Mémoire

Auteur : Mafrica, Cléa
Promoteur(s) : Dequiedt, Franck; Fettweis, Grégory
Faculté : Faculté des Sciences
Diplôme : Master en biochimie et biologie moléculaire et cellulaire, à finalité approfondie
Année académique : 2022-2023
URI/URL : http://hdl.handle.net/2268.2/18437

University of Liège – Faculty of Science
Life Sciences Department

Laboratory of Gene Expression and Cancer –
GIGA Molecular Biology of Diseases

Characterizing the Role of the Cohesin Loader NIPBL in Ewing Sarcoma

Cléa MAFRICA

Master's thesis presented in fulfilment of the requirements for the master's degree in Biochemistry, Molecular and Cellular Biology

Academic Supervisor: Pr. Franck Dequiedt

Co-Promotor: Dr. Gregory Fettweis

Academic year 2022-2023 – September 2023

Abstract

Ewing's sarcoma tumors are mainly driven by a single genetic mutation, the fusion of a FET protein and an ETS transcription factor, which in 85% of cases is the oncogenic fusion EWS-FLI1 (EF1). Emerging evidence demonstrated that the EWS-FLI1 fusion protein promotes some 3D structure change of the genome of Ewing sarcoma cells at multiple scales, including enhancer-promoter chromatin loops.

The cohesin complex has several roles in the cell cycle: cohesion of the sister chromatids, DNA damage repair and as well contribution to the 3D genome organization, which plays a role in gene regulation by bringing enhancers and promoters into close proximity. Interestingly, in 20% of cases, patients were identified to carry, a mutation in the gene *STAG2* coding for the STAG2 protein. STAG2 is a subunit of the cohesin complex, highlighting its relevance in the disease. The subject of this thesis focuses on this last essential role of the cohesin complex.

In a previous study, a member of the laboratory demonstrated that the glucocorticoid receptor, GR, can recruit NIPBL after glucocorticoid induction by direct physical interaction at the glucocorticoid element to promote enhancer-promoter loops.

The aim of this project is to find out whether the cohesin loader, NIPBL, directly interacts with EF1. In addition, if the interaction occurs, the dependence of NIPBL on EF1 to be recruited at EF1 target genes cis-regulatory regions will be investigated. Furthermore, the influence of long-range regulation by NIPBL on EF1 target gene expression after NIPBL knockdown (KD) will be evaluated.

Our results showed that EF1 c-ter domain (FLI1) physically interacts with the cohesin loader NIPBL by GST-pull down. Next, we found that EF1 is determinant for the recruitment of NIPBL at the cis-regulatory regions, such as the ones of the *PRKCB* gene, an EF1 target. We also highlighted that long-range gene regulation mediated by NIPBL can influence the regulation of EF1 target gene expression.

Furthermore, it appears that long-range gene regulation may cooperate or interfere with other enhancer-promoter communication depending on the localization of EF1 binding to cis-regulatory regions compared to the transcriptional start site (TSS) of EF1 target genes. Moreover, the phenotypes observed in NIPBL KD have affected both the EWS-FLI1low state by impairing cell migration and the EWS-FLI1high state by reducing the proliferation of Ewing sarcoma cells (as observed by Dr. Fettweis).

Our research may lead to a new model of transcriptional regulation dependent on the recruitment of NIPBL by oncogenic transcription factors. These results demonstrate the important role of NIPBL in Ewing sarcoma, leading to a potential new therapeutic target.

Mafrica, C. (September 2023*). Characterizing the Role of the Cohesin Loader NIPBL in Ewing Sarcoma* (master's thesis). Academic supervisor: Pr. Franck Dequiedt. Laboratory of Gene Expression and Cancer, GIGA-Molecular Biology of Diseases, ULiège.

Acknowledgments

Parfois, il est difficile de trouver les mots qui se bousculent dans notre tête. Je souhaite vous dire merci de façon sincère, mais les mots ne viennent pas. Merci, n'est composé que de cinq petites lettres. Mais, Il en faudrait infiniment plus pour pouvoir vous remercier à la hauteur de votre contribution pour ce mémoire.

Je tiens d'abord à remercier Pr. Dequiedt pour m'avoir accueilli au sein de son laboratoire, et de m'avoir proposé ce si beau sujet qui je l'espère, et j'y crois aura un bel avenir devant lui.

Merci à toi, Greg, de m'avoir fait découvrir le monde de la structure du génome et surtout pour ton guidage et ton aide, tout au long de ce projet du début jusqu'à l'extrême fin de cette aventure, même en heures supplémentaires.

Je tiens également à remercier mes nombreux collègues pour m'avoir accepté dans votre petite famille pleine de rires, de joies et de sérieux aussi ! Merci d'avoir été les petits soleils de mes journées et de m'avoir apporté votre précieux support pour le succès de ce travail.

Merci à Mimi, John, Loïc, Eva, Laurence, Benedetta, Florence, Christina, Madame Habraken, Anna-Maria, Alexia, Jeremy, Clément, Maximilien, Sharron, tous les membres du Laboratoire GIGA et nos collègues des Labos voisins.
Encore un merci, ils ne sont jamais de trop, à Loïc pour avoir eu la gentillesse de relire une partie de ce travail et d'avoir répondu à mes nombreuses questions.

Un remerciement tout spécial à mes parents, qui m'ont toujours soutenue et encouragée tout au long de mes études.
In fine, merci à mes deux compatriotes futures biologistes, Marie et Kiril, qui illuminent mes journées depuis que je les ai rencontrés.

Table of Contents

List of Abbreviations

3C, 4C, 5C, Hi-C	Chromosome conformation capture methods
3C	(One vs one)
4C	(One vs all)
5C	(Many vs many)
Hi-C	(High-throughput-C) (all vs all): high-throughput chromosome conformation capture
Micro-C	(Micro Capture-C)
-dox/ +dox	-/+ doxycycline
ABC	ATP Binding Cassette ATPase
AFM	Atomic force microscopy
ATCC	American Type Culture Collection
ATAD	Amino-terminal transactivation domain
ATP	Adenosine triphosphate
bp	base pair
BRD4	Bromodomain Containing 4
BRG1	BRG1 protein
BRM	Brahma protein
BSA	Bovine Serum Albumin
CBP	CREB-binding Protein
CdLs	Cornelia de Lange syndrome
cDNA	Complementary DNA
Co-IP	Co-Immunoprecipitation
Cohesin$^{STAG1/2}$	Cohesin complex with STAG1/2 sub-unit
C-ter	C-terminal domain
CREs	Cis-regulatory elements
CTAD	Carboxy-terminal transactivation domain
CTCF	CCCTC-binding factor
DMEM	Dulbecco's Modified Eagle's Medium
DNA	Desoxyribonucleic acid
DTT	Dithiothreitol (reducing agent to break down disulfide bridge)
EF1	EWS-FLI1; EWSR1-FLI1
EM	Electron Microscopy
E-P loops	Enhancer-Promoter loops
eRNA	enhancer RNA
ES	Ewing sarcoma
ESCO1/2	Establishment of sister chromatid cohesion N-acetyltransferase ½
ETS	TF protein family (Erythroblast Transformation Specific)
EWSR1/ EWS	Ewing sarcoma breakpoint region 1
EV	Empty vector refer to the presence of the 3X-FLAG-HALO transposon in our A673 cells
FBS	Fetal Bovine Serum
FET	RNA-binding protein family composed of the *FUS, EWSR1, TAF15* gene
GR	Glucocorticoid Receptor
GRE	Glucocorticoid Response Element
GST	Glutathione S-transferase
GST-FLI1full	GST Recombinant protein with EWS-FLI1 C-ter domain (FLI1)
GST-FLI1pot	GST Recombinant protein FLI1potential
GST-FLI1HC	GST Recombinant protein FLI1Highconfidence
HAT	Histone acetyltransferase
HEAT	Huntington, Elongation Factor 3, PR65/A, TOR domain
qPCR	quantitative Polymerase Chain Reaction
IDPs	Intrinsically Disordered Proteins
IDRs	Intrinsically Disordered Regions
kbs	kilobases

Mbs	Megabases
KD	Knockdown
LB	Lysogeny Broth (bacteria culture medium)
LLPS	Liquid-Liquid Phase Separation
LOF	Loss of function
LSD1	Lysine specific histone demethylase 1
MAU2	Sister Chromatid Cohesion 4 (SCC4) (Yeast)
Mb	Mega base
MCM	Minichromosome Maintenance protein Complex
mESCS	mouse Embryonic Stem Cells
MLL3/4	Lysine (K)-specific methyltransferase 2C
MSCs	Mesenchymal cells
NIPBL	Nipped B-like (Sister Chromatid Cohesion 2 (Scc2) (Yeast), Nipped-B (Drosophila), Delangin)
N-ter	N-terminal domain
NuRD	Nucleosome remodeling and deacetylase
ON	Over night
PBS	Phosphate-buffered saline
PD-1	Programmed cell death protein 1
PD-L1/PD-L2	PD-1 ligands 1/2
PDS5A/B	Sister chromatid cohesion protein PDS5 homolog A/B
PI	Protease Inhibitor
PNT	Pointed domain
qPCR	quantitative PCR
RAD21	SCC1/Mcd1 (Yeast)
RNA	Ribonucleic acid
RNAi	RNA interference
RNA polII	RNA polymerase II
RRM	RNA recognition motif
RT	Room temperature (~20-25°C)
RTqPCR	Reverse transcription quantitative PCR
rtTA	reverse tetracycline-controlled transactivator
SA: STAG1/STAG2	Stromal Antigen 1/Stromal Antigen 2
Scc1,2,3,4	Sister Chromatid Cohesion 1,2,3,4
shEF1	short hairpin RNA to silence EF1 gene expression via RNAi
SMC	Structural Maintenance of Chromosome (SMC1/SMC3)
SMT	Single-Molecular Tracking
SWI/SNF	Switch/Sucrose Non-Fermentable
siCtrl	siControl
siNIPBL	Small interfering against mRNA NIPBL
siRNA	Small interfering RNA
TADs	Topologically Associated Domains
TF	Transcription Factor
TRE	Tet Response Element
TSS	Transcriptional Start Site
YY1	Yin Yang 1
WAPL	Wings apart-like protein homolog
WB	Western blot

I. Introduction

1.1 The 3D conformation of the genome

1.1.1 The 3D genome

The holder of hereditary information is the DNA. It carries genetic information necessary for the development and functioning of an organism. Human DNA molecules stretched end to end are around 2 meters in length for ~3 billion base pairs (bp). In eukaryotic cells during the interphase, DNA is confined within the nucleus, which range from 5 to 20 µm in diameter in mammalian cells [1]. The challenge behind the success of packaging DNA into the cell nucleus is not solely to enable its storage, but also to allow DNA condensation to prepare mitosis or meiosis, DNA damage repair, DNA replication, and gene expression in terms of tissue and cell type.

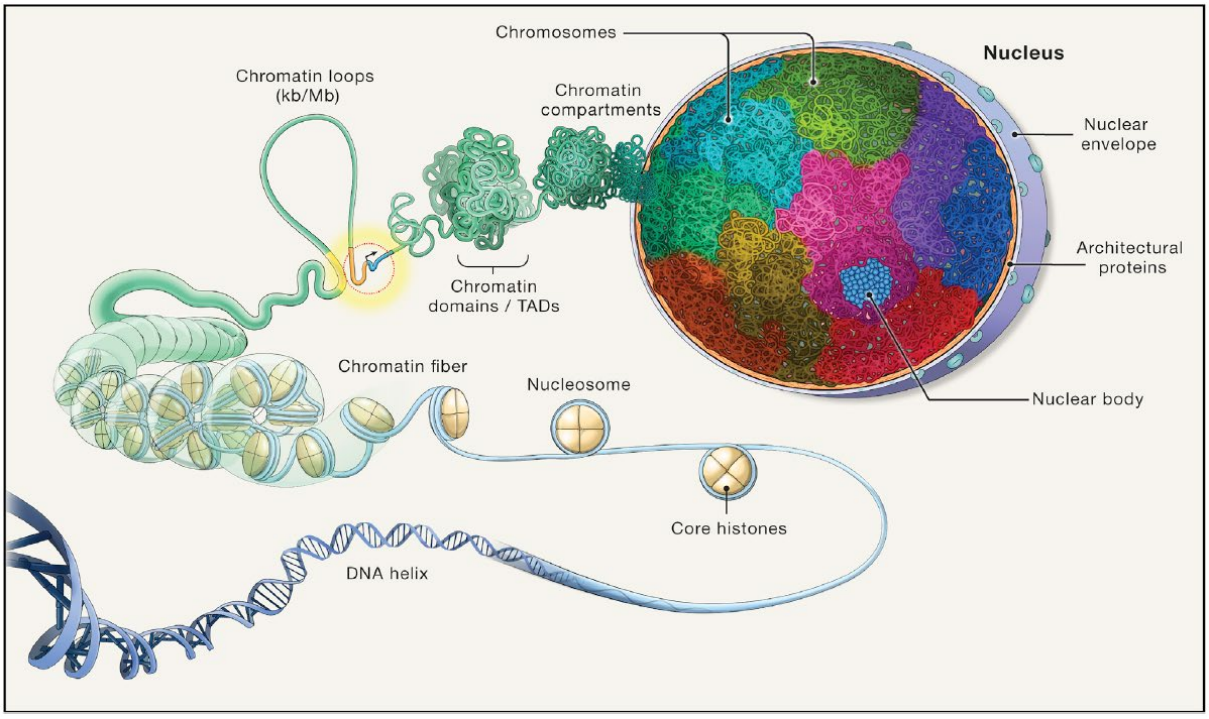


Figure 1. The 3D organization of the eukaryotic genome. The double helix of DNA is wrapped around octamers of core-histones called nucleosomes, to form the chromatin fiber. The fiber can fold to form loops that promote the proximity of regulatory elements such as enhancers (yellow) to the associated promoter (orange) to control gene transcription. Topologically associating domains results in the folding of fiber, interacting with each other more frequently than other surrounding regions. TADS or chromatin domains associate to form chromatin compartments. Chromosome result in the association of chromatin compartments, each chromosome occupying a distinct region defined as a chromosome territory (From [2]).

The genome is organized and condensed at multiple scales in the 3D space [2] (**Fig.1**). The 3D architecture of eukaryotes genomes is defined by conserved architectural features: chromatin fiber, chromatin loops, topologically associating domains (TADs), chromatin compartments and chromosome territories.

Firstly, DNA of 146 base pairs is wrapped around core histones proteins, a histone octamer, to form nucleosomes. Nucleosomes are connected by linkers DNA, are then arranged to constitute the chromatin fiber with a diameter of 5-24 nm [3], [4].

Then, loops initiate self-interaction of chromatin fibers. The range size of the loops is between kilobases (kbs) and megabases (Mbs) [2]. Loops provide spatial close proximity of two distant cis-regions of one chromosome [2], [5]. During interphase, loops are mediated by the loop extrusion mechanism, which involves the passage of DNA through a ring-shaped protein complex called the cohesin complex. Cohesin is an ATP-dependent molecular motor that drives DNA extrusion until reach of two convergent CCCTC-binding factor (CTCF) bound sites [6]–[9], this will be further detailed in section 1.2. CTCFs are the boundaries of topologically associating domains [10]. The TADs confine sequences that interact preferentially with each other rather than with other surrounding sequences present in another TADs, thus defining TADs as structural and functional boundaries [11]. The next level of organization is due to the association of individual chromatin domains into chromatin compartments (12). Genome structure is related to the genome function, which explains the distinction into two compartments, named A and B. Compartment A refers to transcriptionally active euchromatin regions and important gene richness. In contrast, compartment B corresponds to highly condensed heterochromatin defined by transcriptionally inactive regions. Chromatin compartments can then assemble to form the highest organization of the genome, the chromosomes. This ultimate level of organization corresponds to the chromosome territories [12]. Between neighboring chromosomes, chromatin fibers can intermingle without intertwining [2], [12].

1.1.2 The cohesin complex

The cohesin complex has been known firstly to establish the cohesion of replicated DNA from S-phase. Through its role in cohesion, cohesin promotes the correct bi-orientation of sister chromatids on the mitotic or meiotic spindle to allow the correct segregation of sister chromatids between the two sister cells [13].

Nowadays, it is well established that cohesin complexes affect long-range chromatin interaction in trans for sister-chromatids cohesion [14]. Moreover, the cohesin complex has multiple roles in interphase. The cohesin mediates cis-chromatin interaction to allow immunoglobulin class switch recombination [15] and to shape the 3D structure of the genome by forming TADs and loops [16]. Loops influence gene transcription by promoting spatial proximity of two distant regions as enhancer and its associated promoter. [16]. The function of the cohesin complex in the 3D structure of the genome, and in particular its involvement in the formation of enhancer-promoter loops to promote gene regulation, will be described in detail in this work.

1.1.2.1 Structure and composition

The cohesin complex belongs to the SMC family (Structural Maintenance of Chromosome). In vertebrate somatic cells, the cohesin complex is a ring-shaped tetramers composed of a dimer of SMC ATPase, the α-kleisin subunit RAD21 (Scc1/Mcd1 in yeast or the REC8 variants in meiotic cells) and a stable associated HEAT repeat protein STAG1 or STAG2 (the Scc3 in yeast) which interact with the α-kleisin, and the SMC subunits. The SMC subunits SMC1/SMC3 possess an ABC ATPase head domain that binds two ATP molecules, a ~50 nm long coiled coils arm and a hinge domain by which they heterodimerize. On the other end, the amino- (N-ter) and carboxy-terminal (C-ter) domain of the α-kleisin subunit RAD21 interacts with the two ATPase head domains of the SMC to close off the cohesin ring. Cohesin is also associated with accessory proteins: the cohesin loader NIPBL(Scc2) and its stabilizer MAU2(Scc4) or the two release factors WAPL and PDS5A/PDS5B. The binding of kleisin with one of the two accessory HEAT repeats proteins (HAWK): NIPBL or PDS5A/PDS5B is mutually exclusive [17], [18] (**Fig2**).

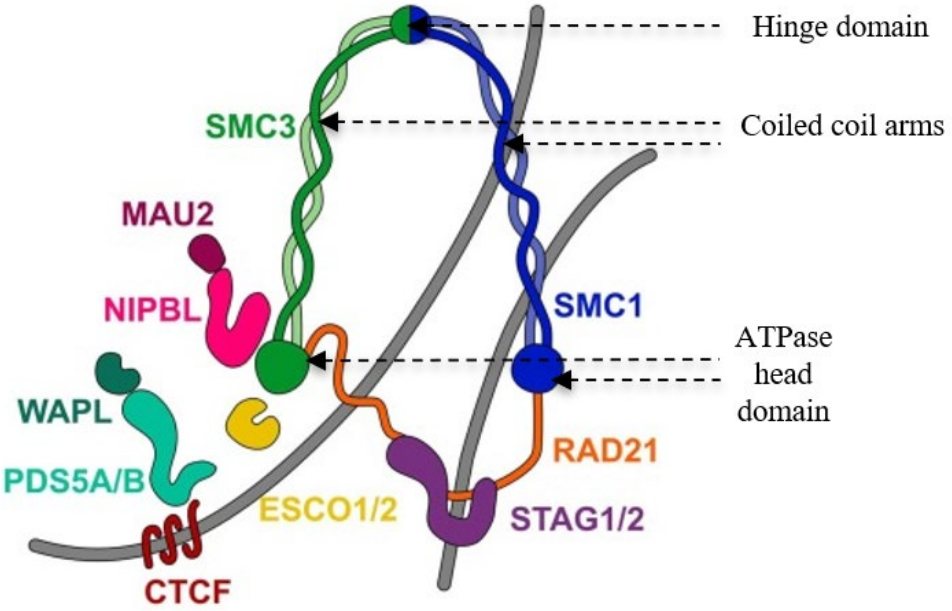


Figure 2. Schematic representation of the cohesin structure in open ring conformation state. The cohesin complex is composed of four subunits: the structural maintenance of chromosomes proteins, SMC1 and SMC3, the kleisin, RAD21, and the kleisin-associating, STAG1 or STAG2. SMC are composed of a long antiparallel coiled coil arm, a hinge domain by which they dimerize, and an ABC ATPase head domain. Accessory proteins are associated with cohesin as the cohesin loader NIPBL and its stabilizer MAU2, the unloading factor WAPL and PDS5 A/B, ESCO1/2, and CTCF (Adapted from Loop extrusion rules: the next generation https://doi.org/10.1016/j.gde.2023.102061).

1.1.2.2 Cohesin mediates loop extrusion

During the G1 phase, the monomeric form of cohesin-ring entraps DNA to promote the loop-extrusion process by its ATP-dependent motor activity mediated by NIPBL [7], [19]. After the cohesin loading on chromatin by NIPBL (the loading function of cohesin by NIPBL is currently reassessed, as discussed in the review of Alonso-Gil and Losada 2023 [20]), DNA is extruded directionally by cohesin at a rate of 0.5-1 kb/sec. The loop extrusion phenomenon ceases when the cohesin is released by WAPL/PDS5α, or when the cohesin complex reaches convergent CTCF binding sites corresponding to the loop anchors [7], [14], [21], [22].

The acetylation or deacetylation of the SMC3 subunits, catalyzed respectively by ESCO1/2 and HDAC8, influence the stability of cohesin on chromatin by extending the length of chromatin loops. PDS5A/B binding at SMC3 following acetylation promotes loop extrusion interruption and cohesin release by PDS5-WAPL by competing with NIPBL for the same binding sites. On the other hand, PDS5 has a positive impact on cohesin residence time by countering WAPL activity in cohesin^{STAG1} stabilization due to SMC3 acetylation by ESCO1/2 at CTCF sites and other DNA loops anchors [23], [24].

During interphase, the loop extrusion phenomenon leads to the formation of spatial proximity between distant chromatin loci to form loops such as enhancer-promoter loops or promoter-promoter loops, and also to the formation of TADs by increasing chromatin contact within a domain [25].

1.1.2.3 Mechanisms of loop extrusion

Cohesin extrudes DNA at a rate of up to 0,5-1 kb/s [7], [19]. Electron microscopy (EM) and atomic force microscopy (AFM) have identified three conformations of the cohesin complex: the SMC open ring conformation (defined by the separation of the two coiled coils), the rod conformation (related to the alignment of the coiled coils) and the bent conformation (described by the folding of the coiled coils to allow the close proximity between the hinge and the ATPase head domains of the SMCs) [25], [26]. Many conformational change theories have been proposed to tent to explain DNA extrusion mediated by the cohesin complex, one of the most convincing theories being the “swing and clamp” model [26]. This model explains that in the absence of ATP, the cohesin complex adopts the nucleotide-free apo state conformation. This state consists of the association of the cohesin/NIPBL nose (the hallow of the U-shape of NIPBL) with the hinge of the SMCs. Thermal motion leads to the alignment of the SMCs, and their bending towards the SMCs heads, which also provokes DNA translocation to the SMC3 heads. Simultaneously, ATP binding to the SMCs' ATPase head domains causes their assembly. NIPBL will be equally disengaged from the SMC hinge to be linked at SMC1/3 ATPase head domains transferring equally DNA to the heads. The disassembly of this structure, the clamp structure, and the disengagement of the ATPase domains are induced by the ATP hydrolysis triggered by DNA binding to the heads. ATP

hydrolysis generates DNA translocation and brings cohesin/NIPBL back into the nucleotide-free apo-state conformation to provoke translocation of another DNA segment [20], [26]. The relevance of STAG in mediating loop extrusion is still debated and could be explained by another theory, the Brownian ratchet model, which will not be discussed here [19], [27].

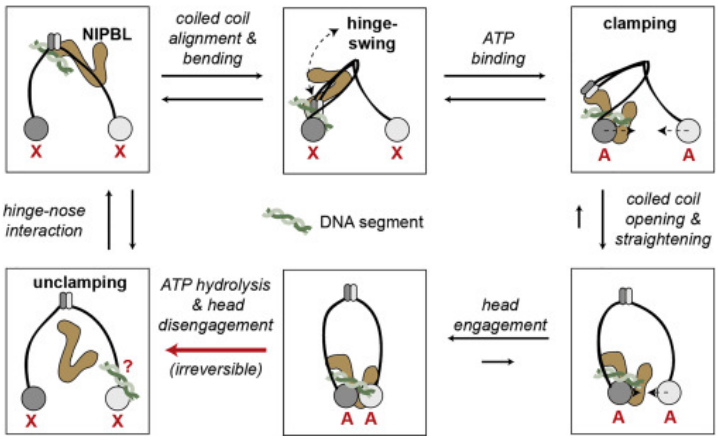


Figure 3. Schematic representation of the 'swing and clamp' mechanism for mediating DNA translocation. In the open conformation of cohesin, the nucleotide-free ("X") state, NIBPL interacts with the SMC hinge. Spontaneously, the coils align and bend to bring the SMC hinge closer to their heads, transporting NIPBL from the hinge to the SMC3 head domain. The binding of ATP causes engagement of the SMC head domains, as well as dissociation of the NIPBL nose of the SMC hinge, resulting in the stretching of the coils. Engagement of the head domains triggers ATP hydrolysis, leading to their dissociation. At the end of a cycle, DNA might be transferred to the head domain of SMC1. The nucleotide-free state allows again the association with a new DNA segment to restart the cycle (according to [41]).

1.1.2.4 STAG1/STAG2 cohesin influences chromatin association dynamics

In human somatic cells, two paralog genes of Scc3 encode for two variants of the STAG subunit (SA): STAG1 and STAG2 [28]. STAG1 and STAG2 present 75% homology in the core region and differ by their N- and C-ter domains. Regarding spatial distribution, both SA subunits are present on CTCF binding sites. Nonetheless, STAG2 is equally found at non-CTCF binding sites as enhancer regions [45].

Competitive interaction between WAPL and CTCF for cohesin influence loops' length and cohesin stability on chromatin, depending on the sensibility of STAG variants for them. Cohesin^{STAG1} display long-lived loops due to longer-residence time on chromatin conferred by ESCO1 and CTCF [24]. Preferential sensitivity of STAG2 for WAPL mediates shorter loops [30], [31].

STAG1 and STAG2 have distinct functions in the 3D organization of the genome. Cohesin^{STAG2} is involved in tissue-specific transcription by enhancer-promoter loops, while cohesin^{STAG1} is required for the establishment of TADS. In addition, STAG1 depletion also impairs transcription of gene closer to the TADs boundary [31], [32].

1.1.2.5 The interplay between transcription and the loop extrusion mediated by the cohesin complex

1.1.2.5.1 Enhancers are distal regulatory regions

Eukaryotic genomes are identical between all somatic cells in an organism although the transcriptomic signature varies widely. Spatiotemporal gene expression patterns differ depending on cell differentiation, cell type, and cell response to environmental change [33]. Cis-regulatory elements (CREs) such as promoters, enhancers, silencers, or insulators are DNA sequences able to modulate gene expression to respond to intrinsic and external signals (Preissl, Gaulton, and Ren 2023). Enhancers are distal regulatory elements that stimulate gene expression regardless of their orientation or genomic distance, within certain limits, from their target promoter or position (upstream or downstream target genes) relative to their associated promoter [34]. Enhancers are located in intergenic regions, introns, or exonic regions of unrelated genes [35] and measure few hundred bp in length and serve as binding sites for transcription factors [36]. The binding of pioneer transcription factors may recruit methyltransferases such as MLL to mediate H3K4me1 and H3K4me3. Histone tails undergo other post-translational modifications as the lysine acetylation of histone 3 or 4 by histone acetyltransferase (HAT). CBP/p300 catalyzes the acetylation of H3K27, which promotes the neutralization of the positive charge of the lysine, leading to a decrease in DNA affinity. Acetyl-lysins are recognized by other HATs to re-enforce the signal. They are also identified by TF or other bromodomain proteins such as the SWI/SNF complex to promote chromatin remodeling[36]. In addition, active enhancers are associated with the presence of histone variants H3.3 and H2A.Z, which are responsible respectively for maintaining high levels of H3K27ac or reducing nucleosome stability to facilitate the binding of TFs and the transcriptional machinery [37]–[39]. In conclusion, the histone variants H3.3 and H2A.Z and post-translational modifications are required for enhancer activation, which is determined by an open and active chromatin state[38]. Then, transcription factors may recruit other co-activator proteins such as the Mediator complex which role is to coordinate activation signal between enhancer and RNA polymerase II (RNA PolII) at the promoter site [36]. RNA PolII is equally requested at the enhancer site for the bidirectional transcription of short-lived enhancer RNA (eRNA) [34]. Enhancers stimulate many steps of transcription, including initiation and elongation by directly communicating with target promoters [40].

1.1.2.5.2 Enhancer-promoter communication

Various models have been proposed to tempt to explain enhancer-promoter communication, which is essential owing to the large genomic distance separating the enhancer of its cognate promoter. Transcriptional regulation depends therefore on long-range chromatin interactions to physically link enhancers and promoters, which in most cases are separated by around 50 kb in vertebrates, but the

distance can exceed 1 Mb [34], [41], [42]. Enhancer-promoter communication could be described by several models as tracking (or “scanning”), “linking” (or “chaining”), “conformation” and enhancer-promoter looping [40], [43] (**Fig.4**). The scanning model is based on the consideration that the enhancer acts as a bidirectional entry site for RNA PolII, which then "scans" the DNA until the finding of transcriptional start site (TSS) at promoter sites. The linking models proposed the idea that proteins scaffold could mediate the communication between enhancers and promoters. In addition, “conformation” models suggest that, like the propagation of action potential along the axon, the open state of active enhancers could transmit along the chromatin fiber up to the promoter. Furthermore, the loops theory originated from the 3D vision of the genome. The contact frequency of DNA sequences will be more frequent inside topological associating domain than outside. Loops might be formed within TADs to facilitate close spatial proximity of two genetically distant loci [40]. TADs might create spatial delimitation by the CTCF boundary to isolate related enhancers and promoters within the same structural unit, which underlie spatial and functional isolation mediated by CTCF [44], [45]. There is yet no universal theoretical method governing enhancer/promoter communication. Several methods might be involved, depending on enhancer and the context. [40].

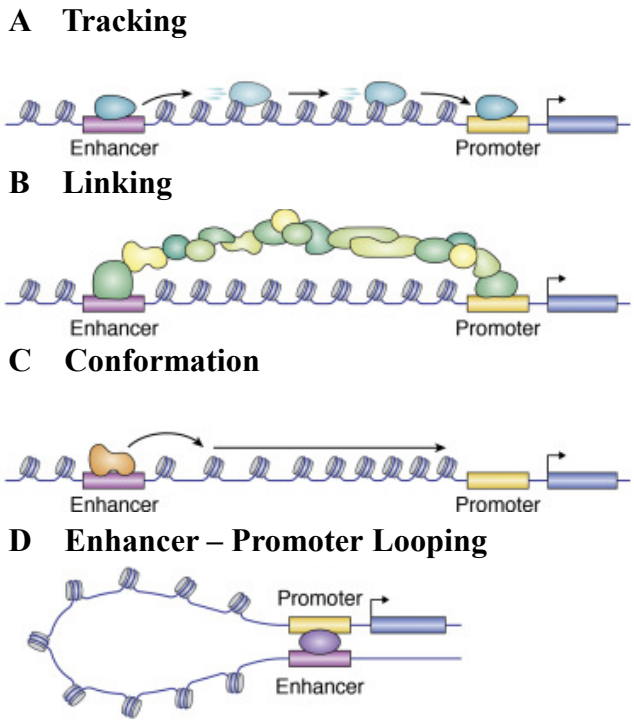


Figure 4. Enhancer-promotor communication models: A. Tracking, B. Linking, C. Conformation, D. Enhancer-Promoter Looping [40]

1.1.2.5.3 Cohesin implication to mediate E-P loops by loops extrusion

3D conformational capture experiments such as 3C, 4C, 5C, Hi-C and Micro-C have demonstrated that the loops extrusion-mediated by the cohesin complex play a key role in enhancer-promoter communication [35], [36]. The role of cohesin-CTCF in gene expression is a result of direct physical contacts between the enhancer and the related promoter [46], but as well indirectly by the increase of contact frequency within the TADs to create enhancer-promoter loops [47]. Nonetheless, acute depletion of CTCF or cohesin (by auxin induce degron system) did not display major gene expression defect, loss of compartment domains or histone mark modifications despite the significant loss of loops [48], [49]. From the transcription point of view, solely neighboring genes of superenhancers have shown a strong downregulation following acute cohesin or CTCF depletion [49]. In addition, the acute depletion (3H) of cohesin, CTCF, WAPL or YY1, which dimerize to form loops, in mESCS showed only a weak E-P loss [48].

Nevertheless the presence of NIPBL, the cohesin loader, at active enhancer/promoter sites rather than at CTCF sites strengthen the hypothesis that the cohesin complex might be involved in long-range gene regulation via enhancer/promoter loops [32], [46]. Furthermore, NIPBL not solely co-occupy Mediator sites at active promoters [46], [50], but also NIPBL and the Med23 subunits of the Mediator complex appear to physically interact. This phenomenon links the chromatin structure, mediated by the NIPBL/cohesin complex through the formation of loops between enhancer-bound transcription factors and core promoters, to gene expression via mediator interaction with RNA PolII, thereby controlling the initiation and elongation phases of transcription [46], [51]. In addition, NIPBL seems to physically interact with the epigenetic reader BRD4, which is involved in the detection of lysine acetylation at chromatin accessible sites of active enhancers [52]. In neutrophils, calcium induction has been shown to be involved in differential recruitment of NIPBL to enhancer or promoter active sites by the chromatin remodelers BRG1, BRM of the SWI/SNF complex. Degradation of the co-activators p300, CBP by the molecular degron inducible system leads to the decreased occupancy of NIPBL at enhancer, promoter sites following the reduction of H3K27 acetylation [53]. An additional study reveals that NIPBL is recruited to the Glucocorticoid response element (GRE) enhancer by the glucocorticoid receptor (GR) transcription factor following glucocorticoid induction. This interaction leads to the enrichment of the cohesin complex via NIPBL at GRE to mediate loops between GRE and GR target gene promoters [54]. Furthermore, Cornelia de Lange syndrome is mostly caused by heterozygous splice site, frameshift, missense mutations or deletions in frame of the *NIPBL* gene, which varies in cell-type population, characterizing NIPBL mutations as mosaic in CdLs [55]. This multisystem disorder did not present an aneuploidy phenotype, but revealed dysregulation of developmental gene expression associated with the role of NIPBL in DNA damage repair, transcriptional activity, cohesin loading and mediator of DNA extrusion through the cohesin complex [56]–[58], [19]. These several studies highlighted the role of the

recruitment of NIPBL by transcription factor, chromatin remodeler, the Mediator complex or co-activator at enhancer sites to further recruit the cohesin complex to mediate enhancer-promoter loops to promote gene regulation.

Recent studies suggested that NIPBL is not loaded at active promoters, in spite of that NIPBL-cohesin could be enriched at TSS by the RNA polymerase [59]–[61] which plays a “barrier role” to lead to cohesin enrichment at, for example, active promoters [61]. The connection between the cohesin-loop extrusion and the transcription is thought to depend on a "moving barrier mechanism". According to this hypothesis, RNA PolII could be a dynamic extrusion barrier that mediates the translocation of cohesin and the relocalization of the complex from the promoter to the 3’ end of transcribed genes [59]–[61]. Like the RNA PolII, the MCM complex has also been characterised as an extrusion barrier [62].

Several studies highlighted the minimal effect on overall steady-state gene transcription following acute depletion of cohesin and CTCF, however solely the basal transcription levels are analyzed[49]. Furthermore, cohesin appears to be involved in long-range gene regulation, as indicated by the preferential enrichment of NIPBL cohesin/STAG2 within enhancer-promoter regions by various transcriptional effectors [46], [52], [54], [55], [58]–[64]. The importance of cohesin in mediating the E-P interaction directly or indirectly might also be related to the distance of the enhancer from the associated promoters, the formation of E-P loops may be favored by cohesin for enhancers that are more than 100kb distant away from their target promoters [44], [65].

1.1.2.6 Chromosome conformation capture strategies

The development of chromosome conformation capture (3C, 4C, and 5C) and the related chromatin conformation capture Hi-C (dependent on restriction enzyme digestion) or Micro-C (based on micrococcal nuclease digestion) methods enabled is the spearhead associated with topologically associating domains (TADs) and loops. These techniques rely on the chemical crosslinked by paraformaldehyde of distant genomic sequences brought into close proximity thanks to protein complexes such as the cohesin complex. Crosslinking sequences are digested and then ligated for further analysis by qPCR or next-generation sequencing approach which allows the construction of a contact probability map of the genomic DNA [4], [5].

When the cohesin subunit is depleted, the peaks (refer as loops) and stripes (corresponding to interaction of the same anchor with multiple other sequences) in the Hi-C or Micro-C contact map disappear, indicating that the interaction of distant sequences as enhancer-promoter depends on the motor activity of cohesin to mediate DNA extrusion [66].

1.2 Ewing sarcoma

1.2.1 Ewing sarcoma and the EWS/FLI1 (FET/ETS) fusion

In the 1920s, James Ewing described a tumor found in a 14-year-old girl that was initially diagnosed as an osteosarcoma tumor [67]. However, the morphological features of the cells - undifferentiated round cells - and the sensitivity of the tumor to radiotherapy suggested that it was not osteosarcoma [67]. Ewing therefore defined the tumor as a "diffuse endothelioma of bone" [67]. Ewing sarcoma (ES) is nowadays described as a highly aggressive bone and soft tissue tumor that occurs predominantly in children and young adults [68], [69].

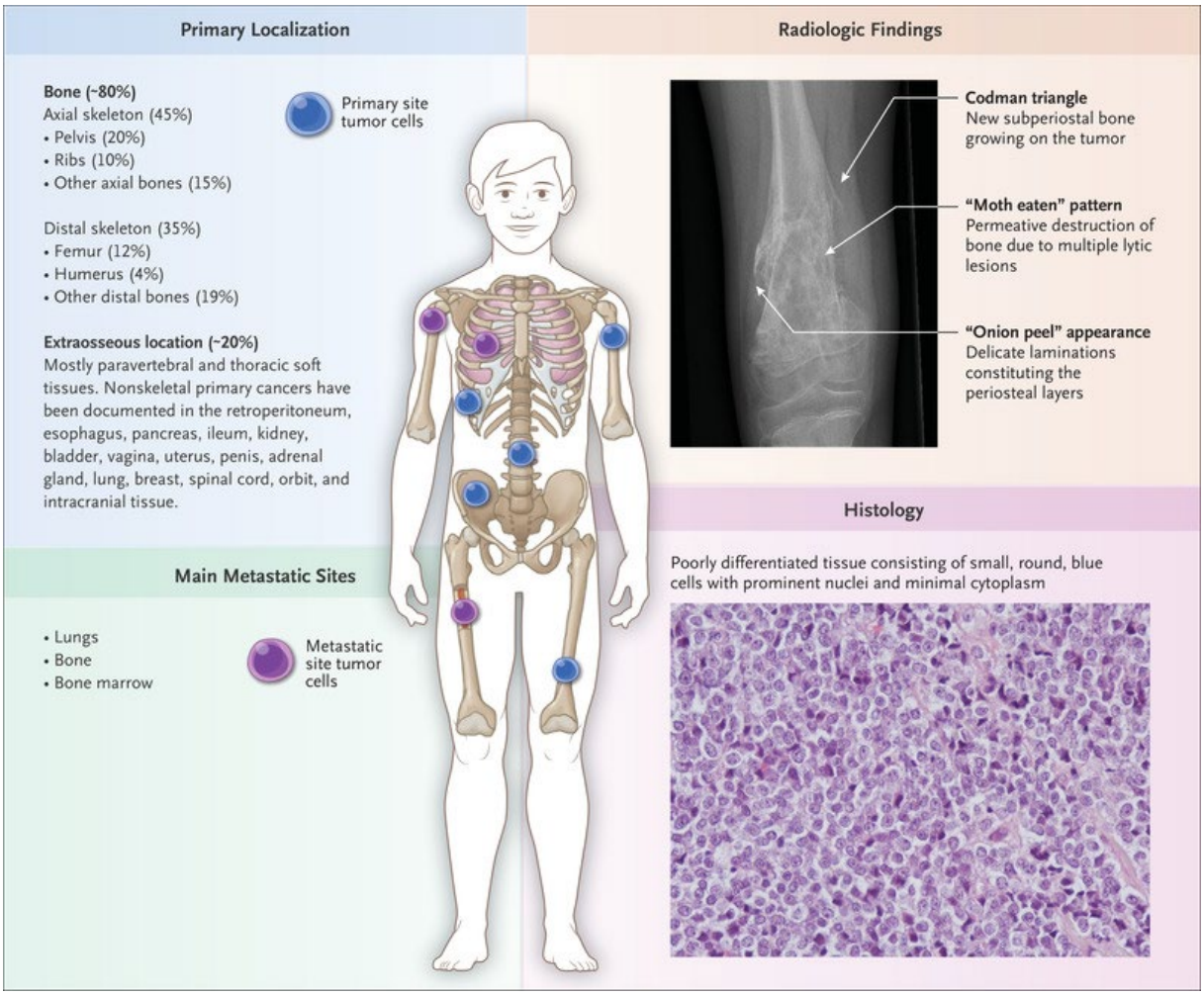


Figure 5. Ewing Sarcoma: a highly aggressive bone and soft tissue pediatric tumor. In the illustration, are represented primary and secondary sites where the tumor develops, histologic *characteristic* features and the most common metastatic sites. [68]

In 80% of cases, the primary localization is bones (pelvis, ribs, or other axial bones and femur, humerus, or other distal bones) [68]. In 20% of cases, ES occurs in extraosseous soft tissues (mostly paravertebral and thoracic)[68]. Ewing sarcoma is a malignant tumor, which metastatic sites could be lungs, bone

marrow, bone, lymph nodes, liver and brain [68]. The incidence rate is 1 case per 1.5 million population with a predilection for males (male to female incidence ratio of Ewing's sarcoma is 1.6/1) and European descendants individuals [68], [69]. ES has a peak of incidence at the age of 15 years [68], [69]. Survival rates for 5 years are greater than 70% for patients with local disease amenable to multimodal therapy. However, the presence of metastases leads to worst outcome with a survival rate of less than 30% for 5 years [68], [69].

Seventy years after Ewing sarcoma's first description, the most common chromosomal rearrangement, t(11;22)(q24;q12) at the root of most ES tumor was discovered [70]. The chromosomal translocation gives rise to an oncogenic transcription factor leading to a fusion between genes of the FET and ETS families [70], [71]. The FET family of RNA-binding proteins, FUS, EWSR1, TAF15, is characterized by a low complexity domain (LCD) containing the SYGQ-rich transactivation domain in N-ter, a RNA recognition motif (RRM), two or three RGG domains, a zinc finger (ZnF) domain, and a nuclear transactivation signal (NLS) domain in C-ter [68], [72]. The ETS (E26-transformed specific) family of transcription factor contain a winged-helix-loop-helix DNA binding domain [73]. In N-ter, ETS are composed of a transactivation domain (ATAD : Amino-terminal transactivation domain) and a PNT domain. The c-ter domain is composed of the ETS-DNA binding domain and the transactivation domain (called CTAD : Carboxy-terminal transactivation domain). The ETS-DNA binding motif of the oncogenic FET/ETS transcription factor corresponds to the 5'-GGAA/T-3' DNA sequence [74].

In ~85% of cases, ES is genetically characterize by the chromosomal rearrangement, t(11,22)(q24,q12), of *EWSR1* and *FLI1* genes, resulting in the sequence encoding the neomorphic transcription factor EWS-FLI1. In other cases the fusion is formed by the fusion of EWS with other members of the ETS family, which the most frequent is the EWSR1-ERG fusion (approximatively 10 %) [69]. Nonetheless, FUS and TAF15 could also formed a fusion protein with ETS proteins [75].

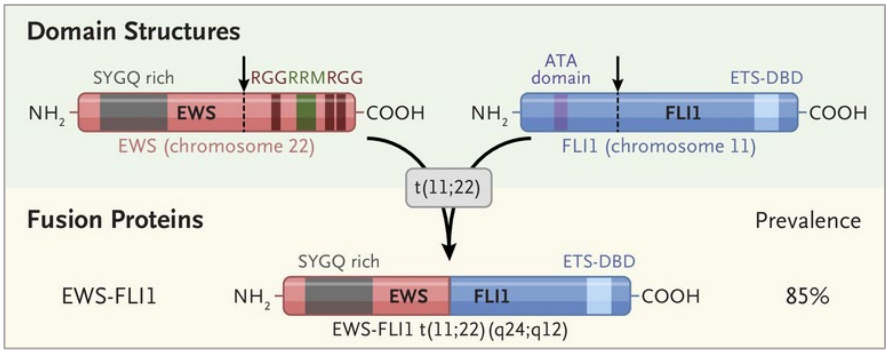


Figure 6. Chromosomal translocation t(11;22)(q24;q12) leading to the most common FET-ETS fusion, the oncogenic transcription factor EWS-FLI1 in Ewing sarcoma [68]. The domains annotations are more condensed compared to the explanations of ETS and FET fusions domains in the text.

Ewing sarcoma possess a single cellular lineage origin, possibly arising from bone-marrow derived mesenchymal cells (MSCs) [76], neural-crest-derived stem cells [77] or epithelial origin [78]. Besides the clonal origin of ES, this tumor is constituted by cells subpopulations with distinct phenotypes. As in most pediatric cancers, ES transcriptional heterogeneity is mainly due to epigenetic modification, in contrast to adult cancers for which somatic mutations are the main driver of inter- and intra-tumor heterogeneity [81]. EWS-FLI1 transcriptional activity and its expression levels is not constant within the tumor, which enables EWS-FLI1 to promote tumor progression and metastasis in addition to its initial role of promoting tumorigenesis [80]. EWS-FLI1 protein levels vary between individual tumor cells, leading to two different outcomes. Decease EWS-FLI1 concentration is leading to shift from cell-cell to cell-matrix interaction by cytoskeleton and adherence change. Low levels of EWS-FLI1 are associated with a strong capacity to migrate, invade, and metastasize. Meanwhile, high levels of EWS-FLI1 are associated with strong cell proliferation [79], [80].

Albeit genetic mutations are rare in ES, additional mutations were identified at diagnosis as the somatic mutations in the *STAG2* gene (will be discussed in section 1.3) or in the *TP53* gene or in *CDKN2A* genes respectively in 15-21%, 5-7%, and 10-22% of cases [69]. CDKN2A and STAG2 mutations are mutually exclusive, unlike STAG2 and TP53, whose association may increase genetic instability leading to aggressive subtypes of ES tumors [81].

In terms of clinical treatments, FET-ETS fusion proteins is difficult to target with the actual technologies due to the absence of enzymatic activity and their prion-like structure [68]. Currently primary Ewing's sarcoma tumor are treated with multimodal therapy including chemotherapy with cytotoxic drugs, and local therapy as radiotherapy or surgery [82]. However, these treatments are not without risks, as they can lead to long-term outcomes such as some disabilities like infertility, cardiac toxicity or other cancers like acute myelogenous leukemia, radiation-induced sarcomas and malignant neoplasms [83]. Several research groups are working on other targets as for example: effector molecules of oncogenic FET-ETS fusion, or molecules directly interacting with the FET-ETS fusion, … [68]

1.2.2 EWS-FLI1 remodeling the chromatin to activate EF1 target genes enhancer or repress ETS target genes enhancer

EWS/FLI1 binds to isolate ETS-binding sites and microsatellite repeats of the sequence 5'-GGAA-3', which conduct respectively to gene repression and oncogenic gene transcription. The silencing of ETS target genes depends on the binding of single EWS-FLI1 at 5'-GGAA-3' ETS-DNA binding motif which lead to the recruitment of the nucleosome remodeling and deacetylase (NuRD) complex, histone deacetylase, lysine specific histone demethylase 1 (LSD1) to reduce DNA accessibility. In addition, the binding of EWS-FLI1 at this site avoid the binding of ETS transcription factor, and therefore the recruitment of coactivator as p300 leading to chromatin inactivation state and target ETS gene repression. In addition, at 5'-GGAA-3' microsatellites, EWS-FLI1 is considered as a pioneer

transcription factor, which is due to its IDR allowing EWS-FLI1 multimerization. The binding of EF1 at repeat 5'-GGAA-3' allows the initiation of open chromatin state thanks to the recruitment of the ATP-dependent chromatin remodeling complex BAF (or SWI/SNF2), the histone acetyl transferase p300 and the mixed-lineage leukemia methyltransferase (MLL).

In conclusion, EWS-FLI1 is able to induce de novo enhancer at GGAA repeat element by the recruitment of chromatin modifier and histone-modifying complexes. EWS-FLI1 also promotes the chromatin repression of ETS target genes enhancer by preventing ETS transcription factor binding and the recruitment of chromatin modifier and coactivators. The repression at GGAA site is also induced by the recruitment of HDAC and the NuRD complex [74].

1.3 The cohesin complex implication in Ewing sarcoma

Genes encoding proteins of the cohesin complex or associated cohesin proteins are frequently mutated in a wide range of cancers, such as the *STAG2* gene in bladder urothelial carcinoma, glioblastoma, melanoma, acute myeloid leukemia, and Ewing's sarcoma [69]. The *STAG2* gene is located on the X-chromosome, which results in the absence of allelic compensation in male and female by X-inactivation if the gene is mutated. *STAG2* inactivation is mostly the consequence of somatic mutation as a truncating mutation (e.g nonsense, frameshift, or splice site mutation) or a missense mutation, which usually results in loss of STAG2 function (LOF)[84].

As previously mentioned, in Ewing sarcoma, *STAG2* mutations occur in 15-20% of Ewing sarcoma cases [79]. STAG2 loss of function (LOF) has been primarily associated with aneuploidy. However, several studies demonstrated that *STAG2* mutation is not significantly associated with impaired chromosome segregation due to cohesion dysfunction [84]. Two studies showed that STAG2 LOF in Ewing sarcoma are related to the cohesin role in gene regulation [96], [97].

As highlighted in clinical results, the presence of *STAG2* mutation in the genome of ES patients is associated with poor prognosis and metastasis [81]. Both research are concurring that STAG2 LOF increases metastasis occurrences. These two studies together demonstrated different transcriptome alterations leading to poor outcome [96], [97]. After STAG2 LOF in Ewing sarcoma cells, the level of its paralog, STAG1, increased. Nonetheless STAG1 did not demonstrate sufficient association with activated enhancers (H3K23ac marks)[1]. One study reveals that H3K27ac marks, EF1 binding, CTCF and cohesin sites are not altered by *STAG2* KO [85]. The other study suggests nonetheless a decrease in EF1 binding at H3K27ac promoter or enhancer regions following STAG2 KO [86]. Both studies agree

[1] As previously mentioned, cohesin^{STAG2} and cohesin^{STAG2} binding colocalize at CTCF binding sites, solely cohesin^{STAG2} is associated with active enhancer.

on the alteration of EWS/FLI1 anchored long-range chromatin contacts after STAG2 LOF, which affects the transcription of some EF1 target genes [96], [97]. In addition, STAG2 KO leads to transcriptional changes in neurodevelopmental programs [86]. In conclusion, transcriptional changes directly connected to EF1 E-P loops or independent of EF1, following STAG2 mutation, result in a migration and invasive phenotype, similar to the EF1low state phenotype [96], [97].

Additionally, the oncogenic transcription factor itself reprograms the 3D genome in terms of A/B compartments, TADS and E-P loops. Precisely, EF1 can engender a B-A compartment shifting, but rarely an A-B shifting. In addition, EF1 binding at *de novo* enhancers, GGAA microsatellites, promotes looping to other EF1 binding sites resulting in upregulation in gene expression. The colocalization of the cohesin subunit near gained loops anchors in EF1 WT/EF1KD suggests the recruitment or stalling of the cohesin at EF1 binding sites to mediate EF1 3D chromatin reprogramming [87]. The cohesin^{STAG2} may be the linker of the cohesin mediation to reprogram 3D structure as E-P loops, nonetheless it might not be the only one involved.

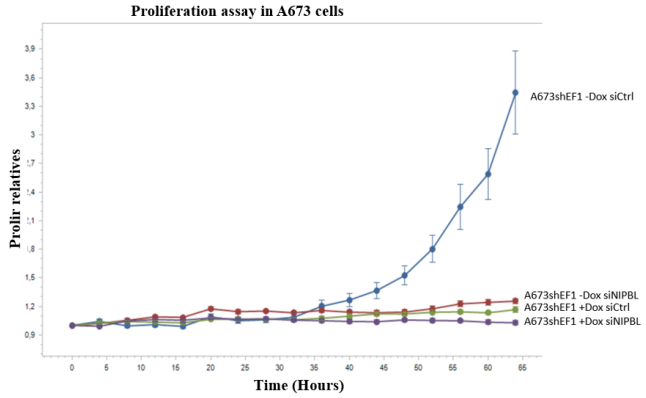
II. Objectives

Recently, the oncogenic transcription factor EWS-FLI1 (EF1) has been shown to promote a change in the 3D structure of the genome, including loop enhancer-promoter loops [88]. So far, two other studies have associated this role with STAG2, a subunit of the cohesin complex [85], [86]. Concomitantly, a study driven by a member of our laboratory has shown that the transcription factor, glucocorticoid receptor (GR), recruits the cohesin loader, NIPBL, by direct physical interaction at glucocorticoid response elements (GREs). Subsequently, NIPBL recruits the cohesin complex and activates the ATP-motor activity of the cohesin complex to promote the DNA extrusion through the cohesin complex [19], [26]. The loop extrusion will contribute to the spatial proximity between the GRE, the cis-regulatory regions of the GR target genes and the associated promoter [54]. Dr Fettweis, who carried out this pioneering study, is now collaborating with Professor Dequiedt team who is studying the role of EWS-FLI1 or other FET fusions at different levels of molecular biology. Their aim is to determine whether the EWS-FLI1 transcription factor could follow the same pattern as GR to regulate the transcription of its genes and to potentially find if this mechanism might be targetable by a small molecule in this cancer that remains mostly without a solid and targeted therapeutic strategy.

2.1 Preliminary results

To explore this potential role, Dr. Fettweis previously performed cell proliferation assays on A673 treated by an siRNA anti-NIPBL (NIPBL KD cells) in combination or not with the activation of a EF1 shRNA (induced by doxycycline) (**Fig.7A.**). He was able to show that while, A673 shEF1cells treated with siCtrl and no doxycycline treatment, show an exponential growth, A673 shEF1 -dox NIPBL knockdown exhibited low proliferation rate over time as did EWS-FLI1 knockdown for A673 shEF1 +dox siCtrl. In addition, the proliferation rate of A673 shEF1 +dox siNIPBL cells remained constant, indicating that as many cells were dividing as were dying. The proliferation rate of A673 cells is similar after NIPBL or EWS-FLI1 knockdown. The down expression of NIPBL causes A673 cells to decrease in proliferation.

A.



B.

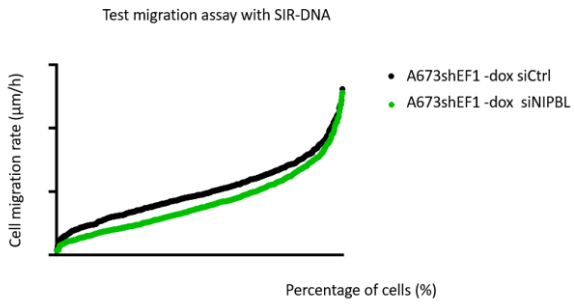


Figure 7. A. Proliferation assay in A673 cells - A673 cells were transfected with siCTR or siNIPBL for 4 days (3 independent experiments). After nuclei staining with SIR-DNA for 2 hours (500 nM), data were acquired on an automatic Incucyte SX5 microscope for 64 hours with an image taken every 4 hours for an exposure of 400 ms per frame. Data were generated by nuclei counting using the inmate Incucyte image analysis software.

7. B. Tracking avec SIR-DNA - A673 cells were transfected with siCTR or siNIPBL for 4 days (3 independent experiments). After nuclei staining with SIR-DNA for 2 hours (500 nM), data were acquired on an automatic Incucyte SX5 microscope for 48 hours with an image taken every 20 minutes for an exposure of 400 ms per frame. Images analysis was performed tracking nuclei movement using the plugin TrackMate of the Fiji (ImageJ) software. Data are displayed as the speed cumulative distribution function for both conditions.

[2] In the A673shEF1 -dox possess a native expression of EF1 because the expression of shEF1 is not induced in lack of doxycycline.

Finally, Proximity Ligation assay (PLA) (**Fig.8**), realized by the Dr.Fettweis, in U2OS cells revealed physical proximity between NIPBL and EWS-FLI1 in vivo. Which suggests a potential direct role of NIPBL in EF1 target genes expression through its function in long-range gene regulation thanks to the loop extrusion mechanism.

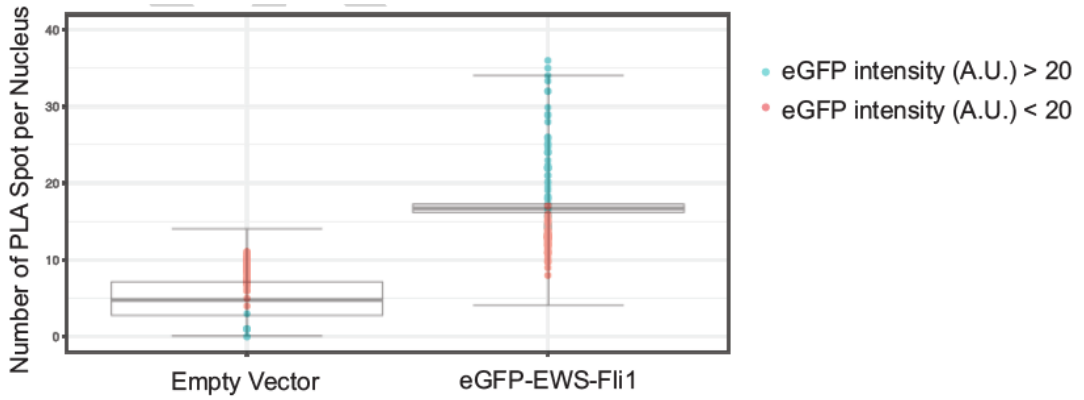


Figure 8. Interaction between eGFP-EWS-FL1I and NIPBL was performed using a Proximity Ligation Assay.

After cell transfection with eGFP empty vector and eGFP -EWS-FLI1 expressing plasmids, cell was fixed with 4% PFA, and the Duo-link (Sigma) protocol followed. Data were acquired on a CellVoyager 8000 automatic microscope and foci (highlighting interactions) count achieved on the Colombus Imaging platform. Data analysis was then finalized with a custom R script. Data represent 2 independent experiments. eGFP-EWS-FLI1 expressing cells were selected through the total eGFP intensity of each nucleus with a threshold set at 20 (A.U.) and are showing more interaction signal than the untransfected cells.

In conclusion, these preliminary results strongly suggest that the downregulation of the cohesin loader, NIPBL, impairs Ewing sarcoma cells proliferation, and migration properties. They also highlight a potential direct interaction between EF1 and NIPBL.

2.2 These Master thesis objectives

In the course of this work, we would like to investigate:

- The direct interaction between EWS-FLI1 and NIPBL by GST-pull assay. The interaction would be tested between exogenous FLAG-HALO-NIPBL and three constructions of the EWS-FLI1 c-terminal domain (FLI1).
- Determine if the KD of NIPBL through the loss of long-range gene regulation could influence EF1 gene regulation.
- Investigate if EF1 is determinant to recruit NIPBL at cis-regulatory regions of its target genes to influence gene regulation.

Altogether, the discovery of a direct physical interaction of NIPBL and EF1, as well as its implication in transcriptional regulation of EF1 target genes, could lead to potential new therapeutic strategy. Especially since, NIPBL down regulation impair both cell proliferation and cell migration.

Furthermore, the confirmation of this mechanism of transcriptional regulation where transcription factor could recruit the cohesin by physical interaction to promote gene regulation, could be a new principle of long-range gene regulation used by different family of transcription factors.

III. Materials and Methods

3.1 Plasmid insertion in host bacteria and colony selection by PCR reaction

3.1.1 Transformation

3.1.1.1 Transformation by heat shock

Escherichia coli DH5α chemically competent cells conserved at -80°C were thawed on ice before the addition of 5 µL of the ligation product to 30 µL of DH5α E. coli bacteria. Then, samples were incubated on ice for 30 min. The vector was inserted in bacteria by heat at 42°C for 30 sec. Transformed bacteria were then incubated 2 min on ice, then 200 µL of warm LB was added. Samples were incubated for 1 hour at 37°C on constant agitation before plating on LB medium plate supplemented with the appropriate selection agent (ampicillin (1:1000) or spectinomycin (1:500)). Plates were finally incubated overnight (ON) at 37°C.

3.1.1.2 Transformation by electroporation

DH5α electroporation competent cells conserved at -80°C were thawed on ice. 50 µL of bacteria and 3 µL of dialysis ligation product were gently mixed and inserted into an electroporation cuvette. Electroporation was realized with a pulse of 1,8 kV during 4 sec. Transformed bacteria were retrieved with LB, and were centrifuged at 4000 rpm for 5 min. After the supernatant was discarded, the pellet was resuspended with 200 µL of LB. Samples were incubated for 1 h at 37°C on constant agitation before plating on LB medium plate supplemented with the appropriate antibiotic (ampicillin (1:1000) or spectinomycin (1:500)). Plates were finally incubated overnight (ON) at 37°C.

3.1.2 Colony selection

Colony PCR is a screening method to determine the presence of the insert DNA in the plasmid of bacterial colony. Single colonies were selected with a sterile tip, set at the bottom of the PCR tubes then incubated in a mix of PCR reaction (**Table 1,2**). Following clone selection by colony PCR and electrophoresis on 2% agarose gel, positive clones were maintained in LB at 37°C with constant agitation and were then selected and resuspended in 5 ml of LB supplemented with 1/1000 of the adequate antibiotic. Liquid cultures were incubated ON at 37°C with constant shaking. The next day, 500 µL of the liquid culture was mixed with 500 µL of 100% glycerol. The cultures were stored at -

80C°. The 4500 µL of liquid culture was used to purify the plasmid using the NucleoSpin® Plasmid EasyPure kit (Macherey-Nagel). The plasmid was sequenced to verify the presence of the insert.

Mix PCR
0,25 µL - GoTaq polymerase
5 µL - GoTaq Buffer 5X
1µL - dNTP 10mM
0,0625 µL - Primer F 100 µM
0,0625 µL - Primer R 100 µM
H_2O

Table 1. Colony PCR mixture to amplify the interest sequence (Primers list are in the annex)

	PCR cycle	
Pre-denaturation	10 min at 95°C	
Denaturation	1 min at 95°C	35X
Annealing	45 sec at _°C	
Extension	_ sec à 72°C	
Final extension	10 min at 72°C	
	∞ at 4°	

Table 2. Colony PCR cycles to amplify the sequence of interest (Tm depends on the annealing temperature of the appropriate primers and the extension time rely on the length of the amplified sequence ~1min/kb)

3.2 Generation of Pbac-3XFLAG-HALO-EV-PuroR or Pbac-3XFLAG-HALO-msNIPBL-WT-PuroR vectors by restriction, ligation to create A673 shEF1 EV/NIPBL stable cell lines

The Pbac-3XFLAG-HALO-EV and Pbac-3XFLAG-HALO-msNIPBL-WT vectors, which give rise to the stable A673 ATCC EV/NIPBL cell lines, are maintained in the genomes of these cells thanks to their resistance to blasticidin. However, the transposon that contain shEF1 sequence is also maintained in the A673 shEF1 genomes by blasticidin antibiotic pressure. In order to create A673 shEF1 (short-hairpin RNA against EF1 mRNA) EV/NIPBL cell lines, the resistance of the Pbac-3XFLAG-HALO-EV-BafR and Pbac-3XFLAG-HALO-msNIPBL-WT-BafR vectors was modified for puromycin resistance.

3.2.1 Nested Polymerase chain reaction (Nested PCR)

Nested PCR allowed amplification of the puromycin sequence and an additional sequence containing Apa1, KspA1 and Mfe1 restriction sites. Three forward primers and one reverse primer allowed amplification of this sequence from the GVV1898#7 plasmid. (**Table 3,4**).

1st PCR	2nd PCR	3rd PCR
0,5 µL - polymerase HF	0,5 µL - polymerase HF	0,5 µL - polymerase HF
10 µL - HF Buffer 5X	10 µL - HF Buffer 5X	10 µL - HF Buffer 5X
1 µL - dNTP 10mM	1 µL - dNTP 10mM	1 µL - dNTP 10mM
2,5 µL - Primer F 10µM: DF4966	2,5 µL - Primer F 10µM: DF4967	2,5 µL - Primer F 10µM: DF4968
2,5 µL - Primer R 10µM: DF4969	2,5 µL - Primer R 10µM: DF4969	2,5 µL - Primer R 10µM: DF4969
1,5 µL - DMSO	1,5 µL - DMSO	1,5 µL - DMSO
50 ng - Plasmid GVV1898#7	5 µL -1st PCR product	5 µL -2nd PCR product
H_2O	H_2O	H_2O

Table 3. Nested PCRs mix to amplify sequences containing Hpa1, Mfe1 restriction sites and puromycin resistance sequence

	1ST AND 2ND PCR		3RD PCR	
PRE-DENATURATION	45 sec at 98°C		45 sec at 98°C	
DENATURATION	10 sec at 98°C	15X	10 sec at 98°C	35X
ANNEALING	30 sec at 55°C		30 sec at 55°C	
EXTENSION	45 sec à 72°C		45 sec at 72°C	
FINAL EXTENSION			5 min at 72°	

Table 4. Nested PCRs cycle conditions to amplify sequences containing Hpa1, Mfe1 restriction sites and the puromycin resistance sequence

3.2.2 Electrophoresis on 2% agarose gel and purification of bands of interest

PCR products were loaded in a 2% agarose gel in TAE 1X buffer. DNA amplification was revealed by Midori green DNA Stain 1000x. DNA has migrated in an electrophoresis chamber at ~80-100 volts. Bands of interest were extracted from the gel and purified using the Gel and PCR Clean-Up Nucleospin kit.

3.2.3 Digestion of puromycin sequence and Pbac-3XFLAG-HALO-EV, BlastRPbac-3XFLAG-HALO-msNIPBL-WT-BlastR vectors and ligation of digested products

The purified puromycin resistance sequence and the 3XFLAG-HALO-EV-BafR plasmid or the purified puromycin resistance sequence and 3XFLAG-HALO-msNIPBL-WT-BafR plasmid were digested respectively with the Apa1 and KspA1 or Mfe1 and KspA1 restriction enzymes at 37°C for 2 hours (**Table 5**). After the 3XFLAG-HALO-EV-BafR and 3XFLAG-HALO-msNIPBL-WT-BafR digest vectors were purified with the Gel and PCR Clean-Up NucleoSpin™ kit (Macherey-Nagel™). The ligation was performed with 100 ng of 3XFLAG-HALO-EV-BafR and 3XFLAG-HALO-msNIPBL-WT-BafR vector, 35,91 ng or 15,66 ng of the puromycin sequence (digested puromycin sequence concentration to add for the ligation with EV or NIPBL vectors depend on the length in bp of the host vectors), 1 µL of Buffer T4 DNA ligase (10X), 0,5 µL of T4 DNA ligase (New England Biolab T4 DNA ligase), 1 µL of ATP 10 mM and H_2O. Samples were incubated ON at 16°C. Ligation reactions were stopped by samples melting at 65°C for 20 min. Then, 5 µl of ligation products were added to 30 µL of chemically-competent DH5α or electrocompetent DH5α. Due to the size of 15 kb of the NIPBL ligation product, electroporation transformation was required.

Puromycin digestion to insert the vector in the 3XFLAG-HALO-EV-BlastR	**3XFLAG-HALO-EV-BlastR plasmid digestion**	**Puromycin digestion to insert the vector in the 3XFLAG-HALO-msNIPBL-WT-BlastR**	**3XFLAG-HALO-msNIPBL-WT-BlastR plasmid digestion**
2,5 µL – Apa1	2,5 µL– Apa1	2,5 µL – Mfe1	2,5 µL – Mfe1
2,5 µL – KspA1	2,5 µL – KspA1	2,5 µL – KspA1	2,5 µL – KspA1
5 µL – Fastdigest buffer 10X	5 µL – Fastdigest buffer 10X	5 µL – Fastdigest buffer 10X	5 µL – Fastdigest buffer 10X
30 µL - Purified puromycin sequence	6 µg - EV Plasmid	30 µL Purified puromycin sequence	6 µg - NIPBL Plasmid
H_2O	H_2O	H_2O	H_2O

Table 5. Digest mix of amplified sequences containing Apa1, KspA1, Mfe1 restriction sites and puromycin resistance sequence, and digest mix of 3XFLAG-HALO-EV-BlastR, and 3XFLAG-HALO-msNIPBL-WT-BlastR vectors

3.3 Cell Culture

3.3.1 Cell lines

A673 is a cell line isolated from the muscle of a 15-year-old female patient suffering from Ewing sarcoma. A673 cell line was obtained from the American Type Culture Collection (ATCC). By transfection, the PiggyBac plasmids containing 3X-FLAG-HALO (corresponding to the Empty-vector (EV) cell line) or 3X-FLAG-HALO-msNIPBL were inserted in the A673 cells genome. Blasticidin selection allows the preservation of the transposon in the cells' genome. Additionally, A673 shEF1 cell lines contain rtTA and shEF1 construct for the Tet-ON system, respectively conserved by zeocin and blasticidin resistance. Generation of A673 shEF1 EV and A673 shEF1 NIPBL cell lines has required the substitution of the blasticidin resistance into puromycin resistance in order to maintain the transposons: 3X-FLAG-HALO and 3X-FLAG-HALO-NIPBL in A673 shEF1 cells (**Table 6**).

A.

Stable transposed cell lines	Stable transposons inserted by Piggy bac system	Selection Agent
A673 ATCC EV **A673 ATCC NIPBL**	Insertion of transposon containing 3X-FLAG-HALO for Empty vector (EV) cell lines or 3X-FLAG-HALO-msNIPBL for NIPBL cell lines	blasticidin (1:1000)
A673 shEF1	Insertion of two transposons containing the Tet-ON vectors: rtTA (reverse tetracycline-controlled transactivator) and shEF1	blasticidin (1:1000), zeocin (1:500)
A673 shEF1 EV **A673 shEF1 NIPBL**	Insertion of stables transposons containing the Tet-ON vectors: rtTA (reverse tetracycline-controlled transactivator), shEF1, and 3X-FLAG-HALO (EV) or 3X-FLAG-HALO-msNIPBL	blasticidin (1:1000), zeocin (1:500), puromycin (1:10000)

B.

Selection Agent	Initial Concentration	Final Concentration
blasticidin	10 mg/ml	10 µg/ml
zeocin	100 mg/ml	200 µg/ml
puromycin	10 mg/ml	1 µg/ml

Table 6. A. List of described stable cell lines, including selection agent added at DMEM (Dulbecco's Modified Eagle's Medium) medium to maintain transposons in A673 genomes - blasticidin (InvivoGen) – Zeocin InvivoGen) – puromycin (InvivoGen) - **B. Concentration of selection agent used in cell culture**

3.3.2 Cell culture conditions

Cells were maintained with Dulbecco's modified Eagle's medium (DMEM) with L-Glutamine and Sodium Pyruvate supplemented with 10% of Fetal Bovine Serum (FBS) and 1% of Penicillin-Streptomycin to prevent bacterial contamination (DMEM full). According to the type of cell lines, cells were selected continually for blasticidin and/or zeocin and/or puromycin resistance.

All cell lines were sustained at 37°C in a humidified incubator with 5% CO_2. Several controls have been performed to ensure mycoplasma-free culture.

3.3.3 Doxycycline treatment of A673 shEF1 cell lines

A673 shEF-1 cells have been cultivated in a 6-wells plate (10 cm^2) and transfected at 10% of confluence with Pbac-3XFLAG-HALO-EV-PuroR or Pbac-3XFLAG-HALO-msNIPBL-WT-PuroR. For this purpose, 500 µL of DMEM w/o FBS and w/o Pen-Strep has been mixed with 0,4 µg of 1:10 diluted transposase (1,75 µg/µL) and 1,6 µg of plasmids. After the solution has been vortexed quickly and 4 µL of Pei 40 00 has been added. Then, the solution has been vortexed for 10 s and spin. Then the mix has been incubated 15min at RT before being added to the cells medium. Finally, four hours later, the medium has been changed for DMEM, then the antibiotic selection started 48 hours post transfection. full supplemented with 1:1000 of blasticidin, 1:500 of zeocin, and 1:10 000 of puromycin (see Table 6B).

3.3.4 Transfection

3.3.4.1 Stable transfection of Pbac-3XFLAG-HALO-EV-PuroR or Pbac-3XFLAG-HALO-msNIPBL-WT-PuroR vectors in A673 shEF1 cells to create A673 shEF1 EV/NIPBL cell lines by Piggybac system

A673 shEF-1 cells have been cultivated in a 6-wells plate (10 cm^2) and transfected at 10% of confluence with Pbac-3XFLAG-HALO-EV-PuroR or Pbac-3XFLAG-HALO-msNIPBL-WT-PuroR. For this purpose, 500 µL of DMEM w/o FBS and w/o Pen-Strep has been mixed with 0,4 µg of 1:10 diluted transposase (1,75 µg/µL) and 1,6 µg of plasmids. After the solution has been vortexed quickly and 4 µL of PEI 40 000 has been added. Then, the solution has been vortexed for 10 s and spin. Then the mix has been incubated 15min at RT before being added to the cells medium. Finally, four hours later, the medium has been changed for DMEM, then the antibiotic selection started 48 hours post transfection. full supplemented with 1:1000 of blasticidin, 1:500 of zeocin, and 1:10 000 of puromycin (see Table 6B).

3.3.4.2 Transient transfection of siCtrl and siNIPBL on A673 shEF1 -/+ doxycycline

In 6-wells plate (10 cm^2), A673 shEF-1 – dox cells or A673 shEF-1 + dox cells have been cultivated and transfected at a confluence of 10% or 40% respectively with siCtrl (Eurogentec SR-CL000-005-siRNA Controls) or siNIPBL (Eurogentec). Cells were harvested 72 h after the transfection for the RNA extraction and the RTqPCR analysis. Because of the huge differences between the doxycycline treated and non-treated A673 cells, different transfection reagents have to be used.

3.3.4.2.1 Transient transfection by JetPrime for A673 shEF1 – doxycycline

1 µL of siRNA (100µM) was added to 200 µL of JetPrime buffer (Polyplus, jetPRIME®) and mixed thoroughly. Following this, 4 µL of lipofectamine Transfection Reagent (Polyplus, jetPRIME®) was added to this mix, after the mix was vortexed 5 s and spun. Before supplementing DMEF Medium with the JetPrime reaction mix, the mix has been incubated 10 min at RT. A673 shEF1 – dox cells medium has been changed with DMEM supplemented with 1:1000 of blasticidin and 1:500 of zeocin (see Table 6B), 4 h after the transfection.

3.3.4.2.2 Transient transfection by RNAiMax for A673 shEF1 + doxycycline

Two solutions of 250 µL of Opti-MEM medium have been supplemented respectively with 1 µL of siRNA and 5 µL of RNAiMAX before being quickly vortexed and spin. The two reactions have been gently mixed in one tube and were incubated at RT for 15 min. The mix has been then dropped on A673 shEF1 + dox cells. Four hours after, the medium has been changed with DMEM supplemented with 1:1000 of blasticidin, 1:500 of zeocin, and 1:500 of doxycycline (see Table 6B).

3.4 RNA expression quantification

The influence of NIPBL KD on the transcription of EWS-FLI1 unknown and known direct EWS-FLI1 targets (see the primers listed in the annex) was measured in A673 shEF1 – doxycycline and A673 shEF1 + doxycycline cells. As mentioned above, doxycycline treatment allows the expression of anti-EF1 shRNA (shEF1) to downregulate EF1 by RNA interference (see the doxycycline treatment of A673 shEF1 cell lines for more details on EF1 KD and transient transfection of siCtrl and siNIPBL in A673 shEF1 -/+ dox cell lines).

3.4.1 RNA extraction

Cell medium was first discarded and, the cells were washed once with cold PBS. The plates were then scraped to harvest the cells. The RNA extraction was realized according to the manufacturer protocol of the NucleoSpin® RNA extraction kit (Macherey-Nagel). The RNA was quantified using the NanoDrop 1000 Spectrophotometer (ThermoFisher) at 280nm. Finally, RNA solutions were aliquoted and stored at -80°C.

3.4.2 Generation of cDNA from extracted RNA by reverse transcription

The amount of RNA used for reverse transcription varied between 200 ng and 1000 ng for the different technical replicates, depending on the lowest RNA concentration of the sample analyzed. However, 1µg of total RNA was preferred whenever possible. The generation of cDNA by reverse transcription was performed using the cDNA FastGene® Scriptase II kit (Nippon Genetics). The cDNA was stored at -20°C.

3.4.3 qPCR

A volume of 380µl of dd H_20 was added to the 20 µL of the cDNA sample. For each pair of primers (the primers list is in the table), a mix was prepared containing 75 µL of SYBR™ Green PCR Master Mix (Takyon™), 6 µL of 10 µM of primers pair mix, and 9 µL of dd H_20. A volume of 6 µL of the mix and 4 µL of diluted cDNA were loaded into a 384 plate. The plate was then centrifuged at 1000 rpm for 1 min.

		PCR cycle	
Preincubation	Pre-denaturation	3 min at 95°C	
Amplification	Denaturation	10 sec at 95°C	45X
	Annealing + Extension	45 sec at 60°C	
Melting curve		10 sec at 95 °C	
		10 sec at 50°C	
		50°C until 95°C (0,11°C/sec)	
Cooling		95°C until 40°C (20 sec)	

Table 7. qPCR cycles

The qPCR analyses were performed using the ΔΔCt method. The average Ct was corrected by the mean of the average Ct of five housekeeping genes: *RPL32, NDUFA12, GADPH, HPRT*. ΔΔCt corresponds to the difference between the tested conditions (siNIPBL) and the control conditions (siCtrl). Finally, the $2^{-\Delta\Delta Ct}$ was calculated to measure the fold change between the tested and the control conditions.

3.5 Protein quantification by BCA protein assay

Protein concentrations were measured by BCA protein assay (Thermo Scientific™ Pierce™ BCA Protein Assay Kits). The absorbance was measured at 562 nm. The emitted signal was detected by the SpectraMax® Plus Absorbance Plate Reader).

3.6 Subcellular fractionation

Cells were harvested in 150 cm^2 dished when they reached 90% of confluence. The medium was collected and conserved at 4°C, then the cells were washed once with PBS. 3 mL of trypsin were then added to detach the cells. Trypsin was neutralized with the conserved medium and the cells were

collected. The cells were counted (Bio rad TC 20™ Automated Cell counters) and an equal number of cells were used for each condition for the rest of the protocol. Cells were centrifuged at 350 g for 5 min and then the supernatant was removed. The pellet was resuspended with PBS. Next, the cells were centrifuged at 350g for 5 min and the pellet was resuspended with a specific volume of Buffer A at a concentration 6.10^6 cell/mL.

Cell nuclei were then extracted using buffer A adjusted with 0,06% of NP40 (v/) of and 1X PI and were incubated on ice for 10min before centrifugation at 350 g for 5min at 4°C. The supernatant was discarded, and the nuclei resuspended in 5ml of NSB buffer (Glycerol 25%, MgAc2 5Mm, HEPES pH7,5 5mM, EDTA pH 8 0,08mM, PI 1X, 0,5 mM DTT). Nuclei the incubated with 1,3 mL of Co-IP Lysis buffer (Glycerol 10%, HEPES pH7,5 20 mM, NaCl 150 mM, MgCl2 1,5mM, 0,5 mM DTT, NP40 0,8%, PI 1X, 0,5mM DTT) and the samples were rotated for 30 min at 4°C. Then, they were centrifuged at 12 000g for 5 min at 4°C and the nucleoplasmic fraction was conserved as the fraction N. Finally, the pellet was resuspended in 1,3 ml of Co-IP lysis buffer supplemented with 100 u/ml of Turbonuclease (HS-Nuclease, recombinant Endonuclease by Mo Bi Tec). They were rotated for 1 h at RT, then the solution was triturated and centrifuged at 12 000g for 5min at 4°C. The supernatant was the considered as the chromatin fraction (P) and conserved. For each fraction, 50 µl of solution was collected for the input.

3.7 Protein separation by SDS PAGE

Protein extracts, co-IP eluates or GST pull-down eluates were mixed with LDS loading buffer before heating at 95°C for 5 minutes. Proteins were separated by SDS-PAGE (Sodium Dodecyl Sulphate-PolyAcrylamide Gel Electrophoresis), which means that the proteins were separated according to their size thanks to SDS. For this purpose, the proteins were loaded onto a Tris-acetate polyacrylamide gel or a Tris-glycin polyacrylamide gel. The gel consists of a 4% stacking gel and a 9% or 12% running Gel. The 9% Tris-acetate gel was used to separate Co-IP eluates or GST pull-down eluates to analyze, NIPBL, a protein with a high molecular weight. In contrast, 12% Tris-glycin gel was used to determine the expression of GSTtag-recombinant proteins or protein extracts. Proteins were migrated in 1X Tris-acetate or 1X-Tris-glycin running buffer at 110 V for 1-2h. In addition, 10 µL of Spectra™ Multicolor High Range Protein Ladder and 4 µL of PageRuler™ Plus Prestained Protein Ladder were loaded into polyacrylamide gel to evaluate protein sizes.

3.8 Western blot

3.8.1 Western blotting transfer

Proteins were transferred to a nitrocellulose membrane in transfer buffer at 385 mA for 1h30 at RT or at 40 mA ON at 4°C.

3.8.2 Blocking and Antibodies

The nitrocellulose membrane was saturated for 1H with TBS-T/5% milk blocking solution. The membranes were then incubated for 4 h at RT or ON at 4°C with the appropriate diluted primary antibodies solutions (**Tables 8. and 9.**). After incubation, the membranes were washed 3 times for 5 min with TBS-T. The membranes were then incubated with the appropriate 1:5000 diluted secondary antibodies for 1h at RT in 3% BSA (VWR, Bovine Serum Albumin) TBS-T. The membranes were washed 3 times with TBS-T.

Target	Animal source	Clonality	Supplier	Reference	Antibodies dilution	Buffer
GST	Mouse	Monoclonal	San	B-14:sc-138	1/2000 1/4000	TBS-T, BSA 5%
FLAG-M2	Mouse	Monoclonal	Sigma-Aldrich	F3165	1/1000	TBS-T, BSA 5%
HSP90	Rabbit	Monoclonal	Cell Signaling	C45G5	1/2000	TBS-T, BSA 5%

Table 8. Primary antibodies

Target	Animal source	Clonality	Supplier	Reference	Antibodies dilution	Buffer
Rabbit IgG	Goat	Monoclonal	Cell Signaling	7074S	1/2000 1/4000	TBS-T, Milk 5%
Mouse IgG	Horse	Monoclonal	Cell Signaling	7076S	1/1000	TBS-T, Milk 5%

Table 9. Secondary antibodies

3.8.3 Chemiluminescence

The horse radish peroxidase coupled to the secondary antibodies allowed the visualization of proteins of interest following the addition of its substrate mix: 2 mL of A solution, 200 µL of B solution, and 0,6 µL of C solution. The emitted chemiluminescent signal was detected by the Amersham ImageQuant 800 imaging system (Cytiva).

3.9 GST pull down assay to determine a potential interaction between the C-terminal domain of EWS-FLI1 (FLI1) and NIPBL

3.9.1 The Gateway cloning method

Recombinant proteins were tagged with GST owing to the insertion of their coding sequences into an expression clone according to two recombination reactions using Invitrogen's Gateway® cloning system (**Fig.9**). The BP reaction depends on the recombination of attB and attP sites to give rise to an entry clone containing the gene fragment of interest flanked by attL sites. The second step, the LR reaction is based on the recombination of attL and attR sites to create the expression clone with the GSTtag-insert flanked by two attB sites. The positive and negative selections allow the selection of the target by-product of the recombination reactions. The donor vector of the BP reaction or the Destination vector of the LR reaction contains the *ccdB* gene flanked by recombination sites and respectively spectinomycin and ampicillin resistance in their backbone. *CcdB* gene is a cytotoxic gene coding for a helicase inhibitor while conducting to the death of the bacteria. The antibiotics select bacteria that have inserted the plasmid containing the sequence coding for antibiotic resistance [89].

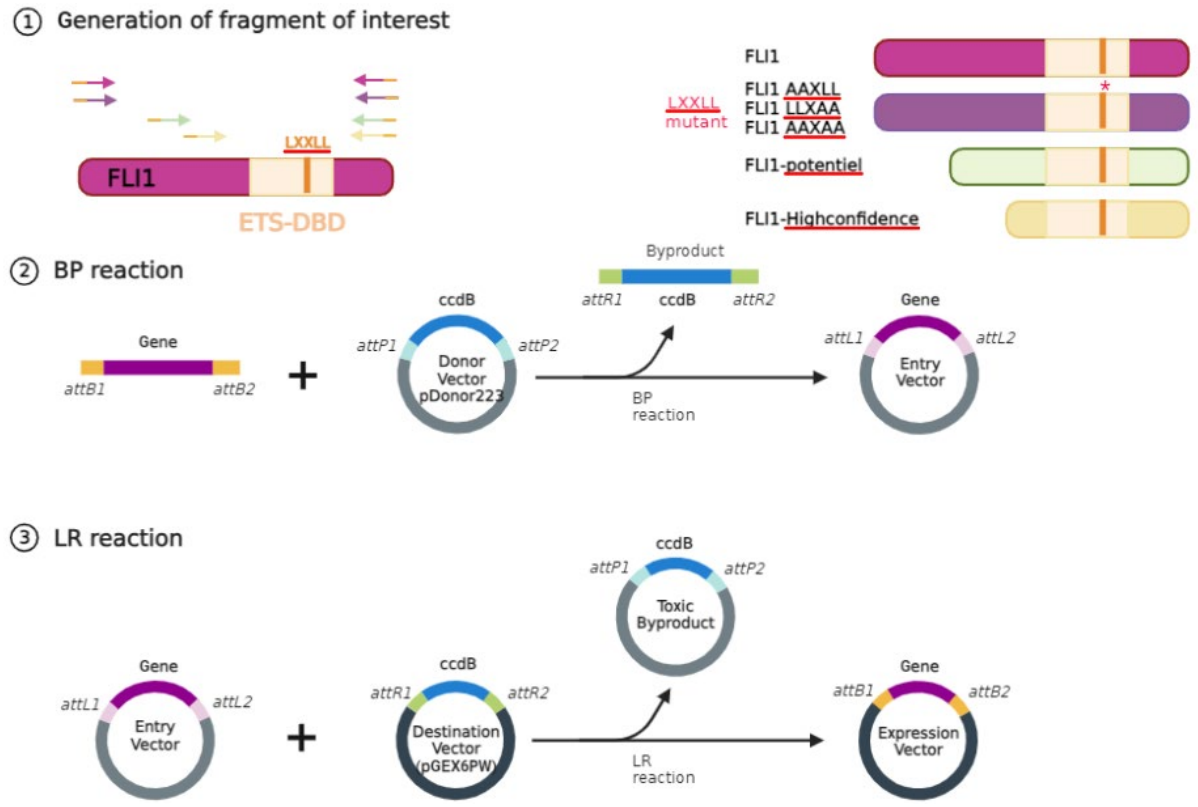


Figure 9. PCR reaction to amplify and add attB sites at the interest constructs (FLI1full, FLI1 potential, FLI1Highconfidence, FLI1-AAXLL mutant, FLI1-LLXAA mutant, FLI1-AAXAA mutant). Following the Gateway cloning reactions, BP and LR, create an expression vector containing GST-tag sequence coupled with coding sequences of the interest fragments

3.9.2 GST-FLI1full, GST-FLI1potential, GST-FLI1highconfidence, GST-FLI1AAXLL, GST- FLI1LLXAA, GST-FLI1AAXAA generation

For the FLI1potential, FLI1Highconfidence constructions, as well as the three LXXLL mutants, overhang PCR was adopted to add attB sites and amplify the appropriate FLI1sequence from the FLI p3 pN-NIL and pN-MSE FLI1 C1NC5 plasmids, while the FLI1ms2 plasmid was used for the FLI1pot and FLI1HC constructions. For the LXXLL mutant, touchdown PCR was used to modify and amplify the sequence from the FLI1-Δ-c-ter plasmid. Then, nested PCR enabled to add attB sites to the mutant sequences (**Table 10, 11, 12**). PCR products were loaded on 2% agarose gel. After electrophoresis, appropriate bands were conserved and purified using the gel and PCR clean-up kit.

Mix PCR	Mix Nested PCR
0,5µl – HF polymerase	0,5µl – HF polymerase
10 µl – HF Buffer 5X	10 µl – HF Buffer 5X
1µl - dNTP 10mM	1µl - dNTP 10mM
2,5 µl - Primer F 10 µM	2,5 µl - Primer F 10 µM DF 5065
2,5 µl - Primer R 10µM: DF5066	2,5 µl - Primer R 10µM: DF5066
0,5 µl de DMSO	0,5 µl de DMSO
50 ng - Plasmid	1 µl of the PCR purification
H_2O	H_2O

Table 10. PCRs mix to add amplify sequences containing the FLI1pot sequence with the DF5054 forward primer, the FLI1HC sequence with the DF5065 forward primer, the FLI1-AAXLL sequence with the DF5067 forward primer, the FLI1-LLXAA sequence with the DF5068 forward primer, the FLI1-AAXAA sequence with the DF 5069 forward primer and the DF5066 reverse forward primer for all the sequences. The DF 5065, DF 5066 primers enabled to add attB sites at the FLI1-LLXAA, the FLI1-AAXLL and the FLI1-AAXAA sequences by nested PCR

	Overhang PCR cycle – Touchdown PCR	
Initial Denaturation	at 98°C	
Denaturation	20 sec at 98°C	10X
Annealing	30 sec at 68°C – 58°C	
Extension	30 sec at 72°C	
Denaturation	30 sec at 98°C	30X
Annealing	20 sec at 65°C	
Extension	30 sec at 72°C	
Final extension	15min at 72°C	
	∞ at 4°	

Table 11. Touchdown PCRs cycles conditions to amplify FLI1-AAXLL, FLI1-LLXAA, and FLI1-AAXAA sequences mutants

	Overhang PCR cycle	
Pre-denaturation	10 min at 98°C	
Denaturation	10 sec at 98°C	35X
Annealing	30 sec at 57°C	
Extension	45 sec à 72°C	
Final extension	5 min at 72°C	
	∞ at 4°	

Table 12. Overhang PCRs cycles conditions to amplify FLI1pot and FliHC sequences and to add attB sites for the FLI1pot, FLI1HC and the FLI1 LXXLL mutant constructions

For the BP reaction, mix 7 µL of the fragment, 2 µL BamHI-linearized pDONOR223, and 2 µl of BP clonase mix (Invitrogen's Gateway® Gateway™ BP Clonase™ Enzyme Mix) and incubated the solution at RT ON. On the following day, 1 µL of Proteinase K was added to stop the reaction. Samples were then incubated at 37° C for 10 min. The heat shock protocol was used to introduce the entry clone into DH5α and the colony selection protocol was used to select the clones. The FLI1full was already present in the pDONR223.

Subsequently, the FLI1full, FLI1potential, and FLI1Highconfidence and the LXXLL mutants were coupled with GSTtag coding sequence by LR reaction consistent of the mix of 100 ng/µL of the expression vector, 75 ng/µL pGEX-6P-1, 1 µL of LR Mix clonase II (Invitrogen's Gateway® Gateway™ LR Clonase™ Enzyme Mix) and H_2O. Samples were incubated for 4h at RT, then 1 µL of Proteinase K was added. Samples were incubated at 37°C for 10 min. As the BP reaction the Heat shock and the clone selection protocols have been followed for the plasmid insertion into DH5α and then the clone selection.

3.9.3 Liquid pre-culture

Following the sequencing results (Sanger sequencing by the GIGA-Genomics platform or Eurofins Genomics), 5 mL of LB was added to 10 µL of the appropriate glycerol clone and 1/1000 of ampicillin. The liquid preculture was incubated for ~ 4h with constant agitation. Overnight, a liquid preculture of 50 mL was incubated at 37°C with constant shaking. The next day, a culture of 500 mL was incubated for 2h at 37°C with constant agitation.

3.9.4 Induce by IPTG GSTtag-recombinant protein production

During exponential phase growth, 100 mM of IPTG (Thermo Scientific™ IPTG, dioxane-free) was added to induce the expression of the coding sequence by the lac operon mechanism. The liquid culture was incubated for 4H at 30°C with constant agitation.

3.9.5 GSTtag-recombinant protein extraction

Following the induction, the bacteria were centrifuged at 4000 rpm for 5 min. The pellet was resuspended in 11,25 mL of lysis solution (1X Protease Inhibitor, Triton 1%, PBS). The pellet was sonicated for 30 sec on ice at a power of 30% Bioblock scientific Vibra cell 75041). The lysate was centrifuged at 17 000g for 10 min at 4°C. The supernatant was then transferred on 375 µL of 50% slurry glutathione coated agarose beads (Cytiva glutathione Sepharose™). The beads were previously washed 4 times with PBS and were centrifuged for each washing 5 min at 500g. The beads and the supernatant solution were incubated for 30 min at 4°C on a rocker. Finally, they were washed twice with PBS-1%Triton solution and then twice with PBS. For each wash, the beads were centrifuged for 5 min at 500 g. The GST/recombinant proteins beads were stored at -20°C in a solution of PBS and PI 2X solution to obtain a mixture of beads at 50% slurry.

3.9.6 Analysis of the expression of GSTtag-recombinant proteins (GST, GST-FLI1full, GST-FLI1potential, GST-FLI1Highconfidence, GST-FLI1-AAXLL, GST-FLI1-LLXAA, GST-FLI1-AAXAA)

Quantification allowed adjustment of the number of effective beads required to achieve the same expression for each condition. For this purpose, different volumes of 2, 5, 10, 20, and 40 µL of GSTtag-recombinant proteins beads 50% slurry were mixed with loading buffer. The samples were heated at 95°C for 5 min, then the solutions were centrifuged at 17 000g for 5 min. The supernatant was collected and loaded in 12%-Tris glycin gel (cf. to the SDS-PAGE protein separation section for further details).

The proteins were transferred to a nitrocellulose membrane to allow the quantification of the expression of the different GSTtag-recombinant proteins beads by Western blot analysis (cf.SDS-PAGE protein separation section for details). The quantification could be also performed by Coomassie blue stained SDS-PAGE. Following the electrophoresis, the SDS-PAGE gel was washed with purified H_2O and then incubated with DestI solution to fix the proteins for 5min with constant shaking. The gel was next stained with blue Coomassie solution for 30 min. Finally, the gel was washed with DestII solution with constant shaking O/N. In both cases by Western blot analysis or by Coomassie blue stained SDS-PAGE gel, the quantification of the expression of the GST-recombinant proteins was carried out using Image Studio Lite.

3.9.7 GST pull down – IP

A673 ATCC EV/NIPBL cells were harvested and lysed to finally obtain their nucleoplasmic and their chromatin fractions (see the subcellular fractionation for further details). Fractionating allows to recover more NIPBL proteins. The analysis of the interaction between NIPBL and EWS-FLI1, depending on the original fraction of NIPBL proteins, was not performed during the time frame of the thesis, but will be done later. For now, the nucleoplasmic and the chromatin fractions were pulled at equal volumes in one tube. A volume of 10µl or 50µl of GSTtag-recombinant proteins beads was prepared for 100µl of lysate. According to the lowest expression of GST-tag recombinant protein beads, glutathione coated beads (Cytiva glutathione Sepharose™) (the stock solution is maintained at 75% slurry) were mixed with recombinant protein beads to obtain the identical final volume of beads of the lowest expression of GST-tag recombinant protein beads. The mixed beads were centrifuged at 500 g for 5 min and the preserving liquid was discarded. The mixed beads were then washed with co-IP lysis solution and centrifuged at 500 g for 5 min, then the supernatant was removed. A solution of 100 µL of pull lysate for 10 µL of beads or 50 µL of beads was then prepared and incubated at 4°C for 2 hours, 4 hours or ON. The solution was centrifuged 5 min at 500g and washed 3 times with co-IP lysis buffer. For each wash, the samples were incubated on a rocker at 4°C and centrifuged at 500 g for 5 min. The beads were mixed with LDS 1X, dtt 1X solution and incubated at 95°C for 5 min before being loaded with the

nucleoplasmic and the chromatin fraction inputs onto a 9% Tris-acetate gel (see the Western blot analysis for more details).

3.10 Chromatin immunoprecipitation - ChIP

3.10.1 Cell harvesting

A673 shEF1 cells expressing 3xFLAG-Halo alone (EV) or 3xFLAG-Halo NIPBL WT and treated or not with doxycycline were collected by trypsinization and counted to equalize the number of cells between the conditions. They were centrifuged at 350 g for 5 min and washed in PBS.

3.10.2 Protein-protein/ADN-protein fixation

Cells were then fixed with 2 mM of DSG diluted in PBS and incubated on a rocker for 30 min at RT. Cells were then washed 3 times with PBS, then fixed with 1% PFA for 7 min at RT on a rocker. 1X glycine was added, and the cells were incubated on a rocker for 5 min at RT. Later, fixed cells were washed 3 times with PBS (the centrifugation were performed at 2000g at 4°C). The pellet was then resuspended in ChIP lysis buffer supplemented with protease inhibitor (1200 µL of ChIP lysis buffer/per 4X150 cm^2) and incubated on ice for an hour. Lysates were then equally distributed in 4 tubes adapted for sonication.

3.10.3 Chromatin Fragmentation

The samples were sonicated 30 sec ON/ 30 sec OFF for 30 cycles, then centrifuged at 20 000g at 4°C for 20 minutes. The supernatant was then collected and 15 µL was collected for the sonification verification step (see section 3.10.4.3)

3.10.4 Chromatin-ImmunoPrecipitation

3.10.4.1Immunoprecipitation

Equal amount of shredded chromatin was diluted with ChIP dilution buffer at least 5 time, then 50 µL was collected for as the input. Washed magnetic FLAG-beads were then added to the samples and incubated ON at 4°C with gentle rocking. The next day, beads were washed 2 times with the Low salt buffer, once with High salt buffer, once with LiCl buffer, and 2 times with TE buffer (for each wash, tubes were gently rotated at 4°C for 10min).
2 identical elution steps were performed on the beads and achieved with their incubation with 200 µL of elution buffer and their incubation for 20 min at RT. 400 µL of supernatant were collected, and 50 µL of reversal mix added. For the input, 350 µL of elution buffer was added. Next, samples were incubated for 2 hours at 50°C then at 65°C for 7 hours for fixation reversal and protein digestion.

3.10.4.2 Phenol chloroform extraction, DNA purification and quantification

400 µL of phenol chloroform isoamyl alcohol were added to the solution. They were then transferred to phase lock tubes and centrifuged at 20 000g for 5min. The aqueous phase was collected and mixed with 1/10 of 3M sodium acetate and 20 µg of glycogen and vortexed. Then, 2,25X of cold 100% EtOH was added, and the samples were stored at -20°C for 1 hour. Next, the samples were centrifuged for 30 min at 4°C at 20 000g. The supernatant was removed, and the pellet washed with 70% EtOH once. The supernatant was removed, and the pellet was dried at RT. The DNA was finally resuspended in 200 µL Finally, qPCRs were performed with primers designed to amplify EF1-binding peaks at cis-regulatory regions of EF1 target genes. (see section 3.4.3 for further details).

3.10.4.3 Chromatin shredding verification

15 µL of the shredded chromatin was supplemented were decrosslinked with a mix of 180 µL of mastermix buffer and 2,5 µL RNAse/15 µL of samples. The samples were incubated for 30 min at 37°C. Then 2,5 µL of proteinase K was added and the samples heated at 50-65°C for 120 min. Finally, 200 µL of H_2O finally were added, and the DNA extracted following the previously described phenol-chloroform extraction and precipitation (see section 1.1.4.2). DNA concentration was dosed by Nanodrop for quantification. Next, 10 µL were then loaded on 2% agarose gel to verify the size of the DNA fragment. The desired length of DNA fragments should range between 150-500 base pairs.

IV. Results

4.1 Determine a possible interaction of EWS-FLI1 and the cohesion loader NIPBL

4.1.1 3XFLAG-HALO or 3XFLAG-HALO-NIPBL expression in A673 EV or A673 NIPBL cells

To determine a possible interaction between the cohesin loader, NIPBL, and the chimeric transcription factor, EWS-FLI1, A673 cells were transfected with the plasmid Pbac-3XFLAG-HALO-EV-BlastR or Pbac-3XFLAG-HALO-msNIPBL-WT-BlastR using the piggy-bac system by Dr. Fettweis. Exogenous proteins were tagged with the FLAG-tag to simplify the immunoprecipitation step with anti-FLAG antibody. Before investigating the interaction, we verified the expression of the recombinant proteins, 3X-FLAG-HALO (EMPTY vector – EV) or 3X-FLAG-HALO-NIPBL (NIPBL) in A673 cells (**Fig.10**).

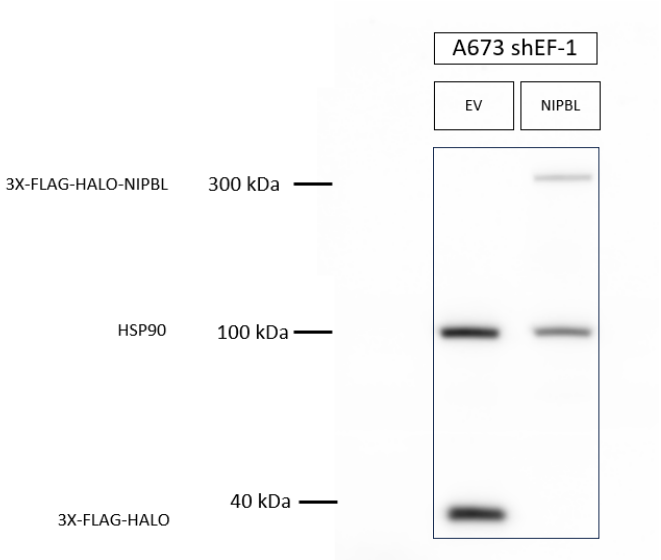


Figure 10. Expression of exogenous proteins 3XFLAG-HALO or 3XFLAG-HALO-NIPBL respectively in A673 EV ATCC cells or A673 NIPBL ATCC cells. HSP90 was used as a loading control.

4.1.2 Identification of NIPBL as a direct interactor of the c-terminal domain of EWS-FLI1 by GST pull down assay

Our hypothesis is that EF1 could recruit NIPBL through a direct interaction at enhancers of EF1 target genes. This initiates E-P loops formation following cohesin loading by NIPBL as already observed with the transcription factor Glucocorticoid Receptor (GR) [54]. In an unpublished study mediated by Dr. Fettweis, they determined that the interaction between GR and the NIPBL proteins depends on three LXXLL motifs of GR. For this purpose, we decided to investigate the potential interaction of NIPBL with the C-terminal domain of EWS-FLI1, FLI1, in vitro by GST-pull-down assay. EWS-FLI1

possesses a unique LXXLL motif within its ETS-DNA binding domain, the sequence of which corresponds to LLELL. The N-terminus of EWS-FLI, EWS, is a prion-like region containing the transactivation domain. The C-terminal domain of EWS-FLI1, FLI1, consists of unstructured and structured regions. The structured region corresponds mainly to the ETS-DNA binding domain (**Fig.11**).

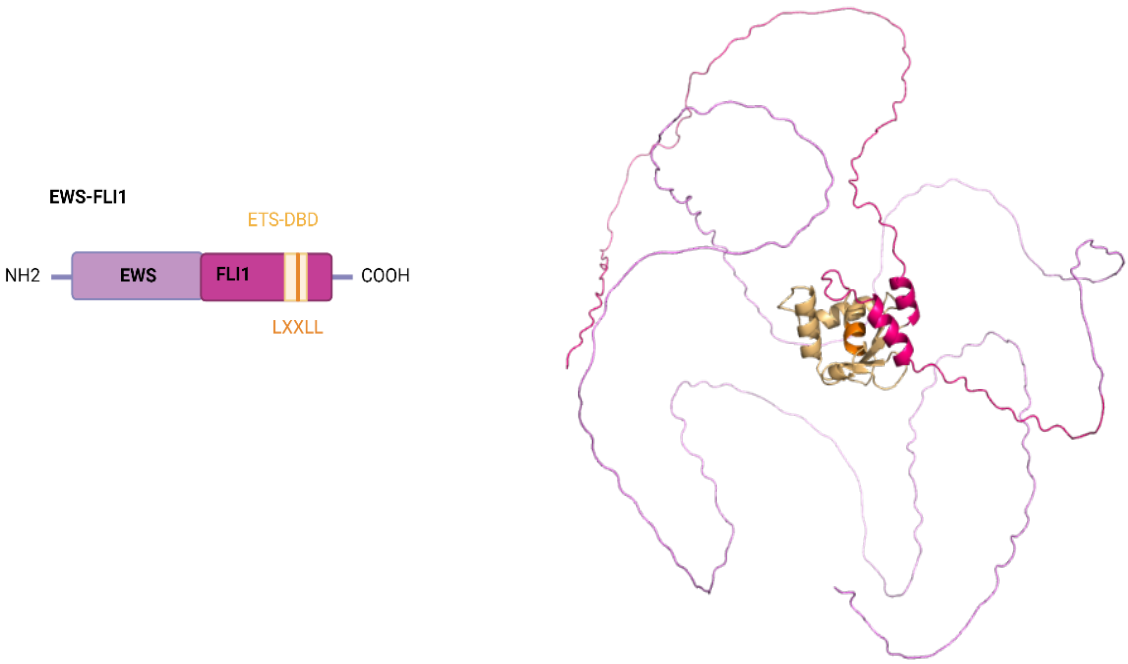


Figure 11. On the left is a schematic view of the EWS-FLI1 fusion protein. On the right, the oncogenic EWS-FLI1 fusion (Alphafold illustration), represented by the prion-like region of the EWS domain (purple), the FLI1 domain (red/purple) including its ETS-DNA binding domain (yellow) and its LLELL motif (orange).

4.1.2.1 Design experiment to determine the interaction between NIPBL and the EWS-FLI1 C-terminal domain (FLI1) by GST pull down assay

Three FLI1 constructs were selected according to the per-residue score of the Alphafold prediction of EWS-FLI1 (**Fig.12**).

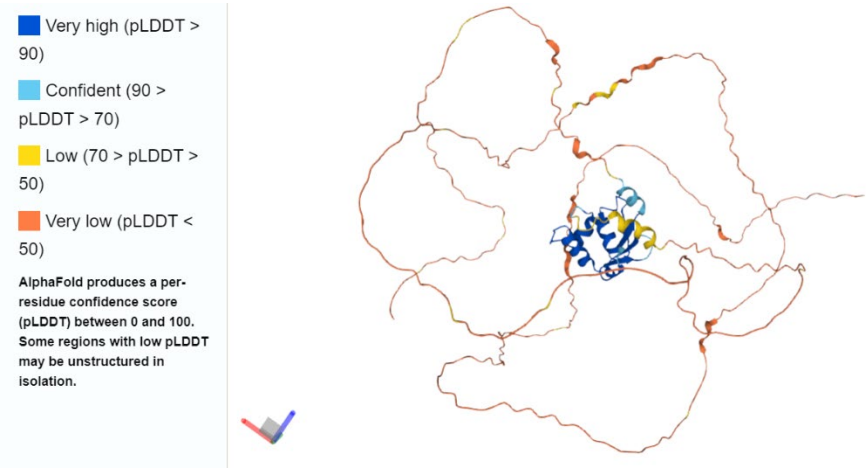


Figure 12. Alphafold prediction of EWS-FLI1 (AF-V9GZ02-F1). (Uniprot data)

The FLI1-cter containing the complete amino acid sequence of FLI1 (233 amino acids), the FLI1 potential (191 amino acids) consisting of structured regions (173 amino acids) and unstructured regions (18 amino acids) of FLI1, and the FLI1Highconfidence (173 amino acids) containing only the structured region of FLI1 (**Fig.13**). We built N-terminal GST-chimera with those three sequences and expressed them in $DH_5\alpha$ E.coli after IPTG induction. GST-tag recombinant proteins were purified with glutathione (GSH) agarose beads. GST-tag recombinant proteins were used as baits to capture their prey following beads incubation with the nucleoplasmic and chromatin fractions of EV or NIPBL A673 cells (**Fig.13**).

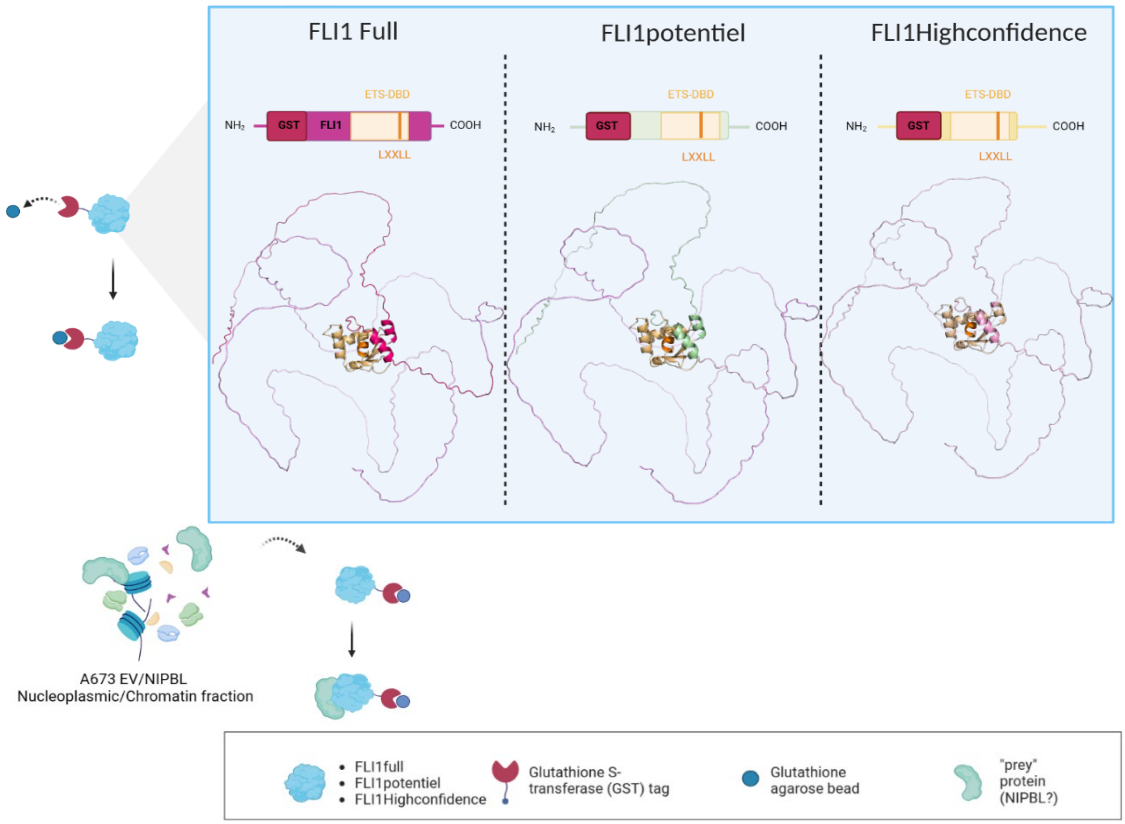


Figure 13. Design experiment to determine the interaction between NIPBL and EWS-FLI1 c-ter (FLI1) by GST pull down assay. The blue square shows the Alphafold prediction of the EWS-FLI1 protein structure. The different sequences of the FLI1full, FLI1potential and FLI1Highconfidence constructs corresponds to the colours of the schematic representation of the GST-recombinant proteins (GST is not represented in the Alphafold illustration)

4.1.2.2 Gateway Cloning to generate FLI1Full, FLI1potential (FLI1pot), FLI1Highconficence (FLI1HC) sequence

The first goal was to clone the FLI1full, FLI1potential, and FLI1highconfidence sequences in a Gateway vector in order to produce in $DH_5\alpha$ E.coli the proteins chimere with a GSTtag in N-ter. The amplification of the GSTtagFLI1full, GSTtagFLI1pot (FLi1potential), GSTtagFLI1HC (FLI1Highconfidence) DNA sequences to mediate clone selection of the expression plasmid is shown below (**Fig.14**).

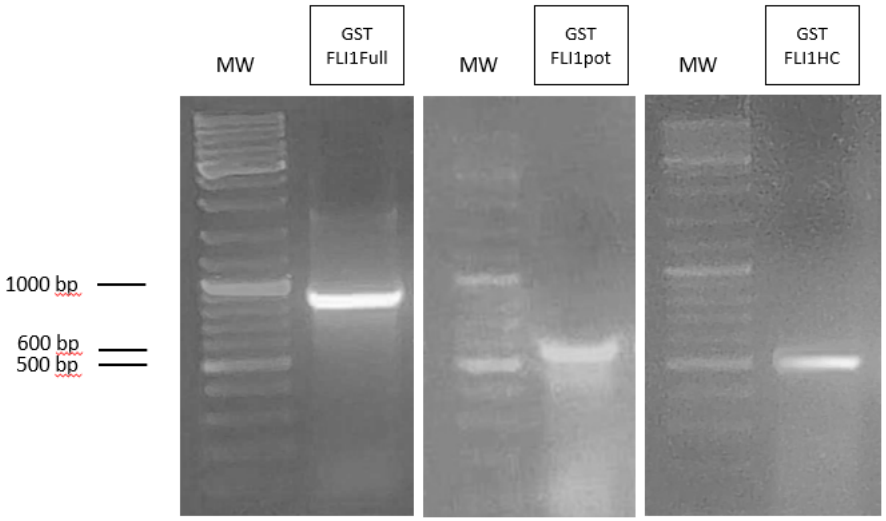


Figure 14. DNA fragment resulting to clone selection following colony PCR of the expression clone: GSTtagFLI1full, GSTtagFLI1pot, GSTtagFLI1HC. Gateaway Cloning-LR reaction to insert the FLI1full, FLI1pot (FLi1potential), FLI1HC (FLI1Highconfidence) BP fragment into the pGEX-6P1W plasmid (pDEST). MW: Molecular weight: GeneRuler DNA Ladder Mix

4.1.2.3 GST proteins and GSTtagFLI1full, GSTtagFLI1pot, GSTtagFLI1HC recombinant proteins production and quantification

GST and GSTtagFLI1full, GSTtagFLI1pot, GSTtagFLI recombinant proteins were produced in $DH_5\alpha$ E.coli after IPTG induction, and purified with glutathione coated agarose beads following bacteria lysis. To obtain a comparative expression level between the four conditions for the IP, SDS-PAGE was performed to quantify protein expression by Coomassie blue or Western blot (**Fig.15**).

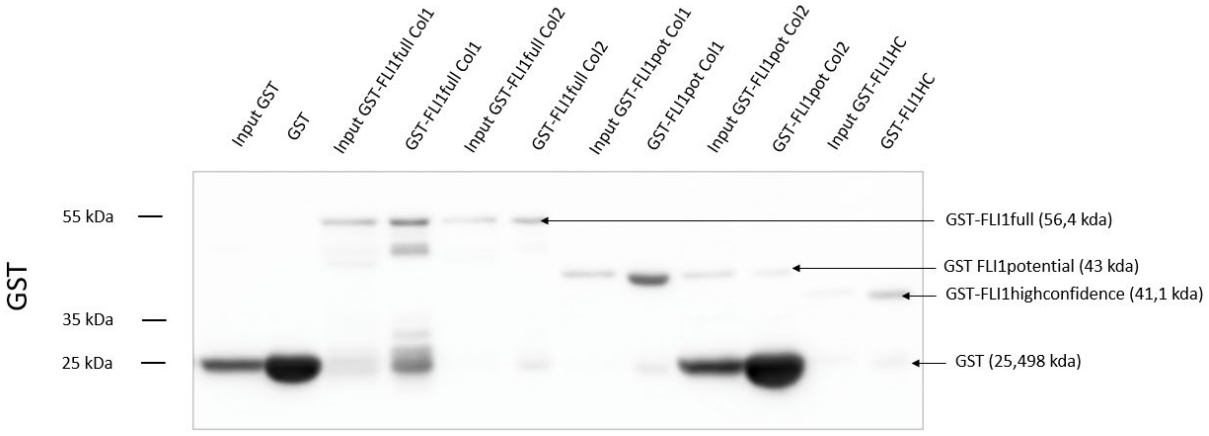


Figure 15. Western blot to quantify GST proteins and GSTtagFLI1full, GSTtagFLI1pot, GSTtagFLI1HC recombinant proteins. The input corresponds to the total proteins extracted from E.Coli DH5α without the purification step with the glutathione coated agarose beads. The weight of the GST proteins and GSTtagFLI1full, GSTtagFLI1pot, GSTtagFLI1HC is respectively 25.498 kDa, 56.4 kDa, 43 kDa, 41,1 kDa. The bands around 25 kDa for each condition correspond to the GST resulting from the degradation of the GSTtag recombinant proteins.

4.1.2.4 Determine NIPBL and EWS-FLI1 c-ter (FLI1) interaction by GST pull down assay

According to the quantification of each condition, GST alone and GSTtagged chimeras were mixed with glutathione coated agarose beads to obtain an equal beads/GST-constructs ratio. The beads were then incubated with A673 EV or NIPBL mixed nucleoplasmic and chromatin fractions, to proceed to the IP. The beads were then washed with Co-IP lysis solution containing 150mM NaCl to eliminate non-specific interaction.

Interaction between NIPBL and FLI1full was successfully revealed in our two replicates (**Fig.16, Supplementary Fig.1-2**). Although a faint band, representing FLAG-HALO-NIPBL, is present in the GST control condition (**Fig.16**), the band is 20X more intense in the FLI1full condition. Furthermore, in **Fig.16**, we determined that the interaction was specific between the EWS-FLI1 complete c-terminal domain, FLI1, and the exogeneous NIPBL proteins, independently of the FLAGtag (absence of bands in the negative ctrl: EV for the FLI1full condition). Our two replicates suggest that NIPBL and the EF1 c-terminal domain, FLI1, interact significantly with each other.

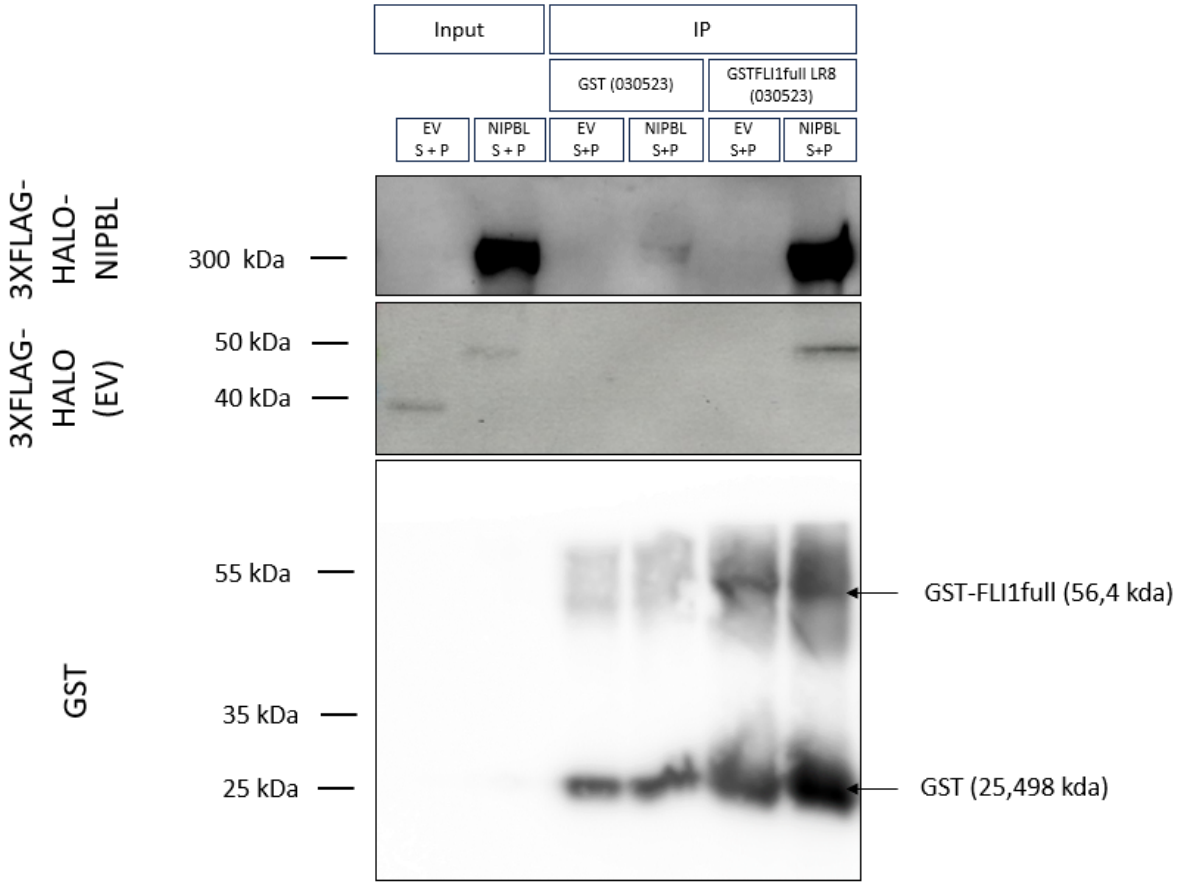


Figure 16. **GST pull down revealed the interaction between NIPBL and the EWS-FLI1 c-ter domain.** Blot showing on the left the input of the nucleoplasmic fraction (S fraction) mix with the chromatin fraction (P fraction) of the A673 3XFLAG-HALO (EV) and 3XFLAG-HALO-NIPBL (NIPBL) cells. On the right, the blot shows the GST, GST-FLIFull,) IP against NIPBL (detected by the FLAG antibody). The A673 EV and NIPBL chromatin fractions were mix volume/volume for the IP. (Second replicat)

In the first replicate (**Supplementary Fig.1**), FLIHC was shown to interact with NIPBL. Unfortunately, we were unable to detect this interaction in the second replicate (**Supplementary Fig.2**), the band was less intense than the faint band present in the GST control. For some unknown reason, the recombinant GSTFLI1, GSTFLI1pot and GSTFLI1HC are highly degraded during the IP even in the presence of 1X and even 2X protease inhibitor (for the second IP). Furthermore, in the WB of the quantification of the GST-recombinant proteins expression, we detected a significant amount of quantity truncated or degraded GST recombinant proteins, corresponding to the GST band (**Fig.15**). Concerning GST-FLI1pot, in our first attempt to perform the GST pull-down assay, we detected a high degradation of GST-FLIpot and GSTFLI1HC. We tried a second time to produce the recombinant proteins with different colony, but after the IP, the new GST-FLI1pot recombinant proteins were completely degraded (data not shown).

In conclusion, we have demonstrated a direct specific physical interaction between the complete c-terminal domain of EF1 and exogenous NIPBL proteins. Nevertheless, we detected a significant degradation of the recombinant proteins during the production, bacterial lysis, purification and IP steps. The degradation was even more important during the IP for the recombinant GST-FLI1pot and GST-FLI1HC. It will be necessary for the future to improve our protocol to address this issue to prevent degradation of the recombinant proteins.

4.1.2.5 LXXLL FLI1 mutant to the sequence to the proteins production

Succeeding the successful determination of the interaction between NIPBL and the EWS-FLI1 c-ter domain (FLI1), the FLI1 mutant needs to be performed to determine if we are able to lose the interaction with the LXXLL mutants. The LLXLL mutant will allow us to ascertain whether the unique LLXLL present in the ETS domain of FLI1 is required for the direct interaction between NIPBL and EWS-FLI1 as it is for GR and NIPBL (data not published).

The importance of each leucine is assessed by its replacement with an alanine, three mutants were generated the FLI1AAXLL, FLILLXAA and FLI1AAXAA.

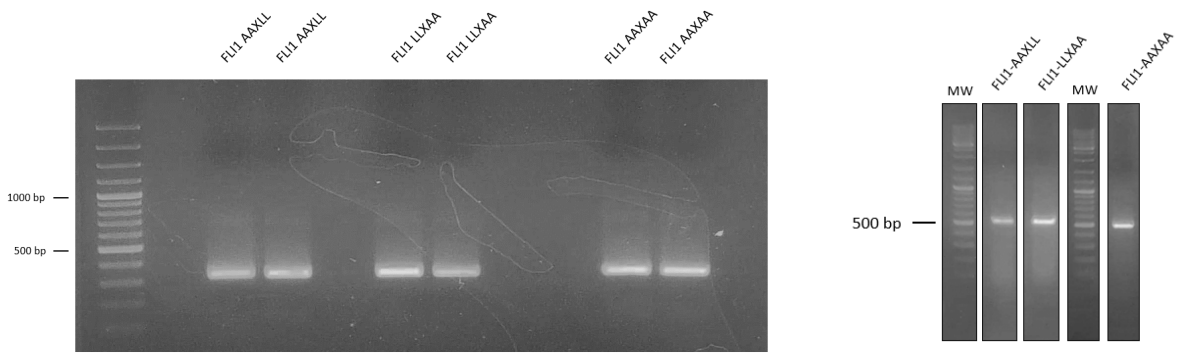


Figure 17. Left. DNA amplicons after the amplification of the FLI1AAXLL, FLI1LLXAA, FLI1AAXAA sequence added to attB sites. Right. DNA fragment resulting from clone selection after colony PCR of the expression clone: GSTtagFLI1HC, GSTtagFLI1AAXLL, GSTtagFLI1LLXAA, GSTtagFLI1AAXAA. Gateway Cloning-LR reaction to insert the FLI1AAXLL, FLI1LLXAA, FLI1AAXAA BP fragment into the pGEX-6P1W plasmid (pDEST). MW: Molecular weight: GeneRuler DNA Ladder Mix

After the first replicate, we started to produce our FLI1-LLXLL mutant. Afterwards, we did not see any significative bands for the FLI1HC, nonetheless GST-FLIHC as GST-FLI1pot were degraded importantly even more than the GST-FLI1Full. For this reason, we decided to continue to produce the GST-FLI1 mutant. We also decided to increase the concentration of the PI during the IP.

Touchdown nested PCR allowed us to create the three FLI1-LLXLL mutants added to the attB sites (**Fig.17.Left**). These three sequences were cloned by Gateaway Cloning to add a GSTtag in N-ter (**Fig.17.Right**).

As with the previous GST proteins and the GST-FLI1full, GSTFLI1HC recombinant proteins, the GSTFLI1-LXXLL mutants were produced in $DH_5\alpha$ E.coli and purified using glutathione coated agarose beads. As an example, **Fig.18** shows the expression of the GST and GST recombinant proteins expression by Coomassie blue. Subsequently, the expression of the proteins volume will be considered for a future IP to test if the interaction will be lost in the FLI1 mutant.

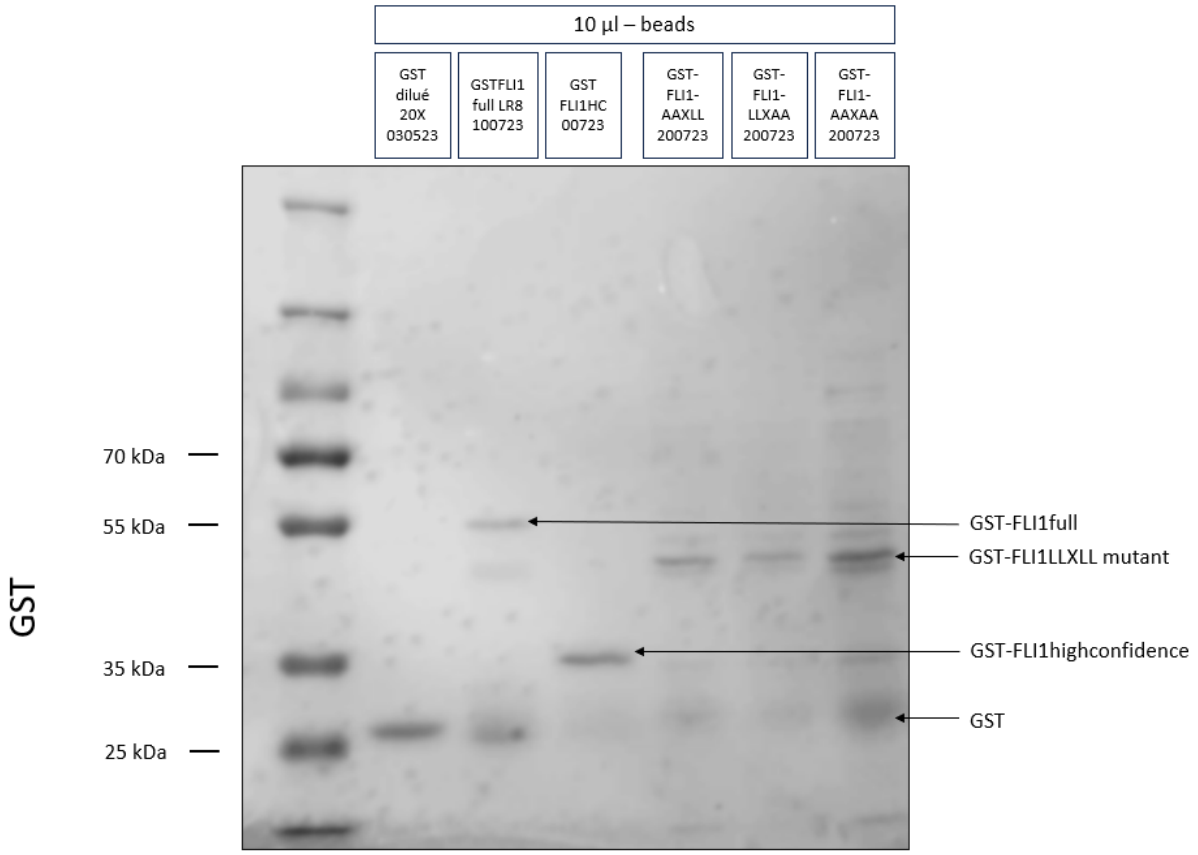


Figure 18. Coomassie blue GST proteins and GSTtagFLI1full, GSTtagFLI1pot, GSTtagFLI1HC, GSTtagFLI1AAXLL, GSTtagFLI1LLXAA, GSTtagFLI1AAXAA recombinant proteins. The weight of the GST proteins and GSTtagFLI1full, GSTtagFLI1HC, GSTtafFLI1 LXXLL mutant is approximatively 25.498 kDa, 56.4 kDa, 41,1 kDa, 46kDa. The bands around 25 kDa for each condition correspond to the GST resulting from the degradation of the GSTtag recombinant proteins. MW : Ladder PagePlusPrest.

4.2 E-P loops mediated by cohesin-NIPBL regulate some EF1 target genes transcription

We would like to investigate whether NIPBL knock-down is affecting EF1 target genes regulation. NIPBL downregulation might impair long-range EF1 target genes regulation subsequent to the reduction of E-P loops mediated by the cohesin complex.

The A673 shEF1 cells treated or not with doxycycline were transfected with a siRNA control (siCtrl) or against NIPBL (siNIPBL) using respectively the JetPrime or the Lipofectamine RNAiMax kit[3]. Cells were harvested 72h after the transfection to proceed to RNA extraction for the RTqPCR analysis. The expression of the EWS-FLI1 target genes (*PRKCB, EZH2, NPY1R, ICAM1, FEZF1, FCGRT, NROB1, PPP1R1A, KIT, MAFB)*[4] was evaluated in A673 shEF1 -/+ dox siCtrl or siNIPBL cells. The average Ct values of these genes were corrected by the mean of the difference between average Ct values of siCtrl and siNIPBL conditions of five housekeeping genes (*RPL32, NDUFA12, MAFB, HPRT*), then $2^{-\Delta\Delta Ct}$ was calculated to obtain the fold induction relative to siCtrl-. The p-value was calculated on the average Ct value corrected of the -dox siCtrl and +dox siCtrl by ratio paired t-test two-tailed analysis. The fold induction did not allow to take into account the Ct variation of the siCtrl between the 4 biological replicates. For this reason, we decided to perform the statistical analysis on the Ct value.

[3] Two different transfection methods were required because transfection of A673 shEF1 – dox using Lipofectamine RNAiMax kit or of A673 shEF1 – dox using JetPrime kit resulted in cells death.

[4] The expression of the *PRKCB, FEZF1, FCGRT, NROB1, PPP1R1A* were analyzed by RNAseq. The upregulation or downregulation of EF1 target genes results were associated with loops gain or loss following EF-WT (Endo or Rescue) versus EF-KD Hi-C analysis (resolution of 20kb) [87]. The genes *EZH2, NPY1R, ICAM1, KIT, MAFB c*orresponding to other known, or unknown direct EF1 target genes were not included in the set of genes identified in the study discussed just above. All the genes are identified in the H3K23ac epigenetic and transcriptional change experiments performed in two A673, SKNMC Ewing sarcoma cell lines [74].

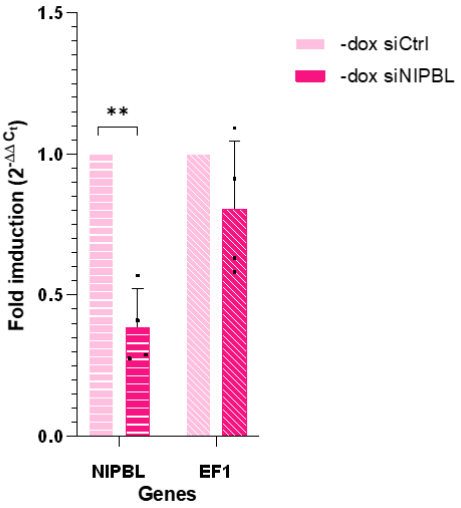
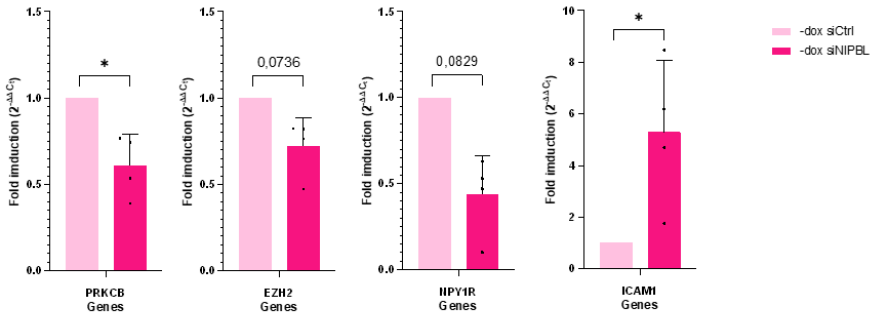


Figure 19. Visualization of NIPBL knockdown by RTqPCR. NIPBL mRNA demonstrate down expression of NIPBL ensue mRNA specific-target decrease by siNIPBL (pvalue = 0,0094) after A673 shEF1 -dox siCtrl vs A673 shEF1 -dox siNIPBL ratio paired t-test analysis two-tailed). EF1 is not affected by NIPBL KD (pvalue = 0,1918).

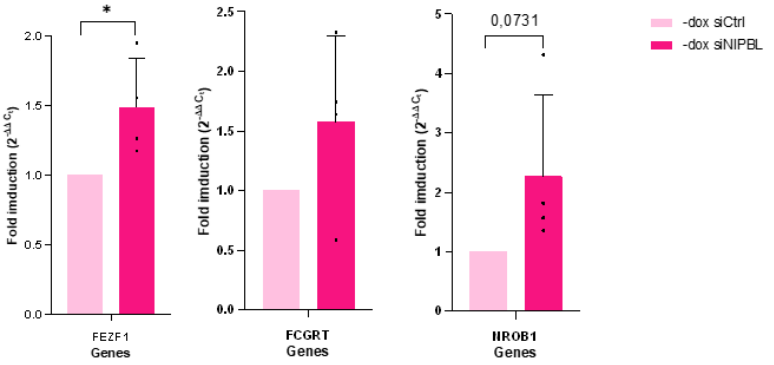
Ensuing NIPBL knockdown by siNIPBL, the NIPBL mRNA levels were checked and indicated a 60% decrease (in terms of fold induction) between A673 shEF1 -dox siCtrl and A673 shEF1 -dox siNIPBL of our 4 biological replicates (pvalue = 0,0094). EF1 transcription does not appear to be affected by NIPBL KD (ns - pvalue = 0,1918), suggesting that its transcription is not dependent on E-P loops mediated by the NIPBL/cohesin complex (**Fig.19**).



EF1 binding sites					
	Proximal MS	✓	?	?	?
	Proximal NMS	✓	?	?	?
	Distal MS	✓	?	?	?
	Distal NMS	x	?	?	?
	≥ 20kb gained loop	✓	?	?	?

Figure 20. Downregulation and upregulation of EWS-FLI1 target genes following NIPBL KD (analysis mediated by RTqPCR). PRKCB (pvalue = 0,0467), EZH2 (ns - pvalue = 0,0736), NPY1R (ns - pvalue = 0,0829) gene appears to be downregulated after NIPBL KD. ICAM1 is significantly upregulated following NIPBL KD (pvalue = 0,0174). P values were determined by ratio paired t-test analysis of the mean Ct of the 4 technical replicates of siCtrl shEF1 -dox, siNIPBL shEF1 +dox conditions. The table shows the position of the EF1 binding sites at the proximal MS, NMS or distal MS, NMS for each gene. The association of the gene with ≥ 20kb gained loops is also indicated (see the supplementary data of the article *[87]* (S30-32)). Proximal, ≤20kb of TSS; Distal, >20kb from TSS; MS: GGAA-microsatellite; NMS, non-microsatellite, " ? ": Information not available in reference article (*[87]* (Unknown information because the gene is not present in the differential gene expression (DEG) coupled with differential interaction determined by RNAseq and Hi-C analysis in EF WT (Endo or Rescue) - EF KD).

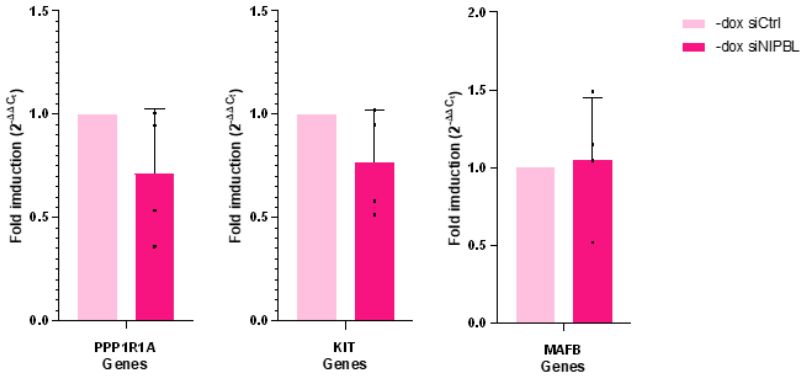
In the **Fig.20**, *PRKCB, EZH2* and *NPY1R* gene expression decreased following NIPBL KD as also observed in EF1 KD (*PRKCB* (pvalue = 0,0467), *EZH2* (ns - pvalue = 0,0736), *NPY1R* (ns - pvalue = 0,0829)). In addition, the ICAM1 expression is significantly upregulated in NIPBL KD as well as in EF1 KD A673 cells (pvalue = 0,0174). In addition, the *PRKCB* gene was also observed to be upregulated following gain of loops in EF1 WT compared to EF1 KD [87].



EF1 binding sites	FEZF1	FCGRT	NROB1
Proximal MS	✓	✓	✓
Proximal NMS	✓	✓	x
Distal MS	x	x	x
Distal NMS	x	x	x
≥ 20kb gained loop	✓	x	x

Figure 21. Abolition of long-range gene regulation effect cause upregulation of some EWS-FLI1 target genes after NIPBL KD (analysis mediated by RTqPCR**)**. mRNA levels of FEZF1 (pvalue = 0,0446), FCGRT (ns - pvalue = 0,3388), NROB1 (ns - pvalue = 0,0731) genes transcription seems to be increased following NIPBL KD. P values were determined by ratio paired t-test analysis of the mean Ct of the 4 technical replicates of siCtrl shEF1 -dox, siNIPBL shEF1 +dox conditions. The table shows the position of the EF1 binding sites at the proximal MS, NMS or distal MS, NMS for each gene. The association of the gene with ≥ 20kb gained loops is also indicated (see the supplementary data of the article *[87]* (S30-32)). Proximal, ≤20kb of TSS; Distal, >20kb from TSS; MS: GGAA-microsatellite; NMS, non-microsatellite, " ? ": Information not available in reference article (*[87]* (Unknown information because the gene is not present in the differential gene expression (DEG) coupled with differential interaction determined by RNAseq and Hi-C analysis in EF WT (Endo or Rescue) - EF KD).

As represented in the **Fig.21**, the *FEZF1, FCGRT and NROB1* genes are upregulated after NIPBL KD (*FEZF1* (pvalue = 0,0446), *FCGRT* (ns - pvalue = 0,3388), *NROB1* (ns - pvalue = 0,0731)). Interestingly, these three genes present EF1 binding sites at less than 20kb of the TSS. It could be feasible that the regulation of these three genes was impaired by long-term gene regulation mediated by the cohesin complex, which interferes with another E-P communication that could regulate gene expression at short distance.



EF1 binding sites				
	Proximal MS	x	?	?
	Proximal NMS	✓	?	?
	Distal MS	x	?	?
	Distal NMS	x	?	?
	≥ 20kb gained loop	✓	?	?

Figure 22. Unaffected EWS-FLI1 target genes after NIPBL KD. (RTqPCR analysis). PPP1R1A (ns - pvalue = 0,1798), KIT (ns - pvalue = 0,1702), MAFB (ns - pvalue = 0,9366) mRNA levels are unaffected by NIPBL KD. P values were determined by ratio paired t-test analysis of the mean Ct of the 4 technical replicates of siCtrl shEF1 -dox, siNIPBL shEF1 +dox conditions. The table shows the position of the EF1 binding sites at the proximal MS, NMS or distal MS, NMS for each gene. The association of the gene with ≥ 20kb gained loops is also indicated (see the supplementary data of the article *[87]* (S30-32)). Proximal, ≤20kb of TSS; Distal, >20kb from TSS; MS: GGAA-microsatellite; NMS, non-microsatellite, " ? ": Information not available in reference article (*[87]* (Unknown information because the gene is not present in the differential gene expression (DEG) coupled with differential interaction determined by RNAseq and Hi-C analysis in EF WT (Endo or Rescue) - EF KD).

NIPBL knockdown does not appear to impair the expression of PPP1R1A, KIT, MAFB (*PPP1R1A* (ns - pvalue = 0,1798), *KIT* (ns - pvalue = 0,1702), *MAFB* (ns - pvalue = 0,9366) (**Fig.22**). In addition, the PPP1R1A gene presents a proximal EF1 binding site, nonetheless this gene is equally associated with 20kb gained loops in EF WT compared to EF1 KD.

Despite some variability, we can observe 3 trends following downregulation of the cohesin loader NIPBL, in the expression of EF1 target genes:

1. NIPBL KD mimics the effect of EF1 KD on its target gene expression (**Fig.20**)

2. The abolition of long-range gene regulation mediated by cohesin/NIPBL upregulates EF1 target genes (**Fig.21**)

3. NIPBL downregulation does not impact EF1 target gene expression (**Fig.22**)

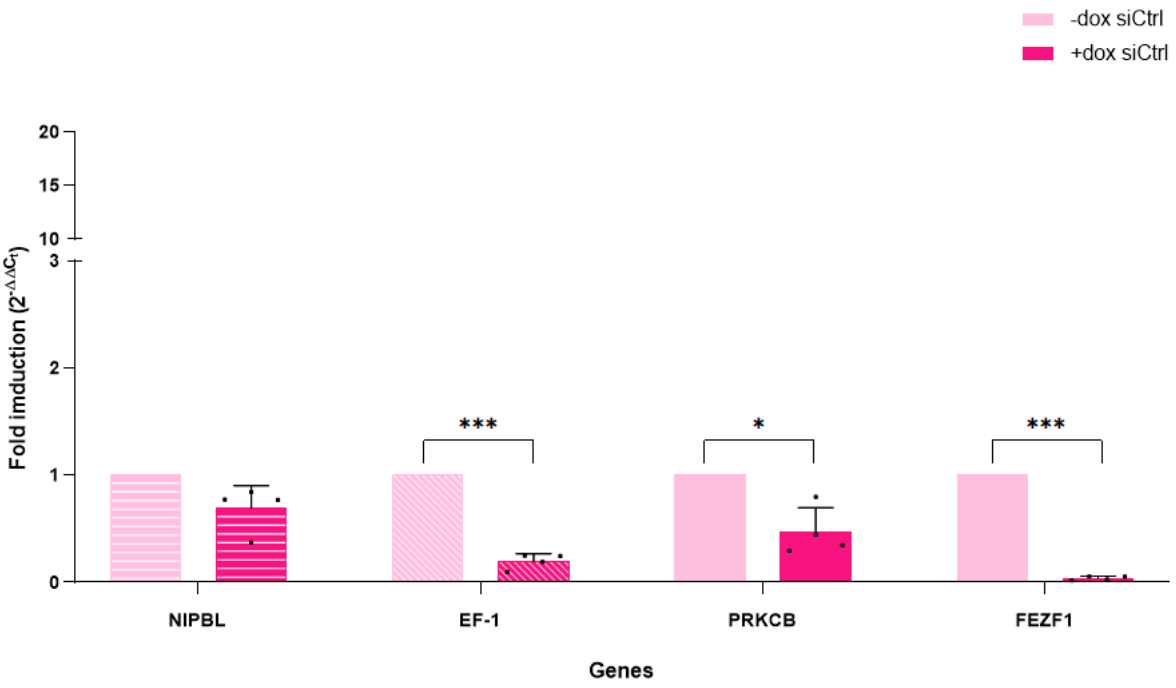


Figure 23. Downregulation and upregulation of EWS-FLI1 target genes following doxycycline induction of shEF1 expression analyzed by RT-qPCR. Measurement of the mRNA levels of the genes: EF1 (pvalue = 0,0003), PRKCB (pvalue=0,0436), FEZF1 (pvalue = 0,0002) showed a significant difference between the two conditions siCtrl -dox/siCtrl +dox. The mRNA level of the NIPBL genes showed no significant difference (pvalue = 0,1839). The p-value was calculated on the average Ct value corrected of the -dox siCtrl and +dox siCtrl by unpaired t-test two-tailed analysis.

The **Fig.23** confirmed the EF1 knockdown after the induction of shEF1 expression by doxycycline (pvalue = 0,0003). The knockdown of EF1 resulted also in significant downregulation of *PRKCB* (pvalue=0,0436) and *FEZF1* (pvalue = 0,0002), two EF1 target genes showed as an example.

Furthermore, it is important to emphasize that we supposed that NIPBL KD, will affect long-range gene regulation because NIPBL is the loader of the cohesin complex and NIPBL promotes the loop-extrusion phenomenon. Nonetheless, in the future we should prove that the cohesin is present at EF1 target genes cis-regulatory regions dependent on NIPBL. Moreover, the presence of loops enhancer-promoter with or without EF1 (or NIPBL) will be evaluated in the near future.

In conclusion, EF1 seems to rely on the loop extrusion mechanism to ensure E-P loops of specific target genes: *EZH2, PRKCB, NPY1R, ICAM1.* In addition, for the *FEZF1, FCGRT, NROB1* EF1 target genes, the long-range gene regulation appears to interfere with short-range regulation. Nonetheless, some EF1 target genes: *PPP1R1A, MAFB, KIT* do not demonstrate expression change upon NIPBL KD, which could be explained by an absence of role of cohesin/NIPBL to mediate these gene expression regulation. Several mechanisms appear to be involved in E-P communication to regulate EF1 target gene expression.

4.3 Determination of the NIPBL recruitment at EF1 target genes cis-regulatory regions

4.3.1 Generation of A673 shEF1 3X-FLAG-Halo (EV) and 3X-FLAG-Halo-NIPBL cell lines

According to our hypothesis, NIPBL is recruited by EF1 to its target gene enhancers to mediate long-range gene regulation. ChIP-qPCR or ChIP-seq against FLAG (3X-FLAG-HALO-NIPBL) will allow us to determine if NIPBL is present at these sites. In addition, doxycycline treatment, which induces shEF1 expression, will allow to determine whether EF1 is a key determinant of NIPBL-cohesin recruitment to EF1 target gene enhancers. This suggests the design of a new A673 cell lines containing either 3x-FLAG-HALO (EV) or 3x-FLAH-HALO-msNIPBL (NIPBL), shEF1 coding sequences and Tet-on machinery stably inserted into the cell genome. However, both transposons containing 3X-FLAG-HALO (EV)/3X-FLAH-HALO-msNIPBL (NIPBL) or shEF1 sequences depend on blasticidin resistance to force A673 cells to maintain these sequences in their genome. The resistance sequence carried into the 3X-FLAG-HALO (EV)/3X-FLAG-HALO-msNIPBL (NIPBL) plasmids was changed to a puromycin resistance sequence because the original plasmids carried a blasticidin sequence as the shEF1 plasmid already stably transfected in A673 cells. Then the plasmid PuroR-3x-FLAG-HALO (EV) or PuroR-3x-FLAH-HALO-msNIPBL (NIPBL) were then transfected into the genome of A673 shEF1 cells using a piggy-bac transposase system.

Overhang and nested PCR allow amplification of the puromycin sequence from the GVV 1898 #7 puro plasmid with the insertion of the Mfe1/Apa1/HspA1 restriction sites (**Fig.24**).

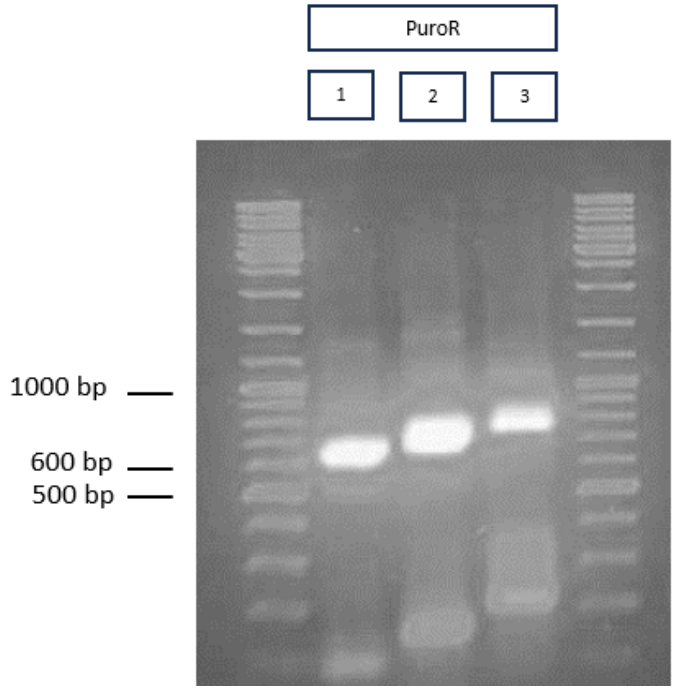


Figure 24. **Overhang and nested PCR to amplify the puromycin sequence and the added Mfe1, KspA1, Apa1 restriction sites.** The amplified puromycin sequence of the first, second, third PCR should have respectively a length of 670bp, 732bp, 783bp corresponding to the ~ 600-700bp, ~ 700-800bp, ~ 750-850bp bands shown in the 2% agarose gel.

The 670bp, 732bp, 783bp bands were purified. The purified bands and the EV/NIPBL plasmids were then digested by the Mfe1/HpA1 restriction enzymes. Then, the EV plasmid and NIPBL digested plasmid were purified to eliminate the blasticidin resistance sequence and preserve the rest of the backbone (**Fig.25**). Then, the puromycin sequence was ligated to the corresponding digested EV or NIPBL plasmid. Heat shock or electroporation was used respectively to introduce the PuroR-3x-FLAG-HALO (EV) or PuroR-3x-FLAH-HALO-msNIPBL (NIPBL) plasmids into DH5α E.Coli to amplify the plasmids (**Fig26**.).

The plasmids were then purified after pre-culture of an adequate clone. Sequencing allowed the selection of the clone 1 PuroR-3X-FLAG-HALO (EV) and the clone 7 PuroR-3X-FLAG-HALO-msNIPBL(NIPBL) in order to transfect these plasmids into A673 shEF1 cells using the piggybac system.

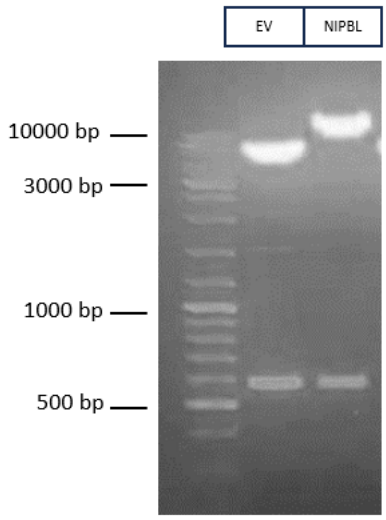


Figure 25. Digestion of the BlastR-3X-FLAG-HALO (EV) and BlastR -3X-FLAG-HALO-msNIPBL (NIPBL) plasmids by the Mfe1/HpA1 restriction enzymes respectively. Bands of ~ 6303 bp or ~ 14700 bp for the BlastR-3X-FLAG-HALO (EV) plasmid digestion and BlastR-3X-FLAG-HALO-msNIPBL (NIPBL) plasmid digestion respectively.

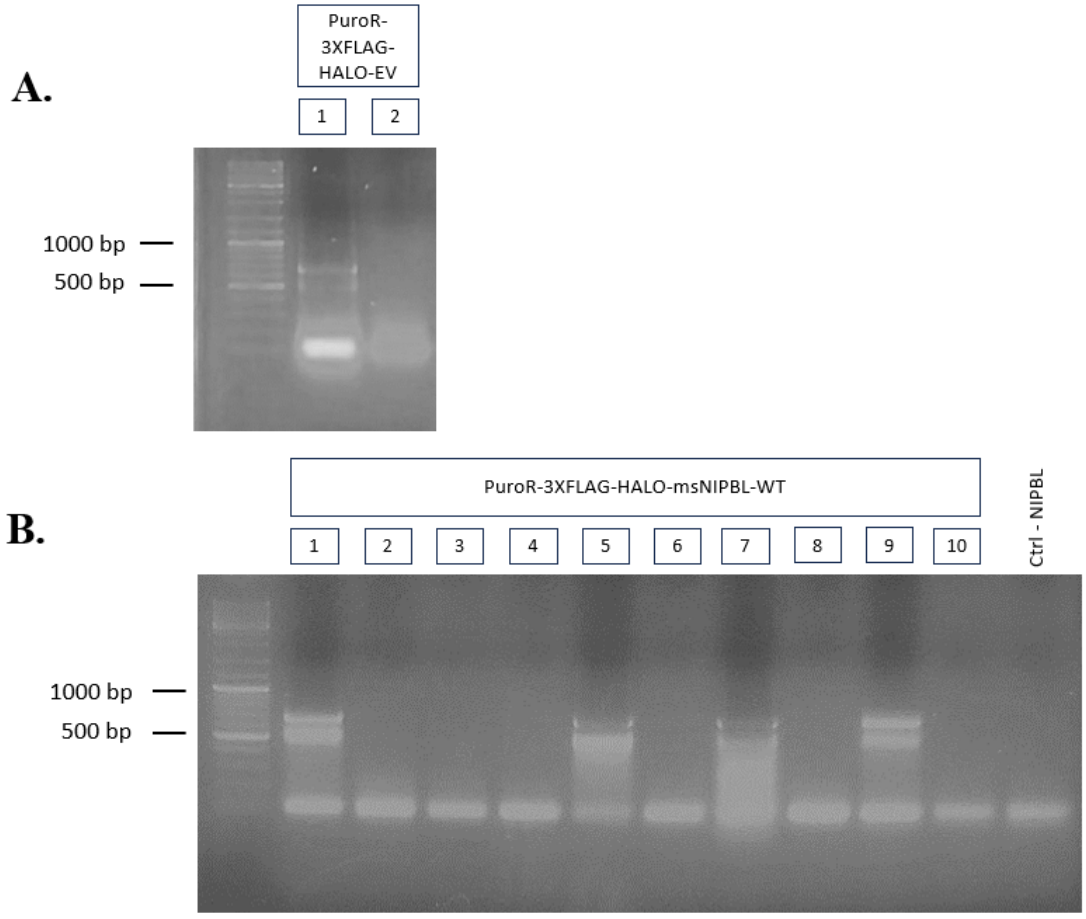


Figure 26. Colony PCR to verify the insertion of the puromycin sequence into the Apa1/HspA1 digested EV plasmid or in the Mfe1/HspA1 digested NIPBL plasmid. A. Successful insertion of the PuroR-3x-FLAG-HALO (EV) plasmid into DH5α E.Coli by heatshock transformation (670pb)**B.** Effective insertion of the PuroR-3x-FLAG-HALO-msNIPBL plasmid into DH5α E.Coli by electroporation transformation in the 1st, 5th, 7th, 9th colonies (670pb). NIPBL negative controls correspond to the BlastR-3X-FLAG-HALO-msNIPBL (NIPBL) plasmid digested with Mfe1/HpA1 restriction enzymes without ligation of the digested puromycin sequence.

The A673 shEF1 puroR-EV and puro-NIPBL were continuously selected with 1.1000 blasticidin, 1:500 zeocyn, 1:10 000 puromycin added in DMSO medium. Cells were then sorted into low, medium, high HALO expression (3X-FLAG-HALO(EV) or 3X-FLAG-HALO-msNIPBL (NIPBL), using the FACS-Flow cytometry GIGA platform (**Fig.27**). Despite the cells express 3x-FLAG-HALO or 3x-FLAG-HALO-msNIPBL exogenous proteins, the Western blot following the protein extraction of these cells did not demonstrate various expression of the recombinant proteins between the A673 EV/NIPBL Low, Medium, High. Nevertheless, the confirmed expression of the exogenous proteins allows us to perform the ChIP-qPCR analysis 10 days after the induction of shEF1 expression by doxycycline treatment.

Figure 27. Verified the 3XFLAG-HALO (EV) or the 3XFLAG-HALO-msNIPBL (NIPBL) expression in the Low, Medium, High A673 shEF1 EV or NIPBL cells. HSP90 is a loading control.

4.3.2 NIPBL recruitment at EF1 target genes cis-regulatory element revealed by ChIP-qPCR

To determine whether NIPBL is recruited by EF1 at EF1 target genes cis-regulatory regions, a ChIP-qPCR was performed using the A673 shEF1 +/-dox EV/NIPBL cells (see cell design above). For this purpose, different primers designed against cis-regulatory regions of EF1 target genes (see ChIP primer lists in appendix) were. The primers enable the amplification, by qPCR, of EF1 binding sites where NIPBL could be recruited.

As mentioned above, shEF1 expression, after doxycycline induction, allows the degradation of EF1 mRNA. EF1 KD will indicate whether the EF1 presence at EF1 target genes cis-regulatory regions is determinant for the recruitment of NIPBL at these sites. NIPBL will then recruit the cohesin complex to form E-P loops in order to mediate long-range gene regulation.

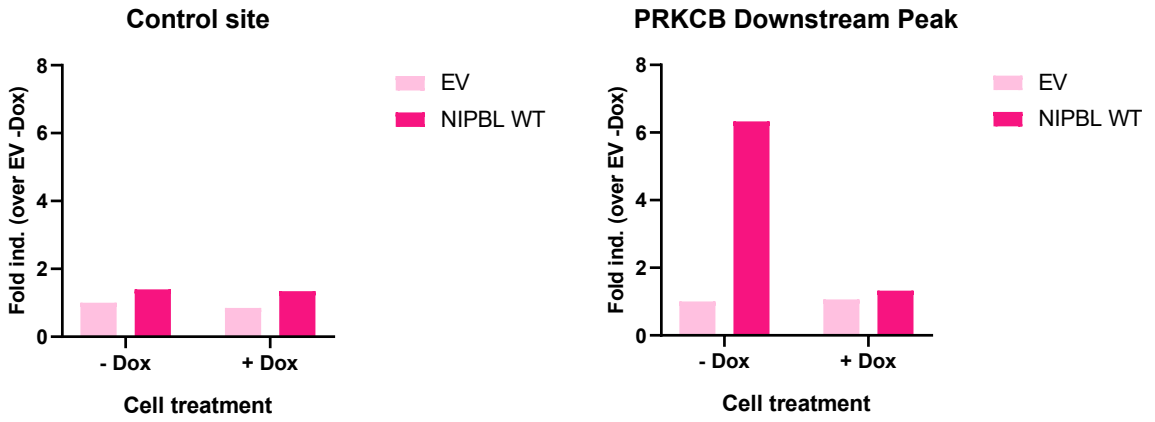


Figure 28. Recruitment of NIPBL by EF1 at the PRKCB cis-regulatory element, at -60kb from the TSS, revealed by ChIP-qPCR. The A673 shEF1 +/- dox EV/NIPBL cells were treated with doxycycline to induce the expression of shEF1 to target EF1 mRNA. The cis-regulatory region of the PRKCB gene, one of EF1 target genes, is localized at -60 kb from the TSS of the PRKCB gene. At the moment, we only have one replicate due to time constraints.

This first replicate is very promising because we are able to find out that NIPBL is present at the downstream cis-regulatory regions of the *PRKCB* genes, at -60kb of the TSS. Furthermore, the binding of EF1 at the enhancer of its target genes allows the recruitment of NIPBL proteins as assessed by the doxycycline condition. Without its presence, NIPBL cannot be recruited to these sites.

In conclusion, further replicates will be performed to confirm this first result and, to discover if the cis-regulatory regions of other genes, such as those discussed in the section 5.1: *EZH2, NPY1R, ICAM1, FEZF1, FCGRT, NROB1,* could be potential sites of NIPBL recruitment by EF1. Determining the recruitment of NIPBL by EF1 at these downstream cis-regulatory regions of the *PRKCB* genes could confirm our hypothesis: “EF1 could recruit NIPBL at cis-regulatory regions by direct physical interaction, in order to engender E-P loops to regulate EF1 target genes expression”.

V. Discussion and Perspective

5.1 Physical interaction between NIPBL and EWS-FLI1 C-ter

Proximity Ligation Assays (PLA) experiment in U2As cells non-transfected or transfected with eGFP-EWS-FLI1, performed by Dr. Fettweis, had already demonstrated spatial proximity between NIPBL and EWS-FLI1 in vivo (see section 2.1). Using the GST pull-down assay, we demonstrated a direct physical interaction between NIPBL and the c-terminal domain of EWS-FLI1 in vitro (see section 4.1.2.4). The FLI1full construct showed direct interaction with the exogenous NIPBL. Which is crucial to establish a new model of transcriptional regulation dependent on the direct interaction of NIPBL with EWS-FLI1 and possibly other ETS/FET oncogenic fusions or by extension other transcription factors, as observed with GR [54]. Furthermore, the project is essential to eventually lead to a potential therapeutic treatment for Ewing's sarcoma tumors.

Regarding the interaction of FLI1HC with NIPBL, we observed the interaction with one replicate, but not as significant for the second. Furthermore, as observed in the WB for the expression quantification, GST-FLI1potential and GST-FLI1HC have been reported to undergo important protein truncation and degradation during the production, bacterial lysis and purification steps. For a reason that we ignored so far, the FLI1potential and the FLI1highconfidence chimeras are highly degraded during the immunoprecipitation despite the presence of 2X protease inhibitor. As a consequence, we cannot assume that the FLI1HC construct interacts less than the FLI1full construct because the FLI1HC protein level is dramatically lower compared to the full version of FLI1. Therefore, improving the production quality and reducing protein degradation will be crucial for the following GST pull down assay.

The production of recombinant GST proteins could be improved to avoid the formation of truncated proteins on different points. Several studies have shown that lowering the temperature to 15-20°C during ON production has improved protein expression [90]. However, pre-incubation at 47°C may be essential to induce chaperone production to allow correct protein folding [90].

As a future perspective, GST pull down assay will be performed again with the FLI1pot and FLI1HC construction in improved IP and production conditions to decrease the truncation or degradation of the GST chimeras. Once the production issues are resolved and the results for the FLI1HC known, GST-pull down assay may be continued with the FLI1 LLXLL mutants (FLI1-AAXLL, FLI1-LLXAA, FLI1-AAXAA) in the case of a confirmed interaction between NIPBL and the FL1HC domain. The objective will be to determine if we are able to lose the interaction with NIPBL, which could mean that LLXLL motif are determinant to realize the interaction between EWS-FLI1 and NIPBL, as already observed between GR and NIPBL by co-IP (data not published).

Finally, since we can observe an interaction between NIPBL and the FL1 domain of EF1, we will confirm this interaction in Hela cells by co-IP. Thus, exogenous expression of *EWS-FLI1* and *NIPBL* will be used to increase the probability of observing this transient complex. Additional in vivo experiments would also need to be performed. However, the different sizes of EWS-FLI1 (68 kDa) and NIPBL (320 kDa) proteins avoid performing several methods such as FRET, PCA (Protein Fragment complementation assays) analysis like BiFC (Bimolecular fluorescence complementation), Split-luciferase complementation or split-ubiquitin system. Nonetheless, we could cut the NIPBL sequence into smaller pieces to test the interaction of these domains with EWS-FLI1 in GPCA. Co-IP experiment in A673 EV/NIPBL cells could also be performed to confirm the interaction in a more relevant model than Hela cells.

5.2 Long-range gene regulation mediated by NIPBL/cohesin loop extrusion influence EF1 target gene regulation

To further investigate the role of the interaction between the cohesin loader NIPBL and the oncogenic transcription factor EF1, the transcriptome changes in A673 -/+ dox cells were analyzed for some EF1 target genes upon NIPBL knockdown. (See section 4.2)

Three trends seemed to emerge:

1. NIPBL KD mimics the effects of EF1 KD on its target gene expression

2. NIPBL KD strengthens EF1 target gene expression, while EF1 KD leads to their downregulation

3. NIPBL downexpression does not affect EF1 target genes expression

The influence of cohesin/NIPBL-mediated long-range gene regulation appears to correlate with the position of the EF1 binding sites at cis-regulatory region relative to the TSS of EF1 target genes. When EF1 binds a GGAA micro-satellite or non-microsatellite site on an enhancer located at more than 20kb from the TSS (referred to as distal in the study), cohesin/NIPBL-mediated long-range gene regulation through E-P loops appears to be required to mediate EF1 target gene transcription[87]. To the contrary E-P communication appears to be independent or negatively affected by long-range gene regulation when the enhancer is close to the TSS (less than 20kb - referred to as proximal according to same manuscript [87]). Further analysis of more genes will be required to confirm our three identified trends (see just above). Following this RTqPCR analysis, an RNAseq experiment will be performed on the same samples. The RNAseq experiment will determine whether the observed pattern can be extended to other EF1 target genes even if we don't anticipate every EF1-dependent gene to rely on NIPBL/cohesin looping regulation. The RNAseq experiment could also be coupled with an in silico analysis of genomic distances between enhancers and associated promoters of EF1 target genes. In addition, these data will be able to be analyzed thanks to the work of Marie Harmel, a Master 1

bioinformatics trainee, who reassembled and validated the EF1 binding sites derived from different ChIP-seq analyses performed by different research groups.

Our results are in some aspects consistent with other research that has analyzed the effect of STAG2 loss on enhancer-promoter loops in A673 cell lines [86]. In this study, they found that the loops governing the E-P contact of chromatin regions located at large genomic distances (> 60 kb) appeared to decrease in frequency after STAG2 Knock Out. Whereas chromatin regions separated by shorter genomic distance (< 60 kb) appeared to increase in frequency after STAG2 KO [86]. As already known and confirmed by their and our own results, STAG2 as NIPBL are essential to mediate long-range chromatin interaction between enhancers and promoters but interferes with the formation of E-P loops at a smaller scale.

Furthermore, according to our results, the E-P communication mediated by the cohesin complex could collaborate or compete with other E-P communications at shorter distances. EWS-FLI1 is known to mediate liquid-liquid phase separation thanks to the LCD of the EWS region [91]. As a consequence, EWS-FLI1 forms highly concentrated hubs at GGAA microsatellite sites [91]. They determined in MSC (mesenchymal stem cells) that the EWS(Y37S)-FLI1 mutant, which is mutated in the tyrosine 37 relevant for the mediation of LLPS, leads to downregulated expression of some EF1 target genes [92]. In their analysis, the genes PRKCB, EZH2, FCGRT, FEZF1, NROB1, PPP1R1A, MAFB, KIT are downregulated after the mutation [92]. Each gene identified above is also present in our analysis. According to our results, it appears that several of these genes are regulated thanks to NIPBL and by extension by a loop extrusion mechanism, with a potential LLPS component. We could speculate that E-P communication mediated by the cohesin complex could collaborate or compete with other E-P communications at longer and smaller distances, such as with the LLPS. In addition, the KD of RAD21, a subunit of the cohesin complex, has demonstrated to reduce the dynamic fraction associated with the intrinsically disordered regions (IDRs), and by extension LLPS [54]. Furthermore, in silico analysis revealed that the combination of phase separation and loop extrusion data could recapitulate more accurately Hi-C analysis compared to solely loop extrusion models [93].

To conclude, long-range gene regulation favorized by the main mediator (NIPBL) of the loop extrusion mechanism, appears to regulate some EF1 target gene expression, but also in collaboration or in competition with other E-P communication**.** According to the literature, the low complexity domains of EF1 could be relevant to mediate this regulation at a short scale by the phase separation mechanism.

5.3 Recruitment of NIPBL by EF1 at distal binding sites regulating EF1 target genes expression

The GST-pull down assay determined the physical interaction of the EWS-FLI1 C-terminal domain and the cohesin loader, NIPBL (see section 4.1.2.4). In addition, NIPBL knock-down in A673 cells revealed that some EF1-dependent gene are regulated by NIPBL (see section 4.2). The next step was then to investigate the recruitment of NIPBL at EF1 target genes cis-regulatory sites. Therefore, we performed a NIPBL ChIP-qPCR experiment at specific EF1 chromatin locations.

New A673shEF1 EV/NIPBL cell lines were generated following the transfection of the PuroR-3XFLAG-HALO (EV) or PuroR-3XFLAG-Halo-NIPBL plasmids into A673shEF1 cells by the piggybac system. The expression of shEF1 induced after doxycycline treatment to downregulate EF1 allowed us to show whether the presence of NIPBL at specific EF1 enhancers is dependent on EF1 to recruit NIPBL to these sites. Due to time constraints, we only have for the moment one replicate. However, the results are extremely promising.

ChIP-qPCR in A673 shEF1 +/- dox EV/NIPBL cells determines that NIPBL is present at the downstream cis-regulatory regions of the *PRKCB* genes, at 60 kB from the TSS, but also that EF1 is essential for the recruitment of NIPBL at these sites (see results section 4.3.2). As already highlighted for GR and NIPBL [54], our results strongly suggest that TF, as EWS-FLI1 or GR, could recruit NIPBL at cis-regulatory regions to regulate the expression of their target genes by long-range gene regulation. NIPBL will then stabilize cohesin at this site and promote loop extrusion to bring enhancers at a closer proximity of its associated promoters.

In the near future, multiple replicates of the ChIP-qPCR analysis will be performed to validate our results. In addition, ChIP-Seq experiments will be performed to determine if NIPBL is present at specific sites and if this recruitment at these sites is dependent on EF1. Additionally, if we coupled them with other techniques such as RNA-seq and micro-C, will enable us to validate, with more certainty, this long-range gene regulation between EF1 target genes enhancers and associated promoters. Furthermore, a ChIP-qPCR against RAD21/SMC1/SMC3 could confirm the presence of the cohesin complex at these sites.

5.4 NIPBL KD proliferation and migration phenotype – promising results leading to a potential treatment

The two analyses previously performed by the Dr. Fettweis, the proliferation assay and the SIR assay, indicated that NIPBL KD impaired cell proliferation and cell migration, two phenotypes, respectively characteristic of EWS-FLI1 high state and EWS-FLI1 low state (See section 2.1).[80]

In the proliferation assay with A673 shEF1 -dox siCtrl/siNIPBL, we noticed an importance of cell proliferation after NIPBL KD compared to the control conditions. The proliferation phenotypes were similar after EW1 KD following shEF1 expression after doxycycline treatment.

The SIR assay, which measures the rate of A673 shEF1 -dox siCtrl/siNIPBL cells in µm/h, indicated that the cell migration rate is consistently lower after NIPBL KD compared to the control (A673 shEF1 -dox siCtrl).

In contrast, STAG2 LOF in A673 cells increase cell migration [85], [86]. To date, we have not understood yet why STAG2, a cohesin subunit, and NIPBL, the cohesin loader, demonstrate two opposite phenotypes in cell migration while they are both required for E-P communications. STAG2 KO present a similar phenotype as EWS-FLI1low state by favoring cell migration [85], [86]. The opposite phenotype of NIPBL KD indicates that the impairment of NIPBL function could be a relevant treatment to interfere at the same time on proliferation and migration.

Furthermore, we are currently unable to understand why STAG2 and NIPBL have opposite effects on cell migration, while both are relevant for mediating E-P loops. In addition, these different phenotypes are unexpected because STAG2 showed its dependence on NIPBL for its presence at CTCF or non-CTCF sites as mentioned in the following article [94]. Another unexpected result is that the *FCGRT* gene decreased after STAG2 LOF compared to our results, because this gene was upregulated after NIPBL KD. However, just like our results, NPY1R decreased after NIPBL KD or after STAG2 LOF for the comparison study [86].

In addition, the RTqPCR analysis mediated on A673shEF1+/-dox siCtrl siNIPBL cells have shown that the expression of the *ICAM1* gene was significantly upregulated after NIPBL KD as observed after EF1 KD. *ICAM1* gene is coding for the intercellular adhesion molecule 1 (ICAM-1) protein belonging to the immunoglobulin superfamily which is a transmembrane protein involved in cell-to-cell interactions. These surface proteins are expressed in leukocytes, endothelial cells and in different types of cancer. ICAM-1 is known to promote T-cell activation by cell tumor/T-cell interaction through its interaction with the LFA1 proteins. ICAM1 or LFA1 are expressed on the surface of cell tumor or leukocytes, respectively [95]. Nonetheless in EWS-FLI1 low cells, ICAM1 upregulation did not enhance T-cell activation. Rather, EWS-FLI1 low cells escape T-cell activation to mediate tumor apoptosis by the

expression of PD-L1 and PD-L2. PD-L1 and PD-L2 are the ligands of the programmed cell death protein 1 (PD-1) receptor, and their binding promotes self-tolerance and suppresses immune response [96]. In contrary to the function of the T-cell activation due to the interaction of ICAM1 and LFA1, ICAM1 is also suspected to promote tumor progression by influencing PD-L1 expression on exosome originating from metastatic melanoma tumor cells [97]. It appears that two opposite roles are observed for ICAM1 to mediate T-cell activation or at contrary to suppress immunity response as observed in metastatic melanoma tumor or EF1 low cells. For the moment, the relation of the expression of ICAM1 and activation or inactivation of T-cell activation seems difficult to interpret without a global view that will be brought by additional RNAseq experiment (see section 5.2). In our results, the significant upregulation of ICAM1 in siNIPBL A673 cells was also unexpected because cell migration rate decreased in A673 cells KD, while ICAM1 expression is often associated with metastatic properties of EWS-FLI1low cells. Further analysis will have to be performed in A673 EF1 low state (shEF1 +dox) to determine by which mechanism the expression of *ICAM1* and *CD274* gene (coding for PD-L1) are regulated.

To conclude, the STAG2 opposite effect in terms of cell migration and the upregulation of ICAM in A673 siNIPBL KD, detected in our RTqPCR analysis, are difficult to interpret in comparison with the cell migration decrease phenotype observed in NIPBL KD.

Nonetheless, the impairment of cell migration in NIPBL KD is a promising result that may lead to a potential treatment of Ewing's sarcoma tumor through its double role to impair proliferation and migration.

VI. Conclusion

Recently, EWS-FLI1 has been identified to reprogram the 3D chromatin structure of Ewing's sarcoma cells at different scales: A/B compartment, topologically associating domains (TADs) boundaries, and enhancer-promoter loops. In the following article [87], it was demonstrated that new chromatin loops are initiated by EWS-FLI1 binding at GGAA or GGAA microsatellite sites, which correlate with enhancer mark activation and EWS-FLI1 target gene upregulation.

In line of this work, we presented some evidence of the role of the cohesin loader, NIPBL, in influencing the transcriptional regulation of EWS-FLI1 target genes. The recruitment of NIPBL at cis-regulatory regions is suggested to be dependent of direct interaction between the cohesin loader, NIPBL, and the oncogenic fusion, EWS-FLI1. Long-range gene regulation would be realized by the suggested formation of enhancer-promoter loops following the DNA extrusion mediated by the cohesin complex. Loops extrusion enabled close proximity between enhancer and promoters to engender transcriptional regulation[40] . Likely other transcription factors, as already observed with the glucocorticoid receptor (GR), could follow the same mechanism to control their gene regulation [54].

We suggest that long-range gene regulation mediated by NIPBL and loop extrusion mechanisms may govern a subset of EWS-FLI1 target gene expression with the collaboration of other enhancer-promoter communication, speculated to be the LLPS (Liquid-Liquid Phase Separation). Our results may suggest that the influence of long-range gene regulation might depend on the distance of EWS-FLI1 binding sites with the transcriptional start site (TSS) of EWS-FLI1 target genes.

From a clinical perspective, the direct physical interaction between NIPBL and EWS-FLI1 could be a potential investigation for a therapeutic treatment. Furthermore, the cell migration rate and the cell proliferation impairment after NIPBL KD could lead to a complete therapy in negatively affecting the proliferation characteristic of EWS-FLI1 high cells, and successively impairing the migration property leading to metastasis of EWS-FLI1 low cells.

Bibliography

[1] J. Lammerding, "Mechanics of the Nucleus," in *Comprehensive Physiology*, R. Terjung, Ed., 1st ed.Wiley, 2011, pp. 783–807. doi: 10.1002/cphy.c100038.

[2] T. Misteli, "The Self-Organizing Genome: Principles of Genome Architecture and Function," *Cell*, vol. 183, no. 1, pp. 28–45, Oct. 2020, doi: 10.1016/j.cell.2020.09.014.

[3] G. Felsenfeld and M. Groudine, "Controlling the double helix," *Nature*, vol. 421, no. 6921, pp. 448–453, Jan. 2003, doi: 10.1038/nature01411.

[4] K. Maeshima, S. Ide, and M. Babokhov, "Dynamic chromatin organization without the 30-nm fiber," *Current Opinion in Cell Biology*, vol. 58, pp. 95–104, Jun. 2019, doi: 10.1016/j.ceb.2019.02.003.

[5] S. Kadauke and G. A. Blobel, "Chromatin loops in gene regulation," *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms*, vol. 1789, no. 1, pp. 17–25, Jan. 2009, doi: 10.1016/j.bbagrm.2008.07.002.

[6] G. Fudenberg, M. Imakaev, C. Lu, A. Goloborodko, N. Abdennur, and L. A. Mirny, "Formation of Chromosomal Domains by Loop Extrusion," *Cell Reports*, vol. 15, no. 9, pp. 2038–2049, May 2016, doi: 10.1016/j.celrep.2016.04.085.

[7] I. F. Davidson, B. Bauer, D. Goetz, W. Tang, G. Wutz, and J.-M. Peters, "DNA loop extrusion by human cohesin," *Science*, vol. 366, no. 6471, pp. 1338–1345, Dec. 2019, doi: 10.1126/science.aaz3418.

[8] Y. Guo *et al.*, "CRISPR Inversion of CTCF Sites Alters Genome Topology and Enhancer/Promoter Function," *Cell*, vol. 162, no. 4, pp. 900–910, Aug. 2015, doi: 10.1016/j.cell.2015.07.038.

[9] E. de Wit *et al.*, "CTCF Binding Polarity Determines Chromatin Looping," *Molecular Cell*, vol. 60, no. 4, pp. 676–684, Nov. 2015, doi: 10.1016/j.molcel.2015.09.023.

[10] J. Dekker and L. Mirny, "The 3D Genome as Moderator of Chromosomal Communication," *Cell*, vol. 164, no. 6, pp. 1110–1121, Mar. 2016, doi: 10.1016/j.cell.2016.02.007.

[11] J. R. Dixon *et al.*, "Topological domains in mammalian genomes identified by analysis of chromatin interactions," *Nature*, vol. 485, no. 7398, pp. 376–380, May 2012, doi: 10.1038/nature11082.

[12] T. Cremer and M. Cremer, "Chromosome Territories," *Cold Spring Harbor Perspectives in Biology*, vol. 2, no. 3, pp. a003889–a003889, Mar. 2010, doi: 10.1101/cshperspect.a003889.

[13] F. Uhlmann, D. Wernic, M.-A. Poupart, E. V. Koonin, and K. Nasmyth, "Cleavage of Cohesin by the CD Clan Protease Separin Triggers Anaphase in Yeast," *Cell*, vol. 103, no. 3, pp. 375–386, Oct. 2000, doi: 10.1016/S0092-8674(00)00130-6.

[14] G. Wutz *et al.*, "Topologically associating domains and chromatin loops depend on cohesin and are regulated by CTCF, WAPL, and PDS5 proteins," *The EMBO Journal*, vol. 36, no. 24, pp. 3573–3599, Dec. 2017, doi: 10.15252/embj.201798004.

[15] A.-S. Thomas-Claudepierre *et al.*, "The cohesin complex regulates immunoglobulin class switch recombination," *Journal of Experimental Medicine*, vol. 210, no. 12, pp. 2495–2502, Nov. 2013, doi: 10.1084/jem.20130166.

[16] I. F. Davidson, "Rapid movement and transcriptional re-localization of human cohesin on DNA," *The EMBO Journal*, Oct. 2016, doi: 10.15252/embj.201695402.

[17] I. Sumara, E. Vorlaufer, C. Gieffers, B. H. Peters, and J.-M. Peters, "Characterization of Vertebrate Cohesin Complexes and Their Regulation in Prophase," *Journal of Cell Biology*, vol. 151, no. 4, pp. 749–762, Nov. 2000, doi: 10.1083/jcb.151.4.749.

[18] C. Hoencamp and B. D. Rowland, "Genome control by SMC complexes," *Nat Rev Mol Cell Biol*, May 2023, doi: 10.1038/s41580-023-00609-8.

[19] Z. Shi, H. Gao, X. Bai, and H. Yu, "Cryo-EM structure of the human cohesin-NIPBL-DNA complex," 2020.

[20] D. Alonso-Gil and A. Losada, "NIPBL and cohesin: new take on a classic tale," *Trends in Cell Biology*, p. S0962892423000478, Apr. 2023, doi: 10.1016/j.tcb.2023.03.006.

[21] Y. Kim, Z. Shi, H. Zhang, I. J. Finkelstein, and H. Yu, “Human cohesin compacts DNA by loop extrusion,” 2019.
[22] J. H. I. Haarhuis *et al.*, “The Cohesin Release Factor WAPL Restricts Chromatin Loop Extension,” *Cell*, vol. 169, no. 4, pp. 693-707.e14, May 2017, doi: 10.1016/j.cell.2017.04.013.
[23] N. L. Arruda, A. F. Bryan, and J. M. Dowen, “PDS5A and PDS5B differentially affect gene expression without altering cohesin localization across the genome,” *Epigenetics & Chromatin*, vol. 15, no. 1, p. 30, Aug. 2022, doi: 10.1186/s13072-022-00463-6.
[24] G. Wutz *et al.*, “ESCO1 and CTCF enable formation of long chromatin loops by protecting cohesinSTAG1 from WAPL,” *eLife*, vol. 9, p. e52091, Feb. 2020, doi: 10.7554/eLife.52091.
[25] I. F. Davidson and J.-M. Peters, “Genome folding through loop extrusion by SMC complexes,” *Nat Rev Mol Cell Biol*, vol. 22, no. 7, pp. 445–464, Jul. 2021, doi: 10.1038/s41580-021-00349-7.
[26] B. W. Bauer *et al.*, “Cohesin mediates DNA loop extrusion by a ‘swing and clamp’ mechanism,” *Cell*, vol. 184, no. 21, pp. 5448-5464.e22, Oct. 2021, doi: 10.1016/j.cell.2021.09.016.
[27] T. L. Higashi, G. Pobegalov, M. Tang, M. I. Molodtsov, and F. Uhlmann, “A Brownian ratchet model for DNA loop extrusion by the cohesin complex,” *eLife*, vol. 10, p. e67530, Jul. 2021, doi: 10.7554/eLife.67530.
[28] A. Losada, T. Yokochi, R. Kobayashi, and T. Hirano, “Identification and Characterization of Sa/Scc3p Subunits in the Xenopus and Human Cohesin Complexes,” *Journal of Cell Biology*, vol. 150, no. 3, pp. 405–416, Aug. 2000, doi: 10.1083/jcb.150.3.405.
[29] A. Kojic *et al.*, “Distinct roles of cohesin-SA1 and cohesin-SA2 in 3D chromosome organization,” *Nat Struct Mol Biol*, vol. 25, no. 6, pp. 496–504, Jun. 2018, doi: 10.1038/s41594-018-0070-4.
[30] F. C. Soardi *et al.*, “Familial STAG2 germline mutation defines a new human cohesinopathy,” *npj Genomic Med*, vol. 2, no. 1, p. 7, Mar. 2017, doi: 10.1038/s41525-017-0009-4.
[31] A. Cuadrado and A. Losada, “Specialized functions of cohesins STAG1 and STAG2 in 3D genome architecture,” *Current Opinion in Genetics & Development*, vol. 61, pp. 9–16, Apr. 2020, doi: 10.1016/j.gde.2020.02.024.
[32] M. Merkenschlager and E. P. Nora, “CTCF and Cohesin in Genome Folding and Transcriptional Gene Regulation,” *Annu. Rev. Genom. Hum. Genet.*, vol. 17, no. 1, pp. 17–43, Aug. 2016, doi: 10.1146/annurev-genom-083115-022339.
[33] S. A. Tammen, S. Friso, and S.-W. Choi, “Epigenetics: The link between nature and nurture,” *Molecular Aspects of Medicine*, vol. 34, no. 4, pp. 753–764, Jul. 2013, doi: 10.1016/j.mam.2012.07.018.
[34] S. V. Razin, S. V. Ulianov, and O. V. Iarovaia, “Enhancer Function in the 3D Genome,” *Genes*, vol. 14, no. 6, p. 1277, Jun. 2023, doi: 10.3390/genes14061277.
[35] A. Panigrahi and B. W. O’Malley, “Mechanisms of enhancer action: the known and the unknown,” *Genome Biol*, vol. 22, no. 1, p. 108, Dec. 2021, doi: 10.1186/s13059-021-02322-1.
[36] F. Spitz and E. E. M. Furlong, “Transcription factors: from enhancer binding to developmental control,” *Nat Rev Genet*, vol. 13, no. 9, pp. 613–626, Sep. 2012, doi: 10.1038/nrg3207.
[37] R. K. Suto, M. J. Clarkson, D. J. Tremethick, and K. Luger, “Crystal structure of a nucleosome core particle containing the variant histone H2A.Z,” *nature structural biology*, vol. 7, no. 12, 2000.
[38] A. Tafessu and L. A. Banaszynski, “Establishment and function of chromatin modification at enhancers,” *Open Biol.*, vol. 10, no. 10, p. 200255, Oct. 2020, doi: 10.1098/rsob.200255.
[39] H. Tachiwana *et al.*, “Structures of human nucleosomes containing major histone H3 variants,” *Acta Crystallogr D Biol Crystallogr*, vol. 67, no. 6, pp. 578–583, Jun. 2011, doi: 10.1107/S0907444911014818.
[40] T. M. Popay and J. R. Dixon, “Coming full circle: On the origin and evolution of the looping model for enhancer–promoter communication,” *Journal of Biological Chemistry*, vol. 298, no. 8, p. 102117, Aug. 2022, doi: 10.1016/j.jbc.2022.102117.
[41] D. Carter, L. Chakalova, C. S. Osborne, Y. Dai, and P. Fraser, “Long-range chromatin regulatory interactions in vivo,” *Nat Genet*, vol. 32, no. 4, pp. 623–626, Dec. 2002, doi: 10.1038/ng1051.
[42] B. Tolhuis, R.-J. Palstra, E. Splinter, F. Grosveld, and W. de Laat, “Looping and Interaction between Hypersensitive Sites in the Active ␣-globin Locus,” *Molecular Cell*.

[43] E. E. M. Furlong and M. Levine, “Developmental enhancers and chromosome topology,” *Science*, vol. 361, no. 6409, pp. 1341–1345, Sep. 2018, doi: 10.1126/science.aau0320.
[44] N. J. Rinzema *et al.*, “Building regulatory landscapes reveals that an enhancer can recruit cohesin to create contact domains, engage CTCF sites and activate distant genes,” *Nat Struct Mol Biol*, vol. 29, no. 6, pp. 563–574, Jun. 2022, doi: 10.1038/s41594-022-00787-7.
[45] M. I. Robson, A. R. Ringel, and S. Mundlos, “Regulatory Landscaping: How Enhancer-Promoter Communication Is Sculpted in 3D,” *Molecular Cell*, vol. 74, no. 6, pp. 1110–1122, Jun. 2019, doi: 10.1016/j.molcel.2019.05.032.
[46] M. H. Kagey *et al.*, “Mediator and cohesin connect gene expression and chromatin architecture,” *Nature*, vol. 467, no. 7314, pp. 430–435, Sep. 2010, doi: 10.1038/nature09380.
[47] O. Symmons *et al.*, “Functional and topological characteristics of mammalian regulatory domains,” *Genome Res.*, vol. 24, no. 3, pp. 390–400, Mar. 2014, doi: 10.1101/gr.163519.113.
[48] T.-H. S. Hsieh, C. Cattoglio, E. Slobodyanyuk, A. S. Hansen, X. Darzacq, and R. Tjian, “Enhancer–promoter interactions and transcription are largely maintained upon acute loss of CTCF, cohesin, WAPL or YY1,” *Nat Genet*, vol. 54, no. 12, pp. 1919–1932, Dec. 2022, doi: 10.1038/s41588-022-01223-8.
[49] S. S. P. Rao *et al.*, “Cohesin Loss Eliminates All Loop Domains,” *Cell*, vol. 171, no. 2, pp. 305-320.e24, Oct. 2017, doi: 10.1016/j.cell.2017.09.026.
[50] P. Garcia *et al.*, “Disruption of NIPBL/Scc2 in Cornelia de Lange Syndrome provokes cohesin genome-wide redistribution with an impact in the transcriptome,” *Nat Commun*, vol. 12, no. 1, p. 4551, Jul. 2021, doi: 10.1038/s41467-021-24808-z.
[51] W. F. Richter, S. Nayak, J. Iwasa, and D. J. Taatjes, “The Mediator complex as a master regulator of transcription by RNA polymerase II,” *Nat Rev Mol Cell Biol*, vol. 23, no. 11, pp. 732–749, Nov. 2022, doi: 10.1038/s41580-022-00498-3.
[52] Deciphering Developmental Disorders Study *et al.*, “BRD4 interacts with NIPBL and BRD4 is mutated in a Cornelia de Lange–like syndrome,” *Nat Genet*, vol. 50, no. 3, pp. 329–332, Mar. 2018, doi: 10.1038/s41588-018-0042-y.
[53] Y. Zhu, M. Denholtz, H. Lu, and C. Murre, “Calcium signaling instructs NIPBL recruitment at active enhancers and promoters via distinct mechanisms to reconstruct genome compartmentalization,” *Genes Dev.*, vol. 35, no. 1–2, pp. 65–81, Jan. 2021, doi: 10.1101/gad.343475.120.
[54] L. Rinaldi *et al.*, “The glucocorticoid receptor associates with the cohesin loader NIPBL to promote long-range gene regulation,” *Sci. Adv.*, vol. 8, no. 13, p. eabj8360, Apr. 2022, doi: 10.1126/sciadv.abj8360.
[55] G. Borck, “NIPBL mutations and genetic heterogeneity in Cornelia de Lange syndrome,” *Journal of Medical Genetics*, vol. 41, no. 12, pp. e128–e128, Dec. 2004, doi: 10.1136/jmg.2004.026666.
[56] D. Dorsett, “Roles of the sister chromatid cohesion apparatus in gene expression, development, and human syndromes,” *Chromosoma*, vol. 116, no. 1, pp. 1–13, Jan. 2007, doi: 10.1007/s00412-006-0072-6.
[57] C. Bot, A. Pfeiffer, F. Giordano, D. M. Edara, N. P. Dantuma, and L. Ström, “Independent mechanisms recruit the cohesin loader protein NIPBL to sites of DNA damage,” *Journal of Cell Science*, p. jcs.197236, Jan. 2017, doi: 10.1242/jcs.197236.
[58] J. Zuin *et al.*, “A Cohesin-Independent Role for NIPBL at Promoters Provides Insights in CdLS,” *PLoS Genet*, vol. 10, no. 2, p. e1004153, Feb. 2014, doi: 10.1371/journal.pgen.1004153.
[59] D. Noordermeer, “RNA Pol II enters the ring of cohesin-mediated loop extrusion,” *Nat Genet*, Jul. 2023, doi: 10.1038/s41588-023-01463-2.
[60] S. Zhang, N. Übelmesser, M. Barbieri, and A. Papantonis, “Enhancer–promoter contact formation requires RNAPII and antagonizes loop extrusion,” *Nat Genet*, vol. 55, no. 5, pp. 832–840, May 2023, doi: 10.1038/s41588-023-01364-4.
[61] E. J. Banigan *et al.*, “Transcription shapes 3D chromatin organization by interacting with loop extrusion,” *Proc. Natl. Acad. Sci. U.S.A.*, vol. 120, no. 11, p. e2210480120, Mar. 2023, doi: 10.1073/pnas.2210480120.

[62] B. J. H. Dequeker *et al.*, "MCM complexes are barriers that restrict cohesin-mediated loop extrusion," *Nature*, vol. 606, no. 7912, pp. 197–203, Jun. 2022, doi: 10.1038/s41586-022-04730-0.
[63] M. Merkenschlager and E. P. Nora, "CTCF and Cohesin in Genome Folding and Transcriptional Gene Regulation," *Annu. Rev. Genom. Hum. Genet.*, vol. 17, no. 1, pp. 17–43, Aug. 2016, doi: 10.1146/annurev-genom-083115-022339.
[64] Y. Zhu, M. Denholtz, H. Lu, and C. Murre, "Calcium signaling instructs NIPBL recruitment at active enhancers and promoters via distinct mechanisms to reconstruct genome compartmentalization," *Genes Dev.*, vol. 35, no. 1–2, pp. 65–81, Jan. 2021, doi: 10.1101/gad.343475.120.
[65] E. De Wit and E. P. Nora, "New insights into genome folding by loop extrusion from inducible degron technologies," *Nat Rev Genet*, vol. 24, no. 2, pp. 73–85, Feb. 2023, doi: 10.1038/s41576-022-00530-4.
[66] C. M. Uyehara and E. Apostolou, "3D enhancer-promoter interactions and multi-connected hubs: Organizational principles and functional roles," *Cell Reports*, vol. 42, no. 4, p. 112068, Apr. 2023, doi: 10.1016/j.celrep.2023.112068.
[67] J. Ewing, "Diffuse Endothelioma of Bone," *CA: A Cancer Journal for Clinicians*, vol. 22, no. 2, pp. 95–98, Mar. 1972, doi: 10.3322/canjclin.22.2.95.
[68] N. Riggi, M. L. Suvà, and I. Stamenkovic, "Ewing's Sarcoma," *N Engl J Med*, vol. 384, no. 2, pp. 154–164, Jan. 2021, doi: 10.1056/NEJMra2028910.
[69] T. G. P. Grünewald *et al.*, "Ewing sarcoma," *Nat Rev Dis Primers*, vol. 4, no. 1, p. 5, Jul. 2018, doi: 10.1038/s41572-018-0003-x.
[70] O. Delattre *et al.*, "Gene fusion with an ETS DNA-binding domain caused by chromosome translocation in human tumours," *Nature*, vol. 359, no. 6391, pp. 162–165, Sep. 1992, doi: 10.1038/359162a0.
[71] G. L. Brien, K. Stegmaier, and S. A. Armstrong, "Targeting chromatin complexes in fusion protein-driven malignancies," *Nat Rev Cancer*, vol. 19, no. 5, pp. 255–269, May 2019, doi: 10.1038/s41568-019-0132-x.
[72] J. C. Schwartz, T. R. Cech, and R. R. Parker, "Biochemical Properties and Biological Functions of FET Proteins," *Annu. Rev. Biochem.*, vol. 84, no. 1, pp. 355–379, Jun. 2015, doi: 10.1146/annurev-biochem-060614-034325.
[73] P. C. Hollenhorst, L. P. McIntosh, and B. J. Graves, "Genomic and Biochemical Insights into the Specificity of ETS Transcription Factors," *Annu. Rev. Biochem.*, vol. 80, no. 1, pp. 437–471, Jul. 2011, doi: 10.1146/annurev.biochem.79.081507.103945.
[74] N. Riggi *et al.*, "EWS-FLI1 Utilizes Divergent Chromatin Remodeling Mechanisms to Directly Activate or Repress Enhancer Elements in Ewing Sarcoma," *Cancer Cell*, vol. 26, no. 5, pp. 668–681, Nov. 2014, doi: 10.1016/j.ccell.2014.10.004.
[75] W. J. Law, "TLS, EWS and TAF15: a model for transcriptional integration of gene expression," *Briefings in Functional Genomics and Proteomics*, vol. 5, no. 1, pp. 8–14, Feb. 2006, doi: 10.1093/bfgp/ell015.
[76] F. Tirode, K. Laud-Duval, A. Prieur, B. Delorme, P. Charbord, and O. Delattre, "Mesenchymal Stem Cell Features of Ewing Tumors," *Cancer Cell*, vol. 11, no. 5, pp. 421–429, May 2007, doi: 10.1016/j.ccr.2007.02.027.
[77] C. Von Levetzow *et al.*, "Modeling Initiation of Ewing Sarcoma in Human Neural Crest Cells," *PLoS ONE*, vol. 6, no. 4, p. e19305, Apr. 2011, doi: 10.1371/journal.pone.0019305.
[78] A. N. Schuetz *et al.*, "Intercellular junctions in Ewing sarcoma/primitive neuroectodermal tumor: additional evidence of epithelial differentiation," *Modern Pathology*, vol. 18, no. 11, pp. 1403–1410, Nov. 2005, doi: 10.1038/modpathol.3800435.
[79] A. A. Apfelbaum, E. D. Wrenn, and E. R. Lawlor, "The importance of fusion protein activity in Ewing sarcoma and the cell intrinsic and extrinsic factors that regulate it: A review," *Front. Oncol.*, vol. 12, p. 1044707, Nov. 2022, doi: 10.3389/fonc.2022.1044707.
[80] G.-A. Franzetti *et al.*, "Cell-to-cell heterogeneity of EWSR1-FLI1 activity determines proliferation/migration choices in Ewing sarcoma cells," *Oncogene*, vol. 36, no. 25, pp. 3505–3514, Jun. 2017, doi: 10.1038/onc.2016.498.

[81] F. Tirode *et al.*, "Genomic Landscape of Ewing Sarcoma Defines an Aggressive Subtype with Co-Association of *STAG2* and *TP53* Mutations," *Cancer Discovery*, vol. 4, no. 11, pp. 1342–1353, Nov. 2014, doi: 10.1158/2159-8290.CD-14-0622.
[82] S. K. Zöllner *et al.*, "Ewing Sarcoma—Diagnosis, Treatment, Clinical Challenges and Future Perspectives," *JCM*, vol. 10, no. 8, p. 1685, Apr. 2021, doi: 10.3390/jcm10081685.
[83] J. P. Ginsberg *et al.*, "Long-term Survivors of Childhood Ewing Sarcoma: Report From the Childhood Cancer Survivor Study," *JNCI Journal of the National Cancer Institute*, vol. 102, no. 16, pp. 1272–1283, Aug. 2010, doi: 10.1093/jnci/djq278.
[84] T. Waldman, "Emerging themes in cohesin cancer biology," *Nat Rev Cancer*, vol. 20, no. 9, pp. 504–515, Sep. 2020, doi: 10.1038/s41568-020-0270-1.
[85] D. Surdez *et al.*, "STAG2 mutations alter CTCF-anchored loop extrusion, reduce cis-regulatory interactions and EWSR1-FLI1 activity in Ewing sarcoma," *Cancer Cell*, vol. 39, no. 6, pp. 810-826.e9, Jun. 2021, doi: 10.1016/j.ccell.2021.04.001.
[86] B. Adane *et al.*, "STAG2 loss rewires oncogenic and developmental programs to promote metastasis in Ewing sarcoma," *Cancer Cell*, vol. 39, no. 6, pp. 827-844.e10, Jun. 2021, doi: 10.1016/j.ccell.2021.05.007.
[87] I. A. Showpnil *et al.*, "EWS/FLI mediated reprogramming of 3D chromatin promotes an altered transcriptional state in Ewing sarcoma," *Nucleic Acids Research*, vol. 50, no. 17, pp. 9814–9837, Sep. 2022, doi: 10.1093/nar/gkac747.
[88] I. A. Showpnil *et al.*, "EWS/FLI mediated reprogramming of 3D chromatin promotes an altered transcriptional state in Ewing sarcoma," *Nucleic Acids Research*, vol. 50, no. 17, pp. 9814–9837, Sep. 2022, doi: 10.1093/nar/gkac747.
[89] J. S. Reece-Hoyes and A. J. M. Walhout, "Gateway Recombinational Cloning," *Cold Spring Harb Protoc*, vol. 2018, no. 1, p. pdb.top094912, Jan. 2018, doi: 10.1101/pdb.top094912.
[90] A. Bhatwa, W. Wang, Y. I. Hassan, N. Abraham, X.-Z. Li, and T. Zhou, "Challenges Associated With the Formation of Recombinant Protein Inclusion Bodies in Escherichia coli and Strategies to Address Them for Industrial Applications," *Front. Bioeng. Biotechnol.*, vol. 9, p. 630551, Feb. 2021, doi: 10.3389/fbioe.2021.630551.
[91] S. Chong *et al.*, "Imaging dynamic and selective low-complexity domain interactions that control gene transcription," *Science*, vol. 361, no. 6400, p. eaar2555, Jul. 2018, doi: 10.1126/science.aar2555.
[92] G. Boulay *et al.*, "Cancer-Specific Retargeting of BAF Complexes by a Prion-like Domain," *Cell*, vol. 171, no. 1, pp. 163-178.e19, Sep. 2017, doi: 10.1016/j.cell.2017.07.036.
[93] M. Conte *et al.*, "Loop-extrusion and polymer phase-separation can co-exist at the single-molecule level to shape chromatin folding," *Nat Commun*, vol. 13, no. 1, p. 4070, Jul. 2022, doi: 10.1038/s41467-022-31856-6.
[94] D. Alonso-Gil, A. Cuadrado, D. Giménez-Llorente, M. Rodríguez-Corsino, and A. Losada, "Different NIPBL requirements of cohesin-STAG1 and cohesin-STAG2," *Nat Commun*, vol. 14, no. 1, p. 1326, Mar. 2023, doi: 10.1038/s41467-023-36900-7.
[95] T. M. Bui, H. L. Wiesolek, and R. Sumagin, "ICAM-1: A master regulator of cellular responses in inflammation, injury resolution, and tumorigenesis," *Journal of Leukocyte Biology*, vol. 108, no. 3, pp. 787–799, Sep. 2020, doi: 10.1002/JLB.2MR0220-549R.
[96] K. M. Bailey, C. M. Julian, A. N. Klinghoffer, H. Bernard, P. C. Lucas, and L. M. McAllister-Lucas, "EWS-FLI1 low Ewing sarcoma cells demonstrate decreased susceptibility to T-cell-mediated tumor cell apoptosis," *Oncotarget*, vol. 10, no. 36, pp. 3385–3399, May 2019, doi: 10.18632/oncotarget.26939.
[97] W. Zhang *et al.*, "ICAM-1-mediated adhesion is a prerequisite for exosome-induced T cell suppression," *Developmental Cell*, vol. 57, no. 3, pp. 329-343.e7, Feb. 2022, doi: 10.1016/j.devcel.2022.01.002.

List of Figures and Tables

Figures

TABLE

APPENDIX – Supplementary Figures

APPENDIX – Supplementary Tables