer, M. Leeb, A. Wutz, and W. A. Bickmore. 2010. «Histone Acetylation and the Maintenance of Chromatin Compaction by Polycomb Repressive Complexes». *Cold Spring Harbor Symposia on Quantitative Biology* 75: 71-78. <https://doi.org/10.1101/sqb.2010.75.053>.
- Farshi, Pershang, Rahul R. Deshmukh, Joseph O. Nwankwo, Richard T. Arkwright, Boris Cvek, Jinbao Liu, and Q. Ping Dou. 2015. «Deubiquitinases (DUBs) and DUB Inhibitors: A Patent Review». *Expert Opinion on Therapeutic Patents* 25 (10): 1191-1208. <https://doi.org/10.1517/13543776.2015.1056737>.
- Fortian, Arola, David Castaño, Gabriel Ortega, Ana Laín, Miquel Pons, and Oscar Millet. 2009. «Uroporphyrinogen III Synthase Mutations Related to Congenital Erythropoietic Porphyria Identify a Key Helix for Protein Stability». *Biochemistry* 48 (2): 454-61. <https://doi.org/10.1021/bi801731q>.
- Fortian, Arola, Esperanza González, David Castaño, Juan M. Falcon-Perez, and Oscar Millet. 2011. «Intracellular Rescue of the Uroporphyrinogen III Synthase Activity

- in Enzymes Carrying the Hotspot Mutation C73R». *The Journal of Biological Chemistry* 286 (15): 13127-33. <https://doi.org/10.1074/jbc.M110.205849>.
- Frank, J., X. Wang, H. M. Lam, V. M. Aita, F. K. Jugert, G. Goerz, H. F. Merk, M. B. Poh-Fitzpatrick, and A. M. Christiano. 1998. «C73R Is a Hotspot Mutation in the Uroporphyrinogen III Synthase Gene in Congenital Erythropoietic Porphyria». *Annals of Human Genetics* 62 (Pt 3): 225-30. <https://doi.org/10.1046/j.1469-1809.1998.6230225.x>.
- Friedel, Roland H., Wolfgang Wurst, Benedikt Wefers, and Ralf Kühn. 2011. «Generating Conditional Knockout Mice». *Methods in Molecular Biology (Clifton, N.J.)* 693: 205-31. [https://doi.org/10.1007/978-1-60761-974-1\\_12](https://doi.org/10.1007/978-1-60761-974-1_12).
- Fritsch, C., K. Bolsen, T. Ruzicka, and G. Goerz. 1997. «Congenital Erythropoietic Porphyria». *Journal of the American Academy of Dermatology* 36 (4): 594-610. [https://doi.org/10.1016/s0190-9622\(97\)70249-4](https://doi.org/10.1016/s0190-9622(97)70249-4).
- Fu, Yanfang, Jeffrey D. Sander, Deepak Reyon, Vincent M. Cascio, and J. Keith Joung. 2014. «Improving CRISPR-Cas Nuclease Specificity Using Truncated Guide RNAs». *Nature Biotechnology* 32 (3): 279-84. <https://doi.org/10.1038/nbt.2808>.
- Gao, Zhonghua, Jin Zhang, Roberto Bonasio, Francesco Strino, Ayana Sawai, Fabio Parisi, Yuval Kluger, and Danny Reinberg. 2012. «PCGF Homologs, CBX Proteins, and RYBP Define Functionally Distinct PRC1 Family Complexes». *Molecular Cell* 45 (3): 344-56. <https://doi.org/10.1016/j.molcel.2012.01.002>.
- Garvin, Alexander J. 2019. «Beyond Reversal: Ubiquitin and Ubiquitin-like Proteases and the Orchestration of the DNA Double Strand Break Repair Response». *Biochemical Society Transactions* 47 (6): 1881-93. <https://doi.org/10.1042/BST20190534>.
- Ged, C., F. Moreau-Gaudry, E. Richard, E. Robert-Richard, and H. de Verneuil. 2009. «Congenital Erythropoietic Porphyria: Mutation Update and Correlations between Genotype and Phenotype». *Cellular and Molecular Biology (Noisy-Le-Grand, France)* 55 (1): 53-60.
- Gonzalez, Inma, Julio Mateos-Langerak, Aubin Thomas, Thierry Cheutin, and Giacomo Cavalli. 2014. «Identification of Regulators of the Three-Dimensional Polycomb Organization by a Microscopy-Based Genome-Wide RNAi Screen». *Molecular Cell* 54 (3): 485-99. <https://doi.org/10.1016/j.molcel.2014.03.004>.
- Gu, Bin, Eszter Posfai, and Janet Rossant. 2018. «Efficient Generation of Targeted Large Insertions by Microinjection into Two-Cell-Stage Mouse Embryos». *Nature Biotechnology* 36 (7): 632-37. <https://doi.org/10.1038/nbt.4166>.
- Haneef, S. A. Syed, and C. George Priya Doss. 2016. «Personalized Pharmacoperones for Lysosomal Storage Disorder: Approach for Next-Generation Treatment». *Advances in Protein Chemistry and Structural Biology* 102: 225-65. <https://doi.org/10.1016/bs.apcsb.2015.10.001>.
- Hartl, F. Ulrich, Andreas Bracher, and Manajit Hayer-Hartl. 2011. «Molecular Chaperones in Protein Folding and Proteostasis». *Nature* 475 (7356): 324-32. <https://doi.org/10.1038/nature10317>.
- Hendriks, Ivo A., and Alfred C. O. Vertegaal. 2016a. «A Comprehensive Compilation of SUMO Proteomics». *Nature Reviews. Molecular Cell Biology* 17 (9): 581-95. <https://doi.org/10.1038/nrm.2016.81>.

- Hendriks, Ivo A., and Alfred C. O. Vertegaal. 2016b. «A High-Yield Double-Purification Proteomics Strategy for the Identification of SUMO Sites». *Nature Protocols* 11 (9): 1630-49. <https://doi.org/10.1038/nprot.2016.082>.
- Hietakangas, Ville, Julius Anckar, Henri A. Blomster, Mitsuaki Fujimoto, Jorma J. Palvimo, Akira Nakai, and Lea Sistonen. 2006. «PDSM, a Motif for Phosphorylation-Dependent SUMO Modification». *Proceedings of the National Academy of Sciences of the United States of America* 103 (1): 45-50. <https://doi.org/10.1073/pnas.0503698102>.
- Hipp, Mark S., Prasad Kasturi, and F. Ulrich Hartl. 2019. «The Proteostasis Network and Its Decline in Ageing». *Nature Reviews. Molecular Cell Biology* 20 (7): 421-35. <https://doi.org/10.1038/s41580-019-0101-and>.
- Hipp, Mark S., Sae-Hun Park, and F. Ulrich Hartl. 2014. «Proteostasis Impairment in Protein-Misfolding and -Aggregation Diseases». *Trends in Cell Biology* 24 (9): 506-14. <https://doi.org/10.1016/j.tcb.2014.05.003>.
- Hochstrasser, Mark. 2009. «Origin and Function of Ubiquitin-like Proteins». *Nature* 458 (7237): 422-29. <https://doi.org/10.1038/nature07958>.
- Hogeling, Marcia, Taizo Nakano, Christopher C. Dvorak, Sheilagh Maguiness, and Ilona J. Frieden. 2011. «Severe Neonatal Congenital Erythropoietic Porphyria». *Pediatric Dermatology* 28 (4): 416-20. <https://doi.org/10.1111/j.1525-1470.2010.01376.x>.
- Ismail, Ismail Hassan, Jean-Philippe Gagné, Marie-Christine Caron, Darin McDonald, Zhizhong Xu, Jean-Yves Masson, Guy G. Poirier, and Michael J. Hendzel. 2012. «CBX4-Mediated SUMO Modification Regulates BMI1 Recruitment at Sites of DNA Damage». *Nucleic Acids Research* 40 (12): 5497-5510. <https://doi.org/10.1093/nar/gks222>.
- Jayaraj, Gopal G., Mark S. Hipp, and F. Ulrich Hartl. 2020. «Functional Modules of the Proteostasis Network». *Cold Spring Harbor Perspectives in Biology* 12 (1). <https://doi.org/10.1101/cshperspect.a033951>.
- Jones, Takako, and Peter L. Jones. 2018. «A Cre-Inducible DUX4 Transgenic Mouse Model for Investigating Facioscapulohumeral Muscular Dystrophy». *PloS One* 13 (2): e0192657. <https://doi.org/10.1371/journal.pone.0192657>.
- Justice, Monica J., and Paraminder Dhillon. 2016. «Using the Mouse to Model Human Disease: Increasing Validity and Reproducibility». *Disease Models & Mechanisms* 9 (2): 101-3. <https://doi.org/10.1242/dmm.024547>.
- Kagey, Michael H., Tiffany A. Melhuish, and David Wotton. 2003. «The Polycomb Protein Pc2 Is a SUMO E3». *Cell* 113 (1): 127-37. [https://doi.org/10.1016/s0092-8674\(03\)00159-4](https://doi.org/10.1016/s0092-8674(03)00159-4).
- Kang, Xunlei, Yitao Qi, Yong Zuo, Qi Wang, Yanqiong Zou, Robert J. Schwartz, Jinke Cheng, and Edward T. H. Yeh. 2010. «SUMO-Specific Protease 2 Is Essential for Suppression of Polycomb Group Protein-Mediated Gene Silencing during Embryonic Development». *Molecular Cell* 38 (2): 191-201. <https://doi.org/10.1016/j.molcel.2010.03.005>.
- Katsumura, Koichi R., Emery H. Bresnick, and GATA Factor Mechanisms Group. 2017. «The GATA Factor Revolution in Hematology». *Blood* 129 (15): 2092-2102. <https://doi.org/10.1182/blood-2016-09-687871>.
- Katugampola, R. P., A. V. Anstey, A. AND. Finlay, S. Whatley, J. Woolf, N. Mason, J. C. Deybach, et al. 2012. «A Management Algorithm for Congenital Erythropoietic



- Porphyria Derived from a Study of 29 Cases». *The British Journal of Dermatology* 167 (4): 888-900. <https://doi.org/10.1111/j.1365-2133.2012.11154.x>.
- Katugampola, R. P., M. N. Badminton, A. AND. Finlay, S. Whatley, J. Woolf, N. Mason, J. C. Deybach, et al. 2012. «Congenital Erythropoietic Porphyria: A Single-Observer Clinical Study of 29 Cases». *The British Journal of Dermatology* 167 (4): 901-13. <https://doi.org/10.1111/j.1365-2133.2012.11160.x>.
- Kauffman, L., D. I. Evans, R. F. Stevens, and C. Weinkove. 1991. «Bone-Marrow Transplantation for Congenital Erythropoietic Porphyria». *Lancet (London, England)* 337 (8756): 1510-11. [https://doi.org/10.1016/0140-6736\(91\)93198-i](https://doi.org/10.1016/0140-6736(91)93198-i).
- Kelberman, Daniel, Lily Islam, Jörn Lakowski, Chiara Bacchelli, Estelle Chanudet, Francesco Lescai, Aara Patel, et al. 2014. «Mutation of SALL2 Causes Recessive Ocular Coloboma in Humans and Mice». *Human Molecular Genetics* 23 (10): 2511-26. <https://doi.org/10.1093/hmg/ddt643>.
- Kiefer, Susan McLeskey, Bradley W. McDill, Jing Yang, and Michael Rauchman. 2002. «Murine Sall1 Represses Transcription by Recruiting a Histone Deacetylase Complex». *The Journal of Biological Chemistry* 277 (17): 14869-76. <https://doi.org/10.1074/jbc.M200052200>.
- Kiefer, Susan McLeskey, Kevin K. Ohlemiller, Jing Yang, Bradley W. McDill, Jürgen Kohlhase, and Michael Rauchman. 2003. «Expression of a Truncated Sall1 Transcriptional Repressor Is Responsible for Townes-Brocks Syndrome Birth Defects». *Human Molecular Genetics* 12 (17): 2221-27. <https://doi.org/10.1093/hmg/ddg233>.
- Klauke, Karin, Višnja Radulović, Mathilde Broekhuis, Ellen Weersing, Erik Zwart, Sandra Olthof, Martha Ritsema, et al. 2013. «Polycomb Cbx Family Members Mediate the Balance between Haematopoietic Stem Cell Self-Renewal and Differentiation». *Nature Cell Biology* 15 (4): 353-62. <https://doi.org/10.1038/ncb2701>.
- Kohlhase, J., A. Wischermann, H. Reichenbach, U. Froster, and W. Engel. 1998. «Mutations in the SALL1 Putative Transcription Factor Gene Cause Townes-Brocks Syndrome». *Nature Genetics* 18 (1): 81-83. <https://doi.org/10.1038/ng0198-81>.
- Kohlhase, Jürgen, Marielle Heinrich, Lucia Schubert, Manuela Liebers, Andreas Kispert, Franco Laccone, Peter Turnpenny, Robin M. Winter, and William Reardon. 2002. «Okhiro Syndrome Is Caused by SALL4 Mutations». *Human Molecular Genetics* 11 (23): 2979-87. <https://doi.org/10.1093/hmg/11.23.2979>.
- Laity, J. H., B. M. Lee, and P. E. Wright. 2001. «Zinc Finger Proteins: New Insights into Structural and Functional Diversity». *Current Opinion in Structural Biology* 11 (1): 39-46. [https://doi.org/10.1016/s0959-440x\(00\)00167-6](https://doi.org/10.1016/s0959-440x(00)00167-6).
- Landecker, H. L., D. A. Sinclair, and H. W. Brock. 1994. «Screen for Enhancers of Polycomb and Polycomblike in Drosophila Melanogaster». *Developmental Genetics* 15 (5): 425-34. <https://doi.org/10.1002/dvg.1020150505>.
- Lauberth, Shannon M., Amy C. Bilyeu, Beth A. Firulli, Kristen L. Kroll, and Michael Rauchman. 2007. «A Phosphomimetic Mutation in the Sall1 Repression Motif Disrupts Recruitment of the Nucleosome Remodeling and Deacetylase Complex and Repression of Gbx2». *The Journal of Biological Chemistry* 282 (48): 34858-68. <https://doi.org/10.1074/jbc.M703702200>.

- Lauberth, Shannon M., and Michael Rauchman. 2006. «A Conserved 12-Amino Acid Motif in Sall1 Recruits the Nucleosome Remodeling and Deacetylase Corepressor Complex». *The Journal of Biological Chemistry* 281 (33): 23922-31. <https://doi.org/10.1074/jbc.M513461200>.
- Lee, Bongyong, and Mark T. Muller. 2009. «SUMOylation Enhances DNA Methyltransferase 1 Activity». *The Biochemical Journal* 421 (3): 449-61. <https://doi.org/10.1042/BJ20090142>.
- Li, Bing, Jing Zhou, Peng Liu, Jialei Hu, Hong Jin, Yohei Shimono, Masahide Takahashi, and Guoliang Xu. 2007. «Polycomb Protein Cbx4 Promotes SUMO Modification of de Novo DNA Methyltransferase Dnmt3a». *The Biochemical Journal* 405 (2): 369-78. <https://doi.org/10.1042/BJ20061873>.
- Li, Jie, Ying Xu, HuiKe Jiao, Wei Wang, Zhu Mei, and GuoQiang Chen. 2014. «Sumoylation of Hypoxia Inducible Factor-1 $\alpha$  and Its Significance in Cancer». *Science China. Life Sciences* 57 (7): 657-64. <https://doi.org/10.1007/s11427-014-4685-3>.
- Lifton, R. P., A. G. Gharavi, and D. S. Geller. 2001. «Molecular Mechanisms of Human Hypertension». *Cell* 104 (4): 545-56. [https://doi.org/10.1016/s0092-8674\(01\)00241-0](https://doi.org/10.1016/s0092-8674(01)00241-0).
- Ma, Yuanwu, Lianfeng Zhang, and Xingxu Huang. 2014. «Genome Modification by CRISPR/Cas9». *The FEBS Journal* 281 (23): 5186-93. <https://doi.org/10.1111/febs.13110>.
- Mardaryev, Andrei N., Bo Liu, Valentina Rapisarda, Krzysztof Poterlowicz, Igor Malashchuk, Jana Rudolf, Andrey A. Sharov, et al. 2016. «Cbx4 Maintains the Epithelial Lineage Identity and Cell Proliferation in the Developing Stratified Epithelium». *The Journal of Cell Biology* 212 (1): 77-89. <https://doi.org/10.1083/jcb.201506065>.
- Matic, Ivan, Joost Schimmel, Ivo A. Hendriks, Maria A. van Santen, Frans van de Rijke, Hans van Dam, Florian Gnad, Matthias Mann, and Alfred C. O. Vertegaal. 2010. «Site-Specific Identification of SUMO-2 Targets in Cells Reveals an Inverted SUMOylation Motif and a Hydrophobic Cluster SUMOylation Motif». *Molecular Cell* 39 (4): 641-52. <https://doi.org/10.1016/j.molcel.2010.07.026>.
- Merino, Anna, Jordi To-Figueras, and Carmen Herrero. 2006. «Atypical Red Cell Inclusions in Congenital Erythropoietic Porphyria». *British Journal of Haematology* 132 (2): 124. <https://doi.org/10.1111/j.1365-2141.2005.05726.x>.
- Merrill, Jacqueline C., Tiffany A. Melhuish, Michael H. Kagey, Shen-Hsi Yang, Andrew D. Sharrocks, and David Wotton. 2010. «A Role for Non-Covalent SUMO Interaction Motifs in Pc2/CBX4 E3 Activity». *PloS One* 5 (1): e8794. <https://doi.org/10.1371/journal.pone.0008794>.
- Minden, Mark D., Donna E. Hogge, Scott J. Weir, Jim Kasper, Debra A. Webster, Lavonne Patton, Yulia Jitkova, et al. 2014. «Oral Ciclopirox Olamine Displays Biological Activity in a Phase I Study in Patients with Advanced Hematologic Malignancies». *American Journal of Hematology* 89 (4): 363-68. <https://doi.org/10.1002/ajh.23640>.
- Misawa, Kiyoshi, Yuki Misawa, Atsushi Imai, Daiki Mochizuki, Shiori Endo, Masato Mima, Ryuji Ishikawa, Hideya Kawasaki, Takashi Yamatodani, and Takeharu Kanazawa. 2018. «Epigenetic Modification of SALL1 as a Novel Biomarker for the Prognosis of Early Stage Head and Neck Cancer». *Journal of Cancer* 9 (6): 941-49. <https://doi.org/10.7150/jca.23527>.

- Mohamed, Fedah E., Lihadh Al-Gazali, Fatma Al-Jasmi, and Bassam R. Ali. 2017. «Pharmaceutical Chaperones and Proteostasis Regulators in the Therapy of Lysosomal Storage Disorders: Current Perspective and Future Promises». *Frontiers in Pharmacology* 8: 448. <https://doi.org/10.3389/fphar.2017.00448>.
- Morey, Lluís, Luigi Aloia, Luca Cozzuto, Salvador Aznar Benitah, and Luciano Di Croce. 2013. «RYBP and Cbx7 Define Specific Biological Functions of Polycomb Complexes in Mouse Embryonic Stem Cells». *Cell Reports* 3 (1): 60-69. <https://doi.org/10.1016/j.celrep.2012.11.026>.
- Nakajo, S., K. Omata, T. Aiuchi, T. Shibayama, I. Okahashi, H. Ochiai, AND. Nakai, K. Nakaya, and AND. Nakamura. 1990. «Purification and Characterization of a Novel Brain-Specific 14-KDa Protein». *Journal of Neurochemistry* 55 (6): 2031-38. <https://doi.org/10.1111/j.1471-4159.1990.tb05792.x>.
- Netzer, Christian, Stefan K. Bohlander, Markus Hinzke, Ying Chen, and Jürgen Kohlhase. 2006. «Defining the Heterochromatin Localization and Repression Domains of SALL1». *Biochimica Et Biophysica Acta* 1762 (3): 386-91. <https://doi.org/10.1016/j.bbadis.2005.12.005>.
- Netzer, Christian, Stefan K. Bohlander, Leonie Rieger, Stefan Müller, and Jürgen Kohlhase. 2002. «Interaction of the Developmental Regulator SALL1 with UBE2I and SUMO-1». *Biochemical and Biophysical Research Communications* 296 (4): 870-76. [https://doi.org/10.1016/s0006-291x\(02\)02003-x](https://doi.org/10.1016/s0006-291x(02)02003-x).
- Ning, Bo, Wei Zhao, Chen Qian, Pinghua Liu, Qingtian Li, Wenyan Li, and Rong-Fu Wang. 2017. «USP26 Functions as a Negative Regulator of Cellular Reprogramming by Stabilising PRC1 Complex Components». *Nature Communications* 8 (1): 349. <https://doi.org/10.1038/s41467-017-00301-4>.
- Ohkuni, Kentaro, Nagesh Pasupala, Jennifer Peek, Grace Lauren Holloway, Gloria D. Sclar, Reuben Levy-Myers, Richard E. Baker, Munira A. Basrai, and Oliver Kerscher. 2018. «SUMO-Targeted Ubiquitin Ligases (STUbls) Reduce the Toxicity and Abnormal Transcriptional Activity Associated With a Mutant, Aggregation-Prone Fragment of Huntingtin». *Frontiers in Genetics* 9: 379. <https://doi.org/10.3389/fgene.2018.00379>.
- Osinalde, Nerea, Anna Duarri, Juanma Ramirez, Rosa Barrio, Guiomar Perez de Nanclares, and Ugo Mayor. 2019. «Impaired Proteostasis in Rare Neurological Diseases». *Seminars in Cell & Developmental Biology* 93: 164-77. <https://doi.org/10.1016/j.semcd.2018.10.007>.
- Phillips, John D., David P. Steensma, Michael A. Pulsipher, Gerald J. Spangrude, and James P. Kushner. 2007. «Congenital Erythropoietic Porphyria Due to a Mutation in GATA1: The First Trans-Acting Mutation Causative for a Human Porphyria». *Blood* 109 (6): 2618-21. <https://doi.org/10.1182/blood-2006-06-022848>.
- Pineault, Kyriel M., Ana Novoa, Anastasiia Lozovska, Deneen M. Wellik, and Moises Mallo. 2019. «Two CRISPR/Cas9-Mediated Methods for Targeting Complex Insertions, Deletions, or Replacements in Mouse». *MethodsX* 6: 2088-2100. <https://doi.org/10.1016/j.mex.2019.09.003>.
- Pirone, Lucia. 2016. «Post-translational modification by SUMO and Ubiquitin-Like proteins; Role of SUMOylation on SALL proteins». Thesis Doctoral.
- Pirone, Lucia, Wendy Xolalpa, Jón Otti Sigurðsson, Juanma Ramirez, Coralía Pérez, Monika González, Ainara Ruiz de Sabando, et al. 2017. «A Comprehensive

- Platform for the Analysis of Ubiquitin-like Protein Modifications Using in Vivo Biotinylation». *Scientific Reports* 7: 40756. <https://doi.org/10.1038/srep40756>.
- Pober, Barbara R. 2010. «Williams-Beuren Syndrome». *The New England Journal of Medicine* 362 (3): 239-52. <https://doi.org/10.1056/NEJMra0903074>.
- Powell, C. M., and R. C. Michaelis. 1999. «Townes-Brocks Syndrome». *Journal of Medical Genetics* 36 (2): 89-93.
- Powers, Evan T., Richard I. Morimoto, Andrew Dillin, Jeffery W. Kelly, and William E. Balch. 2009. «Biological and Chemical Approaches to Diseases of Proteostasis Deficiency». *Annual Review of Biochemistry* 78: 959-91. <https://doi.org/10.1146/annurev.biochem.052308.114844>.
- Ramanujam, Vaithamanithi-Mudumbai Sadagopa, and Karl Elmo Anderson. 2015. «Porphyria Diagnostics-Part 1: A Brief Overview of the Porphyrins». *Current Protocols in Human Genetics* 86 (julio): 17.20.1-17.20.26. <https://doi.org/10.1002/0471142905.hg1720s86>.
- Ren, Xiaoqing, Boqiang Hu, Moshi Song, Zhichao Ding, Yujiao Dang, Zunpeng Liu, Weiqi Zhang, et al. 2019. «Maintenance of Nucleolar Homeostasis by CBX4 Alleviates Senescence and Osteoarthritis». *Cell Reports* 26 (13): 3643-3656.e7. <https://doi.org/10.1016/j.celrep.2019.02.088>.
- Roux, Kyle J., Dae In Kim, Manfred Raida, and Brian Burke. 2012. «A Promiscuous Biotin Ligase Fusion Protein Identifies Proximal and Interacting Proteins in Mammalian Cells». *The Journal of Cell Biology* 196 (6): 801-10. <https://doi.org/10.1083/jcb.201112098>.
- Sánchez, Jonatan, Ana Talamillo, Monika González, Luis Sánchez-Pulido, Silvia Jiménez, Lucia Pirone, James D. Sutherland, and Rosa Barrio. 2011. «Drosophila Sal and Salr Are Transcriptional Repressors». *The Biochemical Journal* 438 (3): 437-45. <https://doi.org/10.1042/BJ20110229>.
- Sánchez, Jonatan, Ana Talamillo, Fernando Lopitz-Otsoa, Coralia Pérez, Roland Hjerpe, James D. Sutherland, Leire Herboso, Manuel S. Rodríguez, and Rosa Barrio. 2010. «Sumoylation Modulates the Activity of Spalt-like Proteins during Wing Development in Drosophila». *The Journal of Biological Chemistry* 285 (33): 25841-49. <https://doi.org/10.1074/jbc.M110.124024>.
- Sato, Akira, Shosei Kishida, Toshiya Tanaka, Akira Kikuchi, Tatsuhiko Kodama, Makoto Asashima, and Ryuichi Nishinakamura. 2004. «Sall1, a Causative Gene for Townes-Brocks Syndrome, Enhances the Canonical Wnt Signaling by Localizing to Heterochromatin». *Biochemical and Biophysical Research Communications* 319 (1): 103-13. <https://doi.org/10.1016/j.bbrc.2004.04.156>.
- Schnütgen, Frank, Nathalie Doerflinger, Cécile Calléja, Olivia Wendling, Pierre Chambon, and Norbert B. Ghyselinck. 2003. «A Directional Strategy for Monitoring Cre-Mediated Recombination at the Cellular Level in the Mouse». *Nature Biotechnology* 21 (5): 562-65. <https://doi.org/10.1038/nbt811>.
- Seeler, Jacob-Sebastian, and Anne Dejean. 2017. «SUMO and the Robustness of Cancer». *Nature Reviews. Cancer* 17 (3): 184-97. <https://doi.org/10.1038/nrc.2016.143>.
- Shen, Tao, and Shile Huang. 2016. «Repositioning the Old Fungicide Ciclopirox for New Medical Uses». *Current Pharmaceutical Design* 22 (28): 4443-50. <https://doi.org/10.2174/1381612822666160530151209>.

- Simon, Jeffrey A., and Robert E. Kingston. 2009. «Mechanisms of Polycomb Gene Silencing: Knowns and Unknowns». *Nature Reviews. Molecular Cell Biology* 10 (10): 697-708. <https://doi.org/10.1038/nrm2763>.
- Soria-Bretones, Isabel, Cristina Cepeda-García, Cintia Checa-Rodríguez, Vincent Heyer, Bernardo Reina-San-Martin, Evi Soutoglou, and Pablo Huertas. 2017. «DNA End Resection Requires Constitutive Sumoylation of CtIP by CBX4». *Nature Communications* 8 (1): 113. <https://doi.org/10.1038/s41467-017-00183-6>.
- Soshnikova, Natalia. 2014. «Hox Genes Regulation in Vertebrates». *Developmental Dynamics: An Official Publication of the American Association of Anatomists* 243 (1): 49-58. <https://doi.org/10.1002/dvdy.24014>.
- Swatek, Kirby N., and David Komander. 2016. «Ubiquitin Modifications». *Cell Research* 26 (4): 399-422. <https://doi.org/10.1038/cr.2016.39>.
- Sweetman, Dylan, Terry Smith, Elizabeth R. Farrell, Andrew Chantry, and Andrea Munsterberg. 2003. «The Conserved Glutamine-Rich Region of Chick Csal1 and Csal3 Mediates Protein Interactions with Other Spalt Family Members. Implications for Townes-Brocks Syndrome». *The Journal of Biological Chemistry* 278 (8): 6560-66. <https://doi.org/10.1074/jbc.M209066200>.
- Szymczak-Workman, Andrea L., Kate M. Vignali, and Dario A. A. Vignali. 2012. «Design and Construction of 2A Peptide-Linked Multicistronic Vectors». *Cold Spring Harbor Protocols* 2012 (2): 199-204. <https://doi.org/10.1101/pdb.ip067876>.
- Taipale, J., J. K. Chen, M. K. Cooper, B. Wang, R. K. Mann, L. Milenkovic, M. P. Scott, and P. A. Beachy. 2000. «Effects of Oncogenic Mutations in Smoothed and Patched Can Be Reversed by Cyclopamine». *Nature* 406 (6799): 1005-9. <https://doi.org/10.1038/35023008>.
- Talamillo, Ana, Orhi Barroso-Gomila, Immacolata Giordano, Leiore Ajuria, Marco Grillo, Ugo Mayor, and Rosa Barrio. 2020. «The Role of SUMOylation during Development». *Biochemical Society Transactions* 48 (2): 463-78. <https://doi.org/10.1042/BST20190390>.
- Tatham, Michael H., Marie-Claude Geoffroy, Linnan Shen, Anna Plechanovova, Neil Hattersley, Ellis G. Jaffray, Jorma J. Palvimo, and Ronald T. Hay. 2008. «RNF4 Is a Poly-SUMO-Specific E3 Ubiquitin Ligase Required for Arsenic-Induced PML Degradation». *Nature Cell Biology* 10 (5): 538-46. <https://doi.org/10.1038/ncb1716>.
- Thomas, Tynisha, Kieran Seay, Jian Hua Zheng, Cong Zhang, Christina Ochsenauber, John C. Kappes, and Harris Goldstein. 2016. «High-Throughput Humanized Mouse Models for Evaluation of HIV-1 Therapeutics and Pathogenesis». *Methods in Molecular Biology (Clifton, N.J.)* 1354: 221-35. [https://doi.org/10.1007/978-1-4939-3046-3\\_15](https://doi.org/10.1007/978-1-4939-3046-3_15).
- Urquiza, Pedro, Ana Laín, Arantza Sanz-Parra, Jorge Moreno, Ganeko Bernardo-Seisdedos, Pierre Dubus, Esperanza González, et al. 2018. «Repurposing Ciclopirox as a Pharmacological Chaperone in a Model of Congenital Erythropoietic Porphyria». *Science Translational Medicine* 10 (459). <https://doi.org/10.1126/scitranslmed.aat7467>.
- Uygun, David S., Zhiwen Ye, Anna AND. Zecharia, Edward C. Harding, Xiao Yu, Raquel Yustos, Alexei L. Vyssotski, Stephen G. Brickley, Nicholas P. Franks, and William Wisden. 2016. «Bottom-Up versus Top-Down Induction of Sleep by Zolpidem Acting on Histaminergic and Neocortex Neurons». *The Journal of Neuroscience*:

- The Official Journal of the Society for Neuroscience* 36 (44): 11171-84. <https://doi.org/10.1523/JNEUROSCI.3714-15.2016>.
- Walsh, Nicole C., Laurie L. Kenney, Sonal Jangalwe, Ken-Edwin Aryee, Dale L. Greiner, Michael A. Brehm, and Leonard D. Shultz. 2017. «Humanized Mouse Models of Clinical Disease». *Annual Review of Pathology* 12 (enero): 187-215. <https://doi.org/10.1146/annurev-pathol-052016-100332>.
- Wang, Xin, Liping Li, Yuanzhong Wu, Ruhua Zhang, Meifang Zhang, Dan Liao, Gang Wang, Ge Qin, Rui-Hua Xu, and Tiebang Kang. 2016. «CBX4 Suppresses Metastasis via Recruitment of HDAC3 to the Runx2 Promoter in Colorectal Carcinoma». *Cancer Research* 76 (24): 7277-89. <https://doi.org/10.1158/0008-5472.CAN-16-2100>.
- Wang, Xin, Ge Qin, Xiaoting Liang, Wen Wang, Zhuo Wang, Dan Liao, Li Zhong, et al. 2020. «Targeting the CK1 $\alpha$ /CBX4 Axis for Metastasis in Osteosarcoma». *Nature Communications* 11 (1): 1141. <https://doi.org/10.1038/s41467-020-14870-4>.
- Webb, Bryn D., Sanjeeva Metikala, Patricia G. Wheeler, Mingma D. Sherpa, Sander M. Houten, Marko E. Horb, and Eric E. Schadt. 2017. «Heterozygous Pathogenic Variant in DACT1 Causes an Autosomal-Dominant Syndrome with Features Overlapping Townes-Brocks Syndrome». *Human Mutation* 38 (4): 373-77. <https://doi.org/10.1002/humu.23171>.
- Wei, Fang, Hans R. Schöler, and Michael L. Atchison. 2007. «Sumoylation of Oct4 Enhances Its Stability, DNA Binding, and Transactivation». *The Journal of Biological Chemistry* 282 (29): 21551-60. <https://doi.org/10.1074/jbc.M611041200>.
- Weir, S. J., L. Patton, K. Castle, L. Rajewski, J. Kasper, and A. D. Schimmer. 2011. «The Repositioning of the Anti-Fungal Agent Ciclopirox Olamine as a Novel Therapeutic Agent for the Treatment of Haematologic Malignancy». *Journal of Clinical Pharmacy and Therapeutics* 36 (2): 128-34. <https://doi.org/10.1111/j.1365-2710.2010.01172.x>.
- Wotton, D., and J. C. Merrill. 2007. «Pc2 and SUMOylation». *Biochemical Society Transactions* 35 (Pt 6): 1401-4. <https://doi.org/10.1042/BST0351401>.
- Yang, Hui, Haoyi Wang, Chikdu S. Shivalila, Albert W. Cheng, Linyu Shi, and Rudolf Jaenisch. 2013. «One-Step Generation of Mice Carrying Reporter and Conditional Alleles by CRISPR/Cas-Mediated Genome Engineering». *Cell* 154 (6): 1370-79. <https://doi.org/10.1016/j.cell.2013.08.022>.
- Yang, Yanfang, Yu He, Xixi Wang, Ziwei Liang, Gu He, Peng Zhang, Hongxia Zhu, Ningzhi Xu, and Shufang Liang. 2017. «Protein SUMOylation Modification and Its Associations with Disease». *Open Biology* 7 (10). <https://doi.org/10.1098/rsob.170167>.
- Yasuda, Makiko, and Robert J. Desnick. 2019. «Murine Models of the Human Porphyrrias: Contributions toward Understanding Disease Pathogenesis and the Development of New Therapies». *Molecular Genetics and Metabolism* 128 (3): 332-41. <https://doi.org/10.1016/j.ymgme.2019.01.007>.
- Zhao, Jicheng, Min Wang, Luyuan Chang, Juan Yu, Aoqun Song, Cuifang Liu, Wenjun Huang, et al. 2020. «RYBP/YAF2-PRC1 Complexes and Histone H1-Dependent Chromatin Compaction Mediate Propagation of H2AK119ub1 during Cell Division». *Nature Cell Biology* 22 (4): 439-52. <https://doi.org/10.1038/s41556-020-0484-1>.











Ana Talamillo, Orhi Barroso-Gomila, **Immacolata Giordano**, Leiore Ajuria, Marco Grillo, Ugo Mayor, Rosa Barrio. *The Role of SUMOylation During Development*. Biochemical Society Transaction. 2020 Apr 29;48(2):463-478. doi: 10.1042/BST20190390.

#### **ABSTRACT**

During the development of multicellular organisms, transcriptional regulation plays an important role in the control of cell growth, differentiation and morphogenesis. SUMOylation is a reversible post-translational process involved in transcriptional regulation through the modification of transcription factors and through chromatin remodelling (either modifying chromatin remodelers or acting as a 'molecular glue' by promoting recruitment of chromatin regulators). SUMO modification results in changes in the activity, stability, interactions or localization of its substrates, which affects cellular processes such as cell cycle progression, DNA maintenance and repair or nucleocytoplasmic transport. This review focuses on the role of SUMO machinery and the modification of target proteins during embryonic development and organogenesis of animals, from invertebrates to mammals.

Laura Bozal-Basterra, Itziar Martín-Ruíz, Lucia Pirone, Yinwen Liang, Jón Otti Sigurðsson, Maria Gonzalez-Santamarta, **Immacolata Giordano**, Estibaliz Gabicagogeascoa, Angela de Luca, Jose A Rodríguez, Andrew O M Wilkie, Jürgen Kohlhase, Deborah Eastwood, Christopher Yale, Jesper V Olsen, Michael Rauchman, Kathryn V Anderson, James D Sutherland, Rosa Barrio. ***Truncated SALL1 Impedes Primary Cilia Function in Townes-Brocks Syndrome***. American Journal of Human Genetics. 2018 Feb 1; 102(2):249-265. doi: 10.1016/j.ajhg.2017.12.017.

## ABSTRACT

Townes-Brocks syndrome (TBS) is characterized by a spectrum of malformations in the digits, ears, and kidneys. These anomalies overlap those seen in a growing number of ciliopathies, which are genetic syndromes linked to defects in the formation or function of the primary cilia. TBS is caused by mutations in the gene encoding the transcriptional repressor SALL1 and is associated with the presence of a truncated protein that localizes to the cytoplasm. Here, we provide evidence that SALL1 mutations might cause TBS by means beyond its transcriptional capacity. By using proximity proteomics, we show that truncated SALL1 interacts with factors related to cilia function, including the negative regulators of ciliogenesis CCP110 and CEP97. This most likely contributes to more frequent cilia formation in TBS-derived fibroblasts, as well as in a CRISPR/Cas9-generated model cell line and in TBS-modeled mouse embryonic fibroblasts, than in wild-type controls. Furthermore, TBS-like cells show changes in cilia length and disassembly rates in combination with aberrant SHH signaling transduction. These findings support the hypothesis that aberrations in primary cilia and SHH signaling are contributing factors in TBS phenotypes, representing a paradigm shift in understanding TBS etiology. These results open possibilities for the treatment of TBS.

Giuseppina Roscigno, Ilaria Puoti, **Immacolata Giordano**, Elvira Donnarumma, Valentina Russo, Alessandra Affinito, Assunta Adamo, Cristina Quintavalle, Matilde Todaro, Maria dM Vivanco, Gerolama Condorelli. ***MiR-24 Induces Chemotherapy Resistance and Hypoxic Advantage in Breast Cancer***. *Oncotarget*. 2017 Mar 21;8(12):19507-19521. doi: 10.18632/oncotarget.14470.

## ABSTRACT

Breast cancer remains one of the leading causes of cancer mortality among women. It has been proved that the onset of cancer depends on a very small pool of tumor cells with a phenotype similar to that of normal adult stem cells. Cancer stem cells (CSC) possess self-renewal and multilineage differentiation potential as well as a robust ability to sustain tumorigenesis. Evidence suggests that CSCs contribute to chemotherapy resistance and to survival under hypoxic conditions. Interestingly, hypoxia in turn regulates self-renewal in CSCs and these effects may be primarily mediated by hypoxic inducible factors (HIFs). Recently, microRNAs (miRNAs) have emerged as critical players in the maintenance of pluripotency and self-renewal in normal and cancer stem cells. Here, we demonstrate that miR-24 is upregulated in breast CSCs and that its overexpression increases the number of mammospheres and the expression of stem cell markers. MiR-24 also induces apoptosis resistance through the regulation of BimL expression. Moreover, we identify a new miR-24 target, FIH1, which promotes HIF $\alpha$  degradation: miR-24 increases under hypoxic conditions, causing downregulation of FIH1 and upregulation of HIF1 $\alpha$ . In conclusion, miR-24 hampers chemotherapy-induced apoptosis in breast CSCs and increases cell resistance to hypoxic conditions through an FIH1-HIF $\alpha$  pathway.



# CURRICULUM VITAE







## Immacolata Giordano

Phone: +34 691924432

Mail: [imma.giordano@gmail.com](mailto:imma.giordano@gmail.com)

LinkedIn: <https://bit.ly/33eeld6>

### Education

- Master's degree in Medical Biotechnology

University of Naples, Italy – 2012/2015

Thesis title: "Analysis of microRNAs expression in breast cancer cell lines: Role of miR-24 in hypoxia survival and tamoxifen resistance".

- Bachelor's degree in Healthcare Biotechnology

University of Naples, Italy – 2008/2012

### Work experience

- 09/2016 - 02/2020 "CICbioGUNE" PARQUE TECNOLÓGICO DE BIZKAIA - 48160 Derio, Bizkaia Spain  
Functional genomic Unit - Ubiquitin-likes and Development Laboratory  
**PhD student** – Proteostasis of rare disease: Townes-Brocks Syndrome and Congenital Erythropoietic Porphyria.
- 02/2016 - 08/2016 "CICbioGUNE" PARQUE TECNOLÓGICO DE BIZKAIA - Derio, Bizkaia – Spain  
Functional genomic Unit - Ubiquitin-likes and Development Laboratory  
**Technician** - Post-translational modifications by SUMO and Ubiquitin-like proteins - Role of SUMOylation on SALL1.
- 07/2015 - 01/2016 "CICbioGUNE" PARQUE TECNOLÓGICO DE BIZKAIA - Derio, Bizkaia – Spain  
Animal facility  
**Technician** - Basic animal care and maintenance: housing, breeding, weaning, treatments, etc.

- 06/2015 “Centro Nacional de Investigaciones Cardiovasculares Carlos III” - Madrid – Spain  
Vascular Pathophysiology Unit  
**Internship** - Role of Lamin A/C in macrophages in the development of atherosclerosis.
- 09/2014 - 03/2015 "CNR Consiglio Nazionale delle ricerche" - Naples - Italy  
Cell biology Laboratory  
**Internship** - Analysis of microRNAs expression in breast cancer cell lines and their role in hypoxia survival and stemness phenotype.
- 03/2014 - 09/2014 "CICbioGUNE" PARQUE TECNOLÓGICO DE BIZKAIA - Derio, Bizkaia – Spain  
Cell biology and Stem Cells Unit  
**Internship** - Study of role of the microRNA-24 in several breast cancer cell lines: analysis of tamoxifen and hypoxia resistance.
- 11/2013 - 01/2014 "CNR Consiglio Nazionale delle ricerche" - Naples - Italy  
Cell biology Laboratory  
**Internship** - Analysis of microRNAs expression in breast cancer cell lines and their role in oncogenic pathways.
- 09/2012 - 10/2013 “IRCCS Istituto Nazionale Tumori G. Pascale” - Naples - Italy  
Cell biology and Molecular Diagnostic Laboratory  
**Internship** - Identification of processes involved in the pathogenesis and neoplastic progression: Analysis of molecular pathways associated with EGFR.
- 03/2012 - 07/2012 “Azienda Ospedaliera Universitaria Federico II”- Naples - Italy  
Human Anatomy Laboratory  
**Internship** - Isolation of cardiac progenitor cells and analysis of their behavior in ischemic microenvironment.

### **Mentoring Experience**

Mentoring graduate and undergraduate students

### Publications & Posters

- *The role of SUMOylation during development.*  
A. Talamillo, ..., **I. Giordano**, et al. (Biochemical Society Transaction. 2020 Apr 29;48(2):463-478.)
- *SALL1 regulates gene silencing through CBX4 stabilization.*  
**I. Giordano** et al. Writing and Editing the Ubiquitin Code (UbiCODE) 2020; Bilbao, 26-27 February 2020.
- *Generating models for the study of Congenital Erythropoietic Porphyria using CRISPR/Cas9.*  
**I. Giordano**, et al. International Congress on Porphyrins and Porphyrins (ICPP) 2019; Milan, 8-11 September 2019.
- *Truncated SALL1 Impedes Primary Cilia Function in Townes-Brocks Syndrome.*  
L. Bozal-Basterra, ..., **I. Giordano**, et al. (American Journal of Human Genetics 1;102(2):249-265.)
- *MiR-24 induces chemotherapy resistance and hypoxic advantage in breast cancer.*  
G. Roscigno†, I. Puoti1†, **I. Giordano**, et al. (Oncotarget 8(12):19507-19521.)

### Technical Skills

- **Cell biology:** Cells culture; Cells transfection; Immunofluorescence microscopy; Confocal microscopy; FACS analysis; qPCR; CRISPR/Cas9.
- **Molecular biology:** Cloning; Plasmid purification; Gel purification; Gene editing.
- **Proteomics:** Immunoprecipitation; Protein purification; Western blot.
- **Drosophila manipulation**
- **Animal care and handler certificate** (B and C).

### Language Skills

Italian: Mother tongue

English: B2

Spanish: C1



## Agradecimientos

Con esa tesis cierro un capítulo importante de mi vida y, por lo tanto, me veo obligada a reconocer a todas las personas que han sido protagonistas de esta historia. Así que aquí van mis gracias.

Quiero agradecer ante todo a Rosa por haberme dado la posibilidad de realizar mi doctorado y haber apostado por mí aun cuando parecía que las circunstancias no estaban de mi parte. Como ya te dije, durante todos esos años has sido un ejemplo para mí. Thank you so much Jim, for everything you have taught me over the years, for your sincerity and for sharing your musical and cultural discoveries with me. Gracias Cora, por ser la amatu de todos nosotros, los niños del laboratorio. Gracias Laura por tu paciencia y compañía. Gracias a todas mis pollitos, Eukene, Paula e Iria por haberme ayudado tanto, científicamente y sobretodo emotivamente. Gracias Orhi por haber sido un compañero durante el tiempo compartido dentro y fuera del labo. Gracias Vero, porque no es la cantidad sino la calidad del tiempo juntas lo que importa. Gracias Leiore por haber sido un soplo de aire fresco. Y gracias Marco, no tengo ni palabras para describir lo importante que ha sido tenerte a mi lado, mi regreso a la superficie ha sido también gracias a ti. Sois unas personas excepcionales y habéis sido un descubrimiento extraordinario. Y porque cualquier persona que pasa por nuestra vida nos influye en nuestro crecimiento, gracias a Ana, a María y a todas las personas que habéis sido, aunque a veces por poco tiempo, parte del RB lab.

Gracias labo por todos los momento geniales que hemos vivido juntos. Escribo eso y me vienen a la cabeza nuestra cueva secreta no secreta y los desayunos de la Denda. Unos momentos espontaneo que muchas veces he sentido como nuestra personal manera de “hacer vestuario”.

Gracias a Oscar Millet, Jokin Castilla and Paco Blaco por sus consejos y ayudas profesionales.

Gracias a las personas del comité de empresa, Ana, Encarni, Itzi, Onin, Miguel, Vidal y Cora. Ha sido una experiencia increíble de la que he aprendido muchísimo a todos los niveles. Gracias a Mada, por su impecable respeto y cuidado.

Unas gracias especiales a todos mis ratoncitos favoritos, Iker, Nahia, Peñi y también Bea, Virginia, Itzi y María. Y a los ratoncitos adoptados Vidal, Nagore y Pili. Habéis sido mi primera familia en bioGUNE y habéis seguido siendo cómplices y refugio durante los años siguientes. Nunca olvidaré los meses pasados con vosotros. Os debo mucho de lo que soy ahora, sobre todo porque sois los pilares de mi sabiduría popular española :D También quiero agradecer a Juan Rodríguez por su ayuda y profesionalidad.

Gracias a aquellas que con su presencia han hecho la diferencia. Gracias Julen, Héctor, Leti, Aize, Lorea, Alice, Cris, Camacho, Inés, Ianire, Ariane, Teresa, Amaya, Ane y Rob. Gracias Nuria, Ana y Egus por ser compañeras premurosas y esenciales.

También hay personas que durante estos años no solo han sido compañeras sino amigas y amarraderos en esta tormenta llamada doctorado. Gracias Amelie, por haber sido la primera en creer que podía hacerlo. Gracias Leticia por ser amiga y mentora, por cada palabra de nuestras larga conversaciones, por el tiempo que me has dedicado y por haber escuchado todas mis penas y compartidos mis victorias. Gracias Miguel por ser tan importante dentro y aún más fuera del trabajo. Gracias Anarose por nuestros momentos de apoyo mutuo y desahogos. Gracias Leire por hacerme sentir menos sola, por entenderme cuando parecía que nadie más lo hacía. Muchísimas gracias Ivana (Txeki), tu afecto y apoyo se escuchaban hasta en los silencios. Gracias Pedro por ayudarme y apoyarme, tenemos muchas milnos que recuperar.

Quiero agradecer a todas mis compañeras y compañeros del CIC porque de una manera u otra me habéis acompañado en este largo camino.

Gracias Giovanna por ser mi persona. No hubiera podido imaginar hacer eso sin ti a mi lado.

Gracias a mis amigas y compañeras Silvia, Carlos, Tini e Igor, a quien debo haber podido depositar. Nada hubiese sido lo mismo sin vosotras. Sois mi columna.

Gracias a mi familia que nunca ha dejado de creer en mí y que pese a la distancia siempre ha sido mi sustento. Grazie papà per avermi insegnato il rispetto, la dignità e l'umiltà. Grazie mamma, per essere esempio di invincibile resistenza. Grazie alle mie sorelle e mio fratello, per essere parte del mio essere.

Gracias Alf, mi compañero de vida, por no haberme dejado atrás, porque la columna avanza al ritmo del más lento. Gracias por haberme apoyado y por haberme enseñado a luchar. Gracias porque siempre has estado aquí.

Gracias a mis compañeras de colectivo, por sostener, acompañar y capacitar. Gracias por hacerme sentir parte de algo, un algo orgánico que sueña con una cierta idea del mundo.

Como escribía Mary Shelley en su obra maestra, solo aquellos que lo han experimentado pueden concebir los atractivos de la ciencia, la maravilla del descubrimiento. El entusiasmo y la pasión por la investigación existen. Salimos de la universidad creyendo en una ciencia pura que está por encima de cualquier ideología y que idolatramos casi como si fuese una religión. La vocación existe pero convive con la raza, la clase y el género. Discriminantes que desafortunadamente cimientan las relaciones desiguales de poder que demasiadas veces se dan en la investigación y que minan la salud y la integridad de las personas investigadoras. No existen las “personas problemáticas” o “conflictivas”, no es un problema personal, sino más bien estructural. Porque la investigación científica también es un engranaje clave de la producción de esa sociedad de hoy.

Esa es tan solo la constatación crítica de quien nunca ha olvidado la posición social de la que procede y nunca dejará de luchar porque la ciencia se ponga al servicio de la misma.

En una carta a sus hijos el Che escribía: “sobre todo sean siempre capaces de sentir en lo más hondo cualquier injusticia cometida contra cualquiera en cualquier parte del mundo. Es la cualidad más linda de un revolucionario.” Os deseo eso. Es mi más grande y valioso agradecimiento.

