Università degli Studi di Milano

Dipartimento di Biotecnologie Mediche e Medicina Traslazionale

Scuola di dottorato in Scienze Biomediche Cliniche e Sperimentali

Corso di dottorato

Biotecnologie Applicate alle Scienze Mediche

(XXV°ciclo)

Neurospecific LSD1 splicing isoform links Epigenetics to mammalian brain physiology

(BIO/13)

DOTTORANDA:

Dott.ssa Leda Paganini

TUTOR:

Dott.ssa Elena Battaglioli

COORDINATORE:

Prof. Alessandro Massimo Gianni

Anno Accademico 2011/2012

INDEX

pag.

ABSTRACT.....1

INTRODUCTION 4

1. EPIGENETICS AND ITS RELEVANCE IN NEURONAL PLASTICITY 4

1.1 The lure of epigenetics in the brain
1.2 Chromatin overview
1.3 The building blocks of the epigenome:
- histone modification
- DNA methylation
1.4 Histone H3 Post-Translational Modifications (PTMs)
1.5 H3K4 methylation pattern

2. HISTONE DEMETHYLASE LSD1/KDM1.....24

2.1 Flavin-dependent and JmjC histone demethylases
2.2 Lysin Specific Demethylase1
2.3 Neuronal LSD1 isoform: expression and activity
2.4 Exon E8a phosphorylation and its effects
2.5 LSD1-related epigenetic mechanisms in the brain

3. ALTERNATIVE PRE-mRNA SPLICING.....38

3.1 Mechanism overview
3.2 *Cis*-acting sequences involved in splicing regulation
3.3 Neurospecific *trans*-acting splicing modulators:
- NOVA1 and NOVA2
- nSR100/SRRM4
- PTBP1 and PTBP2

- RBFOX1 and RBFOX2
- SAM68

3.4 Pre-mRNA secondary structures and splicing regulation

AIM OF THE PROJECT.....65

MATERIALS AND METHODS.....67

1. DNA CLONING.....67

1.1 Polymerase chain reaction
1.2 Agarose gel electrophoresis
1.3 Digestion with restriction enzymes
1.4 Dephosphorylation of linearized plasmid for ligation
1.5 DNA purification from gel slice
1.6 DNA precipitation from PCR reaction
1.7 Ligation
1.8 Generation of competent bacterial cells (TOP10)
1.9 Transformation by electroporation
1.10 Colonies screening by PCR and restriction enzymes digestion
1.11 Colony bacterial culture
1.12 MINI- MIDI- and MAXI-prep
1.13 Sequencing PCR
1.14 DNA purification from sequencing reaction

2. CELL CULTURE.....86

2.1 Transient transfection

3. RNA EXTRACTION AND ANALYSIS.....90

3.1 RNA extraction
3.2 Reverse transcription PCR

3.3 Quantitative real time PCR
3.4 Relative quantity fluorescent PCR (RqfPCR)
and capillary gel elecrophoresis

4. PROTEIN ANALYSIS……..…98

4.1 Western Blot

5. ANIMALS……101

5.1 DNA extraction from mice biopsy
5.2 Pilocarpine treatment and seizures assessment
5.3 Whole animal fixation via transcardial perfusion
5.4 Immunostaining

RESULTS……...……110

1. NEUROSPECIFIC LSD1 ALTERNATIVE SPLICING REGULATION……...110

1.1 Characterization of the exon E8a flanking introns
1.2 Minigene 800 splicing assay
1.3 Role of NOVA1 in regulating exon E8a alternative splicing
1.4 Seizures-dependent coregulation of NOVA1 expression and exon
E8a inclusion
1.5 Role of nPTB in regulating exon E8a splicing
1.7 SAM68: a putative exon E8a splicing regulator
1.8 Exon E8b discovery and its modulation by FOX1

2. GENERATION OF EXON E8a-limited LSD1 KO MICE……..……166

2.1 Generation of pBlueScript II KS targetting vector
2.2 Targetting *Neo* cassette and LoxP sites into PblueScript II KS
plasmid
2.3 Breeding scheme for ES cell chimeras

2.4 Mice genotyping
2.5 Backcross breeding to make a congenic strain
2.6 Accelerated backcross breeding

3. CHARACTERIZATION OF E8a-limited LSD1 KO MICE……..…..184

3.1 Necroscopic examination
3.2 Immunohistochemical characterization
3.3 Pilocarpine treatment responsiveness

DISCUSSION……...216

BIBLIOGRAPHY……226

ABSTRACT

LSD1, the first identified Lysine specific demethylase that removes methyl groups from mono- or di-methylated Histone 3 Lys4 (H3K4), has a mammalian-restricted neuronal isoform (LSD1-E8a), generated by the alternative inclusion of the 12-bp neurospecific exon E8a. LSD1 general function is to inhibit the expression of neuronal genes in non-neuronal cells, but LSD1-E8a isoform is characterized by a less gene repressing action. Indeed, the 4 aa coded by the exon E8a, which form a protruding loop in the LSD1 catalytic domain, contain a Threonine residue that can be phosphorylated and that is required to induce neuronal maturation, neurite outgrowth and the abrogation of LSD1-8a repressive activity.

Using a Minigene reporter assay, we demonstrate that the exon E8a is surrounded by a highly conserved 800-bp intronic region that is sufficient *per se* to regulate exon E8a alternative splicing, containing at least in part the necessary *cis*-acting elements. Among them there are the binding sites for NOVA1 and nPTB, that we found to be exon E8a positive *trans*-acting splicing regulators. Exon E8a, indeed, is very tightly modulated and it is present only in neuronal brain tissues and not in cell line. Along with positive *trans*-acting factors, we identified an exon E8a complementary/inverted sequence with a very strong negative regulatory effect on exon E8a inclusion. Here we show that this sequence acts by forming a double strand pre-mRNA pairing in which exon E8a is masked. Indeed, the deletion of the

12 core nucleotides of such complementary region promotes a strong exon E8a inclusion, allowing binding of *trans*-acting factors that recognize single-strand *cis*-acting motifs. Furthermore, downstream of exon E8a we identified a new 77-bp human-restricted LSD1 alternative exon, that we called “exon E8b”. Its inclusion in mature transcripts is regulated by FOX1 and occurs in many different tissues, although these transcripts are present at a very low level. This is probably due to the fact that exon E8b, by introducing a premature STOP codon inside LSD1 mature transcripts, causes their degradation by the cellular Non-sense mRNA Mediated Decay (NMD) pathway. Exon E8b very low endogenous expression level makes it difficult to study its functional role. At the moment we can only say that its inclusion into LSD1 transcripts could be a tool at cells disposal to finely tune LSD1 RNA amount, providing a new human-restricted LSD1 level of regulation.

Since the epigenetic function of the neurospecific LSD1 isoform has not been completely elucidated *in-vivo*, we generated a knock-out mouse model by replacing the sole LSD1 exon E8a with the Neomycin resistance cassette, flanked by two loxP sites. LSD1-E8a knockout animals are fertile, survive embryogenesis and show no histological differences as well as no obvious developmental defects. Interestingly, specific behavioral differences were detected in the exon E8a-deficient mice in response to a pharmacologically induced epileptic treatment, where they display a longer time of latency and a reduced number of seizures, in addition to a more restrained expression-burst of the two immediate early gene Egr-1 and c-Fos. From the immunohistological point of view, significative

differences were observed in Calbindin, V-GLUT2 and Sox-2 expression and distribution, although the characterization is not completed yet. Our results indicate that, *in vivo*, neuronal LSD1 isoform is not strictly required for normal development but it may became relevant to later functional events inside the Central Nervous System.

INTRODUCTION

1.EPIGENETICS AND ITS RELEVANCE IN NEURONAL PLASTICITY

1.1 THE LURE OF EPIGENETICS IN THE BRAIN

Our phenotype is the product of continuous gene–environment interactions. Environmentally regulated intracellular signals constantly modulate gene transcription in order to obtain cellular plasticity. The interaction of such signals with regulatory DNA sequences is dynamic and cannot explain instances in which extracellular inputs program enduring effects on gene expression and cellular function [1]. This issue is of particular interest in the brain, where persistent alterations in cellular function are essential for normal activity [1]. Indeed, the value of the brain is to guide the function of the organism in accordance with its life history. The capacity to mastermind such adaptation to circumstances relies on the possibility of neurons and glia to change their genomic structure in a replication independent manner (Fig. 1.1.1). But how does the life experience re-program gene expression in the brain [1]?

A great number of accumulating evidences suggests that there is much more than DNA sequence that permits the brain to reorganize itself, and one way that such phenotypic plasticity can be established is through epigenetic mechanisms [2]. Epigenetics refers to heritable changes in gene expression that are unrelated to variation in DNA sequence [1]. The main instruments at epigenetics disposal are DNA methylation and core histone proteins modifications. These epigenetic ‘marks’

can be remodelled by environmentally induced alterations and they are variable over time, for this reason they are considered the best tools to obtain the phenotypic plasticity shown by neurons throughout life [4]. During all the lifehood, indeed, the brain remains flexible and responsive to the outside world [3], so in the brain, heritable maintenance is less of an issue than replication independent methylation change or chromatin marks remodeling [1].

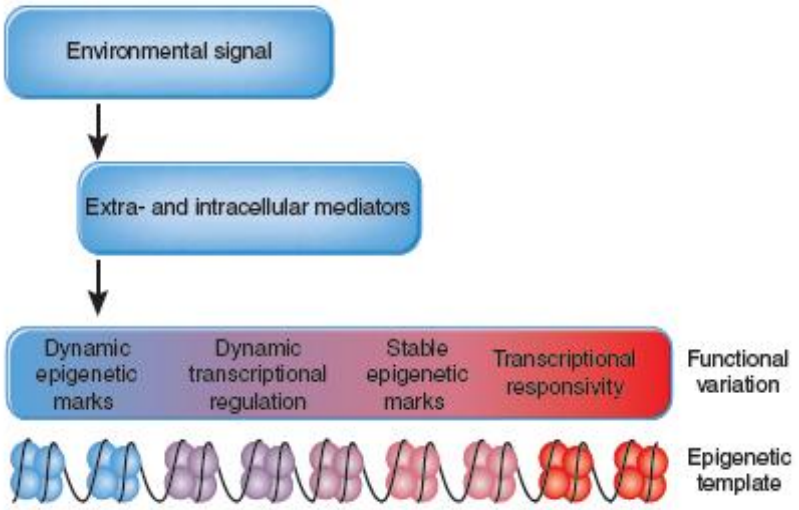


Fig. 1.1.1 The epigenetic interface.

Epigenetic states lie at the interface between environmental signals and genome, serving to govern dynamic changes in transcriptional activity through extra- and intra-cellular mediators. In a multistep process, the epigenetic template attracts specific effectors that determine the responsivity of specific genomic regions to environmentally induced intracellular signaling path ways, thus leading to more stable effects on the potential of transcriptional activation and variation in neural function (color-graded region).

1.2 CHROMATIN OVERVIEW

Chromatin is the intimate association of genomic DNA with histone proteins and represents the physiological form of our genome and the substrate of the epigenetic processes related to gene expression [5].

The fundamental unit of chromatin is the nucleosome: a 146 bp-long supercoiled DNA wrapped around the radial surface of an octamer, composed of two copies of each highly conserved core histone protein H2A, H2B, H3, and H4 [6]. The histones H3 and H4 form a tetramer, whereas the histones H2A and H2B form two dimers. Each of these histone protein consists in a main globular body (that contributes to the central protein mass of the nucleosome) and in a N-terminal tail that contains sites for post-translational modifications that influence chromatin structure and function [7]. Histone H1 has not been isolated within the nucleosomes, because it is located on the top of the structure, keeping in place the DNA wrapped around the core proteins (Fig.1.2.1). In addition, histone H1 binds to the "linker DNA" region between nucleosomes (from 8 to 114 nucleotides in length), helping the stabilization of the chromatin fiber [8].

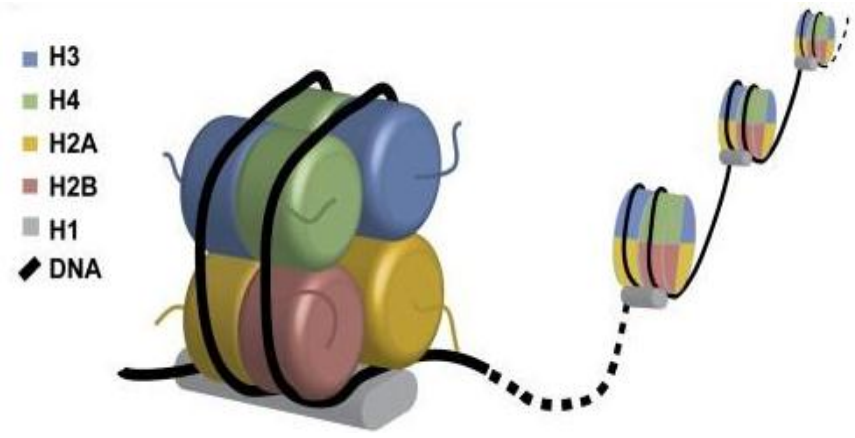


Fig. 1.2.1 Nucleosome Structure.

The arrangement of the eight histone proteins in the nucleosome is shown schematically. 146 base pairs of DNA are wrapped around the histone core. Histone H1 seals the nucleosome separating each nucleosome unit from the other.

In general terms, there are three levels of chromatin organization (Fig 1.2.2):

1. the "beads on a string" structure, consisting of DNA wrapped around histone proteins (euchromatin);

2. the 30 nm fiber: multiple nucleosomes organized in their most compact form (6 nucleosomes for each turn) (heterochromatin);
3. the metaphase chromosome, a higher level of the 30 nm fiber packaging during mitosis and meiosis [9].

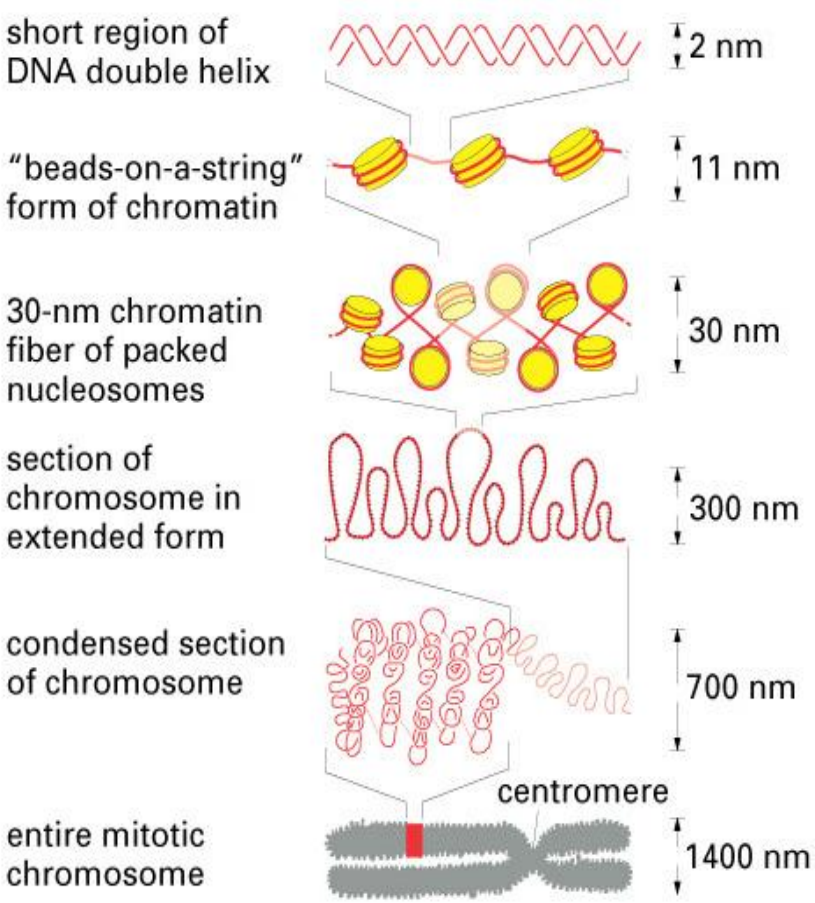


Fig. 1.2.2 Chromatin packing occurs on several levels.

Schematic drawing shows some of the orders of chromatin packing that give rise to the highly condensed mitotic chromosome.

The overall chromatin structure depends on the stage of cell cycle: during interphase the chromatin architecture is lost to allow RNA and DNA polymerases to access, transcribe or replicate DNA [9]. The local structure of chromatin during

interphase depends on the genes present on the DNA: DNA coding genes that are actively transcribed ("turned on genes") is more loosely packaged and is associated with RNA polymerases (euchromatin), while DNA coding inactive genes ("turned off genes") is associated with structural proteins and is more tightly packaged (heterochromatin) [9].

1.3 THE BUILDING BLOCKS OF THE EPIGENOME: HISTONE MODIFICATIONS AND DNA METHYLATION

Within the cells, regulation of biological processes can be achieved via genetic and epigenetic programs. Epigenetic control operates either on DNA, via DNA methylation, and on chromatin, by post-translational modifications (PTMs) of histones or by the exchange and replacement of the major histones with specialized variants. All of this mechanisms act together and define the so called 'epigenetic code'. While in one individual the genetic code in each cell is the same, the epigenetic code is tissue and cell specific (Fig. 1.3.1) [2].

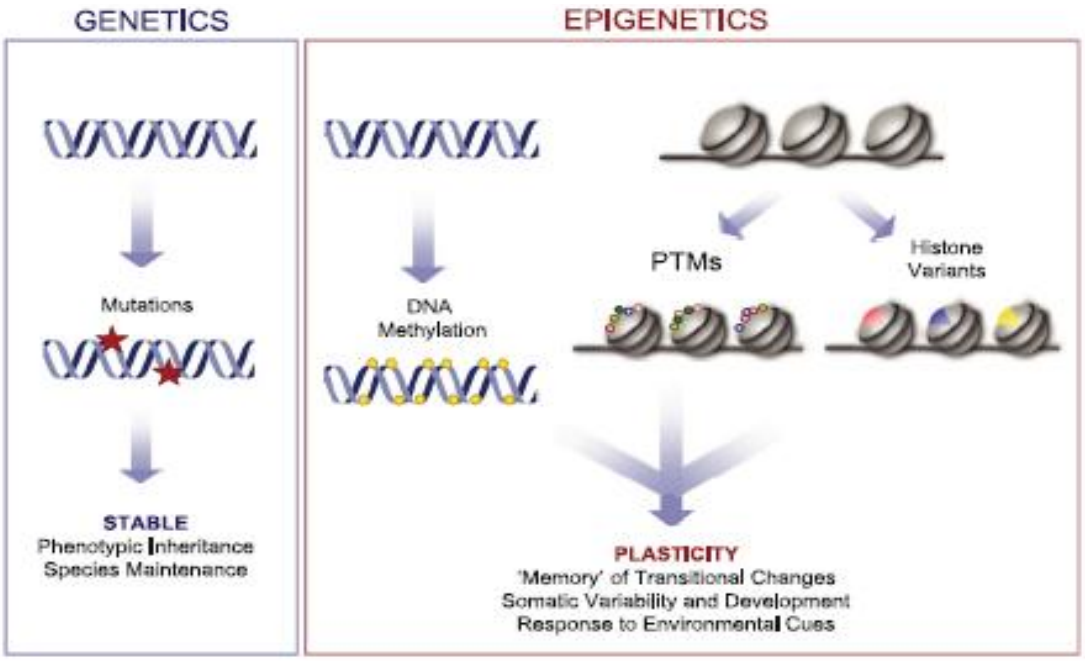


Figure 1.3.1 Genetic versus Epigenetic Control.

Regulation of biological processes can be achieved via genetic and epigenetic programs. Variation in genetic information is obtained by mutagenesis of the DNA sequence that irreversibly changes the encoded message. Epigenetic control operates either on DNA, via DNA methylation, or on chromatin. Variation in the chromatin template can be brought about by posttranslational modifications added to histones, by the exchange and replacement of the major histones with specialized variants or by ATP-dependent nucleosome remodeling, which alters histone:DNA contacts.

HISTONE MODIFICATIONS

The N-terminal tails of histones contain flexible and highly basic 15-30 aa-long sequences that are generally conserved across eukaryotic organisms and it is well established that they act as substrates for several types of post-translational modifications, including acetylation, methylation, ADP-ribosylation, ubiquitylation, and phosphorylation (Fig. 1.3.2) [10]. Histone modifications are proposed to affect chromosome function through at least two distinct mechanisms. The first mechanism suggests that modifications may alter the electrostatic charge of the histone, causing structural changes or alteration of its DNA binding affinity. The

second mechanism proposes that these modifications represent binding sites for protein recognition modules, such as the bromodomains or chromodomains, that bind acetylated lysines or methylated lysine, respectively [11].

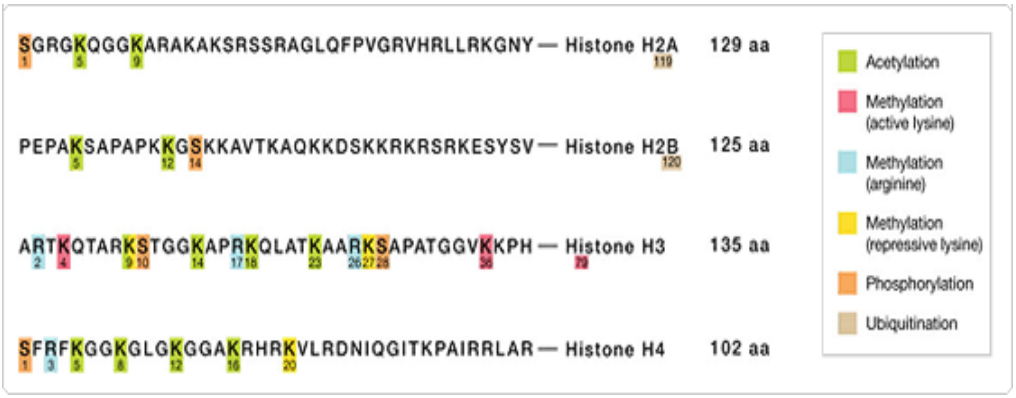


Fig. 1.3.2 Histone modifications.

All histones are subject to post-translational modifications, which mainly occur in histone tails. The main modifications are depicted in this figure: acetylation (green), active lysine methylation (pink), arginine methylation (light blue), repressive lysine methylation (yellow), phosphorylation (orange) and ubiquitylation (grey).

- HISTONE ACETYLATION mainly occurs on ε-amino groups of Lysine residues. The aa surrounding the preferred Lys/Arg are important for the identification of the specific binding site by the modifying enzyme. The enzymes that transfer an acetyl group from an Acetyl- CoA molecule to the Lys/Arg residues are generally called Histone Acetyl Transferases (HATs) while the ones acting on the Lys ammonium group are specifically called Lysine Acetyl Transferases (KATs) [2,7]. Since this modification is reversible, some enzymes that remove the acetyl group have been identified. They are generally called Histone Deacetylases (HDACs). Histone acetylation occurs mainly in two circumstances: during DNA

replication and during active gene expression. During the S phase of the cell cycle, before its incorporation in the nucleosome, the tetramer H3-H4 is transiently acetylated. The modification is probably required for the assembly of the new chromosome. After S phase, histone acetylation is correlated with gene expression. On the contrary, the N-terminal tails of histones H3 and H4 are not acetylated in heterochromatic regions [7]. Histone Lysines acetylation in the promoter region is often associated with gene activation, and the acetylated state is maintained until the gene is active (Fig. 1.3.3). This post-translational modification also represents a platform for binding of other effector proteins that regulate transcriptional activity [14].

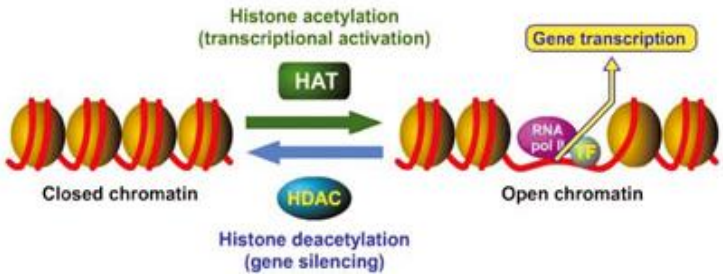


Fig. 1.3.3 Histone acetylation and deacetylation effects.

Acetylation occurs on specific Lys and Arg residues on histones H3 and H4, leading to an open chromatin conformation.

- HISTONE METHYLATION: also the histone methylation occurs on ε-amino groups of H3 and H4 Lysine and Arginine residues. The enzymes that transfer the methyl group from the molecule S-adenosyl methionine to the

Lys/Arg residues are generally called Histone MethylTransferases (HMTs). For many years histone methylation was thought to be a permanent modification, but very recently two families of Histone Demethylase enzymes (HDM) were discovered: the Flavin-dependent and the Jumonji C-containing (JmjC) HDMs. Unlike acetylation, methylation doesn't alter the protein charge, so it probably functions recruiting other effector proteins acting in a *trans*- mechanism. Figure 1.3.4 shows that distinct histone Lys residues have different functions when they are methylated, but also different degrees of methylation may have different functions, relating to positive or negative roles in transcription and other reactions. Also Arg methylated residues can play either a positive or negative role in transcription [15].

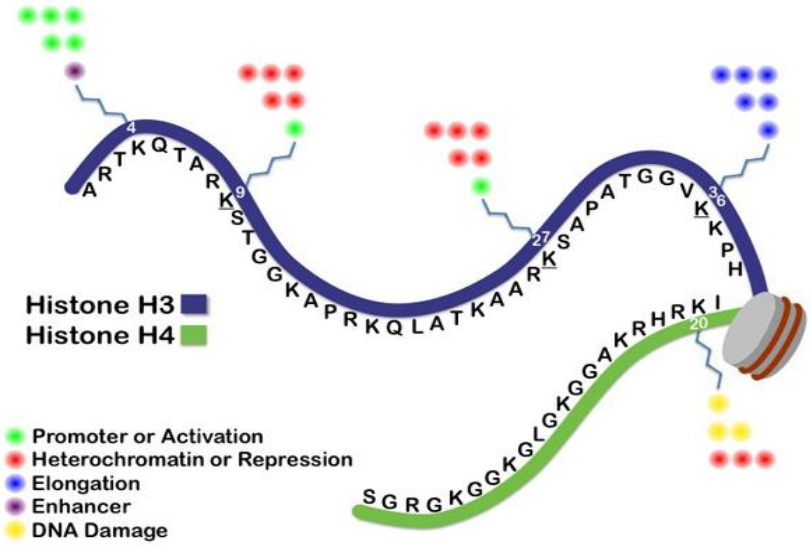


Fig. 1.3.4 Major Lysine methylation points.

The numbers refer to the methylated aa residue on each histone. Depending on the number and the position of the methyl group added, they exert different functions.

- HISTONE SUMOYLATION and UBIQUITINATION have been observed at the same way. Sumoylation appears to be associated with transcriptional repression, whereas ubiquitylation has been suggested to play a role in transcriptional activation and elongation [15].

- HISTONE PHOSPHORYLATION: histone phosphorylation occurs on Serine, Threonine or Tyrosine residues. Also in this case, the modification of different histones or different aa residues in the same histone, has opposite effects on chromatin structure. The phosphorylation of histone H1 has been associated to a particular chromosome conformation during mitosis. This led to speculation that histone phosphorylation might be connected with the condensation of the chromatin. On the contrary, phosphorylation of H3S10 is associated with gene transcription and consequently to chromatin decondensation, probably due to the addiction of negative charges to the DNA environment [15].

In addition to the great number of histone modifications described above, epigenetic studies have led to the discovery of a large number of enzymes wich interact physically and functionally with them (Fig. 1.3.5) [2]. These enzymes are classified in:

- 'writers', that modify specific substrates by adding phosphate, acetyl or methyl groups. They are represented by kinases, histone acetyltransferases

(HATs) and histone methyltransferases (HMTs);

- 'erasers', that catalyze the removal of specific histone modifications, such as histone deacetylases (HDACs) and histone demethylases (HDMs);
- 'readers', that are regulatory proteins containing unique domains that recognize acetyl or methyl or phosphate groups, etc.

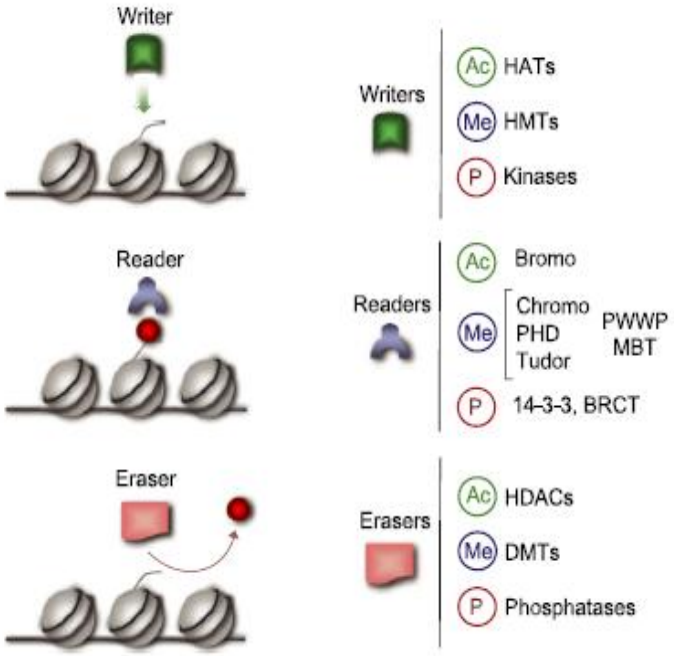


Fig. 1.3.5 Distinct classes of chromatin remodeling molecules.

Specific marks on the N-terminal tails of core histones are PTMs elicited by chromatin remodeling machineries that include a large variety of regulatory molecules, many of which interact physically and functionally. The regulators may be indicated as follows. (1) Writers. These are enzymes such as kinases, HATs, and HMTs that modify specific substrates adding phosphate, acetyl, or methyl groups. (2) Readers. These include a large variety of regulatory proteins that share unique domains implicated in recognizing acetyl or methyl groups. (3) Erasers. These enzymes include phosphatases, HDACs, and DMTs, which directly remove PTMs.

DNA METHYLATION

DNA methylation is a chemical process that adds a methyl group to DNA. It is highly specific and always occurs at the 5-position of Cytosine residue (m5C), in a DNA region called CpG island where a Cytosine nucleotide is located next to a Guanine. CpG islands are methylated by one of three enzymes called DNA methyltransferases (DNMTs), which use S-adenosyl-L-methionine as donor of methyl groups. The CpG islands are mainly located within the first exon and the promoter of many genes. In higher eukaryotes, DNA demethylation is associated with euchromatin. DNA regulatory sequences are generally demethylated. The presence of m5CpG dinucleotides in the promoter can affect gene transcription in 2 ways: in the direct mechanism the methylated dinucleotides interfere with transcription factors binding; in the indirect mechanism the dinucleotides are recognized by a binding protein that blocks the interaction with other transcription factors. It has also been suggested that DNMTs can “read” the histones modifications, recognizing the regions that should be methylated. Moreover, post-translational modifications such as acetylation, phosphorylation and methylation, can influence their catalytic activity. DNA methylation is also used in some genes to differentiate which copy is inherited from father and which one from mother, a phenomenon known as imprinting [16].

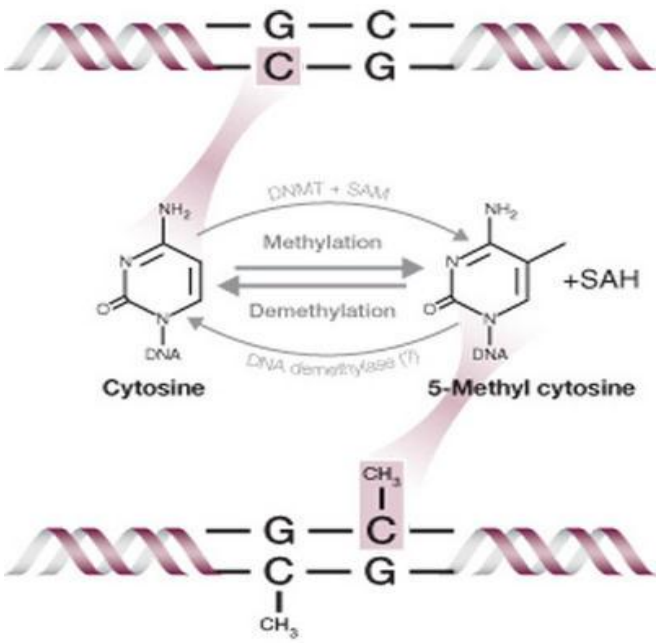


Fig. 1.3.6 Methylation of DNA at the 5-carbon position of cytosine, induced by DNMTs.
Accomplished by the transfer of a methyl group from S-adenosylmethionine (SAM) to the 5-carbon position of the pyrimidine ring of cytosine, DNA methylation contributes to the epigenome by covalently modifying the structure of DNA.

1.4 HISTONE H3 POST TRANSLATIONAL MODIFICATIONS (PTM)

The amino-terminal tail of histone H3 has the highest density of post-translational modifications mapped among all histones [17]. Several enzymes that modify specific residues on the N-terminal tail of histone H3 have been identified and characterized. LSD1 belongs to this category. For this reason, before focusing on the specific residue targetted by LSD1, it may be useful to have a panoramic view of the different histone H3 post-translational modifications and their effects on gene expression.

The major challenge in chromatin biology is connecting specific modifications with distinct biological functions and vice-versa. One of the better-understood histone

modification is the acetylation. It is now generally accepted that hyperacetylated histones are mostly associated with active genomic regions, at both local and global levels. By contrast, deacetylation (leading to hypoacetylation) mainly results in gene repression and silencing. Concerning histone methylation, it appears to have multiple effects on chromatin function in a system- and site-specific manner. Methylation of H3K9 is largely associated with gene silencing and repression while methylation of H3K4 is most often associated with active or permissive chromatin regions. Methylation of H3K36 has been suggested to be involved in transcriptional repression, but the corresponding modifying enzyme, Set2, has been found in complex with actively transcribing RNA pol II. Also H3S10 phosphorylation has a dualistic effect, since it has been implicated not only in transcriptional activation but also in mitotic chromosome condensation. These results suggest that a single histone modification may have distinct biological effects, depending from its epigenetic context. The finding that a particular post-translational modification mediates separate and sometimes opposite physiological effects, led to the suggestion that the multiple readout of a certain covalent mark depends from various combinations of different modifications in the same chromatin region [17] [18].

At the level of N-terminal tail of H3 a great number of PTMs have been identified: serine and threonine residues are well-known phospho-acceptor sites, while lysine and arginine residues have multiple choices of post-translational modification possibilities (Fig. 1.4.1).

Fig. 1.4.1 Schematic diagram of the covalent post-translational modifications (PTMs) of the histone H3 N-terminal tail.

The different Histone H3 modifications are indicated (P, phosphorylation, shown in orange; Ub, ubiquitylation, shown in purple; Ac, acetylation, shown in blue; S, sumoylation, shown in yellow; N, neddylation, shown in pink). Methylation (Me) is shown on top, with green and red indicating the methyl marks that are associated with activation or repression, respectively.

H3-lysines can be modified by acetylation, mono-ubiquitination or mono-, di- and tri-methylation. Similarly, H3-arginines may be mono- or di-methylated. Concerning this, it is well documented that H3-K9 and H3-K14 can be either acetylated or (mono-, di-, tri-) methylated. Obviously, different marks on the same site cannot co-exist and an acetyl group, for example, must be removed before a methyl group can be added. Nevertheless, we now also know that the exact state of methylation (mono-, di- or tri-methylation) of a single lysine residue has an impact on physiological processes. For example, it was recently shown that dimethylation of H3-K4 occurs at both inactive and active euchromatic genes, whereas tri-methylation is present exclusively at active genes [18].

Many of the enzymes that post-translationally modify histones display a high degree of specificity not only towards a particular site, but also towards the pre-existing modification-state of their substrate. So far the N-terminal tail of H3 has the highest intricate combination of marks (Fig. 1.4.2).

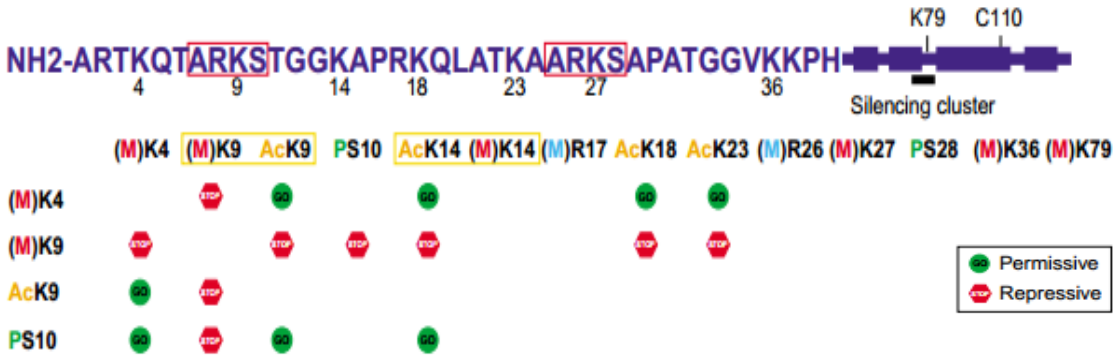


Fig.1.4.2 Local cross-talk on the human H3 amino-terminal tail domain.

The sequence of the amino-terminal tail of H3 (amino acids 1–40) and the four a-helices (represented by boxes) of the globular domain of H3 are shown. Sites of known modifications are listed (M, mono-, di- or tri-methylation). K9 and K14 have been found to be methylated or acetylated (yellow box). ARKS repeats that contain two sites of methylation (K9 and K27), as well as known sites of phosphorylation (S10 and S28), are highlighted. Primary modifications that positively (green, 'go' or permissive) or negatively (red, 'stop' or repressive) influence the modification of other sites in in vitro enzymatic assays are listed on the left.

Methylation on H3-K9, for example, appears to trigger sequential events leading ultimately to transcriptional repression. This mark can inhibit acetylation of the H3 tail (on K14, K18 and K23) mediated by histone acetyltransferases, and methylation of H3 on K4. By contrast, H3-K4 methylation inhibits K9 methylation but promotes acetylation of H3. Remarkably, the choice of methylating H3 on K9 could be dictated by H3-S10 phosphorylation. In mammalian cells, this mark not only inhibits methylation on K9, but also precedes and promotes acetylation on K14, following specific signals [17].

Perhaps, more fascinating than the direct synergism/antagonism or 'communication' of adjacent modifications in the same histone tail ('cis' effects), is the unexpected discovery that modifications on different histones can affect each other ('trans' effects). These effects might be restricted to a single nucleosome or

might affect larger nucleosomal arrays or domains (Fig. 1.4.3).

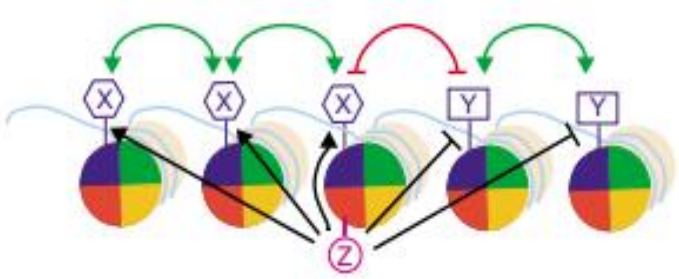


Fig.1.4.3 Cross-talk at the level of nucleosomal domains.

In an array of nucleosomes, different modifications on separate histones (X or Y) might influence each other in a positive or negative way. For example, it has been postulated that methylation of H3 on K9 could be spread over larger domains by recruitment of an HP1-Su(var)3-9 complex to sites of H3-K9 methylation. Similarly, boundaries for modification spreading could be established by inhibition/exclusion of different modifications. On another level, a single modification could regulate the modification pattern of a larger region of nucleosomes ('master control switch', Z). Ubiquitination of H2B in budding yeast could be such a 'master control switch' because of its relatively low abundance in comparison with the methylation on H3-K4 and H3-K79, which are both dependent on this modification. Since histone ubiquitination might be less stable than histone methylation, it is also possible that ubiquitin is removed after a methylation event on the same nucleosome.

For instance, in vitro studies using histone acetyl-transferases p300 showed that this HAT acetylates both H3 and H4 especially in nucleosomes where H3 is methylated on K4. By contrast, methylation of H3 on K9 significantly inhibits the activity of p300 towards nucleosomal histones H3 and H4 [17].

Besides cross-talk between different covalent modifications, another way of 'communication' within the nucleosomes could be the disulfide-bond-mediated dimerization. It may not be a coincidence that H3 is the only core histone containing a single cysteine (C110), which is conserved in all species except for budding yeast. Formation of a disulfide bond between the two H3 molecules of each nucleosome might place severe conformational restraints on the structure of individual nucleosome, nucleosomal arrays or chromosomal domains [17].

It is clear from this roundup on H3 PTMs that we are only beginning to understand and appreciate the far-reaching implications of this non-DNA-encoded information for human biology and disease. The first level of the complexity of chromatin cross-talk originates from the modular organization of chromatin itself. Obviously, each core histone can be post-translationally modified in a remarkably large number of ways, thus generating a vast number of possible combinations of marks for any chromatin domain. Besides a direct input from various signal transduction pathways on local and global chromatin levels, the modifications on a single histones seem to be dependent on each other and to be interconnected via various mechanisms [17, 18]. Deciphering even only a small aspects of the proposed 'histone and chromatin cross-talk' represents a significant and exciting challenge.

1.5 H3K4 METHYLATION PATTERN

In this section I will focus only on the methylation pattern of Histone H3 Lysin 4 (H3K4) since this specific epigenetic mark is directly regulated by LSD1.

H3K4 methylation was first discovered in the trout testes. In subsequent studies, methylation of H3K4 has been linked to transcriptional activation in a variety of eukaryotic species. As I said before Lysine residues can be mono-, di-, or trimethylated at the z-amine *in vivo*. Recent genomic-scale analysis of histone modifications allows for general correlations between different H3K4 methylation states, their genomic loci, and gene expression levels. The emerging consensus is

that high levels of H3K4 trimethylation are associated with the 5' regions of virtually all active genes and that there is a strong positive correlation between this modification, transcription rates, active polymerase II occupancy, and histone acetylation. In contrast, pattern of dimethyl H3K4 differs significantly between yeast and vertebrate chromatin: in S. Cerevisiae, dimethylated H3K4 appears to spread throughout genes, peaking toward the middle of the coding region, and is associated with a transcriptionally ‘‘poised’’ gene as well as active chromatin. Monomethylation, on the contrary, is most abundant at 3' ends of the genes. In vertebrates, the majority of H3K4 dimethylation colocalizes with H3K4 trimethylation in discrete zones, about 5–20 nucleosomes in length, close to highly transcribed genes. On the contrary, H3K4 monomethylation is associated with silenced euchromatin regions of the genome and it has a function in gene repression [20].

The first H3K4 methylase complex, COMPASS, was identified in the yeast S. Cerevisiae and consists of Set1/KMT2 and seven other polypeptides (named Cps60-Cps15). Set1/KMT2 alone is not enzymatically active, but works within COMPASS and is capable of mono-, di-, and trimethylate H3K4. Following the identification in S. Cerevisiae of Set1/COMPASS as H3K4 methylase, it was demonstrated that its mammalian homologues, the MLL proteins, MLL1-4 and hSet1A and B, were found in COMPASS-like complexes capable of methylating the fourth Lysine of histone H3. SET1, SET7/9, Ash1, ALL-1, MLL, ALR, Trx, and

SMYD3 are the other histone methyltransferases that catalyze methylation of histone H3 at lysine 4 (H3K4) in mammalian cells (Fig. 1.5.1) [20].

Fig. 1.5.1 The known enzymatic machineries involved in the methylation of lysine residues of histones H3.

The N-terminal amino acid sequences of histones H3 and H4 are shown along with the positions of specific methylation of lysines (in red) and arginines (in green) sites and the known enzymatic machinery responsible for the corresponding lysine modifications.

For many years histone methylation was thought to be a permanent modification but very recently two families of histone demethylase enzymes were discovered: the Flavin-dependent monoamine oxidase and the Jumonji domain-containing (JmjC) histone demethylases. The first are inactive on tri-methylated lysines while the latter are able to demethylate mono-, di-, or tri-methylated H3K4 [19]. LSD1 belongs the Flavin-dependent family and it represents the first discovered monoamine oxidase .

2.HISTONE DEMETHYLASE LSD1/KDM1

2.1 FLAVIN-DEPENDENT AND JUMONJI DOMAIN-CONTAINING HISTONE DEMETHYLASES

The discovery of Lysine Specific Demethylase 1 (LSD1) introduced a new concept in the field of the histone methylation, clearing the way for the idea that histone demethylation is possible. After LSD1 finding, many studies have been done and mechanisms for direct histone demethylase reactions have been proposed.

Depending on the catalysis reaction, histone demethylases are subdivided in two main families: Flavin-dependent histone demethylases and Jumonji domain-containing demethylases [21].

As shown in figure 2.1.1, the Flavin-dependent histone demethylases, cause the oxidative cleavage of the C-N methyl group bond, coupling it with a two-electron reduction of the Flavin Adenine Dinucleotide (FAD) cofactor. This reaction produces an imine intermediate that is then non-enzymatically hydrolyzed. The methyl group is released as formaldehyde. This proteins exhibit specificity in the Lysine substrate, indeed they preferentially demethylate mono- and di-methyl Lysines. Since LSD1 is a good model for this protein family, their structure and function will be discussed later [21].

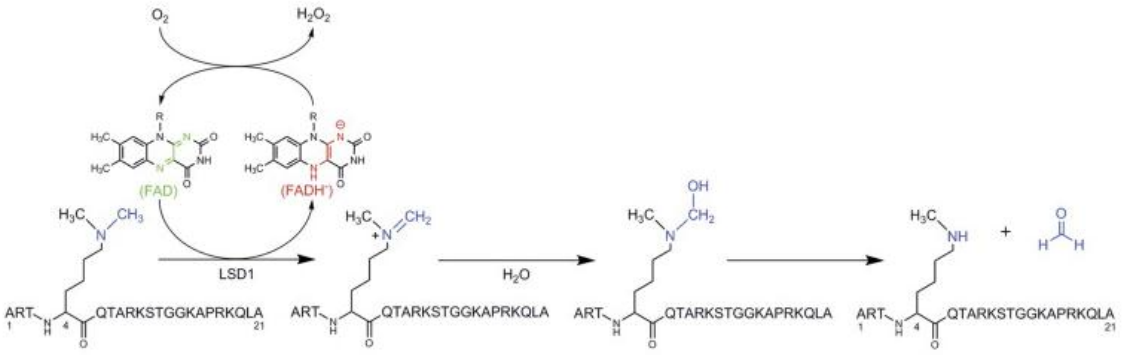


Fig 2.1.1 Catalytic mechanisms of FAD-dependent demethylase enzymes.

The FAD-dependent demethylation of Lys-4 of histone H3 proceeds through the hydrolysis of an iminium ion following a two-electron oxidation of the amine by the flavin. R, ribosyl adenine dinucleotide.

Differently, the JmjC histone demethylases are a group of Fe^{2+}-dependent dioxigenases proteins, that involve the oxidative decarboxylation of α-ketoglutarate coupled to the hydroxylation of the histone methyl group. This reaction creates an unstable hydroxymethyl ammonium intermediate, which is released as formaldehyde. The catalytic centre requires Fe^{2+} and oxygen, as shown in figure 2.1.2 [22].

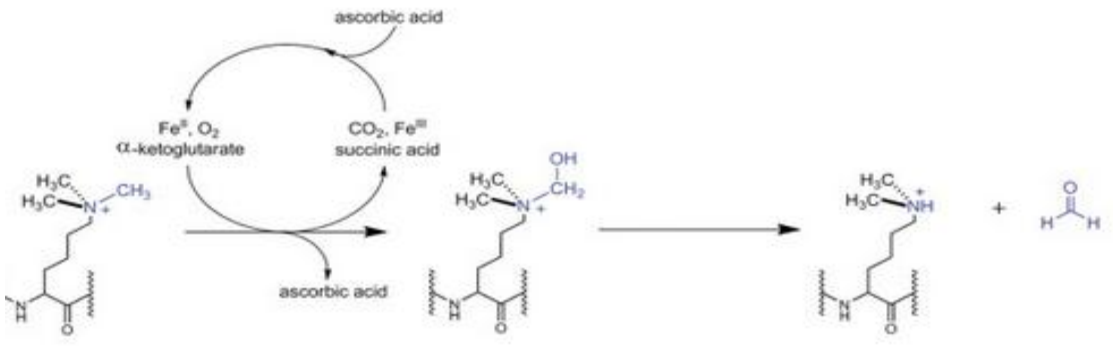


Fig. 2.1.2 Catalytic mechanisms of Jumonji demethylase enzymes.

The iron(II)-dependent demethylation of trimethyl-lysine substrates proceeds through an iron(II), α-ketoglutarate, and O_2 derived hydroxyl radical oxidation of the methyl C–H bond.

These proteins are involved in the demethylation of trimethyl Lysines. The structure of these proteins consists of:

- Jmj N-terminal domain;
- Jmj C-terminal domain, the catalytic core that contains 2 Histidine and 1 Glutamate residue, which chelate the catalytic iron atom;
- C-terminal zinc finger motif;
- a hairpin and a mixed domain to connect JmjN and JmjC domains.

Most of the interactions with the substrate are mediated by hydrogen bonds between the main chains of both the enzyme and the histone tail; for this reason a number of residues is important for the interaction with the peptide. Many Jmj proteins also contain a PHD finger motif. Interestingly, different studies show that the Jmj proteins can hydroxylate histones or other proteins, suggesting their function in alternative reactions.

Every histone demethylases protein has however to be considered as part of a macro-molecular complex that also contains other functionally complementary enzymes that modify histones methyl state, and histone modification “readers”. The presence of auxiliary factors, such as proteins or non-coding RNA, can influence the histone demethylases activity [22].

2.2 LYSIN SPECIFIC DEMETHYLASE 1

Lysine Specific Demethylase 1 (LSD1, also named KMD1A or AOF2), is a flavin-dependent histone demethylases, that specifically acts on mono- or dimethylated lysine 4 on histone 3 (H3K4).

Its function is predominantly related to gene repression, in fact it is required for downregulation of neuronal genes in non-neuronal cells. LSD1 demethylases function can also be related to different biological processes, such as development, cancer and neurological disorders.

The human LSD1 protein is 852 aa long and consists of different domains (Fig. 2.2.1):

- N-terminal domain, composed of the SWIRM domain (involved in proteins interactions) and of an unstructured region made of linear motifs that might represent functional sites responsible for the association of LSD1 with different transcriptional protein complexes;
- C-terminal amine oxidase domain (AOD), composed of a substrate-binding portion and a FAD-binding portion. The enzyme active site is located between these two regions;
- “Tower Domain”, made of two long antiparallel α- helices that project away from the globular AOD. This domain provides the interface for CoREST binding.

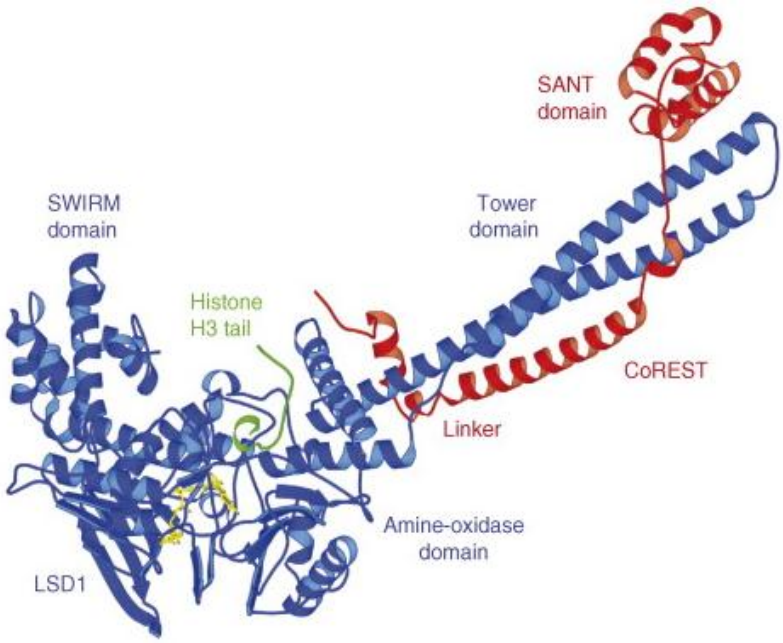


Fig. 2.2.1 Structure of LSD1-CoREST in complex with H3 peptide.

*R*ibbon diagram of the structure. LSD1 is in *blue*, CoREST in *red*, and the peptide in *green*. The FAD cofactor is shown as a *yellow ball-and-stick*. The final model consists of residues 171–836 of LSD1, residues 308–440 of CoREST, and residues 1–16 of pLys4Met peptide.

The histone tail adopts a folded conformation when bound to the enzyme and slides into the substrate-binding domain cavity, establishing a network of specific interactions with the active site residues. These interactions are important to correctly position the Lys4 in front of the FAD [22,14].

The LSD1 core complex contains LSD1, the proteins HDAC1, HDAC2 and CoREST. CoREST is a co-repressor protein that binds the neuronal Responsive Element RE1 Silencing Transcription factor (REST). Indeed LSD1 complex is recruited by REST through the direct binding to CoREST, resulting in the repression of neuronal genes in non-neuronal cells. The presence of CoREST in the complex is also important for the substrate binding and recognition: the C-terminal SANT domain of CoREST facilitates the association with chromatin by

directly interacting with DNA. HDAC 1 and 2 deacetylate histone H3, allowing the binding of CoREST to the nucleosome and the LSD1 recruitment. Thus, the presence of HDAC1 and 2 suggests a coordinate modification of histone tails [23].

Even if the active site of LSD1 is large enough to accommodate also H3K4me3, LSD1 demethylates only mono-or dimethylated H3K4, suggesting that the methyl state selectivity is not structurally inhibited but is chemically constrained, as predicted by its enzymatic mechanism. Some studies demonstrated that certain H3 tail modifications (such as acetylation on Lys 9 or phosphorylation on Ser 10) affect LSD1 activity, suggesting that LSD1 could act after the addiction or removal of other charge-altering histones modifications. Moreover, these findings suggest also that LSD1 complex is capable to read the histone code.

It has also been demonstrated that LSD1 complex function is not limited to REST-regulated neuronal genes, but it can be extended to other contexts. Thus, LSD1 can either repress or activate target genes through interacting with a variety of co-factors. For instance, there are some evidences that LSD1 can demethylate H3K9me2/me1, with an activating function. Indeed, it has been demonstrated that LSD1 directly binds the androgen receptor, and in this molecular complex is able to demethylate H3K9, functioning as a transcription activator. Anyway, it is still unclear how LSD1 changes its demethylase specificity when it engages different interacting partners [22]. It's a given that LSD1 also interacts with other proteins, including Carboxyl-terminal binding protein (CtBP), HMG domain containing protein BRAF35 and PHD-finger containing protein BHC80 [23]. These proteins can

enhance the HDAC1-2 activity, adding further regulatory steps to the LSD1 activity [24].

2.3 NEURONAL LSD1 ISOFORM: EXPRESSION AND ACTIVITY

Given the central role of LSD1 in chromatin remodeling, few years ago in my laboratory the process of LSD1 alternative splicing was investigated in order to find out an additional mechanism of LSD1 regulation. This approach led to the discovery of 4 different LSD1 splicing isoforms (Fig. 2.3.1), deriving from combinatorial retention of two alternative exons: exon E2a and exon E8a. E2a is 60-bp-long and encodes for 20 aa, whereas the E8a is 12-bp-long and is translated into 4 aa with sequence Asp-Thr-Val-Lys [12].

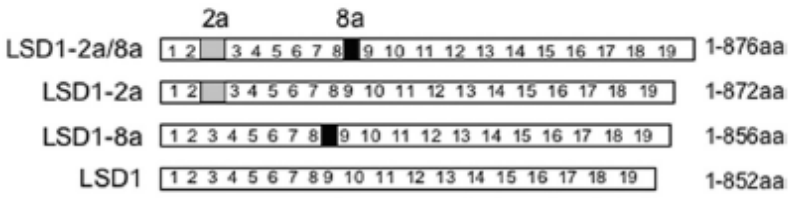


Fig. 2.3.1 Structure of the four LSD1 variants.
Single or double inclusion of two alternatively spliced exons, E2a and E8a, generates 4 different mammalian specific splicing variants.

The inclusion of the two exons does not alter the reading frame and results in a protein of 876 aa. In this isoform, the amino acids coded by the exon E2a localize between the N-terminal disordered region and the SWIRM domain; whereas the four residues of the exon E8a immediately precede the CoREST-binding tower

domain, which is inserted within the amine oxidase domain (Fig. 2.3.2). Both the two alternatively spliced introns present a very high conservation degree between human and mouse, a typical feature of alternatively spliced exons.

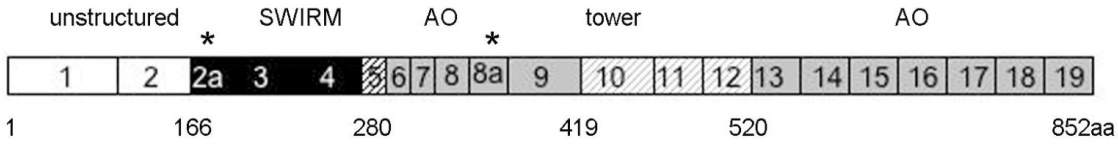


Fig. 2.3.2 Genomic organization of human LSD1 gene.

Schematic representation of the human LSD1 protein domains together with its exons ranging from 1 to 19; asterisks indicate the location of annotated alternative exons (E2a and E8a). Different colors indicate functional domains. N-terminal unstructured region coded by exons 1–2, SWIRM domain coded by exons 2– 4, the SWIRM-oxidase connector coded by exon 5, the amine oxidase domain coded by exons 6–9 and exons 13–19, and the tower domain coded by exons 10 –12.

The isoforms containing exon E2a are ubiquitously expressed, as the wild type isoform is, whereas the isoform containing the sole exon E8a is restricted to the nervous system. Finally, the LSD1-E2a/8a isoform is expressed both in human brain and testis. It is interesting to note that while the E2a isoform is present in different species, such as lizard, chicken and mammals, the E8a isoforms are expressed only in mammals. Moreover, while the epigenetic factors are mostly ubiquitous, LSD1-E8a is one of the very few factors restricted exclusively to neurons. All the 4 splicing variants retain the ability to generate functional proteins and to form an active complex with CoREST and HDAC1-2. Biochemical *in vitro* assays using histone H3 peptides as substrate revealed that all three LSD1 isoforms bound to CoREST can demethylate Lys4 of histone H3 with a catalytic efficiency virtually identical to that displayed by wild type LSD1. Such unchanged enzymatic activity and substrate specificity is supported by a very similar

three-dimensional structure, indeed even in the presence of exon E2a the N-terminal region of LSD1 remains unstructured, at least in the crystalline state. Also the overall conformation of LSD1-E8a is very similar to that of the native protein, since the neurospecific E8a residues Asp-Thr-Val-Lys form a sort of protrusion that emerges from the amino-ossidase domain of the protein but doesn't contact the histone peptide and CoREST (Fig. 2.3.3). From a functional point of view, the presence of the neuronal mini-exon E8a determines a significantly reduced repression of the Luciferase reporter gene activity [12].

Moreover it was demonstrated that a fine balance of LSD1 isoforms allows differentiating neurons to acquire a proper morphology. Indeed, the expression of LSD1 splice variants is dynamically regulated during mammalian brain development and synaptic maturation, in particular during the perinatal stages. At early all the splicing isoforms are detectable, but the preponderal ones are LSD1 and LSD1-E2a. During the perinatal window, however, a rapid inversion of the proportions occurs, with a notable increase of the neuronal exon E8a containing isoforms and a decrease of the exon E2a isoforms. Exon E8a expression increases concomitantly with early stages of synaptogenesis, suggesting its implication during morphogenesis, possibly regulating the proper timing of neurite maturation. Nevertheless, after day P7, all LSD1 isoforms reach comparable levels, as it is possible to detect in adult mouse brain. Thus, LSD1 expression profile suggests a possible implication of neurospecific E8a-containing isoforms in neuronal development.

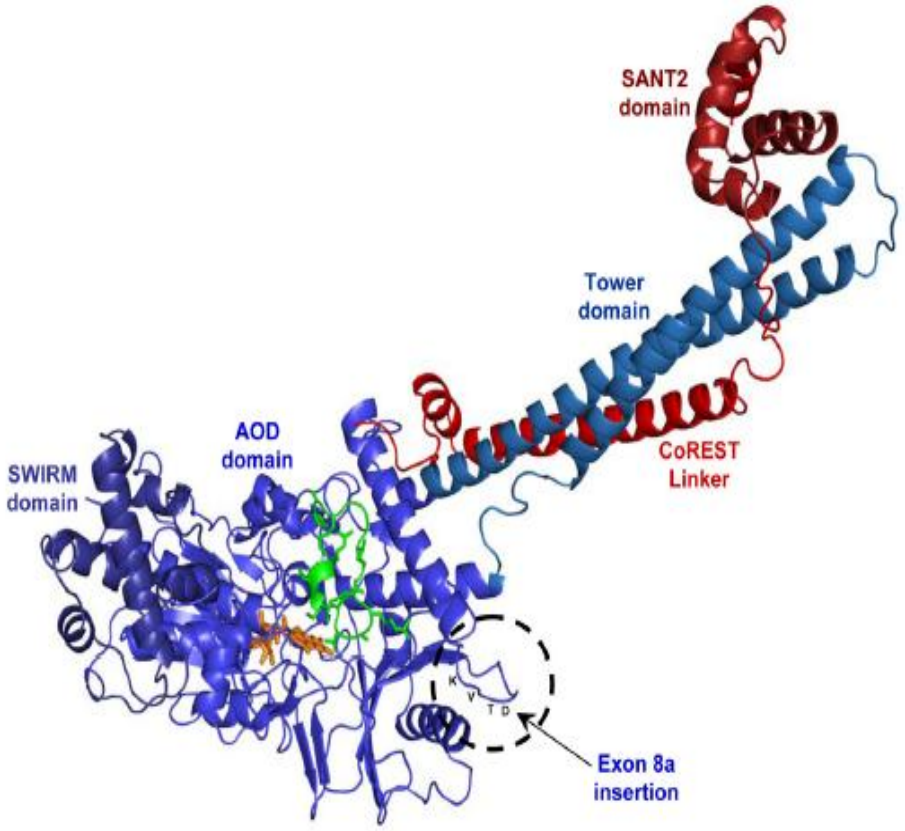


Fig. 2.3.3 Structural analysis of LSD1 and the LSD1-8a splice variant.

Overall crystal structure of LSD1-8a–CoREST in complex with a histone peptide. LSD1-8a (residues 171-840) is in light blue, CoREST (residues 308-440) in red, and te histone H3 peptide (residues 1-16) in green. The FAD cofactor is in the orange ball-and-stick representation. The insertion site of E8a (residues Asp369A-Thr369B-Val369C-Lys369D) is highlighted.

To infer the function of the different LSD1 isoforms within neurons, they were knocked down differentially by short hairpin RNAs (shRNAs) specific for each variant and 3 typical phenotypic parameters of neuronal morphogenesis during *in vitro* maturation were evaluated. Only the change in the amount of neurospecific LSD1-8a isoform is able to alter neurite morphogenesis in terms of cumulative neurite arborization, number of secondary branches and average neurite width [12].

2.4 EXON E8a PHOSPHORYLATION AND ITS EFFECTS

As I said before, mammalian neurons have a specialised chromatin remodelling complex resulting from the presence of a neurospecific LSD1-8a splice variant, whose expression is developmentally regulated. Only this isoform is capable of selectively inducing morphogenic effects in cortical neurons, which indicates that the sole presence of the 4 amminoacids coded by exon E8a is pivotal to acquisition of neuro-specific function.

The mini-exon E8a forms a loop, protruding from the surface of the protein, that is close to the enzyme active site and, more importantly, it is accessible to enzymes for possible post translational modifications (PTM). Indeed Thr 371, inside D-T-V-K aa sequence, can be phophorylated both in neuronal and in non-neuronal cells by a set of unknown enzymes. Although phosphorylation of Thr371 *per se* does not interfere with the intrinsic catalytic activity of the enzyme *in vitro*, this PTM further decreases repression of the Luciferase reporter gene activity much more than unmodified LSD1-8a and enhances morphogenic properties exerted by neuronal isoform. To investigate if the neuronal LSD1-8a repressive activity likewise that one of LSD1, involves the activity of associated factors, demethylase-null isoforms were generated and repressive activity was evaluated. In a condition of total demethylase inactivity, conventional LSD1 shows a slight reduction in repressor strength in the SH-SY5Y cell line while LSD1-8a completely abolished its repressive activity. Since conventional LSD1 is incorporated in multiprotein

complexes consisting of an equimolar ratio of CoREST, HDAC1 or HDAC2, the differences observed in neuronal LSD1 activity could be ascribe to one of such factors. Interestingly, immunocomplexes from the clones expressing HA-LSD1-E8a shows reduced HDAC1 protein levels in comparison with those expressing HA-LSD1, with no significant differences in CoREST and HDAC2. Altogether these results highlight an intrinsic difference in the mechanisms by means of which the two LSD1 isoforms exert their repressive activity and suggest that, in neuronal cells, phosphorylation turns off repressive activity of neuronal LSD1-E8a interfering with demethylase activity directly or mediating recruitment of additional regulatory factors that could interfere with the substrate binding/recognition [13].

2.5 LSD1-RELATED EPIGENETIC MECHANISMS IN THE BRAIN

Modulation of histone methylation and transcriptional repression in the brain is crucial in the control of various neurological functions, however limited knowledge exists regarding the significance of LSD1 in the brain. As I said before, in my laboratory a neurospecific chromatin remodeling enzyme, arising from a splice variant of LSD1, was identified. Its demethylase activity on Lys4 of histone H3 is related to gene repression but so far little is known about neurospecific LSD1 target genes or pathways selectively regulated by this neuronal isoform.

In literature, LSD1 is reported to be involved in a number of neuronal processes. Among them, the regulation of neural stem cells proliferation, the control of pyramidal cortical neurons development and the modulation of long-term memory formation are the most investigated ones.

In 2010 Y. Shy and his group demostrated for the first time that LSD1 is a key regulator of neural stem cell proliferation. Thanks to the recruitment by nuclear receptor TLX to the promoter of its target genes, LSD1 is able to co-repress their expression. TLX, indeed, is an essential neural stem cell regulator and its most well-known tagret genes are pten and p21 [26], which are positive modulators of neural stem cell proliferation. Thus, through the recruitment by TLX and the co-modulation of its repressive activity, LSD1 exerts a strong control on neural stem proliferation [26].

On the side of the cortical neurons development, last year a paper was published that revealed the importance of epigenetic control in the execution of neural development programs, specifically in the cerebral cortex. In fact the transition of newborn cortical pyramidal neurons between multipolar and bipolar stage results markedly deleyed by depletion of CoREST, a corepressor strongly associated to LSD1. The function of CoREST in this process is independent from REST but requires the activity of LSD1. Indeed, the loss of CoREST/LSD1 complex profoundly affects the onset of radial migration of the cortical pyramidal neurons and also perturbs the dynamic of neuronal precursor cell populations, transiently increasing the fraction of cells that remain in progenitor states [27].

Finally, also the issue of long-term memory consolidation mediated by LSD1 is worthy of note. One years ago J.M. Hooker generated two selective LSD1 inhibitors, RN-1 and RN-7, that exhibited good blood-brain-barrier penetration and retention. Systemic administration of RN-1 to rodents revealed for the first time that the inhibition of LSD1-mediated histone demethylation in the brain is implicated in long-term memory formation. Indeed, a memory test, called “novel object recognition”, demonstrated that rodents treated with RN-1 were significantly impaired in long-term memoryzation but not in short-term one [28].

Together these results highlight the importance of LSD1 demethylase activity in the brain in different pathways, but they also reveal that a lot of work has still to be done in order to deeply understand all the neuronal processes regulated by LSD1 and its splicing isoforms.

3. ALTERNATIVE PRE-mRNA SPLICING

3.1 MECHANISM OVERVIEW

Alternative splicing is the major contributor to protein diversity in metazoan organisms. Estimates of the minimum number of human gene products that undergo alternative splicing are as high as 60% [29].

In a typical multiexon mRNA, the splicing pattern can be altered in many ways (Fig. 3.1.1). Most exons are constitutive, this means that they are always spliced or included in the final mRNA. Besides, there are the so called “cassette exons”, which are regulated exons that are sometimes included and sometimes excluded from the mRNA, in a tissue- or condition-specific manner (Fig. 3.1.1 A, B). This is the case of the LSD1 exon E8a. Exons can also be lengthened or shortened by altering the position of one of their splice sites (Fig. 3.1.1 C, D). The 5'-terminal exons of an mRNA can be switched through the use of alternative promoters and alternative splicing (Fig. 3.1.1 E). Alternative promoters are primarily an issue of transcriptional control. Similarly, the 3'-terminal exons can be switched by combining alternative splicing with alternative polyadenylation sites (Fig. 3.1.1 F). Finally, some important regulatory events are controlled by the failure to remove an intron, a splicing pattern called intron retention (Fig. 3.1.1 G) [29]. This is what we observed with the intron E8b-E9. Particular pre-mRNAs often have multiple positions of alternative splicing, giving rise to a family of related proteins from a

single gene (Fig. 3.1.1 H). Changes in splice site choice can have different effects on the encoded protein. Small changes in peptide sequence can alter ligand binding, enzymatic activity, allosteric regulation, or protein localization. Genetic switches based on alternative splicing are important in many cellular and developmental processes, including sex determination, apoptosis, axon guidance, cell excitation and contraction, and many others. Errors in splicing regulation have been implicated in a number of different disease states [29].

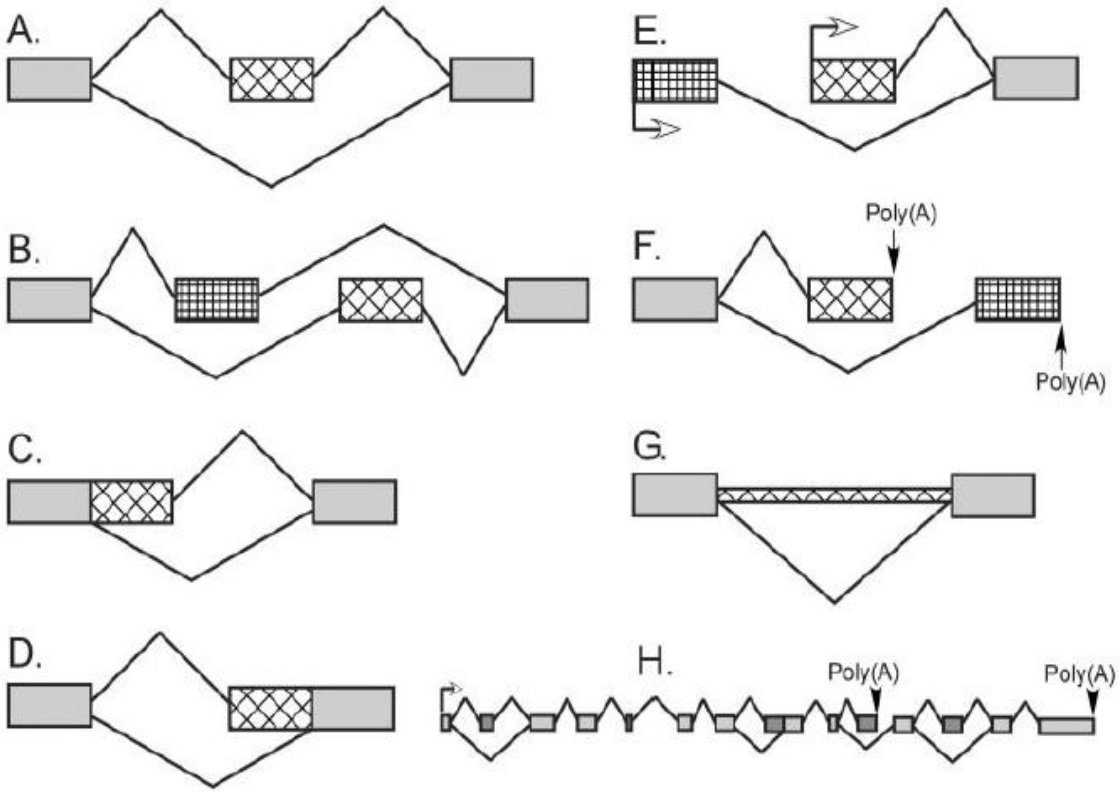


Fig. 3.1.1 Patterns of alternative splicing.

Constitutive sequences present in all final mRNAs are gray boxes. Alternative RNA segments that may or may not be included in the mRNA are hatched boxes. (*A*) A cassette exon can be either included in the mRNA or excluded. (*B*) Mutually exclusive exons occur when two or more adjacent cassette exons are spliced such that only one exon in the group is included at a time. (*C, D*) Alternative 5'and 3' splice sites allow the lengthening or shortening of a particular exon. (*E, F*) Alternative promoters and alternative poly(A) sites switch the 5- or 3-most exons of a transcript. (*G*) A retained intron can be excised from the pre-mRNA or can be retained in the translated mRNA. (*H*) A single pre-mRNA can exhibit multiple sites of alternative splicing using different patterns of inclusion. These are often used in a combinatorial manner to produce many different final mRNAs.

The excision of the introns from a pre-mRNA and the joining of the exons is directed by special sequences at the intron/exon junctions called splice sites. The 5' splice site marks the exon/intron junction at the 5' end of the intron. This includes a GU dinucleotide at the intron end encompassed within a larger, less conserved consensus sequence. At the other end of the intron, the 3' splice site region has three conserved sequence elements: the branch point, followed by a polypyrimidine tract, followed by a terminal AG at the extreme 3' end of the intron (Fig.3.1.2) [29].

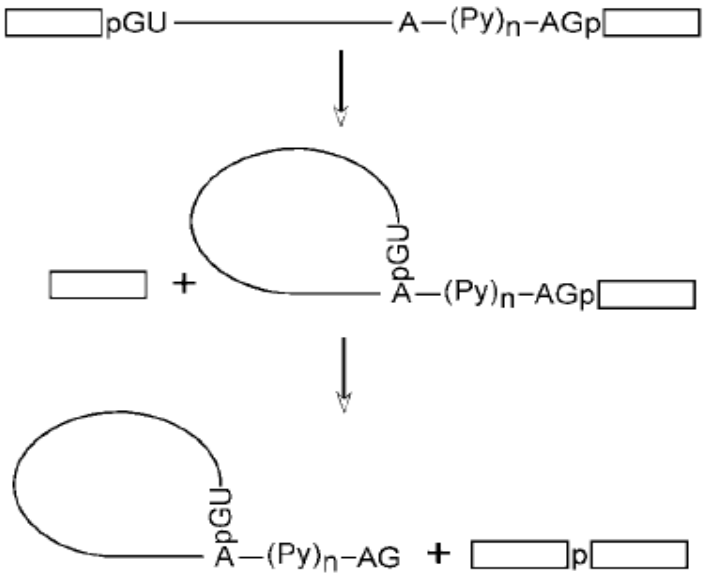


Fig. 3.1.2 Splicing takes place in two trans-esterification steps.

The first step results in two reaction intermediates: the detached 5' exon and an intron/3'-exon fragment in a lariat structure. The second step ligates the two exons and releases the intron lariat.

Splicing is carried out by the spliceosome, a large macromolecular complex that assembles onto these sequences and catalyzes the two trans-esterification steps of the splicing reaction. In the first step, the 2-hydroxyl group of a special A residue at the branch point attacks the phosphate at the 5' splice site.

This leads to cleavage of the 5' exon from the intron and the concerted ligation of the intron 5' end to the branch-point 2-hydroxyl. The second trans-esterification step is the attack on the phosphate at the 3' end of the intron by the 3'-hydroxyl of the detached exon. This ligates the two exons and releases the intron, still in the form of a lariat (Fig. 3.1.2) [29].

The spliceosome assembles onto each intron from a set of five small nuclear ribonucleoproteins (snRNPs) and numerous accessory proteins (Fig. 3.1.3). During assembly, the U1 snRNP binds to the 5' splice site via base pairing between the splice site and the U1 snRNA. The 3' splice site elements are bound by a special set of proteins. SF1 is a branch-point binding protein. The 65-kDa subunit of the dimeric U2 auxiliary factor (U2AF) binds to the polypyrimidine tract. In at least some cases, the 35-kDa subunit of U2AF binds to the AG at the intron/exon junction. The earliest defined complex in spliceosome assembly, called the E (early) or commitment complex, contains U1 and U2AF bound at the two intron ends. The E complex is joined by the U2 snRNP to form the A complex. The A complex is joined by the U4/U5/U6 tri-snRNP to form te B complex. The B complex undergoes a complicated rearrangement to form the C complex, in which the U1 snRNP interaction at the 5' splice site is replaced with the U6 snRNP and the U1 and U4 snRNPs are lost from the complex. It is the C complex that catalyzes the two chemical steps of splicing (Fig. 3.1.3) [29].

Changes in splice site choice arise from changes in the assembly of the spliceosome. In most systems, splice site choice is thought to be regulated by

altering the binding of the initial factors to the pre-mRNA and the formation of early spliceosome complexes. By the time the E complex is formed, it appears that the splice sites are paired in a functional sense and the defined intron is committed to being spliced.

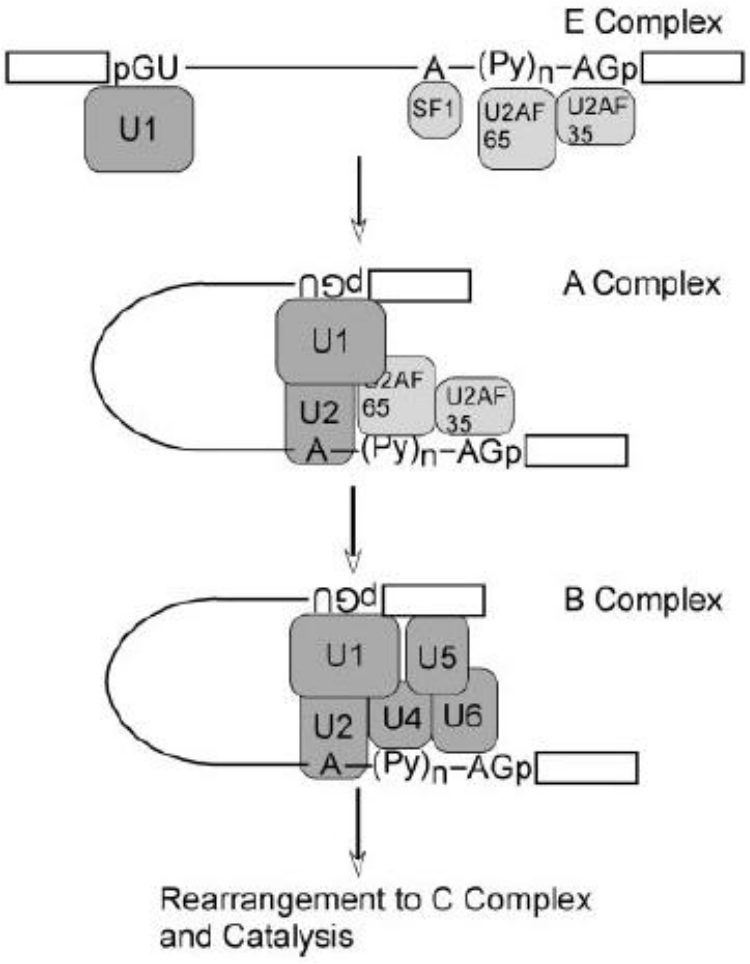


Fig. 3.1.3 The spliceosome contains five small nuclear ribonucleoproteins that assemble onto the intron.

The Early (E) complex contains the U1 snRNP bound to the 5' splice site. Each element of the 3' splice site is bound by a specific protein, te branch point by SF1 (BBP), the polypyrimidine tract by U2AF 65, and the AG dinucleotide by U2AF 35. This complex also apparently contains the U2 snRNP not yet bound to the branch point. The A complex forms when U2 engages the branch point via RNA/RNA base-pairing. This complex is joined by the U4/5/6 Tri-snRNP to form the B complex. The B complex is then extensively rearranged to form the catalytic C complex. During this rearrangement the interactions of the U1 and U4 snRNPs are lost and the U6 snRNP is brought into contact with the 5' splice site.

The splice site consensus sequences are generally not sufficient informations to determine whether a site will assemble a spliceosome and function in splicing. Other information and interactions are necessary to activate their use. Introns can range in size from less than 100 nucleotides to hundreds of thousands of nucleotides. In contrast, exons are generally short and have a fairly narrow size distribution of 50–300 nucleotides. Commonly, spliceosomal components binding on opposite sides of an exon can interact to stimulate excision of the flanking introns. This process is called “exon definition” and apparently occurs in most internal exons. On top of this process, there are many non–splice site regulatory sequences that strongly affect spliceosome assembly. RNA elements that act positively to stimulate spliceosome assembly are called splicing enhancers. Exonic splicing enhancers are commonly found even in constitutive exons. Intronic enhancers also occur and appear to differ from exonic enhancers. Conversely, other RNA sequences act as splicing silencers or repressors to block spliceosome assembly and certain splicing choices. Again, these silencers have both exonic and intronic varieties. Some regulatory sequences create an RNA secondary structure that affects splice site recognition, but most seem to be protein binding sites [29].

In the following paragraphs I will consider the role of cis-acting sequences and neurospecific trans-acting factors, as well as the putative function of the RNA secondary structure, in the alternative exon choise.

3.2 CIS-ACTING ELEMENTS INVOLVED IN SPLICING REGULATION

Accurate and efficient removal of introns from pre-mRNAs is essential to ensure correct gene expression. However, the information content present in the canonical splice signals (5′ splice site, branch site and 3′ splice site) is not sufficient to precisely define exons, indeed a large excess of sequences that conform to these weakly defined consensus elements is present in introns but these sequences are never used [30].

Studies of the molecular basis of splicing reveal the existence of exonic and intronic cis-acting regulatory sequences, which bind trans-acting factors and thus influence splice-site selection. These cis-acting elements are relatively short, usually 4–18 nucleotide, and they are classified as exonic or intronic splicing enhancers (ESE/ISE) and silencers (ESS/ISS) (Fig. 3.2.1). These elements become particularly important in the presence of weak splice sites or when alternative splicing is involved. An important goal toward defining the functions of AS events is to understand the mechanisms by which these cis-acting sequences in pre-mRNA combine with each other and with trans-acting splicing factors to define global splicing patterns.

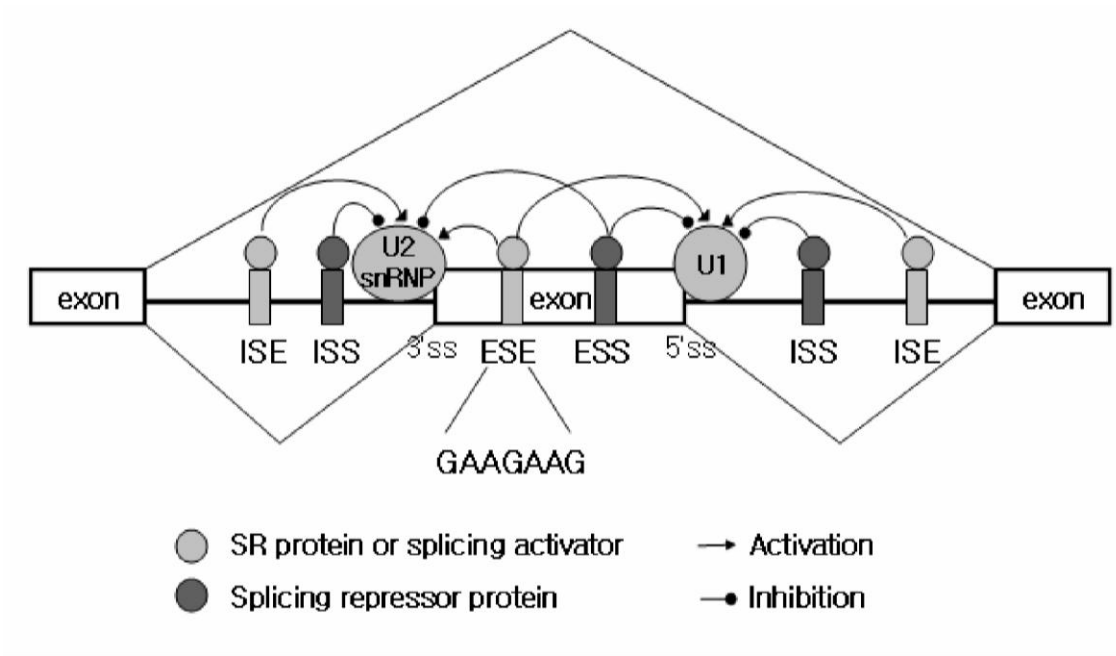


Fig. 3.2.1 Alternative splicing regulatory elements.

In addition to the splice-site consensus sequences, a number of auxiliary elements can influence alternative splicing. These are categorized by their location and activity as exon splicing enhancers and silencers (ESEs and ESSs) and intron splicing enhancers and silencers (ISEs and ISSs). Enhancers can activate adjacent splice sites or antagonize silencers, whereas silencers can repress splice sites or enhancers. Exon inclusion or skipping is determined by the balance of these competing influences, which in turn might be determined by relative concentrations of the cognate RNA-binding activator and repressor proteins.

Speaking about exonic splicing regulators, it is now well established that ESEs include a diverse range of sequences, and many if not all exons contain internal ESE sequences. Most ESEs function by recruiting members of the SR protein family. These factors usually regulate splicing by binding ESEs through their N-terminal RRM domains and mediating protein–protein interactions that facilitate spliceosome assembly through C-terminal RS domains [31]. ESSs, on the contrary, are often bound by splicing repressors of the hnRNP class, a diverse group of proteins containing one or more RNA-binding domains and sometimes splicing inhibitory domains such as glycine-rich motifs. hnRNPs function by a wide variety of mechanisms. For example, PTB (hnRNP I) can block essential interactions

between U1 and U2 snRNPs, whereas hnRNP A1 can inhibit splicing by binding on either side and “looping out” exons or by directly displacing snRNP binding [31]. Beyond molecular genetics, global approaches, including both computational and experimental methods, have been developed to identify ESEs and ESSs on a large scale. ESEs have been identified experimentally by *in vitro* and *in vivo* SELEX approaches. ESEs have also been computationally identified based on their enrichment in authentic exons versus introns and in exons with weak splice sites, and by their enrichment in authentic exons versus pseudoexons and 5′ UTRs of intronless genes. Additional exonic SREs have been predicted based on sequence conservation, but the activities of these elements as either splicing enhancers or silencers or neutral sequences were observed to depend heavily on their exonic context [31].

A number of intronic SREs are also known but fewer large-scale screenings have been conducted for intronic elements, and many more intronic elements likely remain to be identified. One well characterized ISE is the G triplet (GGG) or G run (Gn; n ≥ 3), which often occur in clusters and can enhance recognition of adjacent 5′ss or 3′ss. This ISE is common in GC-rich introns and is conserved between human and mouse. Intronic CA repeats in several cases can enhance splicing of upstream exons, probably through binding of hnRNP L. UGCAUG hexanucleotides or slight variations often occur downstream of neuron-specific exons and function as ISEs by binding to the brain- and muscle-specific splicing factors Fox-1/Fox-2. Pairs of YCAY motifs (Y = C or U) are recognized by the neuron-specific Nova

family of splicing factors to regulate a large number of splicing events in the brain. Interestingly, depending on their relative location in pre-mRNA, YCAY pairs can also function as either ESSs or ISSs [31]. Such context dependence will be discussed in more a detail way below. Intronic elements (ISS and ISE) are likely of primary importance in regulating AS events, as the intronic regions surrounding alternative exons are far more conserved in mammals than those surrounding constitutive exons, out to a distance of 150 bp or more. Such increased conservation has been used to predict unannotated alternative exons, and to predict intronic SREs. Splicing enhancers and silencers often function additively, with additional copies increasing their effect on splicing regulation , either because they increase the affinity of the associated factor or because they increase the copy numbers of the factor that are recruited, sometimes in synergistic fashion. Different SREs may also function cooperatively to regulate alternative splicing. For example, exonic UAGG motifs and intronic GGGG motifs overlapping the 5′ss can function cooperatively to silence the brain-specific CI cassette exon (exon 19) of the glutamate NMDA R1 receptor gene [31].

3.3 NEUROSPECIFIC TRANS-ACTING SPLICING MODULATORS

Alternative pre-mRNA splicing has the potential to greatly diversify the repertoire of transcripts in multicellular organisms. This expansive layer of gene regulation plays a particularly important role in the development and function of the nervous system. The nervous system, indeed, exhibits particularly high levels of alternative splicing [32].

Even if there are many ways to regulate alternative splicing, protein-RNA interaction is considered the primary elements of splicing regulation [33]. Several neuro-specific splicing factors are involved in the inclusion of neuronal alternative exons. Few of them will be described below.

NOVA1 AND NOVA2

The Nova RNA binding proteins are among the first tissue-specific regulators of alternative splicing to be identified. They were initially identified as the target antigens in the POMA disease, a rare immune-mediated disorder characterized by abnormal motor inhibition [40]. Studies in knock-out mice indicate that Nova-1 plays a critical role in the maintenance of brain stem and spinal cord neurons, likely through the regulation of alternative splicing in these cells [34]. Nova expression is indeed restricted to the central nervous system: Nova1 acts primarily in the spinal cord while Nova2 is expressed mostly in the brain [41].

Studies utilizing splicing-sensitive microarray profiling in mice lacking Nova proteins identified a network of brain-specific splicing events coordinated by these factors. Importantly, transcripts with Nova-regulated exons encoded proteins that were significantly enriched in functions associated with the synapse [35]. The integration of such alternative splicing regulatory network with genome-wide cross-linking and immunoprecipitation (CLIP) studies identifying *in vivo* Nova binding sites in the transcriptome, provided two key advances [36, 37] First, these data sets have enabled the formulation of "RNAmaps" correlating cognate Nova YCAY *cis*-element locations in pre-mRNA transcripts with effects on splicing regulation, leading to mechanistic insights into how the Nova proteins modulate alternative splicing [38]. Second, these networks identify target isoforms and pathways that likely contribute to aspects of neuronal physiology.

Nova factors recognize the *cis*-acting nucleotide motif YCAY (where Y indicate a pyrimidine), usually in clusters of multiple tetramers [42]. The positions of protein-RNA interactions play a major role in splicing regulation. Indeed, if the YCAY cluster lays inside an alternatively spliced exon or in the upstream intron Nova triggers the skipping of that exon, whereas if the YCAY cluster lays in the intron downstream of the regulated exon, Nova would promote its inclusion [39]. Nova effects on alternative splicing are due to its ability to physically block or promote the assembly of the spliceosome. One possible mechanism for the enhancing exon inclusion is that Nova can enhance the U1 snRNP recruitment to the 5' splice site of the preceding exon. This would restrict the temporal window

available for splicing inhibitory proteins to assemble to the alternative exon. On the contrary, Nova binding to an exonic YCAY cluster blocks U1 snRNP binding and exon inclusion. These observations suggest that Nova interacts with the pre-mRNA before snRNPs binding, to alter the composition of the pre-spliceosome complex [39].

It was also demonstrated that Nova1 can regulate the alternative splicing of the exon E4 in the Nova1 pre-mRNA itself, promoting both exon inclusion or exclusion. This exon encodes for a domain that contains a number of potential sites for phosphorylation by serine/threonine protein kinases. Indeed, Nova regulated alternative exons encode putative phosphorylation sites more often than constitutive or other alternative exons [43].

Numerous studies have revealed that tissue specific splicing factors act in concert to regulate alternative splicing. In fact, Nova can also work in a cooperative manner with other splicing factors; for instance, it was suggested a functional interaction between the neuronal splicing factor Fox and Nova proteins when they both bind the two sides of an intron [39]. Another example can be the interaction between Nova1 mRNA and nELAV splicing factors. nELAV proteins bind Nova1 mRNA at the AU-rich elements (ARE) in the 3' UTR region, increasing its stability and translation. Thus, the nELAV proteins are able to indirectly influence the splicing activity of Nova1, by controlling Nova1 gene expression and protein content at the post-transcriptional level [44].

Several studies demonstrated that splicing may be mechanistically coupled to

mRNA localization in the cytoplasm. It was indeed observed a significant amount of Nova in both nuclei and cytoplasm of mouse neurons. It has been also demonstrated that Nova and its nuclear RNA targets co-localize in dendrites closed to synaptic contacts. Within dendrites, Nova was mainly located peripherally, along the plasma membrane, in the proximity of synaptic contacts. In fact, Nova is able to bind 3' UTR elements and also contains NES and NLS elements, necessary to the transport through the nucleus. These observations suggest that the role of Nova in RNA regulation goes beyond its action in the nucleus [45].

nSR100/SRRM4

The neural-specific SR-related protein of 100 kDa (nSR100/SRRM4) was found to regulate alternative splicing decisions. Microarray profiling experiments in mouse neuroblastoma cells and tissues revealed that depletion of nSR100 results in increased skipping of alternatively spliced exons normally included in the brain, suggesting that it mainly acts to promote the inclusion of alternative exons. A significant fraction of genes containing these regulated exons are known to be important for neuronal differentiation, raising the possibility that specific splice variants modulated by nSR100 could contribute to this process. Consistent with this notion, nSR100 was found to play a critical role in neuronal differentiation and neurites extension *in vitro* as well as nervous system and sensory organ development in zebrafish embryos *in vivo* [46]. The specific mechanism by which

nSR100 regulates alternative splicing in the nervous system remains to be elucidated. It was found that the introns flanking alternative exons regulated by nSR100 are enriched in pyrimidine rich motifs [46]. The majority of these motifs are likely recognized by the polypyrimidine tract binding protein PTBP1 and its tissue-specific paralog PTBP2 (also called neural- or brain-enriched n/brPTB). Consistent with a link between these regulators in modulating neural-specific alternative splicing, many nSR100-dependent alternative splicing events are also regulated by PTBP1 and PTBP2.

More recently, it was discovered that nSR100 indirectly controls the steady-state abundance of a network of transcripts in neuronal cells distinct from the population of mRNAs that it regulates at the level of splicing. Depletion of nSR100 in mouse neuroblastoma cells led to decreased levels of hundreds of transcripts, and a subset of these changes were shown to be dependent on repressor element 1 silencing transcription factor (REST, also known as NRSF), a transcriptional repressor of genes involved in neurogenesis [47]. In neuronal cells, REST transcripts include an additional exon that results in the introduction of a STOP codon and production of a truncated protein lacking domains required for its repressive activity. Raj et al. found that nSR100 plays a critical role in promoting the inclusion of this alternative exon, suggesting that the expression of nSR100 in neurons contributes to the reduced activity of REST upon differentiation to the neural lineage. Importantly however, REST was also found to directly repress nSR100 transcription in non-neuronal cells and thus indirectly inhibit neural-specific

alternative splicing. This negative feedback loop between two gene regulatory levels was found to be important for developmental outcomes in the nervous system, as inhibiting nSR100 expression in mouse brain disrupted cortical neurogenesis, preventing neuronal precursor cells from committing to a neuronal fate [47]. These results are in agreement with previous studies showing that loss of REST derepresses neuronal transcripts in non-neuronal tissues, while REST overexpression inhibits the expression of transcripts in neuronal tissues, which in one study was shown to result in axon pathfinding errors in chick embryos [48].

PTBP1 AND PTBP2

PTBP1 and PTBP2 display mutually exclusive patterns of expression in the developing brain, with PTBP1 found in glial and non-neuronal cells, and PTBP2 in neurons [49]. This non-overlapping pattern of expression is established by an elegant cross-regulatory network where PTBP1 normally suppresses the inclusion of an exon in PTBP2 transcripts, leading to a non-functional isoform degraded by the non sense-mediated mRNA decay (NMD) pathway [49, 50]. In neurons however, PTBP1 is silenced by miR-124, a neuron-specific microRNA, leading to the derepression of PTBP2. [51] The consequences of modulating the relative levels of PTBP1 and PTBP2 in neuronal cells have been initially revealed through splicing-sensitive microarray profiling of mouse neuro-blastoma cells depleted of these factors. Analogous to the Nova-regulated alternative splicing network,

PTBP1-and PTBP2-dependent alternative splicing events are frequently found in transcripts expressed from genes with known roles in neuronal differentiation and physiology. [49] A role for PTBP1 and PTBP2 in regulating the expression of PSD-95, an important scaffolding protein essential for synaptic maturation and plasticity of excitatory neurons, has recently been identified. Overexpression of PTBP1 and PTBP2 in cultured hippocampal neurons was shown to repress synaptic activity, dendritic spine formation, and reduce levels of PSD-95 transcripts. This reduced mRNA abundance is caused by PTBP1 and PTBP2 binding to a pyrimidine rich *cis*-element upstream of PSD-95 exon18, leading to increased exon skipping and the production of a transcript containing a premature termination codon that is targeted for degradation by the NMD pathway. Importantly, the increased expression of PSD-95 in developing neurons in the cortex was found to correlate with three distinct phases of PTBP1 and PTBP2 expression. At the neural progenitor stage, when PTBP1 levels are high, PSD-95 expression is at its lowest. In embryonic neurons, the weaker repressor PTBP2 is more highly expressed while PTBP1 expression is lost, leading to intermediate levels of PSD-95. Finally, in post-natal cortical neurons, PTBP2 is no longer expressed, allowing PSD-95 abundance to reach its highest levels. These results indicate that the sequential changes in relative expression of PTBP1 and PTBP2 can allow for distinct splicing regulatory programs to be established at different stages in neuronal maturation [52].

RBFOX1 AND RBFOX2

Members of the Rbfox family of RNA binding proteins display enriched or highly specific expression patterns in the neuro-muscular system, and regulate alternative splicing decisions through interactions with the highly conserved *cis*-element (U)GCAUG. Focused biochemical studies and several genome-wide analyses have demonstrated that the Rbfox proteins can function as activators or repressors of splicing, depending on the location of (U)GCAUG elements in target pre-mRNA transcripts. When Fox proteins bind intronic regions downstream from the alternative exon they promote its inclusion; on the contrary, when the binding sites are present in upstream intronic regions, Fox acts as an alternative splicing repressor, blocking the formation of the early-spliceosome complex and inhibiting the recruitment of splicing factors. Together, these studies have begun to shed light on the relevant networks of transcripts modulated by these factors, although the role of the Rbfox proteins in nervous system development and function *in vivo* has remained somewhat unclear [53].

Two recent studies from the Black laboratory using Rbfox knockout mice have provided further insight toward the functional importance of this protein in the nervous system [54, 55]. Deletion of *Rbfox1* specifically in the nervous system of transgenic mice did not seem to have any effects on neuronal development or morphology in the brain. However, loss of *Rbfox1* did lead to spontaneous seizures, increased sensitivity to induced seizures, and increased excitability in

neurons of the dentate gyrus. Integration of splicing-sensitive microarray profiling and CLIP-Seq data sets identified alternative splicing events differentially regulated in the brains of *Rbfox1-/-* mice. Several of which were linked to genes known to be associated with epilepsy and others with roles in synaptic function [54].

In contrast to loss of Rbfox1 in the nervous system, deletion of the gene encoding Rbfox2 in the nervous system led to pronounced defects in cerebellar development. *Rbfox2 -/-* animals have much smaller cerebella than wild-type littermates, defects in Purkinje cell migration and dendritic arborization, and reduction in the migration and number of granule cells [55]. Again, splicing-sensitive microarray profiling experiments were performed, revealing alternative splicing events displaying significant changes upon loss of Rbfox2. Genes with affected exons were associated with neuronal developmental disfunctions, and a subset of Rbfox2-dependent alternative splicing events were also regulated by Rbfox1, suggesting partial redundancy between the two factors. In agreement with these data, double knockout mice displayed far more severe phenotypes than those observed in either single knockout mutant [55]. Finally, in an attempt to separate a possible role for both Rbfox proteins in the mature nervous system from their collective role in development, transgenic animals were generated that deleted these two factors specifically in Purkinje cells. Intriguingly, these double knockout mice possess no gross morphological or developmental abnormalities, but display impaired motor skills and significant reductions in spontaneous firing frequency of Purkinje cells, demonstrating that the Rbfox

proteins also play an important role in mature neural circuitry in addition to their contribution to development [55].

SAM-68

Sam68 is a nuclear RNA binding protein implicated in various aspects of mRNA metabolism, including splicing, nuclear export, somatodendritic transport, and translation. Sam68 belongs to the family of GSG (*G*RP33, *S*am68, *G*LD1) or STAR (*s*ignal *t*ransduction and *a*ctivation of *R*NA) domain proteins. This domain includes a central KH (hnRNP K homology) RNA binding domain flanked by conserved N- and C-terminal. The GSG domain in Sam68 binds to RNA motifs that are rich in A or U, such as UAAA or UUUA, and also mediates homodimerization. Sam68 contains a variety of other protein domains that allow its interaction and modification by multiple signaling pathways. These many regulatory interactions and posttranslational modifications affect the RNA binding activity and localization of Sam68 and make it an appealing molecule for transducing information from signaling systems to pathways of mRNA metabolism [56]. Sam68 is localized primarily in the nucleus as observed by immunofluorescence, consistent with its role in alternative splicing. Sam68 helps regulate the splicing of CD44 variable exon v5 in response to phosphorylation by extracellular signal-regulated kinase in T lymphoma cells [57]. Sam68 also regulates the alternative splicing of the apoptotic regulator Bcl-x, where it cooperates with hnRNPA1 to induce a switch from the antiapoptotic (Bcl-xL) to the proapoptotic (Bcl-xS) isoform [58].

The small number of confirmed Sam68 pre-mRNA targets has limited the

understanding of both its mechanisms of action and its cellular role. A recent paper shows that changes in Sam68 expression during neuronal differentiation affect a large number of specific splicing events. Such genes, carrying Sam68 target exons, are involved in a variety of cellular processes important in neurogenesis, including cytoskeletal organization (*Numa1, Clasp2*, and *Sgce*), organellar biogenesis and transport (*Bin1, Ktn1, Kifap3*, and *Opa1*), and synaptogenesis (*Cadm1, Dlgh4*, and *Sorbs1*). Sam68 controls splicing of its targets through direct binding to RNA elements but also through its interactions with other factors, such as PTB or hnRNP A1 [58]. Loss of Sam68 in P19 cells eliminates their ability to form neurons. Similarly, Sam68 knockdown in neuronal progenitor cells reduces the number of TuJ1-positive neurons that differentiate in culture. However, Sam68 null mice do not show a neurogenesis defect but instead show defects in bone mesenchymal cell differentiation. Only about one-third of Sam68−/−mice survive to adulthood, with the survivors showing defects in motor coordination [59]. Sam68 is widely expressed and plays roles in multiple cell types. It is possible that the neurogenesis defect we observed in vitro results from the loss of a general function that is not specific to neurons, possibly through Sam68 effects on the cell cycle [60]. Alternatively, it may be that in vivo other Sam68 family members, such as Slm-1 and Slm-2, can substitute Sam68 during neurogenesis. In the future it will be important to understand how this protein can coordinate splicing of certain pre-mRNAs in the nucleus and then follow these RNAs to the cytoplasm to see their interactions with a variety of signaling pathways.

3.4 PRE-mRNA SECONDARY STRUCTURE AND SPLICING REGULATION

There are two properties of RNA molecules that cannot be denied: their natural tendency to form highly stable secondary and tertiary structures *in vitro* and *in vivo*, and the observation that alterations in these structures represent a well-known mechanism to regulate the alternative splicing of many pre-mRNA molecules [61]. Three are the possible experimental situations in which pre-mRNA secondary structures are engaged. The first one is represented by the view that hnRNP proteins bind the RNA as it gets transcribed by RNA polymerase II and keep it in a largely linear conformation. In this case, binding of specific factors is regulated only by the competitive advantage provided by sequence-specific interactions over the generic RNA binding affinities of all hnRNP proteins (Fig. 3.4.1 A). The opposite situation is when the drive to form RNA secondary and tertiary structures is stronger than the ability of RNA-binding proteins to prevent it. In this case, the role played by generic RNA-binding proteins is severely reduced and specific complexes can bind through a mix of sequence-specific and structure-specific recognition (Fig. 3.4.1 B). Between these two models there is a situation that should encompass many cellular RNAs. In this case, the potential “ironing” of the RNA by its weak or aspecific interactions with hnRNPs can indeed maintain the mRNA in a largely linear conformation. However, in particular regions the mRNA is still able to form localized RNA structures which might represent, together with the nucleotide sequence, preferential binding sites for specific nuclear complexes

(Fig. 3.4.1 C). Hence, the predictive ability in the search for novel RNA binding targets for well-known proteins can be greatly enhanced if secondary structure is taken into consideration [61].

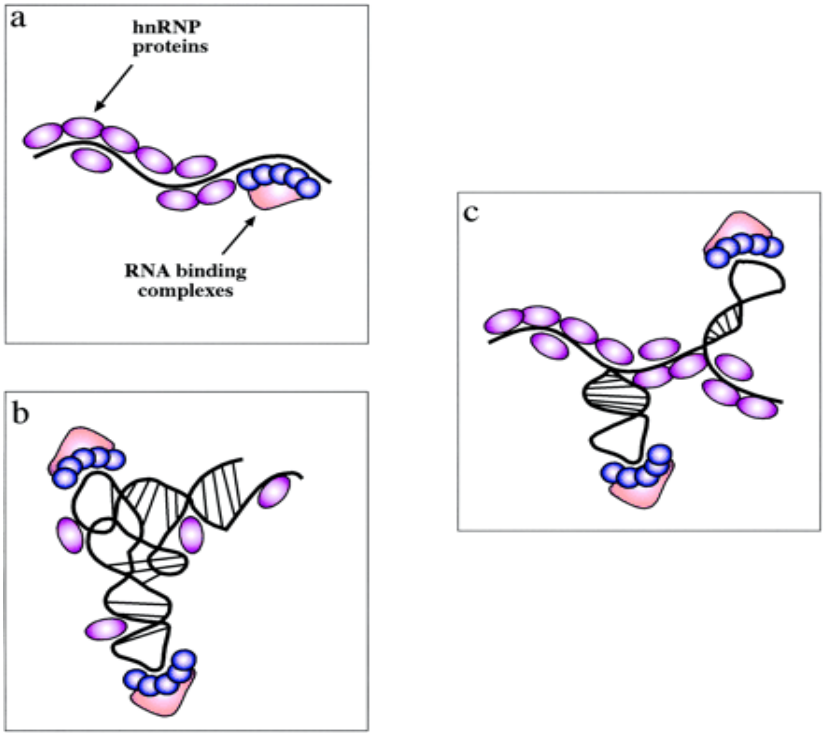


Fig. 3.4.1 Experimental models of pre-mRNA secondary structure.

(a) hnRNP proteins bind the RNA as it gets transcribed by RNA polymerase II and keep it in a largely linear conformation. (b) RNA secondary and tertiary structures are stronger than the ability of RNA-binding proteins to prevent it. (c) In particular regions the mRNA is able to form localized RNA structures which might represent, together with the nucleotide sequence, preferential binding sites for specific nuclear complexes.

With regard to specific factors capable of affecting the splicing process, it has to be noted that the binding of several positive (B52, SRp55, and NOVA-1) and negative (hnRNP) regulators of splicing have been shown to depend on RNA secondary structures as well as on the target nucleotide sequences. For obvious reasons, the earliest and most numerous reports regarding the ability of RNA secondary structures to affect the splicing process concern conserved key regions that define

an exon: 5′ splice site, 3′ splice site, and branch site. The most common case is represented by the presence of structural elements which may hinder the accessibility of such conserved sequences by basic splicing factors. Depending on the system analyzed, this alterated accessibility has been observed to involve only the acceptor site, the donor site, or both. With special regard to the 3′ splice site, it should be noted that recent attempts to correlate the presence of loosely defined secondary structures in 3′ splice site definition have resulted in a small (5 to10%) but significant improvement in predictive ability, indicating that this region may be particularly sensitive to the presence of structured RNA. Another mechanism by which pre-mRNA secondary structures may indirectly affect pre-mRNA splicing involves a change in the relative distance between conserved cis-acting splicing sequences. These changes can indeed determine considerable variation in splice sites usage or efficiency (Fig. 3.4.2 A). Finally, structural constraints may also have an effect on the branch site, indirectly promoting its use by keeping it in a single-stranded configuration (Fig. 3.4.2 B).

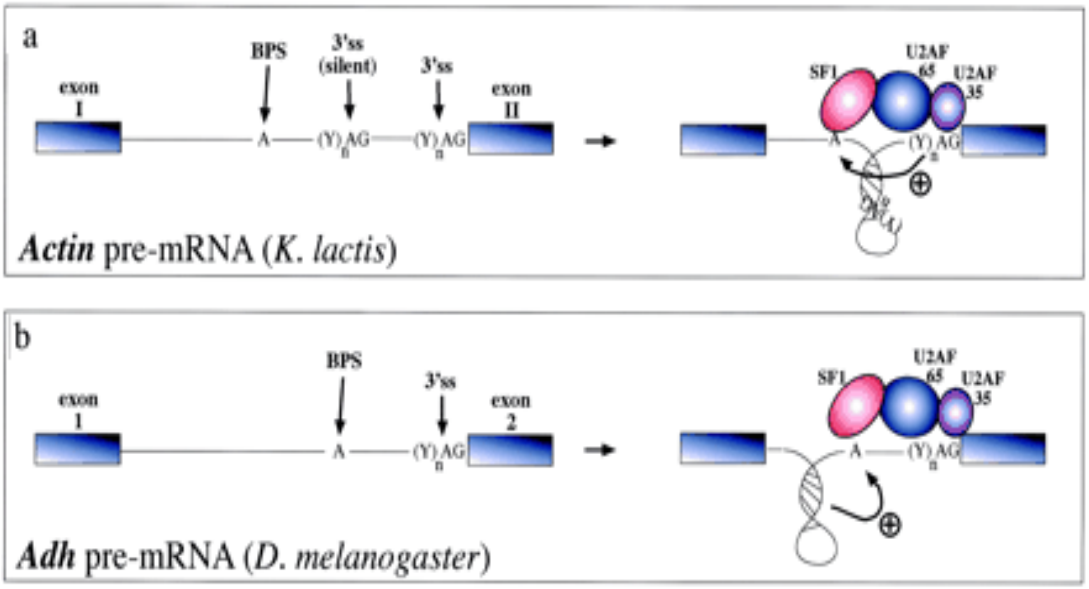


Fig. 3.4.2 RNA structural elements and splicing efficiency.

(a) Hairpins can change the relative distances of splicing regulatory elements and thus affect the final outcome. The function of this structure would be two fold: to bring the BPS into working range of the correct 3'ss and to sequester the silent 3'ss, preventing its use by the splicing machinery. (c) Hairpin structure formation near the branch point of the gene has been recently proposed to play an active role in splicing through a distinct mechanism. In this case, hairpin formation forces the branch point sequence (BPS) into an unpaired conformation that is better recognized by the splicing machinery.

In addition to splicing consensus sequences, there is also cases where structural constraints affect less-defined cis-acting sequences such as exonic/intronic splicing enhancers (ESE/ISE) or silencer elements (ESS/ISS). For example, in the case of fibronectin EDA exons, secondary structural elements can stabilize the conformation of the ESE sequence and enhance its SR protein binding capabilities (Fig. 3.4.3 A). Alternatively, RNA secondary structures can also function on ESE/ESS regulatory regions indirectly. In FGFR2 gene, the function of the stem structure formed by the intronic activating sequence 2 (IAS2) and intronic splicing activator and repressor (ISAR) elements would be that of approximating an inhibitory intronic sequence (a GCAUG-rich sequence) relative to the distal intronic splicing silencer (DISS) element, which would normally repress exon IIIb inclusion.

Inactivation of this element would then lead to activation of exon IIIb splicing in epithelial cells (Fig. 3.4.3 B) [61].

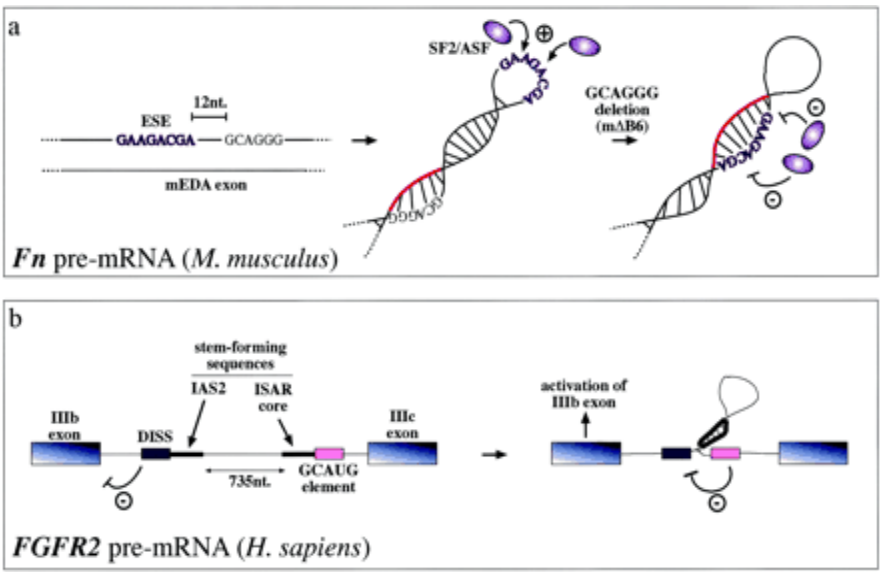


Fig. 3.4.3 Effects of RNA secondary structure on exonic/intronic enhancer or silencer elements.

(a) Mutations that do not directly affect the ESE sequence cause a conformational change in this region (from a loop to a stem) which hinders SF2/ASF protein binding abolishing exon recognition. (b) In *FGFR2* gene the function of the stem structure formed by the intronic activating sequence 2 (IAS2) and intronic splicing activator and repressor (ISAR) element is that of approximating an inhibitory intronic sequence (a GCAUG-rich sequence) relative to the distal intronic splicing silencer (DISS) element, which would normally repress exon IIIb inclusion. Inactivation of this element would then lead to activation of exon IIIb splicing in epithelial cells.

Considering the advancement in the study of RNA structure within the last few years, it is becoming increasingly apparent that RNA structure provides important supplementary information for any RNA function, that needs to be taken into account alongside the primary sequence information.

It is now evident that non-coding RNAs are functionally relevant, and it is very likely that the secondary and tertiary structures they adopt play an important role. Furthermore, there are several examples showing that mutations affecting mRNA structural features are related to human disease, and it is likely that more will

follow. A common effect is that binding of regulatory RNA binding proteins (RBP) is altered, and it has recently been shown that mRNA secondary structure can serve as a prediction tool for RBP binding sites. While this is useful in understanding the interaction between RBPs and mRNA in general, it will be particularly helpful in providing novel therapeutic targets in human disease. This potential has already been explored in tauopathies, and further small molecule drugs that can stabilize or destabilize specific mRNA structures could provide a powerful therapeutic tools [62]. RNAs are already used as drug targets and are generally preferred over DNA drug targets since they are more accessible. These drugs are the so-called “antisense drugs”, which target single-stranded mRNA sequences and bind them through complementary base pairing, thus blocking translation. In conclusion we can say that a full understanding of the functionality of mRNA structures will allow a finer tuned manipulation of gene expression, providing a tool for targeted and personalized therapy [63].

AIM OF THE PROJECT

Alternative splicing is the major source of proteome diversity in humans. In our laboratory a new splicing isoform of LSD1 was identified containing the mini-alternative exon E8a. This isoform is selectively expressed in the nervous system during the whole life, but plays an important role in brain development and synaptic maturation. In such critical phases, a fine modulation of LSD1-E8a expression allows differentiating neurons to acquire proper morphology [12]. Furthermore, the presence of the mini-exon E8a introduces a new phosphorylation site at Thr371 that modulates neuronal LSD1 epigenetic function. Indeed, in cortical neurons Thr371 phosphorylation was found in our group to be pivotal in enhancing neurite outgrowth and in hampering the transcriptional repressive activity of neuronal LSD1 [13].

In the light of these data, the aim of my project is on the one hand to study how the alternative splicing of exon E8a is regulated and on the other hand to clarify the biological function of LSD1-8a isoform *in vivo* and its possible involvement in neurological disorders. Concerning exon E8a inclusion into LSD1 mature transcript, we want to investigate the function of the *cis*-acting regulatory sequences present in the introns flanking the exon E8a, as well as the role of the *trans*-acting factors directly acting on splicing regulation.

On the side of neuronal LSD1 biological relevance, we intend to study *in-vivo* the function of LSD1-E8a isoform by generating a knock-out mouse model for the sole exon E8a of LSD1. By different approaches, ranging from anatomical and behavioural studies to biochemical assays and transcriptional profiles, we propose

to characterize the mutants versus wild-type mice in order to identify physiologic or pathologic pathways involving neurospecific LSD1- E8a isoform.

MATERIALS AND METHODS

1. DNA CLONING

DNA cloning, also known as recombinant DNA technology, refers to the process of creating multiple copies of a DNA fragment by *in vitro* methods. There are various procedures of DNA cloning, but some steps are constant for all. The process begins with the isolation of a DNA stretch, using restriction enzymes or Polymerase Chain Reaction (PCR) with chemically synthesized oligonucleotides. The isolated fragment is then linked to a molecule that is capable of replicate and propagate both itself and the fragment linked to it. To this aim, a restriction enzyme cuts the self-replicating molecule so that the isolated DNA can be linked by a ligation procedure. The DNA stretch artificially linked to a backbone vector is called recombinant DNA. The new-born plasmid is then inserted into host bacterial or mammalian cells to obtain its replication and the increasing in the number of copies. To this aim different techniques like chemical sensitivation or electroporation can be used. The plasmid usually contains selectable antibiotic resistance markers and/or color selection markers which make it easier to know if the cells have been successfully transfected with the plasmid. The antibiotic resistance markers allow only cells in which the plasmid has been transfected to grow up. Transfected cells are then cultured, and proliferation of the recombinant DNA takes place. The resulting clones are genetically identical organisms containing the recombinant DNA. This can be confirmed by using PCR, restriction fragment analysis, or DNA sequencing methods. Cloning DNA is helpful in getting an insight into an organism genetic make-up and how this affects and influences the organism life processes.

1.1 POLYMERASE CHAIN REACTION (PCR)

PCR is one of the best known and useful technique that allows any stretch of DNA to be amplified *in vitro.* It requires a DNA template, two DNA primers, and a particular DNA polymerase isolated from *Thermus aquaticus,* a bacterium that grows at 70°C. The main properties of this polymerase are that it is capable of synthesizing DNA at 70°C and that it is stable at even higher temperatures.

The template used in PCR is a double-stranded DNA but the two strands have to be separated before DNA synthesis taking place. To denature DNA, the PCR reaction mix is first placed at 90–95°C, then temperature is lowered to ~50–68°C to allow primers to anneal to the single-strand template and finally the temperature is raised again to 70-72°C to allow DNA to be elongated. This cycle of denaturation, annealing and elongation is repeated between 25 and 35 times for each PCR reaction. The DNA primers used in PCR are chosen so that only the stretch of DNA of interest is amplified and they are chemically synthesized *in vitro*. If the PCR products have to be cloned, the 5' ends of each primer contain a restriction site so that the resulting fragment is digested with the corresponding enzyme and cloned into the appropriate vector. The template DNA can derive from any source and for each reaction only a small amount is needed. DNA replication that takes place in PCR, like *in vivo* DNA replication, is not 100% accurate and occasionally a mistake occurs: if the amplified fragment has to be cloned, the resulting clones must be sequenced to ensure that they carry the wild-type copy of the gene, while if the

amplified fragment has to be used as a probe, a few mutant copies in a mixture containing a large number of wild-type copies will not represent a problem.

The thermal procedure we perform is the standard one, even if it may change in the annealing temperature (*Ta*) and in the number of cycles, according to the amplified sequence. *Ta* temperature is calculated by decreasing 8-10°C from the melting temperature (*Tm*). T*m* = [4*(G+C)+2*(A+T)]. In every cycle, each new DNA fragment works as template for Taq polymerase. Final amplicons number will be 2^n (n=cycle number). PCR products are then loaded on agarose gel, in order to verify the correct amplification.

For each PCR reaction we used the following reagents:

	Final Concentration			
DNA	< 50 ng			
Buffer 5X	1X	Cycle 1	95°C x 5'	Initial. Denaturation
$MgCl_2$ 25 mM	2 mM		95°C x 30"	Denaturation
Forward Primer 10 µM	0,2 µM	Cycle 2	Tm x 30"	Annealing
Reverse Primer 10 µM	0,2 µM	x 30	72°C x 30"	Elongation
dNTPs 10 mM	0,2 mM	Cycle 3	72°C x 7'	Final Elongation
Taq Polymerase 5 U/µl	1,25 U			
H_2O to 25 or 50 µl				

Primers used in PCR reactions:

hLSD1 cloning intron FW
GGAATTCCATATGTCACTGCTTTCTGACACTTCCCC
hLSD1 cloning intron RV
GGAATTCCATATGACACAAGTGTGCAAATGGTACAGAT
FW AseI hLSD1-MG

GGGAATTCATTAATTCACTGCTTTCTGACACTTCCCCCTCACTCC

RV AseI hLSD1-MG

CCCTTAAGATTAATCTGAAAATCCGAAGTATAGCAGAGTTGGTCC

New 2 del pal hLSD1 RV

AGAGAAATATTTAAAGGGTTAGAGGGACCAAATTTG

New 3 del pal hLSD1 FW

CAAATTTGGTCCCTCTAACCCTTTAAATATTTCTCT

A2-3 FW Fluo

[6-FAM] CAACTTCAAGCTCCTAAGCCACTGC

BraII RV

CACCAGGAAGTTGGTTAAATCA

FN1 FW

TGGAATTATTAAAAGCATTGCTGA

FN1 RV

GTTGCCACACAGCAAAGAGA

hB-actin FW

GCGGGAAATCGTGCGTGACATT

hB-actin RV

CTAGAAGCATTTGCGGTGGA

NOVA1 BclI FW

TCAGAATGATCAACATGATGGCGGCAGCT

NOVA1 HindIII RV

TCGAAGCTTCAACCCACTTTCTGAGG

FOX-1 cloning FW

GCTCTAGAATGGCTCAGCCTTACGCTTC

FOX-1 cloning RV

CGGGGTACCTTAGTATGGAGCAAAACGGTTGTAT

P3

ACGCGTCGACTCTTCAGTGCTTTCTCACTCCCA

P6

ATAGTTTAGCGGCCGCCCTCTATTTTCTGAGCAGCC

Neo PL451 BamHI screening (PB)

CAGCTGGGGCTCGACTAGAGCTTGC

hLSD1 8b FW

CTTTGAGGGGAAGCCAGATACC

hLSD1 8b RV

CTGAGGACCTTCCAAGAATAAGG

mLSD1 8b FW

TCATCCTTGAGCAGGTAAC

mLSD1 8b RV

CAGGTTTATTATTGAGGACG

m/rRT-Bact-FW

GCCTTCCTTCTTGGGTAATGG

m/rRT-Bact-RV

AATGCCTGGGTACATGGTGG

hmrFAMex8 LSD1FW

[6FAM] GAAAAGGAAACTATGTAGC

hFAMex2 LSD1FW

[6FAM] GTGAGCCTGAAGAACCATCG

hLSD1ex9 RV

CTACCATTTCATCTTTCTCTTTAGG

mrFAMex2 FW

[6FAM]AGTGAGCCGGAAGAGCCGTCTG

mrLSD1ex9 RV

CTACCATTTCATCTTTTTCTTTTGG

m/rNOVA1 FW

CATTGAACTATATGGACTCTTCA

m/rNOVA1 RV

TGAGTCCTACTCTTACAGTTTG

hmr PTBP2 FW

TAAGTTGGTGCTGTATTGAAG

hmr PTBp2 RV

AGATCAGGTCGAGTATAATCC

1.2 AGAROSE GEL ELECTROPHORESIS

Electrophoresis is a process which enables the sorting of DNA molecules depending on their size. Using an electric field, DNA can be made to move through a gel composed of agarose. DNA molecules are dispensed into different wells in the gel material together with 1X loading dye. The most common dyes are bromophenol blue (Sigma B8026) and xylene cyanol (Sigma X4126): the former migrates at a rate equivalent to 200–400bp, while the latter migrates at approximately 4kb. DNA density is provided by glycerol or sucrose. The gel is placed in an electrophoresis chamber, which is then connected to a power source. We always run agarose gels in an electrical field of about 100 mV, in Tris-Borate-EDTA (TBE) buffer 0,25X (Tris-borate 0,089M, Boric Acid 0,089M and EDTA 0,002M). When the electric current is applied, the larger DNA fragments are slower to migrate than small molecules, which in turn move faster. Thus, molecules of different weight form distinct bands on the gel. Molecular weight markers are available that contain a mixture of molecules of known sizes. We always use about 500 ng of pUC9 vector, digested with HaeIII enzyme, or Lambda DNA, digested with PstI enzyme. After the electrophoresis is complete, DNA molecules in the gel can be stained using ethidium bromide (EtBr) 0,5 mg/ml, that becomes fluorescent under ultraviolet light. We use 3 µl of EtBr (Sigma Aldrich) in 40 ml of final gel volume. We make agarose gels with 0.7-2% of agarose powder dissolved in TBE electrophoresis buffer. 0.7% gels perform good separation (resolution) of large DNA fragments (5–10kb) while a 2% gels are able to better resolve small fragments (0.2–1kb).

1.3 DIGESTION WITH RESTRICTION ENZYMES

Once DNA has been purified, it must be cut to ligate the ends of the DNA fragment to the ends of the DNA plasmid. A group of enzymes, called R*estriction Enzymes*, are used to this purpose. The restriction enzyme recognizes a specific DNA restriction site, typically four, six, eight, ten, or twelve nucleotides-long, and it cleaves the DNA by introducing breaks in specific phosphodiester bonds. The cleavage is on both the DNA strands so that a double-stranded break is made. 1 U of enzyme is necessary to digest 1 µg of DNA, but to improve the performance of the reaction it is better to use an excess of 3-4 U of enzyme for each µg of DNA. However, it is important not to use more enzyme than 10% of the final reaction volume because the enzyme storage buffer contains glycerol that inhibits the digestion. Each enzyme requires a specific reaction buffer, with a peculiar salt and Magnesium concentration (high or low) and pH. These 10X buffers are sold with the restriction enzyme. Most companies have about 4 different kinds of buffer (called A,B,C,D or 1,2,3,4 etc.) and the enzyme/buffer compatibility has to be check in the enzyme company informations. A few enzymes require special conditions, such as BSA (bovine serum albumin) added to the mixture. This is usually provided with the enzyme at 100X concentration.

A typical reaction is conducted for 2-3 hours or overnight and the temperature varies, according to the enzyme. If suggested by manufacters the enzyme has to be heat-inactivated, generally at 65°C or 80°C for 20 min.

1.4 DEPHOSPHORYLATION OF LINEARIZED PLASMID DNA FOR LIGATION

Dephosphorylation of linearized plasmids is a method to prevent self-ligation of vectors with blunt-ends or with ends digested with a single restriction enzyme. This property can be used to decrease the vector background in cloning strategies but after ligation and transformation, it is necessary to perform a diagnostic experiment (e.g. colony PCR with suitable primers) to ensure that clones contain DNA fragment in the correct orientation.

Calf Intestinal alkaline Phosphatase (CIP) is the most frequently used enzyme to remove 5' phosphate groups from cloning vectors. The reaction is conducted at 37 °C for 3-4 hours, using the following protocol.

DNA		100-1000ng
Buffer 10X		1X
CIP	10.000 U/ml	10 U
H_2O to 20 ul		

1.5 DNA PURIFICATION FROM GEL SLICE

Agarose contains various impurities which may inhibit downstream reactions if not efficiently removed from DNA. Spin-column kits are very efficient in removing these kind of contaminants. To this purpose we use *The Wizard® SV Gel and PCR Clean-Up System* (Promega) which extracts DNA fragments from 100-bp to 10kb from the agarose gels, or purifies PCR products directly from amplification reactions. This membrane-based system can bind up to 40µg of DNA, taking advantage from the ability of DNA to bind silica membranes in the presence of chaotropic salts. Purified DNA can be used for different applications without further manipulation.

This method requires the band to be cut from the gel. To visualize DNA band the trans-illuminator has to be set to long-wavelength UV (or low-power) and the time of exposure has to be minimized, because UV mutagenesis DNA in a measurable rate. Using a scalpel blade to cut around the band, it can be removed from the gel and placed in a 1.5mL microfuge tube. The basic protocol of DNA extraction requires:

- Addition of Membrane Binding Solution to the gel slice in a ratio of 10 µl of solution per 10 mg of gel slice:

- Incubation at 60 °C to dissolve the gel:

- Transfer of the mixture to the SV Minicolumn and centrifugation to discard the liquid solution;

- Washing of the column for 2 times with the Membrane Wash Solution (containing Ethanol), ensuring that no Ethanol residue remains in the column;

- Elution of the DNA in at least 30µl of Nuclease-Free Water;

1.6 DNA PRECIPITATION FROM PCR REACTION

Another commonly used protocol to purify PCR products from the amplification reaction makes use of salt and ethanol added to the aqueous solution to force the nucleic acids to precipitate. The positively charged salt neutralizes the DNA negatively charged backbone and in particular the $(PO_3)^-$ groups of the nucleic acid, making it far less hydrophilic and much less soluble in water. On the other hand, ethanol has a much lower dielectric constant compared to the water, making the interaction of the salt with $(PO_3)^-$groups easier. Thanks to this property, ethanol shields acid nucleic charges and causes its dropping out of the solution. Centrifugation of the precipitated nucleic acids isolates it from the rest of the solution. The pellet is then washed with cold 70% ethanol to remove all the salt. A second centrifugation step is used to separate the nucleic acid from the ethanol, allowing it to be removed. The nucleic acid pellet is then dried and re-suspended in aqueous TE (Tris EDTA) buffer or nuclease-free water.

The classic PCR products precipitation protocol involves:

1 Volume of Ammonium Acetate (2M final conc.)

2 Volume of cold isopropanol.

In this step it is possible to add also 1 µl of glycogen, making the nucleic acid

heavier and visible.

- Incubation at -20 °C over night or 30' in ice;

- Precipitation by centrifugation;

- Washing with 100 ul of 70% cold EtOH;

- Resuspension in 30 ul of TE or water.

1.7 LIGATION

DNA ligation allows two DNA molecules to be joined together, through the generation of a phosphodiester bond between the 3' hydroxyl group of a nucleotide and the 5' phosphate of another. This reaction is usually catalyzed by a DNA ligase enzyme which is able to join DNA fragments having blunt or 'sticky' ends. T4 DNA Ligase (New England Biolabs), from T4 bacteriophage, is the most commonly used enzyme for this purpose. It is supplied with its specific 10X reaction buffer supplemented with 1mM ATP, which ensures the optimal activity of the enzyme. Ligation may be done at room temperature (20-25°C), in 10 minutes for cohesive ends and in 2 hours for blunt ends, or over-night at 16°C. At the end of the reaction T4 DNA Ligase has to be inactivated at 65°C for 20' min. Before using DNA for bacterial transformation it has to be dialyzed to remove buffer salts. To this aim we use Millipore Membrane Filters.

We always perform ligation protocol like this:

DNA	100 ng
Buffer 10X	1X
T4 Ligase 400U/ul (NEB)	400U
H_20 to 20 ul	

1.8 GENERATION OF COMPETENT BACTERIAL CELLS (TOP10)

The procedure to prepare competent cells can sometimes be tricky. Bacteria are not very stable when they have holes in their membrane and they easily die. A poorly performed procedure can result in cells that are not very competent to take up DNA, while a well-performed procedure will result in very competent cells. The competency of a stock of cells is determined by calculating how many E. Coli colonies are produced per 1 µg of DNA added: a good preparation of competent cells will give at least 10^8 colonies per 1 µg of DNA, a poor preparation will be about 10^4/ug or less. First of all it is necessary to streak out a frozen glycerol stock of *TOP 10* bacterial cells onto an LB plate without antibiotics and let the plate grow overnight at 37°C. To prepare the starter culture, we select a single colony of E. Coli from the fresh *TOP 10* LB plate and we inoculate it overnight in 3-4 mL of LB at 37°C in a shaker. This is the protocol to follow the day after:

- Inoculate 0,5 L of LB media with the starter culture and grow at 37°C in a shaker. Measure the OD600 every hour, then every 15-20 minutes when the OD gets above 0,2;

- When the OD600 reaches 0.4-0.5, immediately put the cells on ice. It is also very important to keep the cells at 4°C for the remainder of the procedure. The cells and any bottles or solutions that they come in contact with the cells must be pre-chilled to 4°C;

- Split the 0,5 L culture into two 250 mL centrifuge bottles;

- Harvest the cells by centrifugation at 800*g* for 30-40 minutes at 4°C;

- Decant the supernatant and resuspend each pellet in 200 mL of ice cold ddH2O;

- Harvest the cells by centrifugation at 800*g* for 30 minutes at 4°C;

- Decant the supernatant and resuspend each pellet in 100 mL of ice cold ddH2O;

- Combine resuspensions into one centrifuge bottle and harvest the cells by centrifugation at 800*g* for 20 minutes at 4°C. At this step, rinse two 50 mL conical tubes with ddH2O and chill on ice;

- Decant the supernatant and resuspend each pellet in 40 mL of ice cold 10% glycerol;

- Harvest the cells by centrifugation at 1000*g* for 20 minutes at 4°C;

- Carefully aspirate the supernatant with a sterile Pasteur pipette and resuspend each pellet in 1 mL of ice cold 10% glycerol by gently swirling;

- Aliquot into sterile 1.5 mL microfuge tubes and store in the -80°C freezer;

1.9 TRANSFORMATION BY ELECTROPORATION

Electroporation is a very basic technique used to transform bacterial cells. In this procedure, an electric pulse temporarily disturbs the phospholipid bilayer, allowing molecules like DNA to pass into the cell. Since the phospholipid bilayer of the plasma membrane has a hydrophilic external side and a hydrophobic internal one, any polar molecules, including DNA, are unable to freely pass through the membrane. The concept of electroporation capitalizes on the relatively weak nature of the phospholipid bilayer hydrophobic/hydrophilic interactions and its ability to spontaneously reassemble after disturbance. Thus, a quick voltage shock may temporarily disrupt areas of the membrane, allowing polar molecules to pass through, but after the electric shock plasma membrane may reseal quickly, leaving the cell intact.

The electroporation apparatus is typically commercially produced and purchased. We use *Bio-Rad Gene Pulser II Electroporator ,* set up at these conditions: 2,5kV, 25mF, 200Ω.

Electroporation is performed as described here:

- Add up to 10 ng of plasmid DNA to a tube containing 40 µl of fresh or thawed bacterial competent cells in ice. Mix by swirling with pipette;

- Transfer DNA and cells to a pre-chilled electroporation cuvette (0.2 cm electrode gap) using a narrow pipette tip. Wipe any ice or water from sides of cuvette and place the cuvette into the sample chamber;

- Energize the electroporation apparatus and deliver the pulse by pushing both charging buttons simultaneously and holding until a short beep is heard;

- Remove the cuvette from the sample chamber, add immediately 1 ml of LB medium and transfer the cells to a sterile culture tube;

- Incubate cultures for 60 minutes at 37°C in a moderate shaking to allow bacterial cells to restore their plasma membrane;

- Plate aliquots of the electroporation mixture on agar plates supplemented with the appropriate antibiotics and incubate at 37°C over night.

1.10 COLONIES SCREENING BY PCR AND RESTRICTION ENZYMES DIGESTION

This protocol is designed to quickly screen for plasmid inserts directly from *E. coli* colonies. Even though blue/white screening can be used to determine if inserts are present, these techniques are used to determine insert size and/or orientation in the vector. The procedure of screening by PCR is identical to a standard PCR, but the template is the colony itself instead of a DNA sample. If the restriction sites are maintained in the final plasmid after the cloning procedure, we also perform a control digestion with the same restriction enzymes used to cut both vector and fragment in order to verify if the results is exactly the expected one.

1.11 COLONY BACTERIAL CULTURE

Escherichia Coli may be grown on solid media or in liquid one. Lysogeny Broth (LB) is the most commonly used medium in molecular biology for *E.coli* cell culture. LB broth contains the enzymatic digestion product of casein known as *peptone* (some vendors term it *Tryptone*), yeast extract, and sodium chloride. Peptone is rich in amino acids and peptides. Its composition reflect the casein. In addition to amino acids and peptides, yeast extract also contains nucleic acids, lipids and other nutrients which are needed for bacterial growth.

For our culture we used LB medium with the following composition:

- liquid LB: yeast extract 0,5 %, bacto tryptone 1 %, NaCl 1 %;

- solid LB: yeast extract 0,5 %, bacto tryptone 1 %, NaCl 1 %, agar 1,5%.

The medium is then added with specific antibiotic, depending on the resistance carried by the plasmid:

Kanamycin 50 µg/ml
Chloramphenicol 50 µg/ml
Ampicillin 50 µg/ml.

1.12 MINI- MIDI- and MAXI-PREP

Plasmids are used to carry a specific DNA sequence inside the cell, but for their use as vectors they need to be isolated from the rest of the bacterial genome.

To purify our plasmids we used *QIAGEN Plasmid Kits*. These protocols are designed for preparation of up to 20 µg of plasmid DNA with *QIAGEN Plasmid*

Mini-Kit, up to 100 µg of high- or low-copy plasmid DNA with the *QIAGEN Plasmid Midi-Kit* and up to 500 µg using the *Maxi-Kit*.

Briefly, all of the QIAGEN Plasmid Kits protocols are based on the same steps. Bacterial lysates are cleared by centrifugation. The cleared lysate is then loaded onto the anion-exchange tip where plasmid DNA selectively binds under appropriate low-salt and pH conditions. RNA, proteins, metabolites, and other low-molecular-weight impurities are removed by a medium-salt wash and ultrapure plasmid DNA is eluted in high-salt buffer. The DNA is then concentrated and desalted by isopropanol precipitation and collected by centrifugation. Table below shows LB culture volume that we use for each kit, depending on the type of plasmid:

	High copy plasmid	Low copy plasmid
QUIAGEN Plasmid mini-kit	3 ml	Not recommended
QUIAGEN Plasmid midi-kit	25 ml	100 ml
QUIAGEN Plasmid maxi-kit	100 ml	500 ml

DNA concentration can be evaluated by both UV spectrophotometer at 260 nm or quantitative analysis on an agarose gel.

1.13 SEQUENCING PCR

DNA sequencing includes different methods and technologies used for determining the order of the nucleotide bases in the DNA molecule. Nowadays chain-terminator method (also called Sanger method) is the technique of choice. The key principle of the Sanger method is the use of dideoxynucleotides triphosphates (ddNTPs) as DNA chain terminators. The ddNTPs lack the 3'-OH terminal group, so when they are randomly incorporated, they block the reaction lengthening.

To obtain our sequences we use *BigDye Terminator v3.1 Cycle Sequencing Kit* (Applied Biosystem) which permits sequencing in a single reaction because each of the four dideoxynucleotide chain terminators is labelled with a dye that emits light at a specific wavelength. All of the newly synthesized and labelled DNA fragments are heat denatured and separated by size with a resolution of one nucleotide, using a capillary electrophoresis instrument called *ABI PRISM® 3130 xl Genetic Analyzer (*Applied Biosystem). The outcome of the electrophoresis automatic sequencing is an electropherogram with unequal peak height, color and shape for each of the four nucleotides. DNA sequences of high quality typically result in a read length of 650 to 750 bases with an accuracy of 99%, characterized by sharp peaks and little to no background.

This is the protocol for a sequencing reaction:

	Final Concentration		
DNA	plasmidic DNA 300-400 ng	Cycle 1	95°C x 2.30'
	PCR products 15-20 ng		95°C x 30"
Buffer 10X	1X	Cycle 2	50°C x 20"
FW or Rv primer 3,2 µM	0,5 µl	x 25	60°C x 4'
Big dye terminator	1 µl	Cycle 3	72°C x 7'
H_2O to 10 µl			

1.14 DNA PURIFICATION FROM SEQUENCING REACTION

Before loading PCR products from a sequencing reaction on the capillary electrophoresis instrument, they have to be purified. Instead of ammonium acetate, which is specific in removing dNTPs, we use sodium acetate (0.3M final conc, pH 5.2) that is suitable for routine DNA precipitations.

This is the protocol we follow:

- 1/10 V NaAc and 2V 100% cold EtOH added to the reaction;
- incubation for 20 min. at room temperature;
- precipitation of DNA by centrifugation;
- washing with 70% cold EtOH;
- resuspension in 12 µl of denaturant fomamide;
- denaturation for 5' at 95°C.

2. CELL CULTURE

Animal or plant cells, removed from their tissues, will continue to grow if supplied with the appropriate nutrients and conditions. When carried out in a laboratory, this process is called cell culture and it occurs *in vitro*. Cells in culture may be genetically identical (homogenous population) or may show some genetic variation (heterogeneous population). A homogenous population of cells, derived from a single parental cell, is called *clone*. Therefore all cells within a clonal population are genetically identical.

Freshly isolated cultures from mammalian tissues are known as *primary cultures*. At this stage, cells are usually heterogeneous but still closely represent the parental cell types in the expression of tissue specific properties. After several sub-cultures onto fresh media, cell line will either die or 'transform' to become a *continuous cell line*. Such cell lines show many alterations from the primary cultures including change in morphology, chromosomal variation and increase in capacity to give rise to tumors in hosts.

For our experiments we used the following cell lines:

HELA cells: human epithelial cells from cervical carcinoma. HeLa are adherent cells which normally grow flat and stuck down firmly on the tissue culture flask.

SH-SY5Y cells: is a human derived neuroblastoma cell line, originally cloned from SK-N-SH. The cells differentiate by extending neurites to the surrounding area. The dividing cells can form clusters which are reminders of their cancerous nature,

but treatments such as retinoic acid and BDNF can force the cells to dendrify and differentiate.

IMR-32 cells: is another human neuroblastoma cell line that when is differentiated mimics large projections of the human neurons.

Neuro-2A cells: is a murine neuroblastoma cell line, established from a spontaneous tumor of a strain A albino mouse.

These cell lines grow in 100 mm Ø cell culture dishes, at 37°C and 5% CO_2, using 10 ml of DMEM or RPMI complete medium.

HELA and IMR-32 and growth medium:	Final concentration
-Foetal bovine serum (FBS) (Euroclone)	10%
-Glutamax (Gibco)	1%
-Penicillin / 0.1 mg/ml streptomycin (Euroclone) 100X	1X
-Fungizone (Euroclone)	1%
-Dulbecco's modified Eagle's medium (DMEM) (Euroclone)	to 500 ml

SH-SY5Y and Neuro2A growth medium:

-Foetal bovine serum (FBS) (Euroclone)	10%
-Glutamax (Gibco)	1%
-Penicillin / 0.1 mg/ml streptomycin (Euroclone) 100X	1X
-Fungizone (Euroclone)	1%
-Roswell Park Memorial Institute medium (RPMI) (Euroclone)	to 500 ml

2.1 TRANSIENT TRANSFECTION

“Transfection” means the introduction of foreign DNA into eukaryotic cells. Experimentally, this is most often done as transient transfection, in which the external gene is expressed only for a short period of time. In transient transfection, the gene introduced can be lost by the cell at any time, depending on environmental factors.

We chose a very efficient transfection method based on the inclusion of the DNA to be transfected in liposomes, small membrane-bounded bodies similar to the structure of the cell and capable to fuse with the plasma membrane, releasing DNA inside the cell.

Our immortalized cell lines were transfected with Lipofectamin 2000 (Invitrogen). One day before transfection, cells were plated to obtain 80% of confluence at the time of transfection.

For each transfection sample, we prepare complexes as follows:

6-WELL VASSELS: 1,25 µg of DNA and 4 µl of Lipofectamin 2000 in 250 µl of final transfection volume

100-mm PLATES: 10 µg of DNA and 32 µl of Lipofectamin 2000 in 500 µl of final transfection volume

First of all DNA was diluted in half of the appropriate transfection medium without serum and mixed gently. Lipofectamine 2000 is diluted in the other half of the appropriate transfection medium without serum, mixed gently and incubated for 5

minutes at 37°C. After 5 minutes of incubation, diluted DNA and diluted Lipofectamine were combined together and incubated for 20 minutes at 37°C to allow the DNA-Lipofectamine 2000 complexes to form. Such complexes were then added to each plate containing cells medium without serum. Transfection medium was replaced after 4-6 hours with complete medium. Transfected cells were grown at 37°C in a CO_2 incubator for 24-48 hours until they were ready to assay for transgene expression.

3. RNA EXTRACTION AND ANALYSIS

3.1 RNA EXTRACTION

Several methods are used in molecular biology to isolate RNA from samples but the most common of these is guanidinium thiocyanate-phenol-chloroform extraction. The specific reagent that we used was TRI Reagent (Sigma-Aldrich).

TRI Reagent solution combines phenol and guanidine thiocyanate in a monophasic solution to rapidly inhibit RNase activity. A biological sample is homogenized or lysed in TRI Reagent solution and the homogenate is then separated into aqueous and organic phases by adding chloroform and by centrifuging. RNA goes to the aqueous phase, DNA to the interphase, and proteins to the organic phase. Then the RNA is precipitated from the aqueous phase with isopropanol and finally it is washed with ethanol and solubilized in DEPC water.

TRI Reagent (Sigma-Aldrich) protocol consists in:

- cell washing with PBS;
- lysis of the cultured cells with TRI Reagent: 1 ml TRI Reagent solution for 5–10x10^6 cells or for 10 cm^2 culture dish area;
- addition of 200 ul of chloroform per 1 ml of TRI Reagent solution;
- 5-10 minutes of incubation at room temperature after vigorously shaking;
-centrifugation for 10 min at 12,000g and 4°C, followed by transfering of the aqueous RNA containing phase to a fresh tube;

- RNA precipitation with 500 ul of isopropanol per ml of TRE Reagent used;

- RNA pellet washing with 1 ml of 75% EtOH per ml of TRE Reagent used;

- Resuspension with 30-50 ul of DEPCT water.

The RNA final concentration was subsequently measured at the spectrophotometer, reading the Absorbance at 260 nm with 1 O.D., corresponding to 40 g/ml, for a cuvette 1 cm large (Lambert – Beer's Law).

Final concentration = absorbance value x dilution factor x 40 g/ml conversion factor

As proteins absorb at 280 nm, measuring the 260/280 absorbance ratio it's possible to value the RNA pureness. The ideal ration is between 1,8 and 2.

The 260/230nm ratio is important to detect the presence of contaminants, such as phenol, ethanol or aromatic molecules. The best ratio is 2,2.

To estimate the RNA integrity, it is possible to run 1 ug of RNA on 1% agarose gel electrophoresis in TAE buffer (0,04M Tris-acetate, 0,001M EDTA).

3.2 REVERSE TRANSCRIPTION PCR (RT-PCR)

The vast majority of cellular RNA molecules are tRNAs and rRNAs. Only 1 to 5% of total cellular RNA is mRNA. Separation of mRNA from the RNA pool is essential for the construction of cDNA. Most mRNA contain a poly(A) tails, while the structural RNAs do not, so poly(A) selection enriches the mRNA.

To convert mRNA into double-stranded DNA we use *ImProm-II Reverse*

Transcription System (Promega). This procedure outlines the synthesis of cDNA for subsequent amplification using PCR.

Reverse transcription reactions of up to 1µg of total RNA were performed in 20µl, comprised of components of the ImProm-II™ Reverse Transcription System.

Briefly:

RNA is combined with oligo(dT) primer that binds to poly(A) tails. The primer/template mix is thermally denatured at 70°C for 5 minutes and chilled on ice. A reverse transcription reaction mix is then assembled on ice to contain nuclease-free water, reaction buffer, reverse transcriptase, magnesium chloride, dNTPs and RNasin Ribonuclease Inhibitor. Following an initial annealing at 25°C for 5 minutes, the reaction is incubated at 42°C for up to one hour. Since no cleanup or dilution is necessary following the cDNA synthesis, the product may be directly added to amplification reactions.

The first part of the reaction includes:

RNA	1 ug
oligo dT	0,5 ug
H_2O nuclease free	to 10 ul

To denature RNA secondary structures, this mix is incubated at 70 °C for 5min.

Then we added:

Final concentration	
ImPromII 5X Reaction Buffer	1X
MgCl2	1,5 mM
dNTPs mix	2 mM
RNasin Ribonuclease Inhibitor	20 U
ImPromII Reverse Transcriptase	1 ul
H_2O nuclease free	to 10 ul

The reaction is then incubated at 25°C for 5 minutes to allow reverse transcriptase binding, and at 42°C for 60 min. to allow elongation of cDNA. Finally, we increased temperature at 70°C for 15min to inactivate RT enzyme.

3.3 QUANTITATIVE REAL TIME PCR

Quantitative real time PCR is a technique used to amplify and simultaneously quantify a targeted DNA molecule. The quantity can be either an absolute number of copies or a relative amount when normalized with an endogenous gene (such as GAPDH or β-Actin).

There are two common methods for products detection in real-time PCR:

(1) non-specific fluorescent dyes (such as SYBR Green) that intercalate with any double-stranded DNA

(2) sequence-specific DNA probes (TaqMan probes) consisting of oligonucleotides labeled with a fluorescent reporter that permits detection only after hybridization with its complementary DNA target.

In our experiments we always use SYBR Green (Biorad) as detection system. An increasing in DNA products during PCR leads to an increase in SYBR Green fluorescence that is measured at each cycle, allowing DNA concentrations to be quantified.

Relative concentrations of DNA present during the exponential phase of the reaction are determined by plotting fluorescence against cycle number on a logarithmic scale (so that an exponentially-increasing quantity will show as a straight line). A threshold for detection of fluorescence above background is determined. The cycle at which the fluorescence from our sample crosses the threshold is called the cycle threshold, Ct. The quantity of DNA theoretically doubles every cycle during the exponential phase, and relative amounts of DNA can be calculated with the formula: $2^{-\Delta CT}$, where ΔCT is the difference between the sample Ct and the housekeeping gene Ct. The reaction was performed with the thermalcycler *iQ5 real-time PCR detection system* (Biorad) and SYBR Green Supermix (Biorad), with following protocol:

Final concentration	
cDNA	50 ng of retrotranscribed RNA
SYBR Green 2X	1X
Primer FW 10uM	0,5 uM
Primer RV 10uM	0,5 uM
H_2O to 25 ul	

cycle 1		95°C x 2'	cycle 4		95°C x 1'
cycle 2		95°C x 30"	cycle 5		74°C x 1'
		95°C x 10"	cycle 6	x 43	74°C x 10"
cycle 3	x 50	59°C x 20"			
		74°C x 20"			

Primers used in real time PCR reactions:

mBactEx2real-FW
GGCTGTATTCCCCTCCATCG
mBactEx3real-RV
CCAGTTGGTAACAATGCCATGT
mLSD1-real-FW
GCCTCAGCAGACACAGAAGG
mLSD1-real-RV
TGTTGTAAGGCGCTTCCAGC
mRPSA FW
ACCCAGAGGAGATTGAGAAGG
mRPSA RV
TGGGGAACTGCTGAATGGCC

mBDNF ORF FW
ATGAGGGTCCGGCGCCACTC
mBDNF ORF RV
CAAGGACTGTGACCGTCCCG
mBDNF exon II FW
ATCCGGGCGATAGGAGTCC
mBDNF exon II RV
CACAGGGTTGGCTTTACAGC
mNOVA1 x real FW
GCCGGACTCGCGGAAAAG
mNOVA1 x real RV
TGAACAATTGTCTGTCCTCC
mC-FOS x real FW
GATGTTCTCGGGGTTGAAGC
mC-FOS x real RV
AGAAGGAGTCTGCGGGTGAG
mEGR1 x real FW
CCTTCAATCCTCAAGGGGAGC
mEGR1 x real RV
AACCGAGTCGTTTGGCTGGGA

3.4 RELATIVE QUANTITY FLUORESCENT PCR (Rqf-PCR) and CAPILLARY GEL ELECTROPHORESIS

To quantitatively measure the relative amount of each LSD1 isoform, we used a PCR-based method in which a fluorochrome-conjugated (FAM) forward primer (Sigma-Aldrich) and a reverse unmodified one were used to amplify all the expected isoforms in a single reaction. PCR products, mixed together with a

suitable internal ROX-conjugated size standard (500ROX Standard; Applied Biosystems), were separated by capillary electrophoresis under denaturing conditions, and the amount of each amplified product was measured as microsatellite sample, based on related fluorescence unit levels using the GeneMapper software.

Capillary gel electrophoresis is an useful technique to accurately establish the length of the PCR product fragments. Since the matrix gel is made of polyacrylamide, it is possible to separate fragments that differ even from 1 bp.

This is the protocol we followed:

	Final Concentration
PCR product	up to 1ul
Formaldehyde	11 ul
Molecular weight marker	0,3 ul

Each sample has to be denatured for 5 min. at 95°C before to be loaded on the capillary gel electrophoresis.

4. PROTEIN ANALYSIS

Total protein extraction has been obtained from different cell lines, using RIPA buffer or Low Stringency buffer (LS) as extraction method. Cells has been scraped from culture plate in 1 ml of PBS and centrifuged at 800*g* for 5' at 4°C. The pellet has been resuspended in 5-6 volume of LS or RIPA buffer, depending from the purpose of the extraction, incubated 30' at 4°C in slow rotation and, after centrifugation at high speed for 30', the supernatant containing soluble proteins was stored at -80 °C.

Protein concentration has been quantified using the Bradford method (Pierce).

Low Stringency lysis buffer: 50 mM Tris-HCl pH 7.5, 120 mM NaCl, 0.5 mM EDTA, 0.5 % Nonidet P-40.

RIPA lysis buffer: 50 mM Tris-HCl, 150 mM NaCl, 1.0 % NP-40, 0.5 % Sodium Deoxycholate, 1.0 mM EDTA, 0.1 % (w/v) SDS and proteases inhibitors).

4.1 WESTERN BLOT

Western blot or immunoblotting allowed us to determine, with a specific primary antibody, the relative amount of a protein in different samples. Samples are prepared from tissues or cells homogenized in a buffer that protects the proteins from degradation. Proteins are separated using SDS-PAGE. SDS is a detergent

that is present in the sample buffer, where along with boiling and together with a reducing agent (normally DTT or B-ME to break down protein-protein disulphide bonds) it disrupts the tertiary structure of the proteins. This brings the folded proteins down to linear molecules. SDS also covers proteins with a uniform negative charge, which masks the intrinsic charges of the R-groups. Indeed, SDS binds uniformly to the linear proteins (around 1.4g SDS/ 1g protein), so that their charge becomes approximately proportional to their molecular weight.

SDS is also present in the gel to make sure that once the proteins are linearized and their charges masked, they stay like this also during the running. Gel matrix used for SDS-PAGE is polyacrylamide, which is a good choice because is chemically inert and, crucially, can easily be made up at a different concentrations to produce different pore sizes, giving a variety of separating conditions that can be changed depending on the needs. Once proteins are separated they have to be transferred to a nitrocellulose membrane for detection. Membrane is then incubated for 1h or over night in NaCl/Tris added with 0,3% Tween (TBST) and 4% (w/v) of skimmed milk that binds to any remaining sticky places. Afterwards, a primary antibody is added (in the same 5 % milk in 1X TBST solution) for 4h at room temperature or over night at 4 °C. Then the membrane is washed 3 times for 5 min. in TBST and incubated with the specific HRP-coniugated secondary antibody. Specific signal of the protein is detected with SuperSignal West Pico (Pierce).

1,5 mm-thick polyacrylamide gel for SDS-PAGE is composed of:

	Stacking (5%)	Separating
40% Polyacrylamide	1,25 ml	variable
1M Tris (pH 6.8)	1,25 ml	3,75 ml
10% Ammonium persulfate	100 ul	150 ul
10% SDS	100 ul	150 ul
TEMED	10 ul	9 ul
H_2O	to 10 ml	to 15 ml

Primary antibodies used:

- monoclonal anti HA-probe (F-7) (Santa Cruz sc-7392)
- monoclonal anti-FLAG clone M2 (Sigma F104)
- polyclonal anti-GAPDH (NB 300-320)

HRP coniugated secondary antibodies used:

- donkey anti-rabbit IgG-HRP (GE Healthcare NA934V)
- goat anti-mouse IgG-HRP (Santa Cruz sc-2005)

5. ANIMALS

The experiments were undertaken in accordance with the guidelines defined by the European Communities Council directive (86/609/EEC) for the Care and Use of Laboratory Animals and every effort was made to limit the number of animals used. The experimental protocol was approved by the Ethics Committee of the 'University of Milan'.

Behavioral experiments were performed with 12 weeks-old backcrossed C57BL/6 LSD1 exon E8a knock-out and wild type male littermates, derived from multiple litters. The majority of tests were conducted using the same cohort for multiple tests, performed in order from least to most stressful. Phenotypic characterizations were performed on 3 weeks- and 3 months-old knock-out and wild type female mice. Mice were sacrificed by cervical dislocation in order to provide a fast and painless death without the effects of anesthesia. Immunohistochemical studies were done on 5 weeks-old knock-out and wild type male or female littermates, after fixation via transcardial perfusion.

C57BL/6 mice were purchased from Charles River (Milano). Mice were housed 2-4 per cage in an animal room with constant temperature (22 ± 1 °C) and a 12 hour light/dark cycle (lights on/off at 6:00 A.M/P.M.), with free access to food and water.

5.1 DNA EXTRACTION FROM MICE BIOPSY

Mice genotyping were performed on DNA taken from fingers biopsy. We first incubated over-night at 58°C such small pieces of mice tissue, in 50 uL of a solution composed of:

– 1 uL of proteinase K (10mg/mL),

– 1 X PCR reaction buffer (10X)

– water to 50 uL

We then ground fingers with the help of a tip and we incubated for 1h at 65°C. After that we denatured proteinase K at 95°C for 10 min. and we centrifugated to obtain a clear supernatant that we used as PCR template.

5.2 PILOCARPINE TREATMENT AND SEIZURES ASSESSMENT

Five weeks-old experimental mice were intraperitoneally injected (i.p.) with a single dose of 270 mg/kg of pilocarpine in sterile saline vehicle (0.9% NaCl). Control mice were age-matched with the treated mice and were administered a comparable volume of vehicle. We also pre-treated all of our mice with lithium chloride (3 mEq/Kg), allowing a reduction of the pilocarpine dose required to induce status epilepticus and resulting in a higher percentage of animals developing seizures.

On the day of injection, behavioral observation lasted minimally 1 h after pilocarpine injection, and pilocarpine-induced seizures were evaluated according to a modified version of the Racine scale, using categories 1–5. Categories one

and two (facial automatisms, tail stiffening and wet-dog shakes) were considered as a group to avoid subjectivity in assessing the seizures. All pilocarpine-treated mice displayed some or all of these behaviors. Category 3 (low-intensity tonic–clonic seizure marked by unilateral forelimb myoclonus in addition to the symptoms above), category 4 (the addition of bilateral forelimb myoclonus and rearing), and category 5 (bilateral fore- and hind-limb myoclonus and transient loss of postural control) seizures were considered to be generalized, convulsive seizures. Category 3–5 seizures were typically of 30–90 sec of duration and were separated by periods of relative inactivity of variable duration. The periods between category 3–5 seizure events were usually marked by continuous category 1 and 2 seizures. A mouse that experienced a minimum of three category 3–5 seizure events within 1 h following pilocarpine injection was considered to have undergone status epilepticus (SE). SE was then terminated in all mice by diazepam (10 mg/kg). Control and pilocarpine-injected mice were housed individually in cages that were uniformly intermixed and monitored simultaneously. Following the pilocarpine treatment, mice were monitored for 1 to 7 hours for the occurrence of spontaneous category 3–5 seizures and then sacrificed. Hippocampus, cerebellum and striatum of treated and control mice was removed and different protein expression as well as transcription profile were evaluated in order to find out differences between Exon E8a Knock-out mice and wild type ones, in a condition of neuronal activation.

5.3 WHOLE ANIMAL FIXATION VIA TRANSCARDIAL PERFUSION

Although immersion fixation allows for the adequate cytological preservation of small pieces of tissue for routine histological stainings, we decided to preserve them with perfusion fixation. This method is especially required for optimal fixation of central neural tissues, particularly if they are to be used in immunohistochemical or other specialized neurohistological procedures.

Before starting with fixation we intraperitoneally administered 0,5 ml of 4% cloralio idrato to anesthetize mice. After one minute we assessed and assured surgical-plane anesthesia by toe pinch. We then wetted the fur of the ventral skin surface with 100% ethanol and we superficially cut the ventral skin with scissors, just below the zyphoid-process, being careful not to perforate the peritoneal membrane. After the exposition of the thoracic and peritoneal membrane surfaces, obtained by manually pulling the opposing skin segments rostrally and caudally, we cut the peritoneal membrane with scissors, just below the zyphoid-process to expose the diaphragm and visceral organs, being careful not to lacerate any significant vasculature. Performing a combination of blunt dissection and scissor-assisted dissection techniques, we opened the thoracic cavity by cutting the diaphragm from one lateral aspect to the other lateral aspect. Using thumb forceps, we gently grasped and held the right ventricular free-wall and visualized the delineation of the darker, blood-filled-coloration of right ventricle and the redder, muscular-coloration of left ventricle. First of all we performed an intracardiac injection with 200 ul of Vister Heparin, a strong blood anticoagulant, then we

insert the needle of the infant infusion set, attached to pre-fix solution, into the left ventricular chamber and at the same time we carefully lacerated the right atrial-chamber with scissors: blood and heparinized pre-fix solution flew from the right atria laceration. The right ventricular chamber remain somewhat darkened in color when compared to the left ventricular chamber. The liver begins to blanch as blood is replaced with pre-fix solution. At the completion of the 5 min. of pre-fixing, we carefully changed-out the solution without dislodging the infusion needle from the left ventricle and we started perfusion with 4% paraformaldehyde/PB for 10 more minutes. At the completion of the paraformaldehyde perfusion, we removed the needle from the left ventricle and we dissected brain. Tissue was immersed in 4% paraformaldehyde/PB and fixation continued by immersion with agitation for 10-16 hours at 4°C.

At the conclusion of fixation, we replaced fixative with PB 0.1 M pH 7.4 and we sliced the brain with microtome (Leica VT 1000 S) in coronal or parasagittal sections of 50 um width. Sections were lied down in PB 0.1 M pH 7.4 and some of them were conserved on a slide at 4°C for the immunostaining. On the contrary to perform post-embedding reactions we included brain in paraffin and then we cut 10 um sections and we conserved them on gelatin-coated slices.

Reagents:

Phosphate Buffered (PB)

To obtain 2X solution:

A) 0.2M NaH_2PO4xH_20 27,59 g/L

B) 0.2M $Na_2HPO4x7H_20$ 56,615 g/L

pH B solution was adjusted to 7.4 using A solution

Eparina Vister (Perfusant)
5.000 U.I./ml solution

Paraformaldehyde Fixative (Perfusant)
Pre-fix solution

1% Paraformaldehyde (Fluka) in PB, final pH 7.4

Fix solution
4% Paraformaldehyde (Fluka) in PB, final pH 7.4

The paraformaldehyde is dissolved in PB with constant stirring. The fix solution is filtered.

Intraperitoneal Anesthetic
Cloralio idrato (Sigma 15307) 0,28 g in 7 ml

Pump Set-Up
5 min. of prefix- and 10 min. of fix-solution

5.4 IMMUNOSTAINING

Immunostaining is a general term that applies to any use of an antibody-based method to detect a specific protein in a sample. Immunohistochemistry (IHC) staining of tissue sections (or immunocytochemistry, which is the staining of cells), is the most commonly applied immunostaining technique. While some cases of IHC/ICC staining uses fluorescent dyes (immunofluorescence), other non-fluorescent methods using enzymes such as peroxidase and alkaline phosphatase are also used.

We performed both immunoperoxidase and immunofluorescent staining on LSD1

wild type and knock-out brain slice, using different antibodies.

Before starting with the immonoperoxidase staining we pre-treated brain slices to allow the antibodies interaction. This is the protocol we followed:

- NH_4Cl 0,05 M/PBS for 20 min. to saturate aldeyde groups deriving from fix-solution and to avoid aspecific electrostatic interactions with the antibodies;
- PBS washing;
- H_2O_2 1% /PBS for 10 min. to block endogenous peroxidase;
- PBS washing;
- BSA 1% and Triton X-100 0,2% in PBS for 30 min. to block aspecific binding sites (with BSA) and to improve tissue permeability to the antibodies (Triton X-100 0,2%);

Afterward we performed immunostaining like this:

- incubation with primary antibodies in BSA 0,1%, over night at room temperature or for 2 nights at 4°C but in constant shaking;
- PBS washing to remove unbounded antibodies;
- incubation with biotinylated secondary antibody for 75 min. at room temperature;
- PBS washing;
- incubation with Avidin-Biotin-Peroxidase complex (ABC/HRP complex) for 75 min. at room temperature in costant shaking. This method allows the weak specif signal to be amplified thanks to the binding between Avidin and Horseradish Peroxidase-conjugated Biotin molecules. Indeed, each Avidin molecules is able to bind 4 Biotin molecules. The solution is prepared 30 min. before the usage to let Biotin molecules interact with Avidin;

- PBS washing;

- incubation in a solution of Tris-HCl 0,05 M pH 7,4 with 3-3' diaminobenzidine (DAB) 0,075% and H_2O_2 0,02%. In the presence of hydrogen peroxide, DAB is converted to an insoluble brown reaction product and water by the enzyme HRP (horse radish peroxidase). The reaction is like this:

DAB + H202 ----------HRP----------> DAB ppt + H20

So if we have HRP attached to the antibody, which in turn is bound to an antigen in a particular area of a section, when we add DAB and H_20_2, the reaction will proceed like any other enzymatic reaction and produce an insoluble brown DAB ppt and H_20 in the area of the section where the antibody has bound antigen. In some reactions, staining was intensified by adding nickel ammonium sulfate that gives reaction products a dark purple color.

Sections were then collected on gelatin-coated slides, dried over night at 37°C and the day after they were dehydrated by an ascending alcoholic solution (alcohol 75º, 96º, 100º), clarified in xilol and finally coverslipped with Permount to be observed with optical microscope.

Primary antibodies used for the immunoperoxidase reactions:

- monoclonal anti-Calbindin D-28K (Swant cat. 300)
- polyclonal anti-Calretinin (Swant cat. CR7699/3H)
- polyclonal anti-LSD1 (AB 17721)
- polyclonal anti-SOX2 (AB5603)

- monoclonal anti-VGLUT1 (MAB5502)
- polyclonal anti-VGLUT2 (AB 72310)

All of them were diluted in BSA 0,1%/PBS 0.01M, pH 7.4

Secondary antibodies used in immunoperoxidase reactions:

- biotinylated goat anti-rabbit IgG antibody (BA1000)
- biotinylated horse anti-mouse IgG antibody (BA2000)
- biotinylated horse anti-goat IgG antibody (BA9500)

All of them were diluted in BSA 0,1%/PBS 0.01M, pH 7.4

RESULTS

1. NEUROSPECIFIC LSD1 ALTERNATIVE SPLICING REGULATION

Splicing efficiency of individual exons is determined by multiple features involving gene architecture, a variety of cis-acting elements within the exons and flanking introns, interactions with components of the basal splicing machinery (spliceosome) and regulatory factors which transiently co-assemble with the spliceosome. The main goal of my project was to analyze the molecular mechanisms underlying the alternative splicing of the LSD1 neurospecific exon E8a. Alternative exons are often regulated according to cell-specific patterns and regulation is mediated by sets of cis-acting elements and trans-acting factors. For this reason, first of all we analyzed the intronic E8a flanking regions to find out the potential cis- and trans-acting factors involved in this process. Such analysis was performed with the assistance of specific browsers that were able to predict with a good level of accuracy the factors involved in splicing regulation. After that, combining data obtained from database prediction with the existing evidences in literature we decided to experimentally validate the most promising regulators. Validation occurred by performing mini-gene splicing assays, that are commonly used *in vivo* to identify the intrinsic features of a gene that control exon usage and to find-out specific cis-acting elements and trans-acting factors that modulate alternative splicing in a cell-specific manner.

1.1 CHARACTERIZATION OF THE EXON E8a FLANKING INTRONS

We know from literature that both in human and in mouse, 77% of the conserved alternatively spliced exons are flanked on both sides by long conserved intronic sequences. The average length of these regions is about 103 bases in the upstream intron and 94 bases in the downstream intron, while the average conservation percentage is about 88% [66]. Since they were found to be unexpressed in microarray experiments, many of these conserved intronic sequences are probably cis-regulatory elements with a function in alternative splicing modulation.

In LSD1 genomic sequence, we observed a very high conservation degree in correspondence of the known 19 constitutive exons and a lower conservation in the intronic sequences, as expected. Nevertheless, we also noted a high conservation degree in correspondence of the alternative exons E2a and E8a, and within the E2a and E8a intronic flanking regions (Fig.1.1.1 A), according to what we found in literature. Using the UCSC Genome Browser (http://genome.ucsc.edu) we delineated a region of about 800-bp flanking the exon E8a, that was highly conserved among mammals. In particular, aligning this sequence using the GenomeVista browser (http://www.pipeline.lbl.gov), we observed 89% of identity with the mouse genome (Fig. 1.1.1 B), whereas with the rat genome we found 84% of identity. This observation immediately led us to think that E8a flanking intronic sequences were probably involved in alternative splicing regulation.

Starting from this observation we analyzed the selected region using SpliceAid2 (http://www.introni.it), a human database of experimental RNA target motifs and splicing proteins, that gave us the possibility to find a great number of splicing factors potentially involved in E8a alternative splicing regulation.

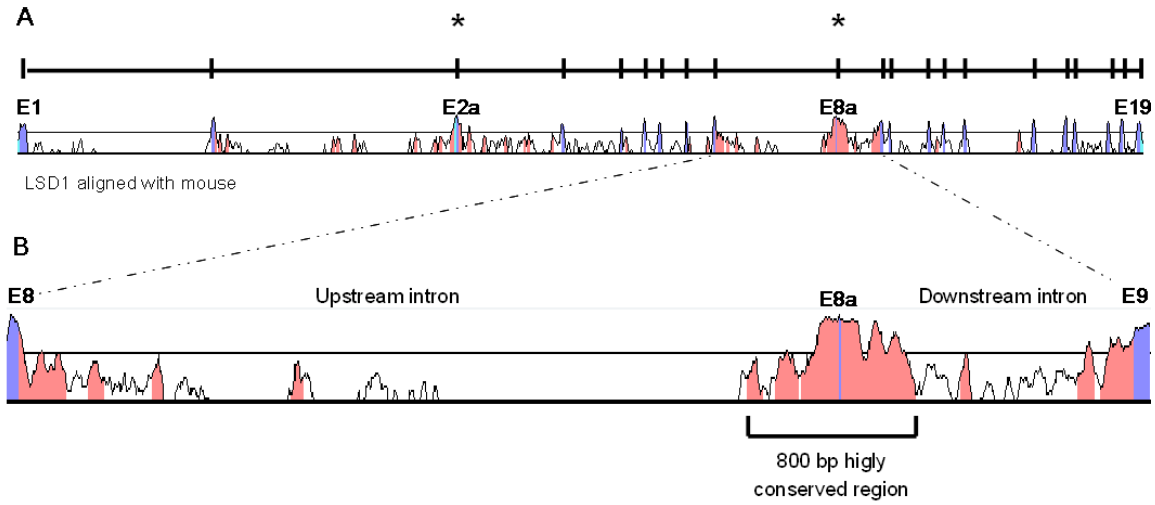


Fig. 1.1.1 Alignment of human genomic LSD1 sequence with mouse genomic LSD1

Fig. 1.1.1A Schematic representation of the human LSD1 introns and exons from 1 to 19; asterisks indicate the location of annotated alternative exons (E2a and E8a). The "peaks and valleys" graphs represent percentage conservation at a given genomic coordinate between the human sequence and the mouse one. The first exon is numbered, as well as the last one and the alternatively spliced ones. Regions of high conservation are colored as exons (purple) or noncoding (pink). Conserved regions are defined as regions with identity of 70% or higher. Fig. 1.1.1 B Enlarged view of the alignment between human and mouse upstream and downstream introns around the alternatively spliced exon E8a.

In particular we were interested in identifying NOVA1 binding sites because we knew that NOVA1 is a neurospecific splicing regulator and in literature there is the evidence of a physical interaction with the exon E8a, obtained by HITS-CLIPS experiments [67]. Thus, we looked for NOVA1 binding sites (YCAY) in the exon E8a flanking introns. Furthermore we were also interested in finding FOX1 binding sites (TGCATG) as well as nPTB ones (CT rich), because of their neuronal

restricted tissue-specificity and their demonstrated interaction with NOVA1 [67, 69]. Fig. 1.1.2 shows that in the highly conserved 800 bp region surrounding the exon E8a, we found NOVA1 binding sites (YCAY) upstream and downstream the exon E8a, with a cluster of 7 sites within the first 100 bp of the downstream intron. Since NOVA1 positively regulates splicing when present in the downstream intron [68], our finding suggests that NOVA1 can have a positive role on exon E8a inclusion. We also found different nPTB binding sequences (CT rich) mainly located in the upstream intron. On the contrary no FOX1 binding sites were found inside this 800 bp conserved region.

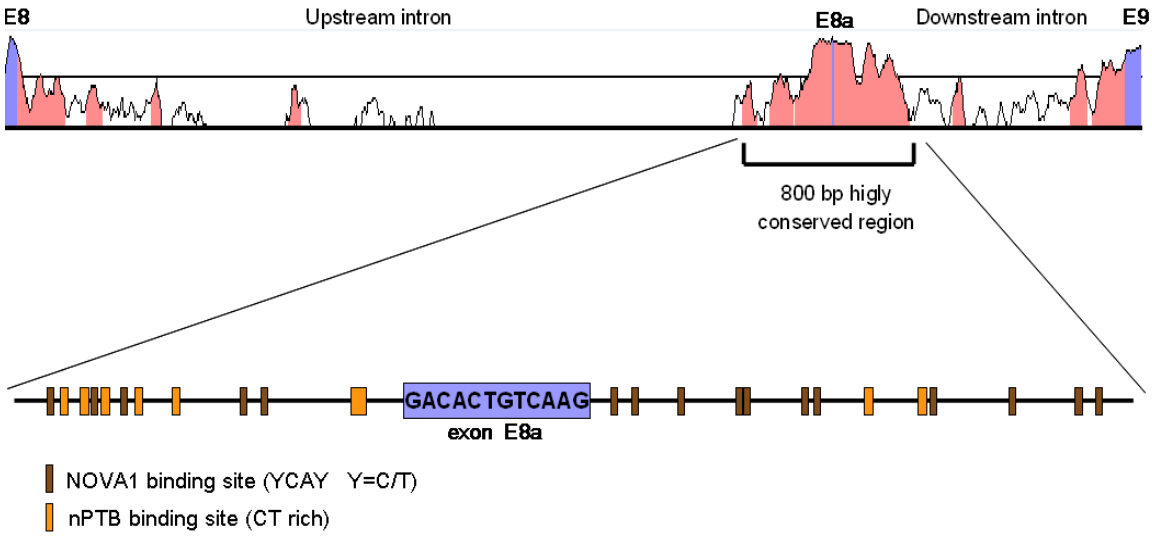


Fig. 1.1.2 Cis-acting sequences involved in E8a splicing regulation

Inside the 800 bp highly conserved region flanking the exon E8a we found several NOVA1 binding sites, YCAY elements in brown in the picture, and different nPTB binding sites, represented by CT rich sequences in orange. No FOX1 binding sites were present inside this restricted portion surrounding the exon E8a.

1.2 MINIGENE 800 SPLICING ASSAY

In order to verify if the *cis*-acting elements that we found were really involved in splicing regulation and if they were sufficient to induce exon E8a inclusion in the mature nLSD1 transcript, we performed functional studies using a Minigene reporter system where we cloned the 800-bp highly conserved portion mentioned above.

Minigene constructs are important tools for the identification and in vivo analysis of the cis-acting regulatory elements and trans-acting factors that establish splicing efficiency and that regulate alternative splicing. Expression of Minigene pre-mRNAs by transient transfection provides a rapid assay for loss-of-function and gain-or-function analyses of cis-elements and trans-acting factors that affect splicing regulation.

Minigene that we generated contained the genomic 780-bp segment from LSD1 gene that included the alternatively spliced small exon E8a and a portion of its flanking intronic regions. Such genomic fragment was generated by PCR amplification directly from genomic DNA of human placenta, using “hLSD1 cloning intron Fw” as forward primer and “hLSD1 cloning intron RV” as reverse primer. These primers contained also the NdeI restriction site in their 5’ ends to allow purified PCR product to be cloned inside the destination vector.

The Minigene reporter vector that we used is an hybrid construct deriving from the low copy number pBlueScript II KS plasmid, where a Minigene cassette was

inserted at the level of its multiple cloning site, between the restriction sites PstI and BamHI (Fig. 1.2.1). This cassette contained exons from α-globin and fibronectin genes under the control of the α-globin promoter and SV40 enhancer sequences, with a 3' polyadenylation site derived from the α-globin gene. In particular, the cassette was composed of the first two exons E1 and E2, deriving only from the α-globin gene, followed by two chimeric exons: one of them composed of the initial part of α-globin exon E3 and the final portion of fibronectin exon E24; the other one deriving from fusion between fibronectin exon E25 and the terminal part of α-globin exon E3. These two chimeric exons were linked together by the standard fibronectin intron, which contained the restriction site NdeI (Fig. 1.2.1). The presence of the chimeric exons allowed the amplification of the Minigene transcripts only and not of the transcript deriving from the genomic fibronectin and α-globin genes. This was possible thanks to the forward primer α2-3 and the reverse one Bra-2, which annealed on the junction of the chimeric exons. Moreover, the forward primer α2-3 was design to anneal with its 5' end on the α-globin exon E2 and with the 3' end on the initial α-globin exon E3. Thus only the mature and spliced transcripts, and not the plasmidic DNA sequences, were amplified.

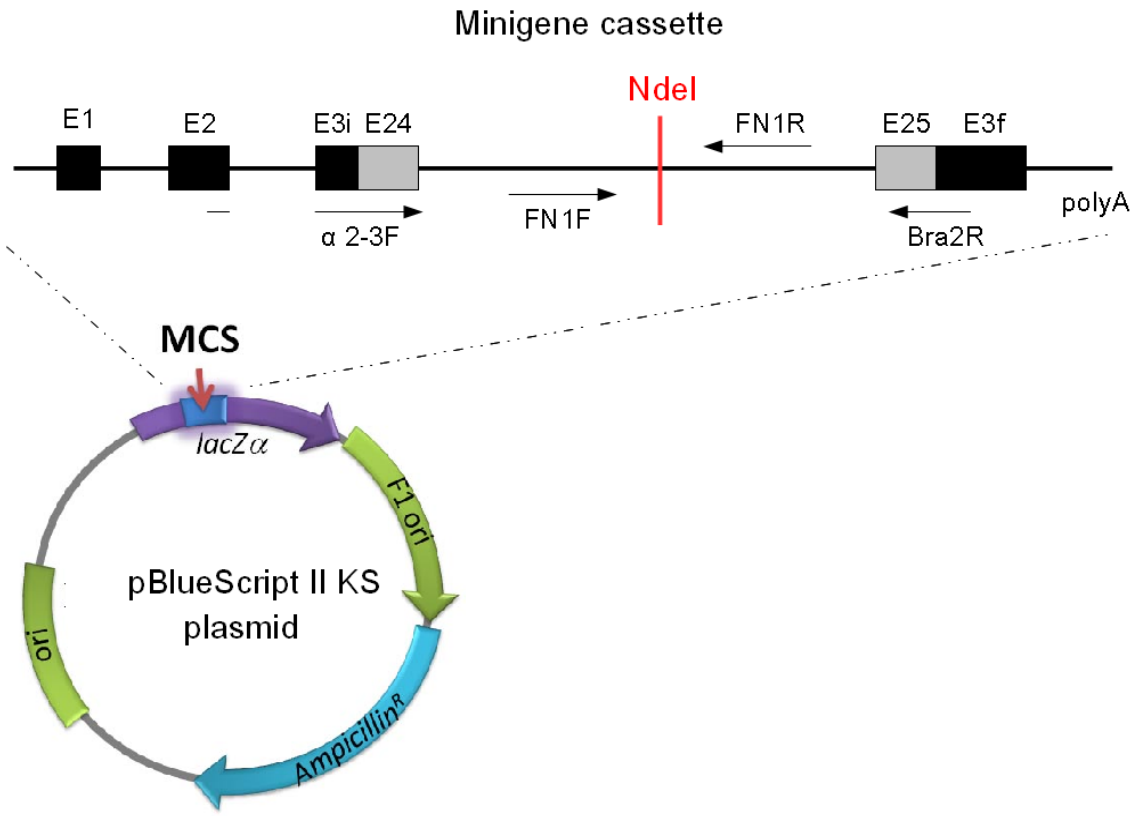


Fig. 1.2.1 pBSplicing plasmid.

pBSplicing plasmid is composed of pBlueScript II KS vector and the Minigene cassette with α-globin (black boxes) and fibronectin (grey boxes) exons. Arrows indicate primers position.

To obtained the final Minigene 800 construct we cloned PCR genomic fragment of 780 bp inside the NdeI restriction site of the pBSplicing plasmid (Fig.1.2.2), screening the insertion and the right orientation with forward and reverse FN1 primers, together with a digestion with the enzymes NdeI and HindIII. Since minigene cassette is normally transcribed and then spliced by the cell machinery, using Minigene 800 reporter system we were able to study exon E8a alternative splicing and its inclusion frequency into the Minigene 800 mature transcripts.

The presence in the 800 bp cloned genomic region of *cis*-acting elements promoting exon E8a inclusion allows the exon E8a to be included in the mature chimeric transcript. Furthermore, Minigene 800 assay gave us the possibility to test specific *trans*-acting factors involved in exon E8a splicing modulation.

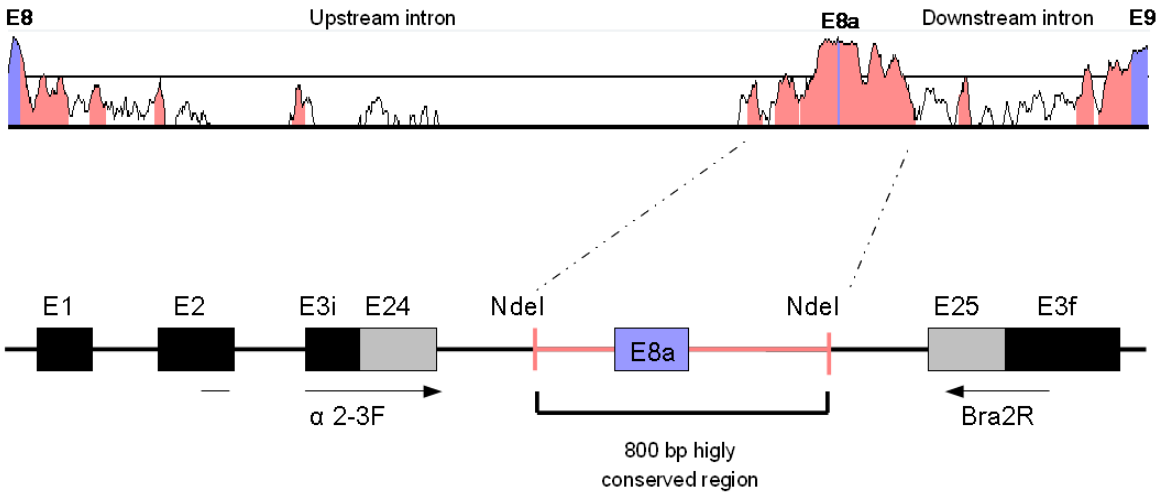


Fig. 1.2.2 Minigene 800 reporter system

Inside the minigene cassette we cloned exon E8a and its highly conserved 800-bp intronic sequence, using NdeI restriction site.

Since *in vivo* the inclusion of the exon E8a is restricted to the nervous system we used different neuronal cell lines as *in vitro* model of neuronal-like cells, and non-neuronal HELA cells as a negative control of inclusion.

For transfection assays we plated the cells in 6-wells (35 mm) plates with 0,5 γ of Minigene and a filling vector to reach a total amount of 1,25 γ of DNA, using Lipofectamine as carrier (see “Materials and Methods” for details). 48 hours after transfection, total RNA was extracted from cells, retro-transcribed and the cDNA

was analyzed by PCR. To amplify the transcripts deriving from Minigene 800 we used fluorescinated primers α2-3F and Bra2R. PCR products were run on the agarose gel, to verify the quality of the reaction. As cDNA positive control we also amplified and run the endogenous β-actin gene, using the primers “hβ-actin Fw” and “hβ-actin Rv”. The output of the agarose gel electrophoresis was a unique band for all the amplicons deriving from the different Minigene 800 spliced transcripts. Indeed, the resolution power of the gel did not allow the discrimination of more than one possible PRC product/s (Fig 1.2.3).

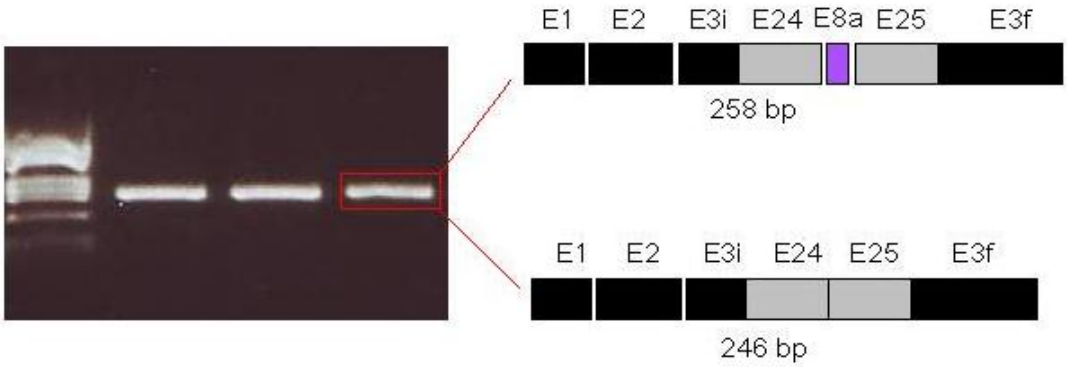


Fig. 1.2.3 Agarose gel output obtained by running Minigene 800 PCR products.

The two possible Minigene 800 splicing transcripts, deriving from the 12-bp exon E8a indusion or not, generate two PCR amplicons that co-migrate in a unique band on the agarose gel.

Thus, to evaluate the presence of the exon E8a in the mature mRNA transcript deriving from Minigene 800, we ran PCR products on capillary gel electrophoresis together with an appropriate molecular weight marker. Since the α23F primer was

fluorescinated, laser was able to read the size of each amplicon and to compare it with the known bp-length of the marker fragments. Therefore we could discriminate the presence or not of the exon E8a into the amplicons because of a length difference of 12 bp (the amplicons containing the exon E8a were 12 bp longer than those without the exon). The final outcome of the capillary gel electrophoresis consisted in peaks of different height, separated 12 bp one from another.

This kind of analysis is also quantitative, because the fluorescence intensity depends on the quantity of the obtained fluorescinated amplicon. Thus, it was possible to evaluate the relative percentage amount of each transcript (containing or not the exon E8a) using the value of the height or the area underlined each peak. In particular, to calculate the percentage of exon E8a-containing transcripts we used the following formula: [height of the exon E8a-containing peak /(height of the exon E8a containing peak+height of the peak without E8a)] X 100.

First of all we transfected pBSplicing vector to demonstrate that the hybrid cassette was correctly spliced by the endogenous cell machinery (Fig.1.2.4).

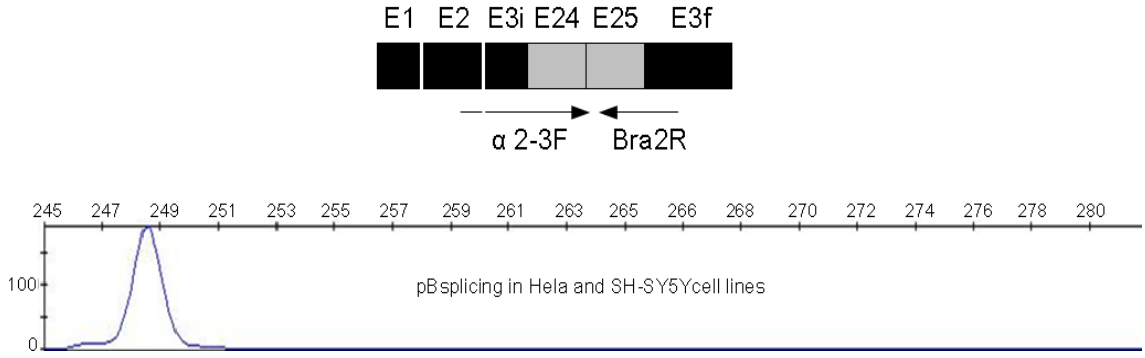


Fig. 1.2.4 Output of the Mnigene constitutive splicing.

pBsplicing outcome in Hela and in SH-SY5Y cells: only one 247bp peak is generated from splicing of the sole Minigene cassette. Numbers above each graph represent length of the amplicons expressed in base pairs. Numbers besides the graph stand for fluorescence intensity.

pBSplicing vector was correctly spliced both in HeLa and in SH-SY5Y cells and the amplicon deriving from the mature mRNA cassette was about 247 bp-long (Fig. 1.2.4).

Then we transfected our Minigene 800 reporter plasmid to investigate exon E8a inclusion.

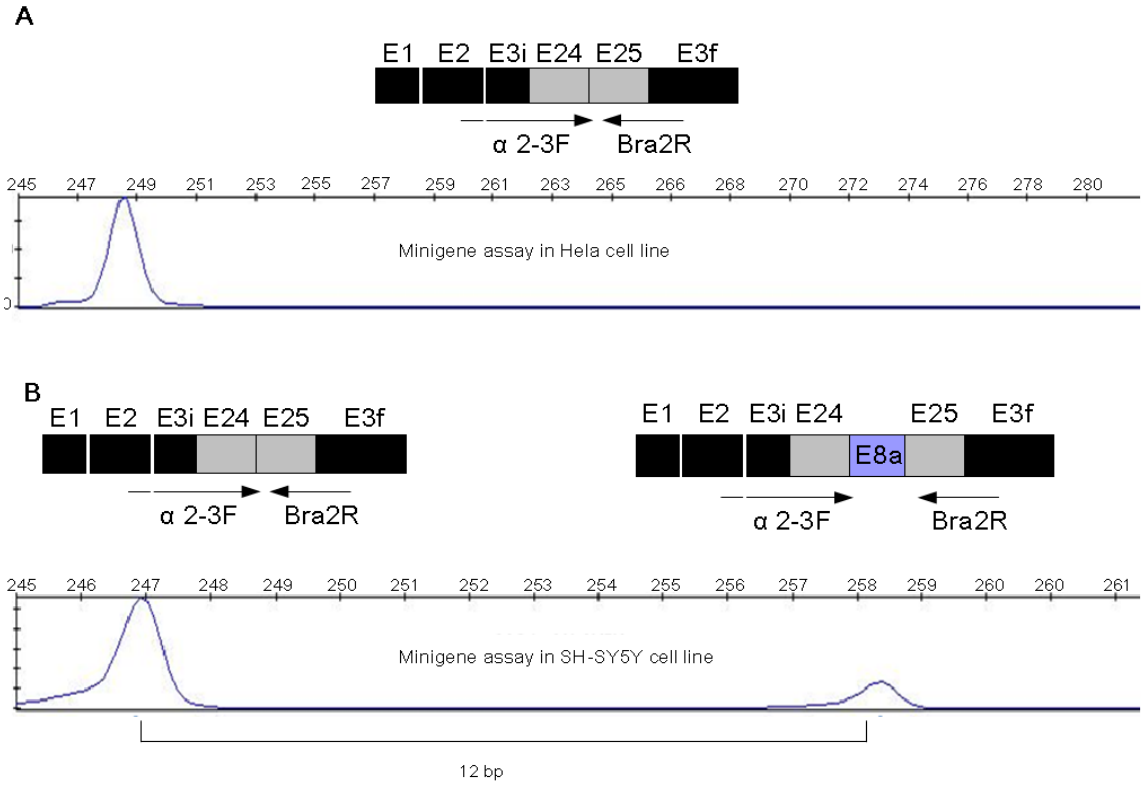


Fig. 1.2.5 Output of the capillary gel electrophoresis in HELA and in SH-SY5Y cells.

A) Minigene 800 splicing outcome in HELA cells: only one 247bp peak is generated from splicing cassette because no exon E8a inclusion occurs. B) Minigene 800 essay in SH-SY5Y cells: 247 bp peak derives from the constitutive splicing while 258bp peak represents transcripts containing the exon E8a. Numbers above each graph represent length of the amplicons expressed in base pairs

From HELA cells transfection with the Minigene 800 reporter we obtained the same result of the transfection with the sole pBSplicing vector, indicating that exon E8a was never included in non-neuronal cells (Fig. 1.2.5 A). On the contrary, when we transfected Minigene 800 reporter plasmid in SH-SY5Y cells, we identified two peaks: one deriving from the minigene cassette costitutive splicing and one, 12 bp longer, deriving from the exon E8a inclusion in the mature cassette transcript (Fig. 1.2.5 B).

We performed Minigene transfection assay in only one non-neuronal cell line (HELA) and in three different neuronal cell lines (SH-SY5Y, Neuro2A and IMR32) in order to evaluate which one of them had the highest basal percentage of exon E8a containing mature transcripts.

Figure 1.2.6 shows the result deriving from at least three independent transfection assays. We calculated the average percentage of exon E8a inclusion and the standard deviation (with the formula [standard error/ √n+ where n is the number of results), for each cell line.

Exon E8a was included in all the neuronal-like cells lines that we used, with the highest percentage in SH-SY5Y (11,08 ± 2,35 %), followed by N2A (4,5 ± 0,36 %) and IMR32 (2,08 ± 0,38 %), but it was never included in HeLa cells.

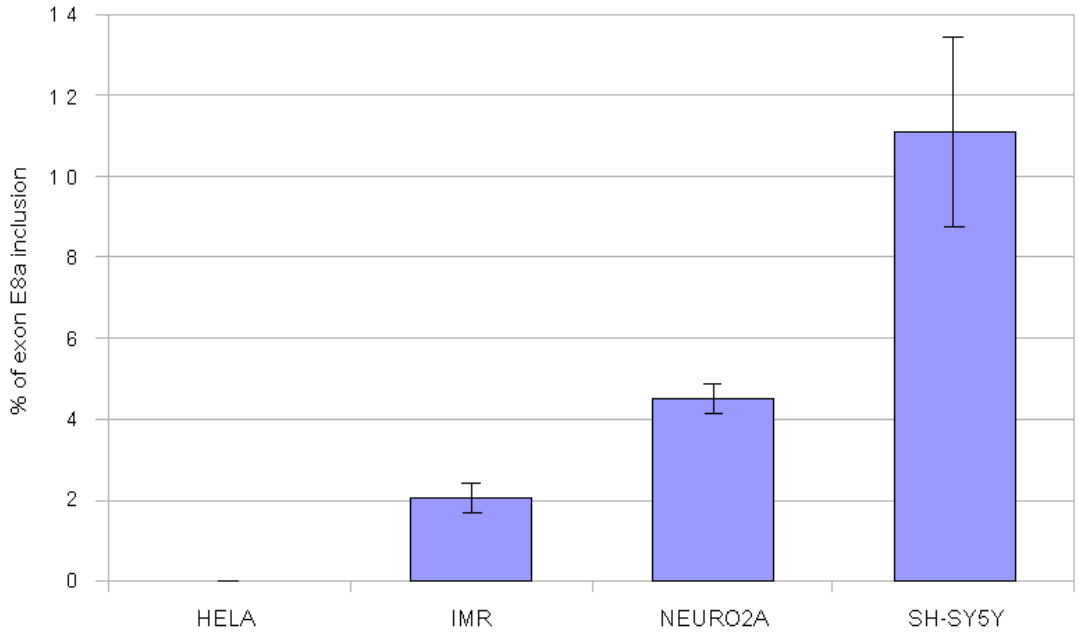


Fig. 1.2.5 Percentage of exon E8a inclusion in Minigene 800 mature transcript.

Overexpressing the same amount of Minigene 800 construct in different neuronal and non-neuronal cell lines we found that SH-SY5Y were the best performing neuronal-like cellular model.

Since Minigene 800 cassette was correctly spliced only in the neuronal context, we can say that the cloned genomic sequence contains the *cis*-acting elements necessary to direct neuro-specific splicing inclusion. Furthermore, exon E8a was included only in a fraction of mature transcript, likewise in the endogenous LSD1 transcripts where LSD1-8a splicing isoform does not replace the ubiquitary one but coexists with it.

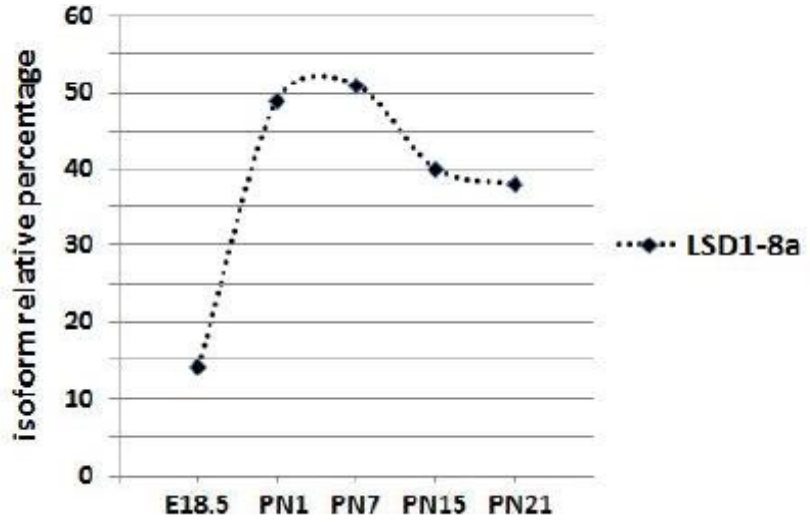
It is also very interesting to note that although neuronal-like SH-SY5Y cells do not express endogenous LSD1-8a, exon E8a is included in the Minigene 800 transcript. Thus, we could conclude that the splicing mechanism regulating exon E8a inclusion in our reporter system is not as tightly regulated as the one acting on the endogenous transcript, possibly due to the lack of additional negative cis-elements or to the increased availability of positive cis-acting sequences.

1.3 ROLE OF NOVA1 IN REGULATING EXON E8a SPLICING

As described above, NOVA1 was shown in HIT-CLIPS experiments to be able to bind *in vivo* exon E8a-containing LSD1 transcripts [67]. Moreover, this neurospecifc pre-mRNA splicing factor has a lot of binding sites in the exon E8a downstream intron and it was reported in literature to preferentially regulate the inclusion of phosphorylated exon, exactly like the exon E8a (unpublished data from our laboratory). Thus, using Minigene 800 reporter system we assayed NOVA1

involvement in the exon E8a splicing regulation.

First of all we searched for NOVA1 endogenous expression during perinatal stage of rat brain development when the exon E8a-containing LSD1 isoform has the highest level of expression. Performing a qrf-PCR (Fig.1.3.1 up) and a RT-PCR (Fig.1.3.1 down) on rat brain cortex at different developmental stages we demonstrated that LSD1-E8a and NOVA1 have a similar expression pattern.



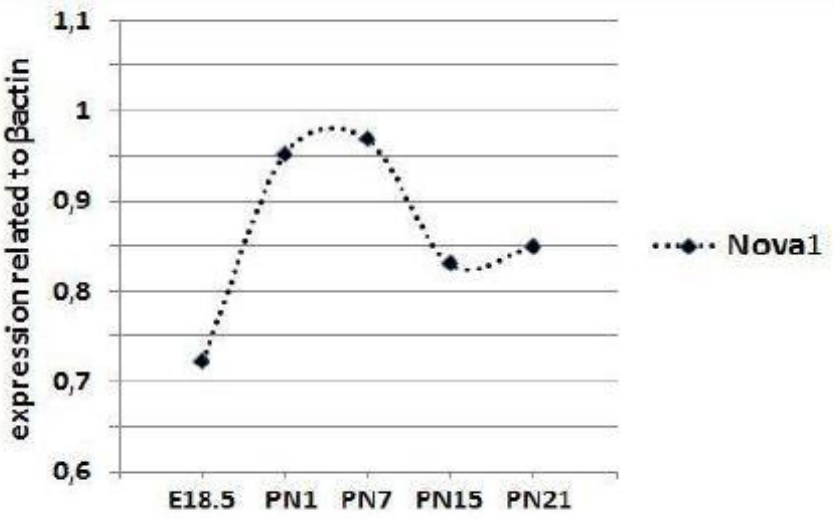


Fig. 1.3.1 LSD1-8a isoform and NOVA1 expression during different perinatal stages of development.

The relative amount of LSD1-8a isoform was measured by rqf-PCR; cDNA were obtained from total RNA of rat embryonic cortex (E18.5) and postnatal rat cortex (PN). The first graph represent the relative percentage of LSD1-8a isoform with respect to the sum of the four. The second graph shows NOVA1 quantification by qRT-PCR on total RNA extract from the indicated rat cortex samples, normalized on endogenous B-actin gene expression.

Since we knew that SH-SY5Y cells endogenously expressed NOVA1 (data not shown) and this neuronal-like cell line showed the best basal percentage of E8a inclusion in LSD1 transcripts deriving from Minigene 800 assays, we investigated if the over-expression of NOVA1 could modulate exon E8a usage. Therefore we cloned mouse NOVA1 cDNA in pCMV-Tag1 expression vector in order to obtain a fusion protein between the amino terminal protein tag (FLAG) and NOVA1 protein.

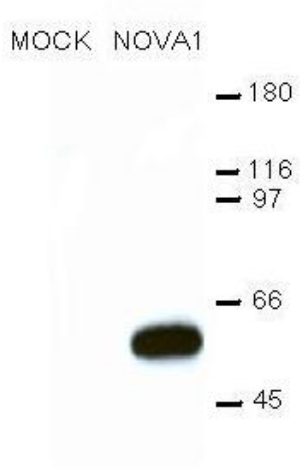


Fig. 1.3.2 Western Blot analysis of FLAG-NOVA1 expression in HELA cells

The molecular weight marker (M) is the Protein SHARPMASS Broad (EuroClone). Western was decorated with anti-Flag antibody.

Starting from the Image clone containing mouse NOVA1 cDNA, we performed a PCR using the primers “NOVA BclI FW” and “NOVA HindIII RV”, which contained in

their 5' ends the specific restriction sites BclI and HindIII respectively. We then cut the PCR product, with the enzymes BclI and HindIII, and the vector with the enzymes BglII (compatible with BclI) and HindIII, in order to ligate the insert in the backbone vector. Finally, to control the expression of the protein we transfected HELA cells and we performed Western Blot analysis, using anti-FLAG Ab (Fig.1.3.2).

In order to evaluate if NOVA1 was able to induce exon E8a inclusion in Minigene 800 reporter transcripts, we co-transfected Minigene 800 with or without NOVA1 expression plasmid both in neuronal (SH-SY5Y) and non-neuronal (HELA) cell line as control.

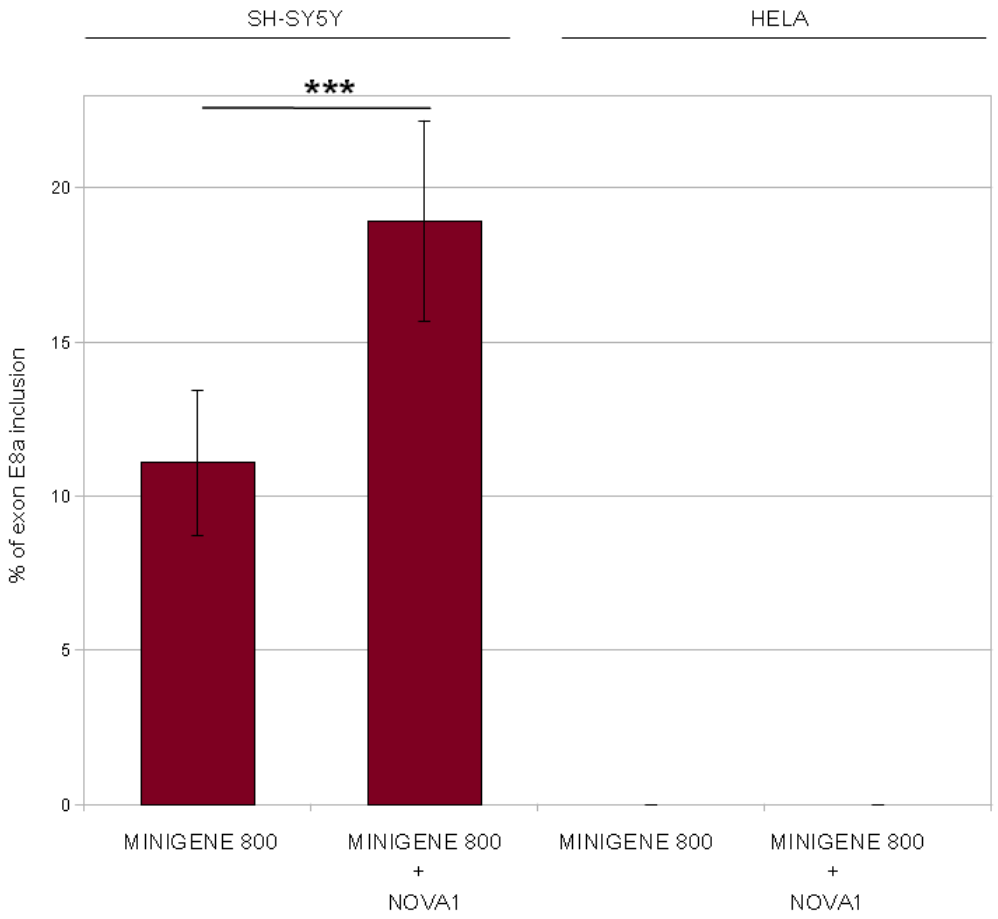


Fig. 1.3.3 Percentage of exon E8a inclusion in Minigene 800 mature transcript. overexpressing NOVA1 protein

Overexpressing NOVA1protein together with the Minigene 800 construct in neuronal (SH-SY5Y) and non-neuronal (HELA) cell lines we found that only in a neuronal-like cells NOVA1 was able to induce a significative increasing of the exon E8a inclusion. A Student's t test (I Stat t I ≥ T p/2) was applied to percentage values of exon E8a inclusion by comparing Minigene 800 with Minigene 800+NOVA1 in SH-SY5Y cells. *p<0.05; **p<0.01; ***p<0.001.

Several experiments of equimolar transfection with NOVA1 and Minigene 800 vectors gave an increase of exon E8a inclusion in neuronal-like cells, from 11,08% ± 2,35 to 18,94% ± 3,25 (Fig. 1.3.3; Minigene 800 vs Minigene 800+NOVA1, 11,08% ± 2,35 vs 18,94% ± 3,25, two tailed *t* test $p=4,4*10^{-13}$). On the contrary we

never observed E8a inclusion in non-neuronal cells (Fig. 1.3.3). Thus we can say that the overexpression of NOVA1 in neuronal-like cells is able to enhance exon E8a inclusion in the mature Minigene 800 transcripts but other neuronal splicing factors, not present in HELA cells, are most probably required.

We also repeated transfection assays increasing NOVA1:Minigene 800 ratio to understand if the inclusion trend reflected the amount of transfected NOVA1. Fig. 1.3.4 shows the results that we obtained about the percentages of exon E8a inclusion in Minigene 800 transcript with different amounts of NOVA1. These data confirm a positive trend of NOVA1 dose-dependent effect in regulating exon E8a inclusion (Fig. 1.3.4; Minigene 800 vs Minigene 800+NOVA1 1x, 11,08% ± 2,35 vs 17,75% ± 2,68, two tailed *t* test *p*=0,000016; Minigene 800 vs Minigene 800+NOVA1 2x, 11,08% ± 2,35 vs 18,83% ± 1,82, two tailed *t* test *p*=0,00000086; Minigene 800 vs Minigene 800+NOVA1 3x, 11,08% ± 2,35 vs 23,35% ± 0,78, two tailed *t* test *p*=0,00000011).

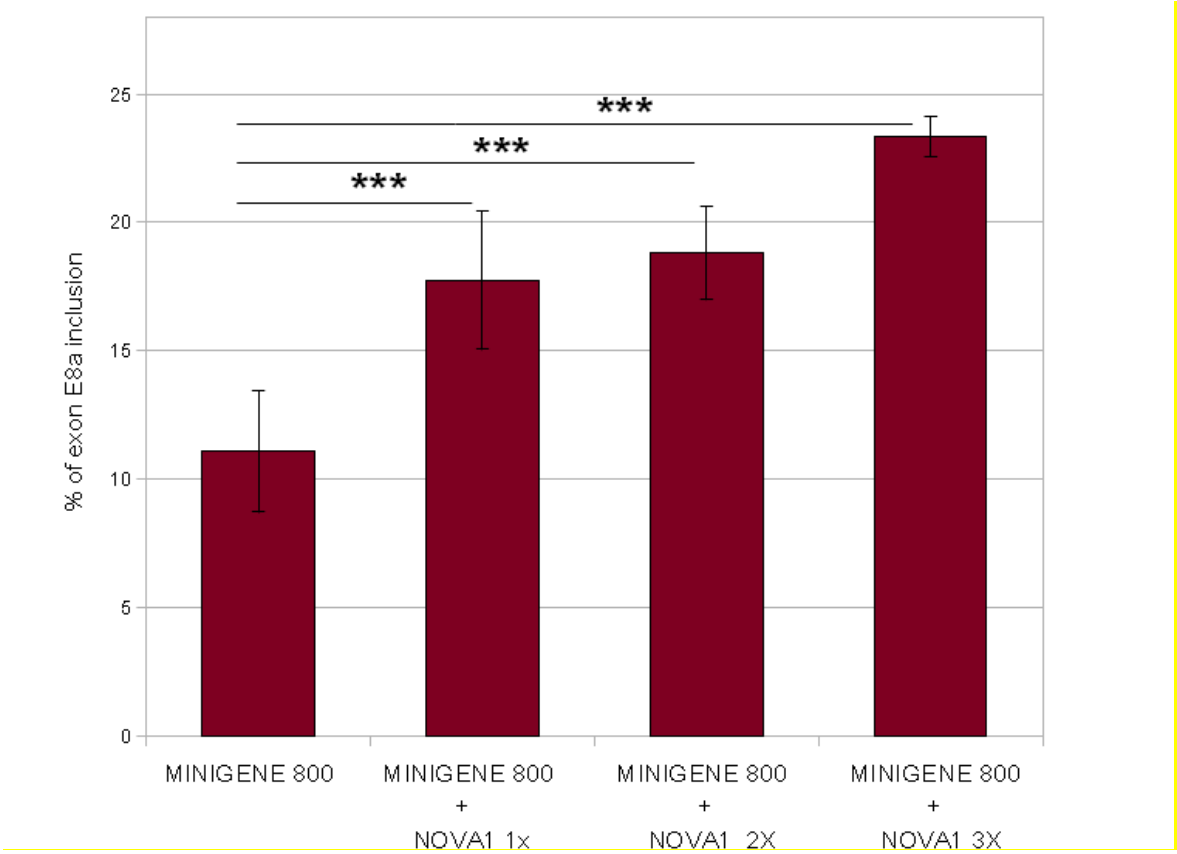


Fig. 1.3.4 Percentage of exon E8a inclusion in Minigene 800 mature transcript. overexpressing different amount of NOVA1 protein.

Overexpressing one-fold, two-fold and three-fold NOVA1protein relative to Minigene 800 amount in neuronal (SH-SY5Y) cell line, we found a positive trend of the dose-dependent effect of NOVA1 in promoting exon E8a inclusion. A Student's t test (I Stat t I ≥ T α/2) was applied to exon E8a percentage of inclusion values by comparing Minigene 8'00 with Minigene 800+NOVA11x/2x/3x. *p<0.05; **p<0.01; ***p<0.001.

Moreover, to further prove NOVA1 involvement in exon E8a splicing regulation we analyzed the effect of knocking down NOVA1 in SH-SY5Y cells.

To this aim we designed and generated a shRNA molecule against NOVA1 in order to interfere with its expression. NOVA1 target sequence was composed of 19 nucleotides but each of the two complementary shRNA strand contained both sense (green) and anti-sense (yellow) sequences, separated one from the other by

the 9 nucleotides of the hairpin loop. This allowed the right folding of the shRNA transcript in a typical harpin structure. Furthermore at the 5' ends of each complementary shRNA strand we inserted restriction sites for BglII and XhoI for the directional cloning inside pSuper.gfp/neo vector. Such vector is suitable for shRNA molecules cloning because of the presence of the H1 promoter that is recognize by RNA-pol III.

(BglII) hairpin (XhoI)

5'-GATCCCC-TGATGGGATCAACTTGCAA-TTCAAGAGA-TTGCAAGTTGATCCCATCA-TTTTTC-3'

3'-GGG-ACTACCCTAGTTCAACGTT-AAGTTCTCT-AACGTTCAACTAGGGTAGT-AAAAAGAGCT-5'

NOVA1 shRNA sequence was designed according to the 8 conventional Reynolds parameters to increase interfering efficiency. Each parameter has a score and the total final score for a shRNA molecule has to be at least 6/8.

The score of our shRNA molecule was 7/8. After the annealing of the 2 complementary strands in buffer STE they were cloned in pSuper.gfp/neo vector linearized with BglII and XhoI.

In order to evaluate if the shRNA was able to induce NOVA1 silencing throught RNAi we co-transfected a 6-wells plate of SH-SY5Y cells with NOVA1 plasmid and two increasing amount of siRNA molecule in order to obtain NOVA1:siRNA ratio of 1:1 and 1:3. Then by real time PCR we quantified NOVA1 transcript amount in each samples after treating them with DNAse to degradate cDNA deriving from

NOVA1 plasmid. NOVA1 level of expression was normalized on endogenous RPSA transcript. Fig.1.3.5 shows that samples overexpressing NOVA1 and siRNA molecule had a reduction of NOVA1 expression of about 55-60% compared to cells transfected only with NOVA1. However siRNA silencing effect was not dose-dependent (Fig. 1.3.5; NT vs NOVA1, 1 ± 0 vs 2,32 ± 0,09, two tailed *t* test p=0,0066; NOVA1 vs NOVA1+siRNA 1:1, 2,32 ± 0,09 vs 0,42 ± 0,02, two tailed *t* test p=0,0014; NOVA1 vs NOVA1+siRNA 1:3, 2,32 ± 0,09 vs 0,44 ± 0,02, two tailed *t* test p=0,0024).

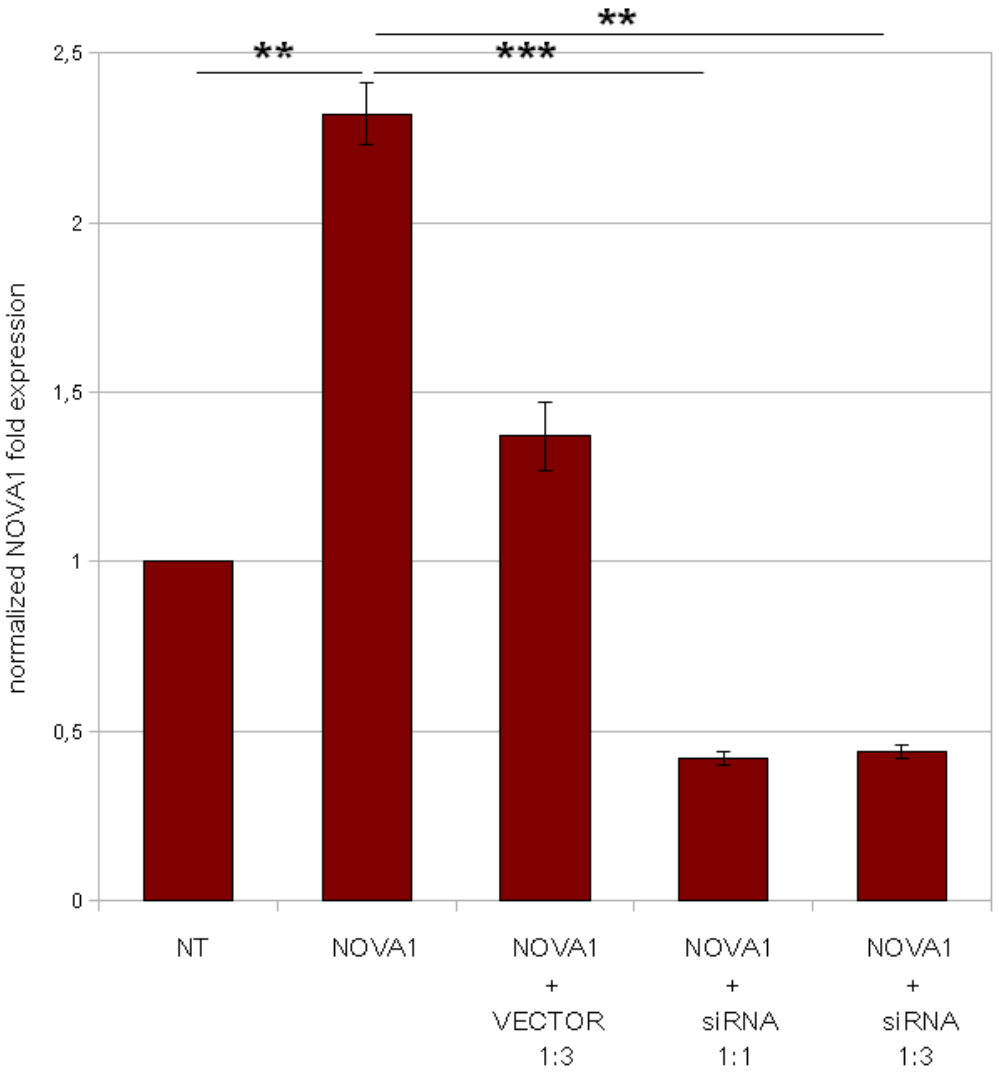


Fig. 1.3.5 Level of NOVA1 expression in SH-SY5Y cells, transfected with different amount of shRNA molecule

NOVA1 expression was normalized over the RPSA expression level. The endogenous amount of NOVA1 transcript (NT) was set to 1 and in all the other conditions NOVA1 level was represented as expression fold change relative to 1. Nova1 expression was significatively increased when it was over-expressed in SH-SY5Y cells (NOVA1 lane) , while it resulted significatively decreased in presence of different amounts of siRNA molecule. No differences in the level of NOVA1 silencencing were found using onefold or threefold siRNA molecule relative to NOVA1. A Student's t test (| Stat t | ≥ T α/2) was applied to normalized Nova1 fold expression change by comparing the endogenous amount of NOVA1 transcript (NT) with the overexpressed NOVA1 amount (NOVA1) . The same Student's t test was applied to normalized Nova1 fold expression change by comparing the overexpressed NOVA1 amount (NOVA1) with the overexpressed NOVA1 and siRNA 1:1 or 1:3. *p<0.05; **p<0.01; ***p<0.001.

Since we confirmed that siRNA molecule was able to strongly reduce NOVA1 expression, we overexpressed NOVA1 and its small interfering RNA together with

Minigene 800 reporter plasmid in SH-SY5Y cells, to verify if NOVA1 was really able to modulate exon E8 inclusion. To this aim, we preserved 1:3 ratio between NOVA1 and siRNA molecule.

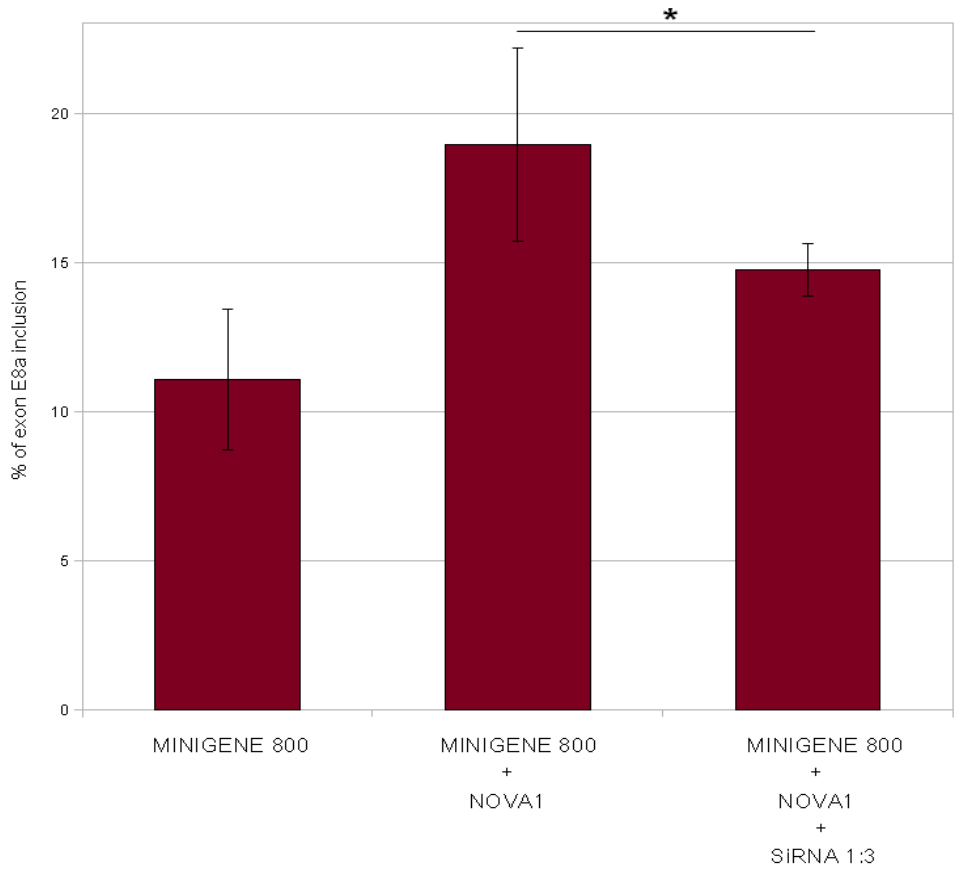


Fig. 1.3.6 Percentage of exon E8a inclusion in Minigene 800 mature transcript. overexpressing NOVA1 protein and siRNA molecule

Overexpressing NOVA1 and three-fold amount of siRNA molecule, exon E8a percentage of inclusion into Minigene 800 mature transcripts was significatively decreased. A Student's t test (I Stat t I ≥ T p/2) was applied to exon E8a percentage values by comparing Minigene 800+NOVA1 with Minigene 800+NOVA1+siRNA 1:3. *p<0.05; **p<0.01; ***p<0.001.

These preliminary experiments of siRNA transfection with Minigene 800 vector revealed that exon E8a splicing is responsive to NOVA1 level of expression (Fig. 1.3.6; Minigene 800+NOVA1 vs Minigene 800+NOVA1+siRNA 1:3, 18,94% ± 3,25 vs 14,73% ± 0,87, two tailed *t* test *p*=0,035). This is in accordance to what we

found about a dose-dependent inclusion of the exon E8a mediated by increasing the amounts of NOVA1.

Taken together the experiments performed so far showed that NOVA1 is involved in regulating exon E8a splicing but it is not sufficient alone to promote exon E8a expression in a non-neuronal cellular context, indicating that other neurospecific splicing factors have to cooperate with NOVA1 to allow exon E8a inclusion.

1.4 SEIZURES-DEPENDENT CO-REGULATION OF NOVA1 EXPRESSION AND EXON E8a INCLUSION

We know from literature that alternative splicing may be affected by some mutations associated with human epilepsy [78], and NOVA1 expression level has been found to be modulated in response to epileptic seizures [77]. This finding suggested the interesting possibility that also the neuronal LSD1 expression might be regulated by seizures, since the neurospecific exon E8a inclusion is under the partial control of NOVA1.

Thus we speculated that farmacologically induced epileptic seizures could be a good model to study the regulation of NOVA1 expression and the neuronal LSD1 alternative splicing responsiveness.

To this aim we performed the well characterized LiCl/Pilocarpine treatment to induce tonic-clonic seizures in 15 C57BL/6 mice of 5 weeks. We used other 15 identical mice as control. Mice were treated by Lithium 24 hours prior to Pilocarpine

(250 mg/Kg, i.p.) and 10-15 minutes after the pilocarpine injection they experienced multiple seizure events. Diazepam was then delivered in order to block tonic-clonic convulsion after one hour from the onset of seizures. Animals were sacrificed 7 hours after Pilocarpine injection and their hippocampus was rapidly collected and liquid nitrigen frozed. We used Pilocarpine because it is a selective muscarinic acetylcholine receptor agonist that provokes seizures indirectly, by over-stimulation of the cortical glutamatergic neurons. We also knew that this was the only farmacological treatment inducing significant changes in the levels of Nova1 mRNAs [77].

We then analysed the exon E8a inclusion in mature LSD1 transcripts of treated mice to investigate if it was modulated by pharmacologically induced neuronal activation (Fig. 1.4.1). Using a couple of fluorescinated primers (mFAMex2 FW and mLSD1ex9 RV), we quantitatively evaluated LSD1 neurospecific splicing thanks to relative quantity fluorescent-PCR (rfq-PCR) on cDNA deriving from hippocampal area of control and Pilocarpine-treated animals. In this brain area, which is the most affected by seizures, we observed a strong reduction of exon E8a inclusion frequency, from 38,92% ± 1,33 to 27,73% ± 1,94 (Fig. 1.4.1 A; control vs pilocarpine treated mice, 38,92% ± 1,33 to 27,73% ± 1,94, two tailed *t* test p= 0,0067). Importantly, the levels of LSD1 gene transcription, measured by qRT-PCR, were not affected by seizures (data not shown).

The same RNAs were analyzed by means of qRT-PCR for the expression level of

NOVA1. We also checked for BDNF expression level, as experimental control of transcriptional response to seizures. As shown in Fig. 1.4.1 B, BDNF transcription underwent the expected transcriptional burst in treated mice compared to controls (BDNF expression level in controls vs Pilocarpine-treated mice, 1 vs 44,32 ± 13,94, two tailed *t* test *p*= 0,0000037), while in the same time window NOVA1 showed a small but statistically significant expression decreasing compared to control (Fig 1.4.1 B, NOVA1 expression level in controls vs Pilocarpine-treated mice, 1 vs 0,77 ± 0,33, two tailed *t* test *p*= 0,022)

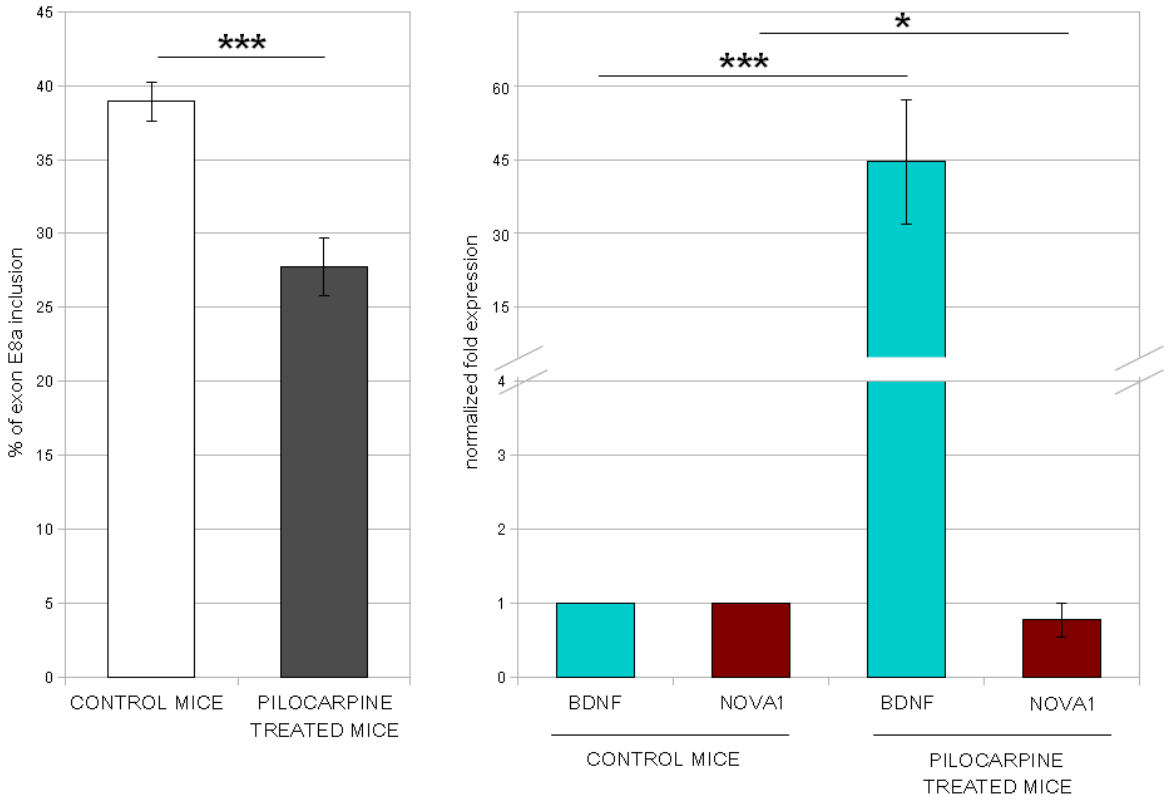


Fig. 1.4.1 Analysis of the exon E8a percentage of inclusion in LSD1 mature transcripts (A) and splicing regulators expression level (B) in pilocarpine-treated mice and controls.

(A) Percentage of exon E8a-including LSD1 transcripts in 15 controls and 15 pilocarpine-treated mice; data obtained by rqf-PCR. A Student's t test (| Stat t | ≥ T p/2) was applied to the exon E8a inclusion percentage by comparing controls with Pilocarpine-treated mice. *p<0.05; **p<0.01; ***p<0.001. (B) Fold expression change of NOVA1 and BDNF in Pilocarpine-treated mice compared to controls. Gene expression was normalized over the endogenous RPSA expression level. The amount of NOVA1 and BDNF

transcripts in not treated mice was set to 1 and their level in Pilocarpine treated animals was represented as expression fold change relative to 1. A Student's t test (| Stat t | ≥ T α/2) was applied to normalized Nova1 and BDNF fold expression change by comparing their expression level in not treated mice with their corresponding level after Pilocarpine treatment. *p<0.05; **p<0.01; ***p<0.001.

This result let us to hypothesize a direct involvement of the alternative splicing mechanism modulating exon E8a inclusion in neuronal LSD1 transcripts as responsive to seizures, through a positive splicing regulator such as NOVA1.

1.5 ROLE OF nPTB IN REGULATING EXON E8a SPLICING

In addition to NOVA1 binding sites in the exon E8a downstream intronic portion, the 800-bp conserved region harbors also many CT-rich sequences mainly located inside the exon E8a upstream intron. By means of yeast two-hybrid system NOVA1 interacting proteins were isolated and among them an RNA binding protein closely related to the polypyrimidine tract-binding protein (nPTB) was identified [71]. The expression of nPTB is enriched in the brain, both in glia and in neurons. Furthermore we know that nPTB interacts with NOVA1 in cell lines and colocalizes with NOVA1 within neuronal nuclei [71].

For all these reasons we decided to investigate a possible role of nPTB in exon E8a splicing regulation.

In the model of *in vitro* generation of neurons from mouse embryonic stem cells we could follow the expression kinetic of exon E8a during neuronal differentiation. In

particular LSD1-E8a transcript appears at DIV7. In the same cell model, it has been observed that the expression of Nova1 and nPTB occurs at DIV5 preceding of about two days the expression of LSD1-E8a isoform [72]. These observations suggested a possible involvement of nPTB in regulating exon E8a splicing.

First of all by RT-PCR we analyzed the kinetic expression of nPTB in rat cortical neurons at different embryonic developmental stages. nPTB expression level was normalized on B-actin one (Fig.1.5.1). We observe that nPTB is expressed in the same developmental stages of LSD1-E8a and NOVA1, even if it follows an opposite temporal kinetic: when the exon E8a-containing isoform and NOVA1 expression increase, nPTB expression decreases.

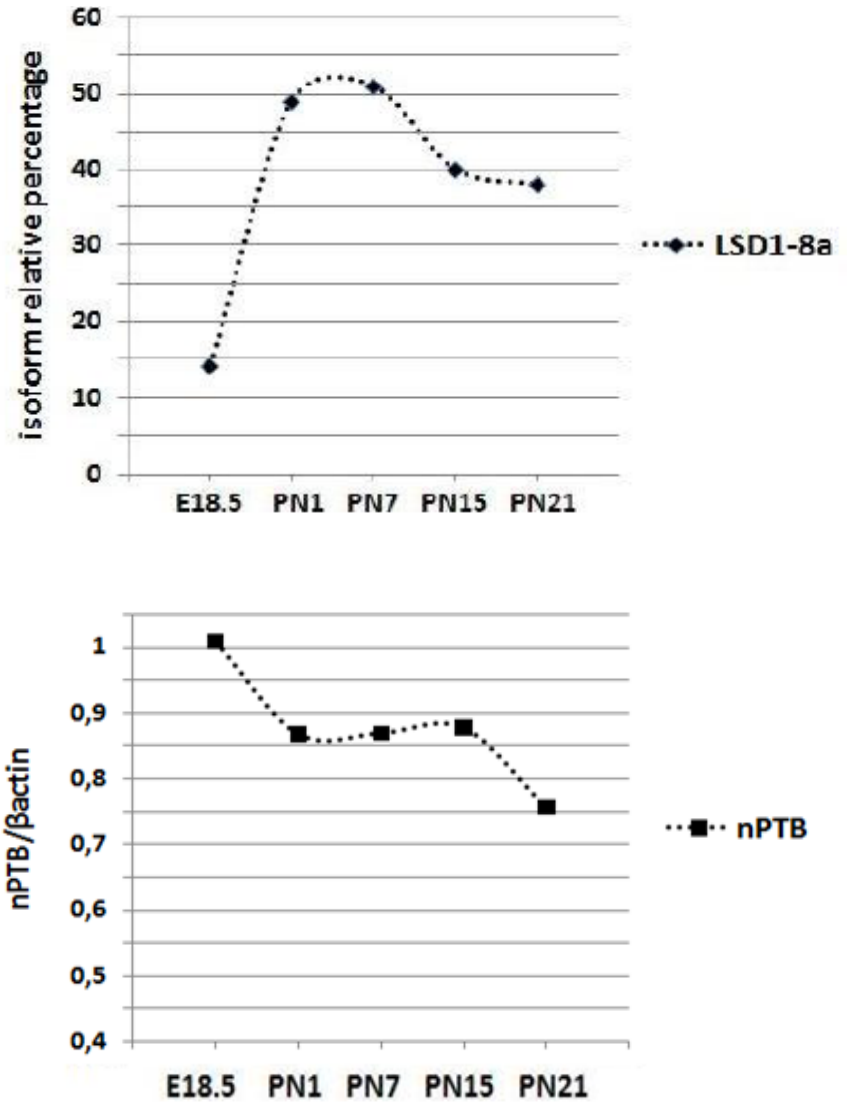


Fig. 1.5.1 LSD1-8a isoform and nPTB expression during different perinatal stages of development.

The relative amount of LSD1-8a isoform was measured by rqf-PCR; cDNA were obtained from total RNA of rat embryonic cortex (E18.5) and postnatal rat cortex (PN). The first graph represent the relative percentage of LSD1-8a isoform with respect to the sum of the four. The second graph shows nPTB quantification by qRT-PCR on total RNA extracted from the indicated rat cortex samples, normalized on B-actin.

To deeply investigate the effective role of nPTB in exon E8a splicing regulation we transfected Minigene 800 reporter vector with or without nPTB expression plasmid, which was kindly sent us from M. Baralle laboratory. The nPTB plasmid that we

used contained indeed a “humanized” version of the wild-type nPTB ORF, obtained by the optimization of the nPTB codons content at the third position. Increasing C+G content at such position they were able to obtain a much more efficient translation *in vitro* and *in vivo* to evaluate nPTB overexpression effects [73].

So far we have performed only few experiments with nPTB expression plasmid and the Minigene 800 reporter system but we have found that nPTB is able to induce exon E8a inclusion into the mature Minigene 800 transcript only in neuronal-like cells, exactly like NOVA1. Fig. 1.5.2 shows that transfecting the same amount of nPTB and Minigene 800, in SH-SY5Y cells we obtained an increasing in the percentage of exon E8a-containing transcript from 11,08% ± 2,35 to 21,2% (Fig. 1.5.2, Minigene 800 vs Minigene 800+nPTB in SH-SY5Y cells, 11,08% ± 2,35 vs 21,2%). On the contrary in HELA cells nPTB had no effect on exon E8a inclusion.

This is a preliminary result deriving from only one experiment, thus it needs to be confirmed; however it sheds a new light on another neurospecific splicing factor that most probably modulates exon E8a alternative splicing in a neurospecific way. We could observe, indeed, that nPTB and NOVA1 had a very similar effect on splicing regulation: they are both necessary but not sufficient to induce exon E8a inclusion. Thus it will be very interesting to investigate their combinatorial effect on splicing regulation in neuronal and non-neuronal cells, in order to find out which is the condition that manages exon E8a alternative splicing.

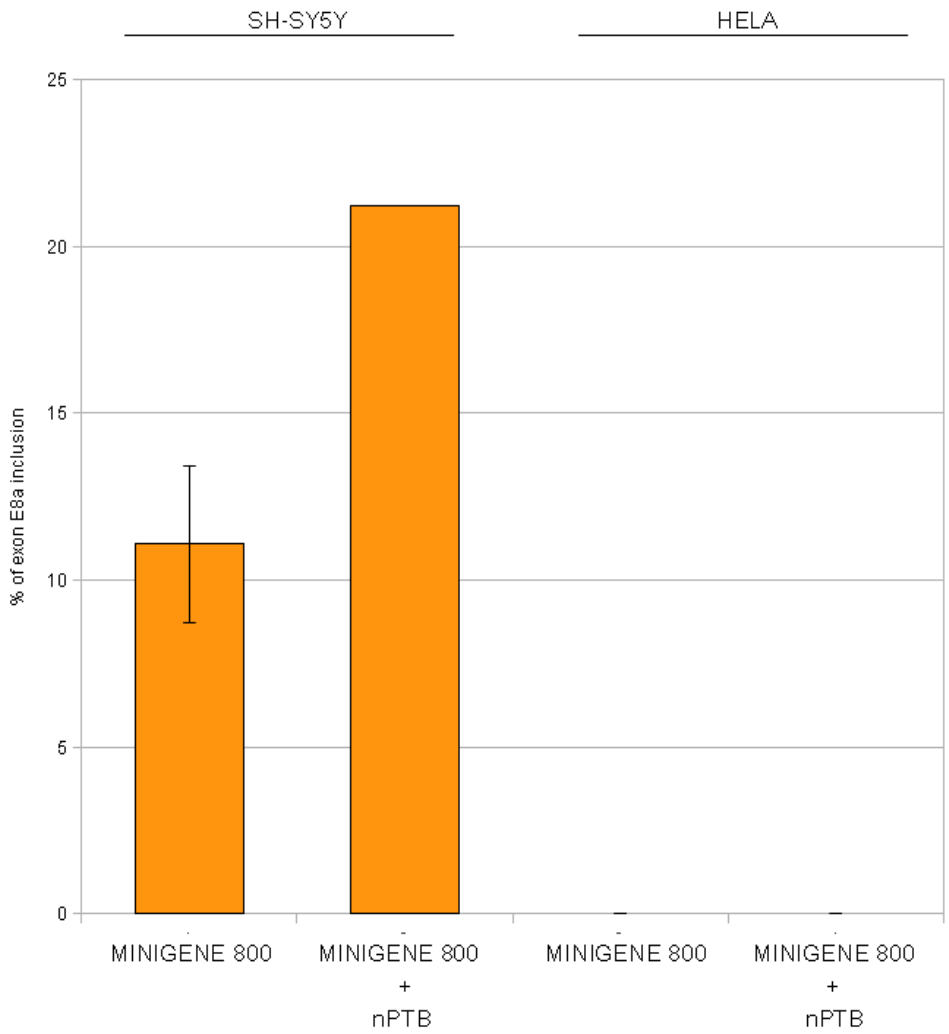


Fig. 1.5.2 Percentage of exon E8a inclusion in Minigene 800 mature transcript. overexpressing nPTB protein.

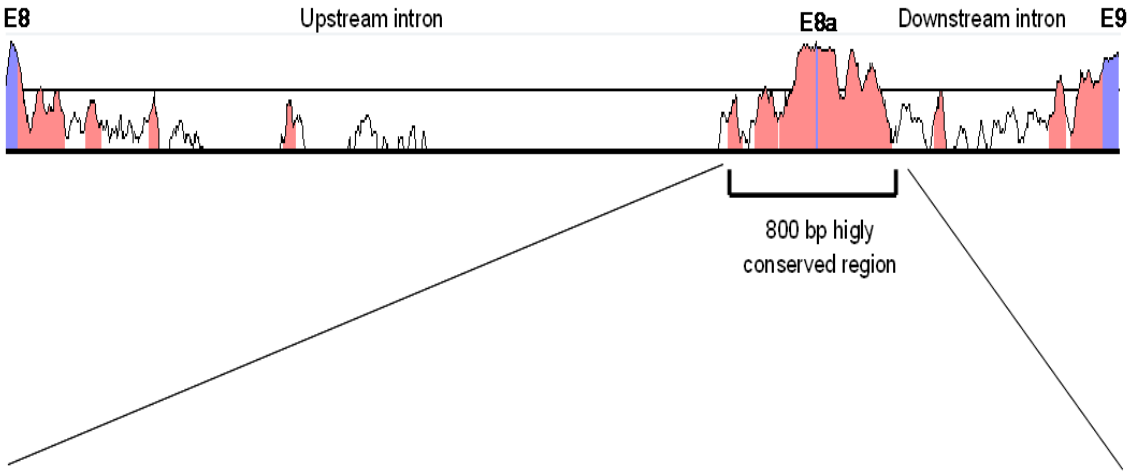
Overexpressing nPTB protein together with the Minigene 800 construct in neuronal (SH-SY5Y) and non-neuronal (HELA) cell lines, we found that only in a neuronal-like cells nPTB was able to induce an increasing of the exon E8a inclusion.

1.6 RNA SECONDARY STRUCTURE AND ITS INVOLVEMENT IN EXON E8a SPLICING MODULATION

As reported in the introduction, several studies [61, 62, 63] so far have demonstrated that also the RNA secondary structure *per se* can affect alternative splicing, mainly through the regulation of the *cis*-regulatory elements availability.

Interestingly, a deep analysis of the 800-bp highly conserved intronic region flanking the exon E8a led us to the identification of a 21-bp sequence that is complementary and inverted to the exon E8a-containing one. This sequence starts about 300 nucleotides downstream the exon E8a and it is higly conserved among mammals as well, even if it is an intronic portion (Fig.1.6.1).

The discovery of this conserved complementary and inverted region is quite intriguing since with a good chance it has a role in LSD1 RNA secondary structure folding and thus it may be involved in splicing regulation.



TCACTGCTTTCTGACACTTCCCCCTCACTCCCCTGTCTTTTGTTAACTGCCTTTCTTTTGGCCATTTCTGTGCCTTGACTTTCCCATTT
GGTCTCTGTACAATCTGTTCCTTATTTCTCTGGTGGTTGCAGCTTCTGAGCTAACAGAATGCTTCTGTCATGCTATGTGTTTAAAGCC
TTGATTTTATTGCTGTCATACAAAACTAAACTTGATGTGGCTGTTGGGCTGTTATTTACAGTTTATTTCTCTCTCTGCTGCACTGGTTA
AAAGGACACTGTCAAGGTATTTTTTTCATAATTTTTCAAATCATTTTCCTCACTGTTTACAATAACCCATTAAAAGCTTGAAACCTAAA
TCTTTTCCTTTCAAAGATCCTCTCCATCATACATTCACACTGGTTTGTTGTTGTAGAGTTGCGCCTATATGTTGGGCTGGTGCCATTT
GAGTTTTTGTCATTGAAAGCACGTGTAAAACATTAGGTTTTTTGTGGGTGGTGTGGGTGGACTAAGGGGGAAAGTGTAATGAATCCA
GTGGGCTGCTCAGCAAACAGAGGGAAGAAACAGGTCAAATTTGGTCCCTCTAACCTTGACAGTGTCCCTTTAAATATTTCTCTACCA
CACACCCTGGACCCCAGTTACTGAGCCTGAGTGCTGTGTACTTTTTCAGTATGAACTTCCCTTTGTGATTGTCATCCTAGTTGCAAA
GATTAAAGCTTGTCTGAAGCAAGCCATCCAGTTGACACTGGAAGAGCTTTAATCTGTACCATTTGCACACTTGTGT

exon E8a

5'....TTAAAAGGACACTGTCAAGGT....3'

3'....AATTTCCCTGTGACAGTTCCA.....5'

Fig. 1.6.1 An exon E8a complementary and inverted 21-nt sequence is present in the 800-bp highly conserved intronic regions, downstream the exon E8a.

Exon E8a and its complementary and inverted sequence are in bolt and highlighted; the exon E8a including sequence is written in black and the exon E8a in purple while its 21-bp complementary downstream portion is red.

To address this issue, using “mfold” web server (http://mfold.rna.albany.edu/?q=mfold/RNA-Folding-Form) we determined the predicted RNA secondary structure deriving from the 800-bp region that we cloned into Minigene 800. From the prediction we observed a perfect match between the exon E8a and its downstream complementary sequence, generating a quite long and stable stem structure (ΔG = 94.96 kcal/mol) (Fig. 1.6.2). Moreover, inside the whole 800-bp cloned sequence there are many different complementary regions that form stem structures interrupted only by few SNPs (not shown). Such tightly packed RNA secondary structure could be relevant in hiding and masking single strand cis-acting sequences recognized by splicing machinery.

To test our hypothesis we generated another Minigene construct containing the same 800 bp region but lacking in the sole 12-bp portion of the downstream intron complementary to the exon E8a.

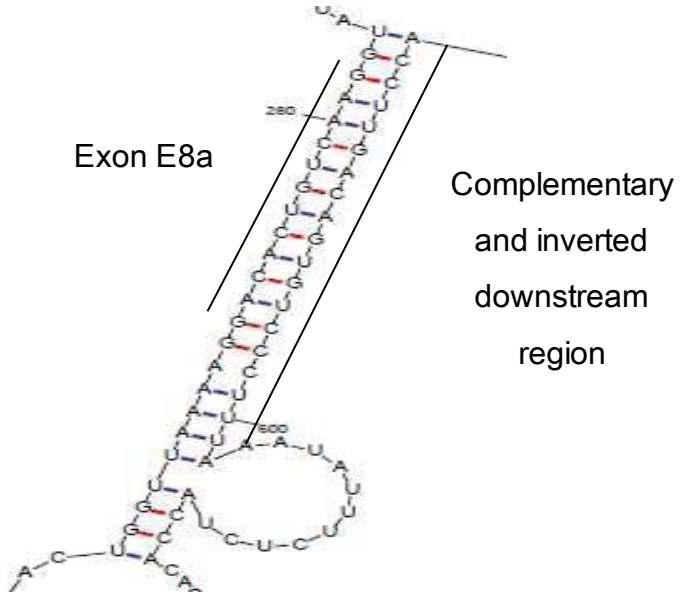


Fig. 1.6.2 "mfold" output.

The 21-bp exon E8a containing sequence and its downstream complementary and inverted regions form a stable stem structure during RNA folding.

In order to obtain the deletion of the 12-bp downstream sequence complementary to the exon E8a we followed the strategy in fig. 1.6.3. Using Minigene 800 as template, we amplified the upstream and downstream regions flanking the deletion by 2 independent PCR reactions, using primers "hLSD1 cloning intron FW" with "new 2 del pal hLSD1 RV" for the up PCR, and "new 3 del pal hLSD1 FW" with "hLSD1 cloning intron RV" for the down PCR. After purification of the two PCR products on the agarose gel, they were mixed together, annealed and used as template for another PCR, obtained with the external primers "hLSD1 cloning intron FW" and "hLSD1 cloning intron RV" (Fig. 1.6.3). These primers had the NdeI restriction site at their 5' end so the resulting amplicon was digested and cloned inside the NdeI unique site of the pBSplicing vector. Then we check for the right

insertion by digesting with HindIII restriction enzyme and by sequencing with forward and reverse FN1 primers.

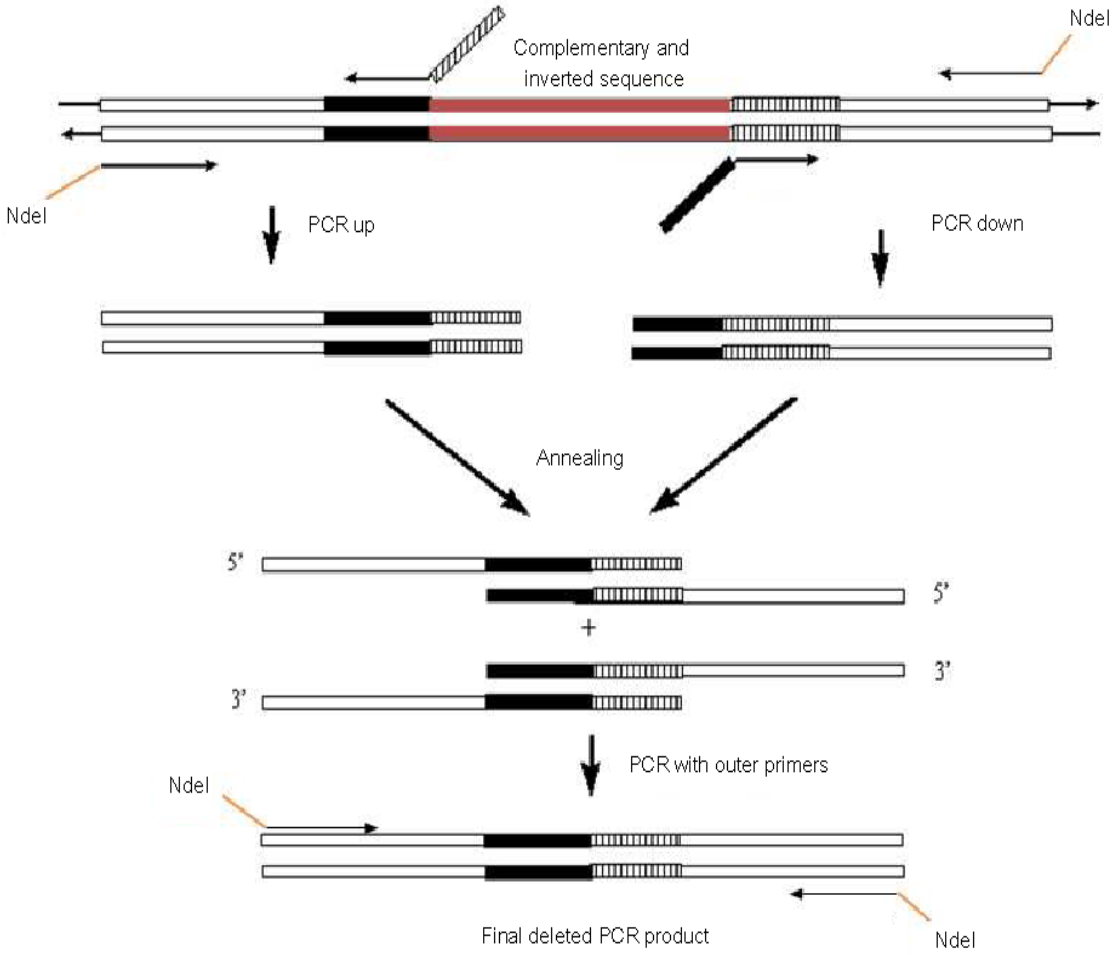


Fig. 1.6.3 Strategy used to obtain the 12-bp core deletion of the exon E8a downstream complementary sequence

Using Minigene 800 as template, we amplified by PCR the upstream and downstream regions flanking the 12bp deletion (in red). We then annealed the 2 PCR products by means of protruding and complementary tails at 5' ends of the internal primers. The annealing was obtained by a slow air cooling of the denatured single strand molecules. Finally by a third PCR on such partially annealed substrate we obtained the deleted fragment with the NdeI restriction sites at the ends.

To understand if the stem structure deriving from the perfect matching between exon E8a and its complementary downstream sequence could really affect exon

E8a splicing we transfected the new deleted Minigene 800 construct (Fig.1.6.4) in neuronal and non-neuronal cells.

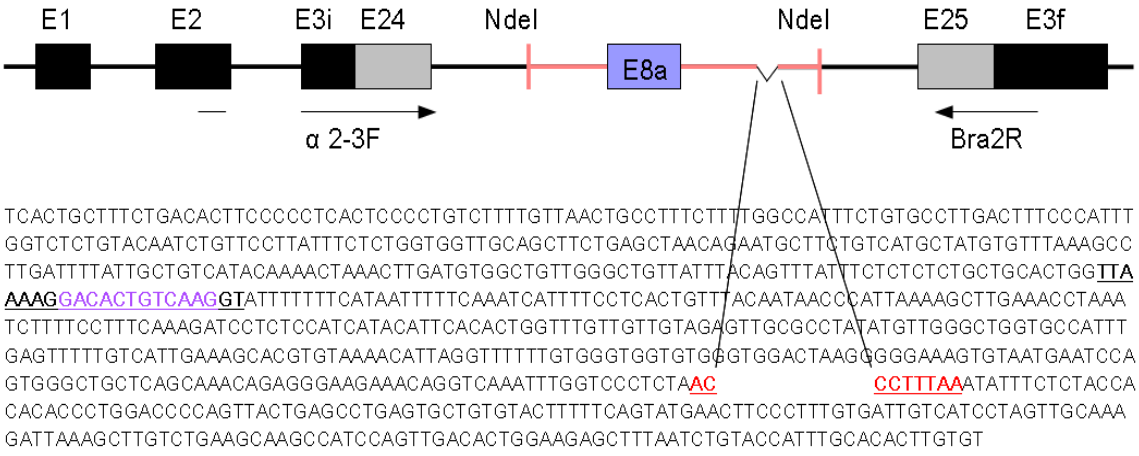


Fig. 1.6.4 Deleted Minigene 800 reporter system

Inside the minigene cassette we cloned same 800-bp highly conserved intronic region used before but lacking in the sole 12-bp downstream core region complementary to the exon E8a

We first overexpressed deleted Minigene 800 alone to set the basal percentage of exon E8a inclusion into the mature deleted Minigene 800 transcript. As Fig.1.6.5 shows, we found that the sole deletion of the exon E8a downstream complementary region was able to highly increase exon E8a inclusion to 85,75% ± 2,26 (Fig. 1.6.5; percentage of exon E8a inclusion into deleted Minigene transcripts in SH-SY5Y, 85,75% ± 2,26; percentage of exon E8a inclusion into deleted Minigene transcripts in HELA, 0%). We observed this effect only in the neuronal-like SH-SY5Y cells, while in non-neuronal cells the inclusion remained

totally absent, exactly like in wild-type Minigene 800 over-expression. This result proves that the downstream complementary and inverted sequence, forming a secondary stem structure with the exon E8a, strongly inhibits its splicing inclusion. The exon E8a “trapping” into a double strand RNA structure most probably prevent the binding of the neuro specific trans-acting factors that normally recognize single strand RNA sequences.

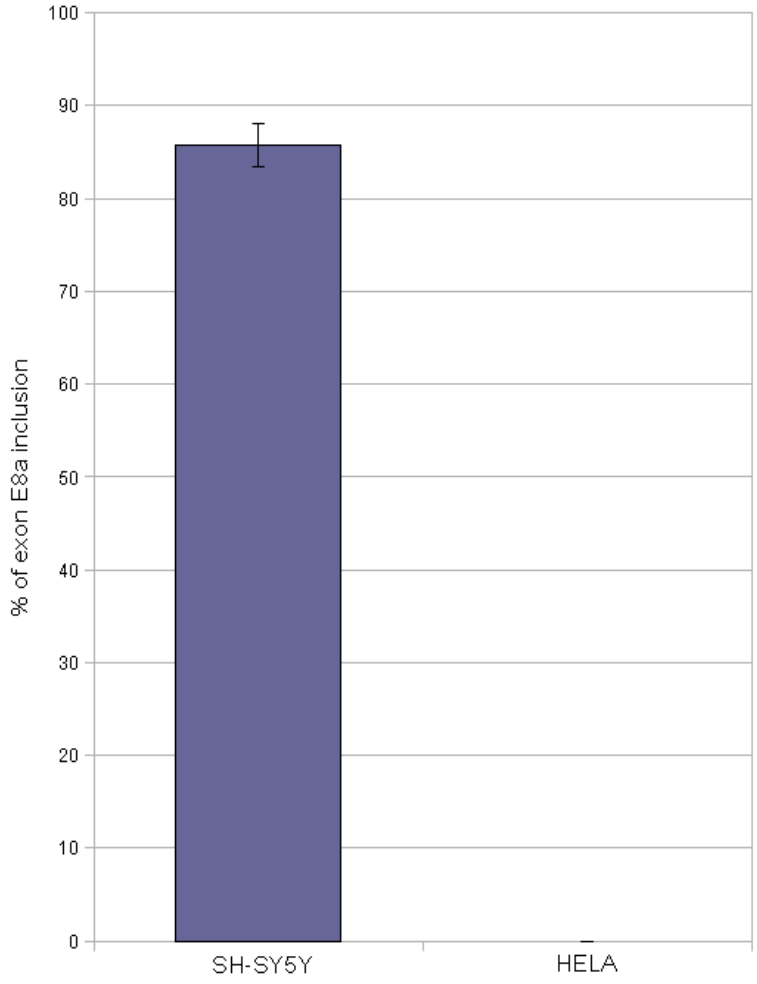


Fig. 1.6.5 Percentage of exon E8a inclusion in deleted Minigene 800 mature transcript.

Overexpressing the same amount of deleted Minigene 800 construct in neuronal (SH-SY5Y) and non-neuronal (HELA) cell lines we found again that inclusion occured only in SH-SY5Y cells, while HELA cells did not promoted exon E8a splicing.

However, *in vivo,* the endogenous exon E8a-containing LSD1 transcripts are non-deleted but the neurospecific splicing occur at the same way. Thus it's clear that in *vivo* the exon E8a inclusion heavily relies on an unknown neurospecific trans-acting factor (or more than one), that promote exon E8a inclusion into the mature transcripts by preventing LSD1 RNA secondary stem structure formation. The 12-bp Minigene 800 deletion is able to recapitulate *in vitro* the best condition of LSD1 RNA structure folding that allows the exon E8a massive inclusion.

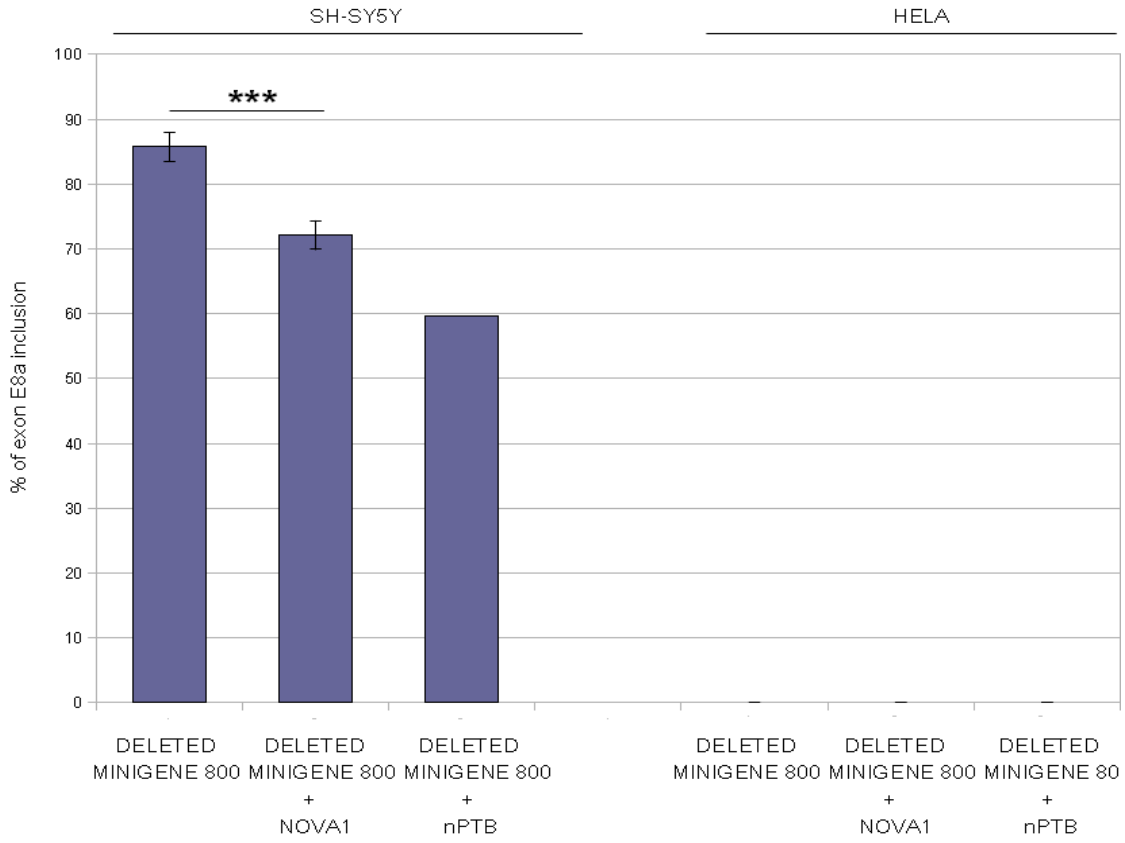


Fig. 1.6.6 Percentage of exon E8a inclusion in deleted Minigene 800 mature transcript. overexpressing NOVA1 and nPTB proteins.

Overexpressing separately the same amount of NOVA1and nPTB together with the deleted Minigene 800, in SH-SY5Y cells we observed a significative decreasing in the percentage of exon E8a inclusion. No effect was seen in HELA cells where the inclusion remained totally absent even upon NOVA1 and nPTB over-expression. A Student's t test (I Stat t I ≥ T α/2) was applied to exon e8a inclusion percentage values by comparing deleted Minigene 800 with deleted Minigene 800+NOVA1in SH-SY5Y cells. *p<0.05; **p<0.01; ***p<0.001.

Together with the deleted Minigene 800, we then overexpressed NOVA1 and nPTB both in SH-SY5Y and in Hela cells to understand if their positive role in exon E8a splicing inclusion could be related to LSD1 RNA structure folding. In non-neuronal cells neither NOVA1 nor nPTB were able to promote exon E8a inclusion inside deleted Minigene 800 transcript (Fig.1.6.6). On the contrary, in SH-SY5Y cells, NOVA1 and nPTB gave the same outcome: both of them did not induce again the expected increasing in the exon E8a inclusion percentage but their over-expression caused a decreasing in exon E8a alternative splicing, from 85,75% ± 2,26 to 72,15% ± 2,21 and 59,7% respectively (Fig. 1.6.6; deleted Minigene 800 vs deleted Minigene 800+NOVA1, 85,75% ± 2,26 vs 72,15% ± 2,21, two tailed *t* test p=0,00000099; deleted Minigene 800 vs deleted Minigene 800+nPTB, 85,75% ± 2,26 vs 59,7%). This may be explained by their partially blocking of the neurospecific unknown factor effect on preventing RNA stem structure folding. In the future it will be interesting also to elucidate the combiantorial effect of NOVA1 and nPTB on the deleted Minigene 800 transcripts, both in neuronal and non-neuronal cell lines.

1.7 SAM68: A PUTATIVE EXON E8a SPLICING REGULATOR

Sam68 (Src-associated in mitosis, 68 kDa) is a nuclear KH-domain RNA binding protein implicated in a variety of cellular processes, including alternative pre-mRNA

splicing. Even if it is not a neurospecific splicing factor, it has been found to be involved in splicing regulation of many different neuronal pre-mRNA, such as Neurexin-1 [74] and neuronal mRNA that become translationally activated in response to electrical activity [75].

Sam68 recognizes and binds A- and T-rich sequences with particular affinity for the AAAATT stretch. This cis-acting site is not present inside the 800-bp intronic conserved region surrounding the exon E8a, but here we found clusters of A and T repetitions that could represent Sam68 binding sites. Moreover Sam68 was predicted by SpliceAid2 to be one of the most relevant trans-acting factor involved in regulating exon E8a inclusion, with a positive score ever higher than NOVA1. For all this reasons we decided to over-express Sam68 (Addgene plasmid 17690) in SH-SY5Y and HELA cells with Minigene 800 and its deleted version. In both the cell-lines Sam-68 did not show any significative effect on exon E8a splicing regulation. We also treated SH-SY5Y and Hela cells with Phorbol ester, which stimulates the Ras/Map kinase pathway, activating Erk and the phosphorylation of Sam68. Indeed, we found in literature that Sam-68 phosphorylation was crucial for its splicing activity [75]. However Phorbol esters treatment did not gave a positive outcome and the inclusion of the exon E8a was the same of the transfection with the sole Minigene 800 deleted or not. This means that Sam-68 is not involved in exon E8a splicing regulation.

1.8 EXON E8b DISCOVERY AND ITS MODULATION BY FOX1

We know from literature that at least 15% of NOVA1 targets may be under NOVA1 and FOX1 combinatorial control, since FOX1-regulated splicing is defined by a position-dependent RNA-regulatory map similar to that of NOVA1 [67].

These observations suggested that additive or synergistic actions of NOVA1 and FOX1 may be favored. FOX1 protein is expressed in mammalian muscle, heart, and brain tissues. In the brain, this proteins is exclusively present in neurons [70], reproducing exon E8a neuronal pattern of expression. In addition, mutation or abnormal expression of FOX1 has been found in patients with severe neurological diseases, including *epilepsy and* mental retardation [70]. Thus, recurring to Minigene reporter system we investigated the putative role of FOX1 in the exon E8a splicing regulation.

Searching for FOX1 binding sites in the 800-bp highly conserved intronic region flanking exon E8a we didn't find TGCATG sequences. On the contrary we identified 2 of these cis-acting elements at the end of the exon E8a downstream intron (Fig. 1.8.1), far from the conserved 800-bp sequence that we cloned into Minigene 800 vector but inside another highly conserved terminal region of the intron.

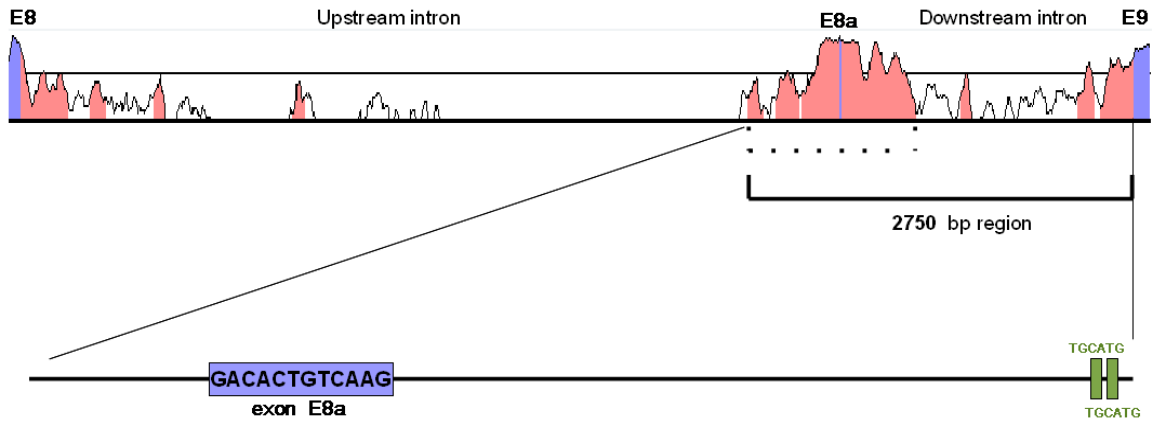


Fig. 1.8.1 FOX1 binding sites in the exon E8a downstream intron

Inside the full length exon E8a downstream intron we found 2 FOX1 binding sites (green boxes) at the level of terminal highly conserved portion (pink).

To investigated if the over-expression of FOX1 could have a role in the exon E8a splicing modulation we extended our Minigene 800 reporter plasmid (Fig.1.8.2), including the full length exon E8a downstream intron and the same conserved portion of the E8a upstream intron, for a total length of about 2750 bp.

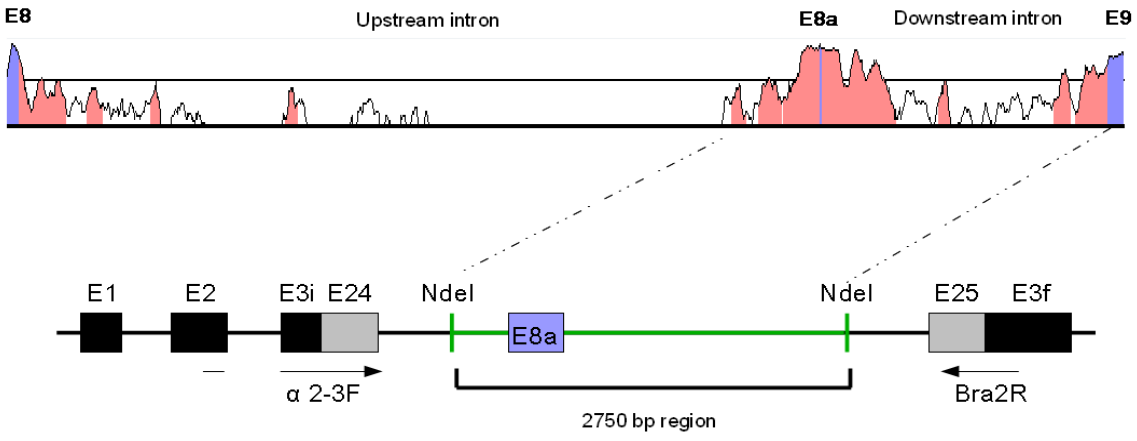


Fig. 1.8.2 Minigene 2750 reporter system

Inside the minigene cassette we cloned the exon E8a with its conserved upstream intronic region and the full length downstream intron, using NdeI restriction site.

Thus we amplified by PCR the genomic fragment of 2750 bp corresponding to the region of interest, using a couple of primers with AseI restriction site at their 5'end (AseI hLSD1-MG FW and AseI hLSD1-MG RV). Then we cloned purified PCR product in NdeI restriction site of pBSplicing vector, thanks to the compatibility between AseI and NdeI, and we check for the right insertion by sequencing with forward and reverse FN1 primers.

In order to obtain FOX1 expression plasmid, starting from the pENTR-A2BP1 clone (Addgene plasmid 16176) containing human FOX1 cDNA, we performed a PCR using primers “FOX-1 cloning FW” and “FOX-1 cloning RV”, containing in their 5’ ends the restriction sites for XbaI and KpnI respectively. We then cut the PCR product and the pCGN vector with the enzymes XbaI and KpnI to ligate the insert in the backbone vector. Finally, we checked for FOX1 expression performing Western Blot analysis with monoclonal anti-HA antibody on a protein lysate deriving from transfected Hela cells (Fig.1.8.3).

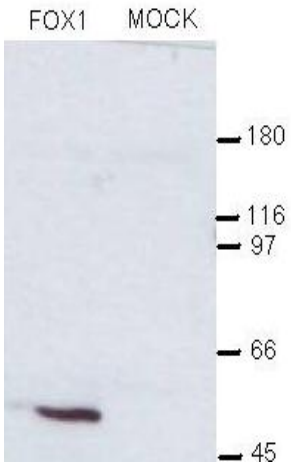


Fig. 1.8.3 Western Blot analysis of FOX1-HA expression in HELA cells

The molecular weight marker (M) is the Protein SHARPMASS Broad (EuroClone). Western was decorated with anti-HA antibody.

In order to evaluate if FOX1 was able to modulate exon E8a inclusion in the new Minigene reporter transcripts, we co-transfected the new version of Minigene 2750 reporter plasmid with or without FOX1 expression plasmid, both in neuronal (SH-SY5Y) and non-neuronal (HELA) cell lines. Several experiments of transfection of the Minigene 2750 reporter alone revealed that the inclusion of the exon E8a occured in 8,06% ± 1,55 of the Minigene 2750 transcripts (Fig. 1.8.4; percentage of exon E8a inclusion in SH-SY5Y cells, 8,06% ± 1,55; percentage of exon E8a inclusion in HELA cells, 0%), a slightly lower percentage compared to the Minigene 800. As we expected exon E8a inclusion occured only in the neuronal-like cell line, while in HELA cells we didn't observe exon E8a containing transcripts (Fig. 1.8.4). Thus we can say that the presence of the full length exon E8a downstream intron has a small negative effect on the neurospecific exon E8a inclusion, may be due to the presence in the Minigene 2750 of cis-acting sequences that negatively compete for the exon E8a inclusion.

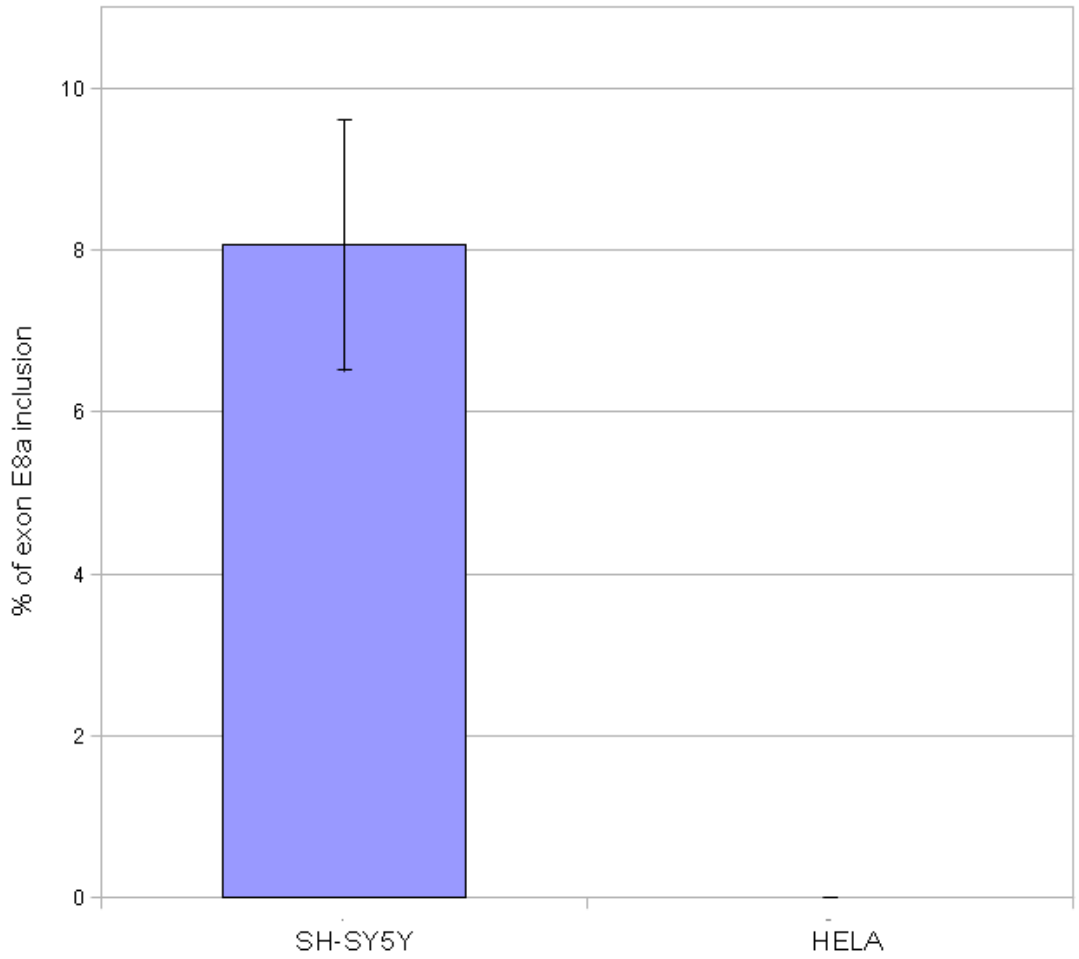


Fig. 1.8.4 Percentage of exon E8a inclusion in Minigene 2750 mature transcript.

Overexpressing Minigene 2750 construct in SH-SY5Y and HELA cells we found that exon E8a was inluded only in neuronal cells but with a lower percentage compared to the inclusion observed in Minigene 800 mature transcripts

When we overexpressed FOX1 protein together with Minigene 2750 in SH-SY5Y and in HELA cells, we obtained a completely unexpected outcome: simply by running Minigene 2750 PCR products on the agorose gel, we didn't find the usual unique PCR band including amplicons with or without the exon E8a, but we observed that FOX1 was able to induce the generation of two distinct bands of different molecular weight (Fig. 1.8.5).

This result was quite intriguing because for the first time it was present both in

HELA and SH-SY5Y cells, and it could have only one explanation: in presence of FOX1 another unexpected splicing event occured that induced the “exonization” of an intronic portion of LSD1.

When we run the samples on the capillary gel electrophoresis we discovered that such new exon was composed of exactly 77 bp and we called it “exon E8b”.

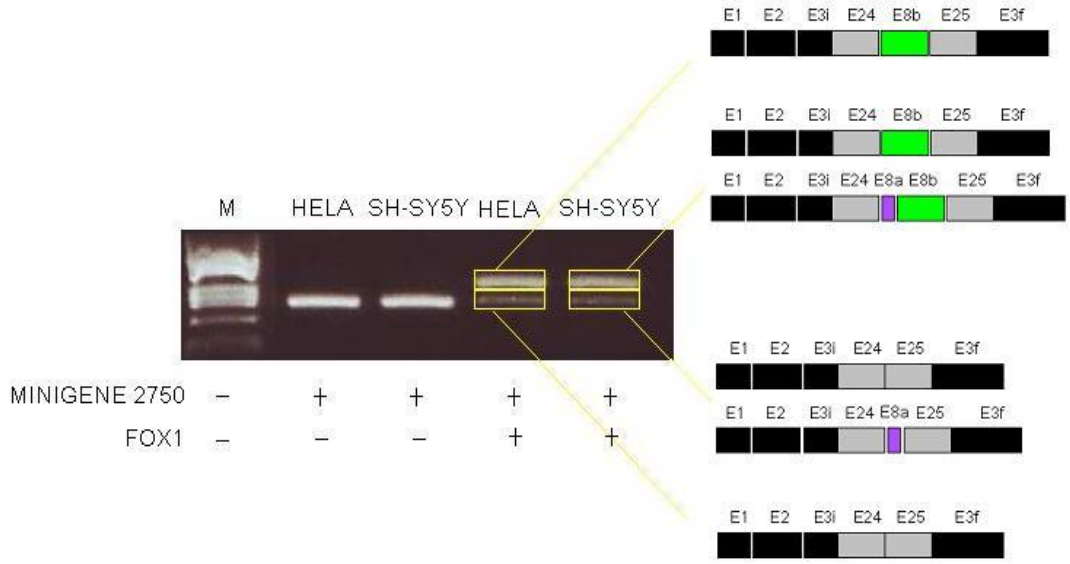


Fig.1.8.5 Effect of FOX1 over-expression on Minigene 2750 splicing products

Both in HELA and SH-SY5Y cells, FOX1 over-expression together with Minigene 2750 reporter plasmid induced the generation of different splicing products which co-migrated in two distinct bands on the agarose gel. Even if the visible outcome was the same in neuronal and non-neuronal cells, by running PCR products on the capillary gel electrophoresis we found which were exactly the splicing products generated in each cell line. The scheme on the right recapitulates the capillary gel electrophoresis output shown in the picture below.

Without over-expressing FOX1, in HELA cells only one amplicon was detectable (246-bp) deriving from the constitutive Minigene 2750 splicing, while in SH-SY5Y cells in addition to this we found also the 258-bp product that proved the basal

inclusion of the alternative exon E8a. These 2 amplicons, with a difference of only 12-bp one from the other, co-migrated on the agarose gel in a unique band (Fig. 1.8.5).

Overexpressing FOX1, we observed the “exonization” of the 77-bp LSD1 intronic fragment that gave rise to the exon E8b. Thus, in HELA cells capillary gel electrophoresis gave two peaks: one coming from the constitutive Minigene 2750 splicing (246-bp) and another longer one (323 bp) including the new exon E8b. In SH-SY5Y cells we even found 4 different peaks because the 77-bp new exon was included both in the constitutive transcript and in exon E8a including one (Fig. 1.8.6). All the amplicons including the 77-bp exon E8b co-migrated in the upper bands of the gel (Fig. 1.8.5).

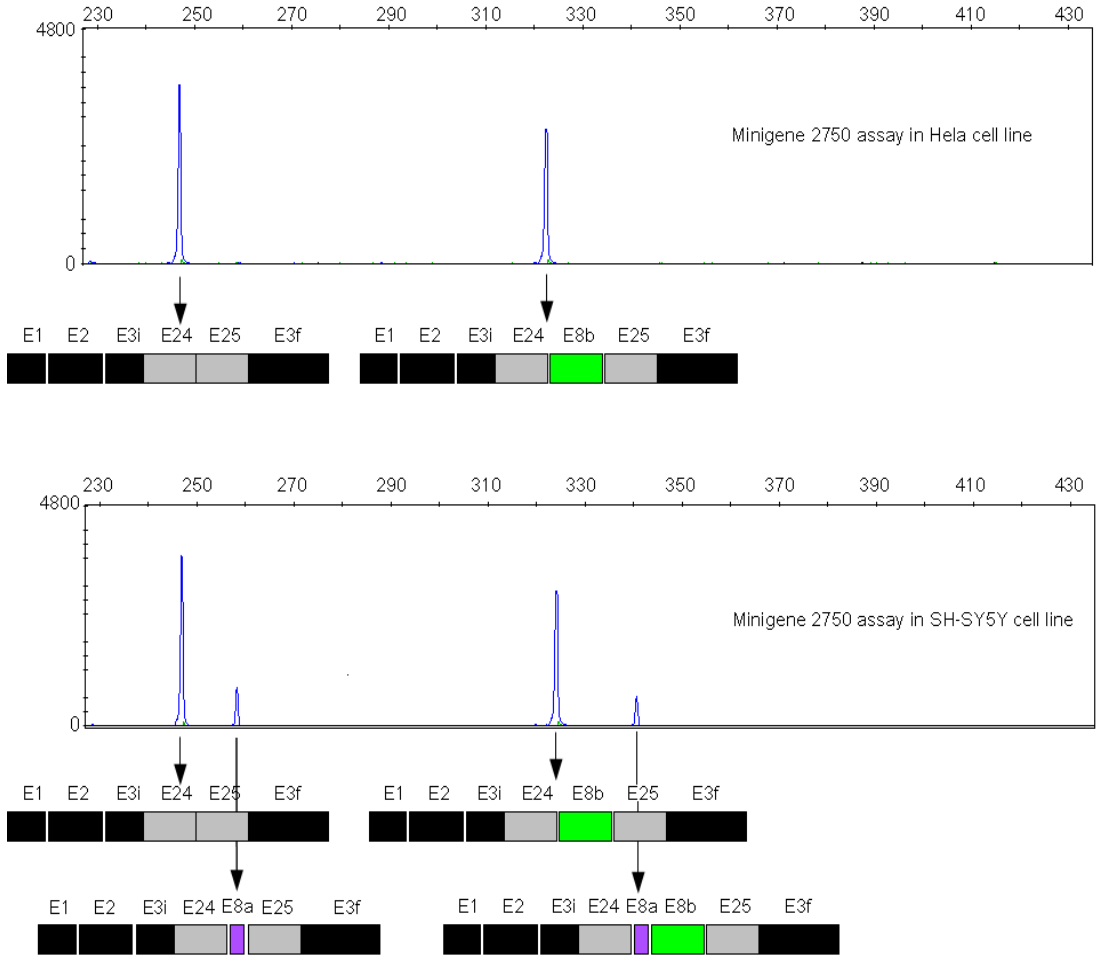


Fig. 1.8.6 Capillary gel electrophoresis output in a condition of FOX1 over-expression

In HELA cells FOX1 overexpression together with Minigene 2750 induced the inclusion of only the exon E8b into the mature transcript. Thus in this cell line we can observe two peaks: one deriving from the canonical splicing event (246-bp) and one belonging to E8b-including transcript (323-bp). In SH-SY5Y cells, alternative splicing generated 4 different transcripts, deriving from single or simultaneous inclusion of the exon E8a or E8b. Black and grey boxes represent Minigene exons; in purple the exon E8a; in green the exon E8b.

Thus, over-expressing FOX1 protein, in SH-SY5Y cells we observed an inclusion of the exon E8b in 42,05% ± 1,95 of the mature LSD1 transcripts, while in HELA cells the percentage was 35,85% ± 1,2 (Fig. 1.8.7; exon E8a and exon E8b in SH-SY5Y cells, 8,06 ± 1,55 and 42,05% ± 1,95; exon E8a and exon E8b in HELA cells, 0% and 35,85% ± 1,2).

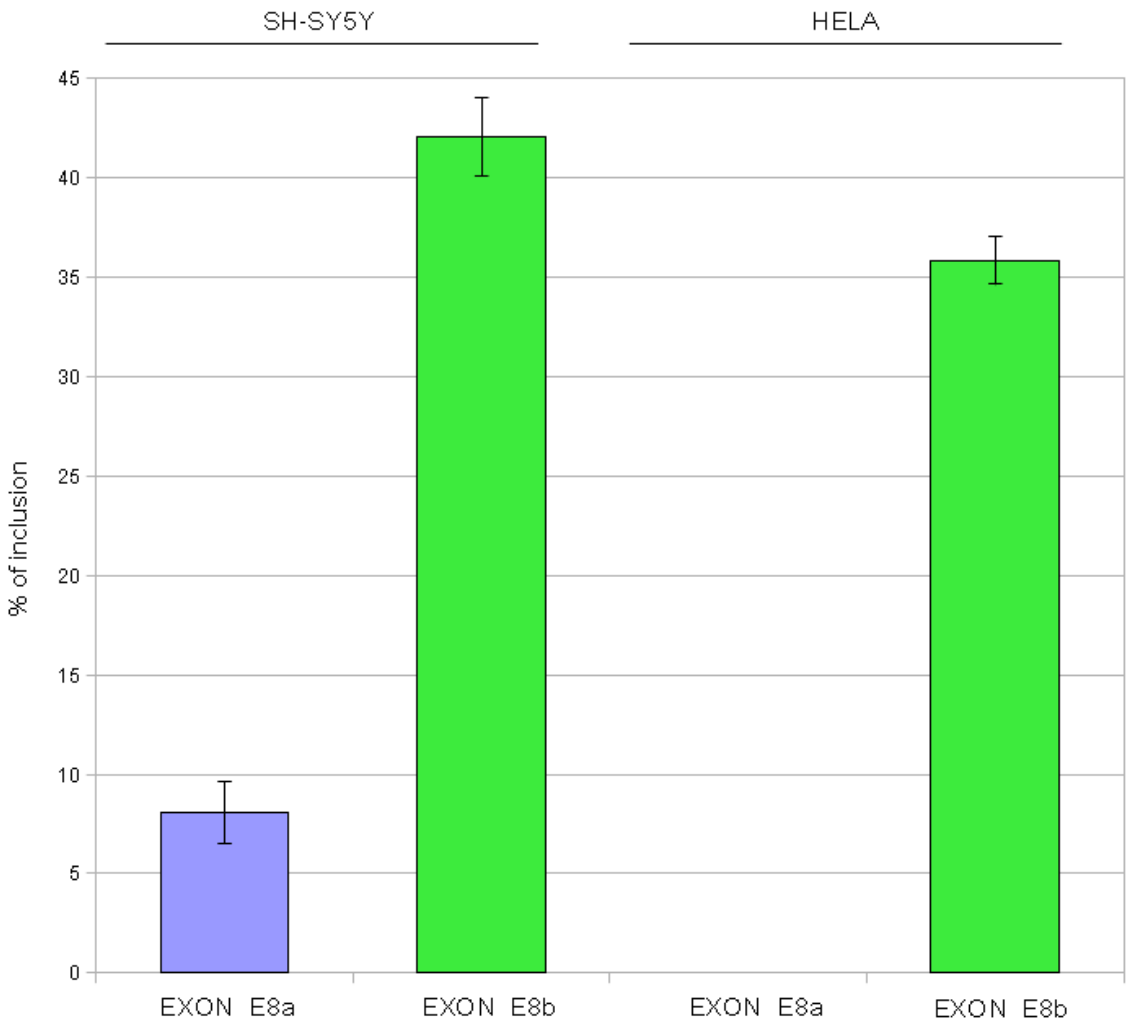
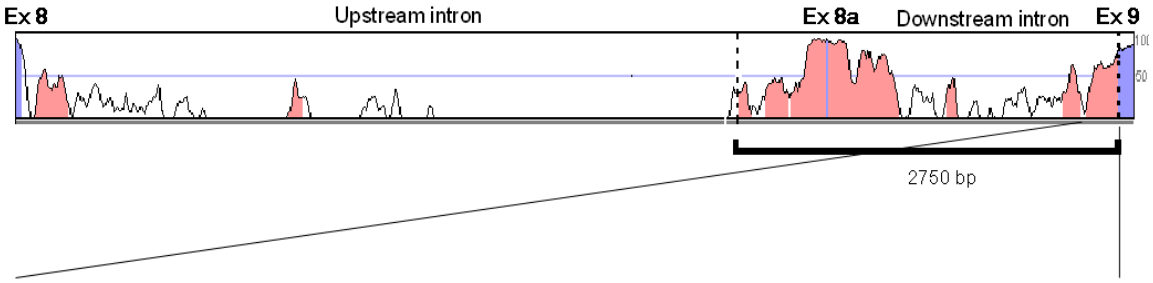


Fig. 1.8.7 Percentage of exon E8a and exon E8b inclusion in Minigene 2750 mature transcript upon FOX1 over-expression

Over-expressing FOX1 together with Minigene 2750 vector in SH-SY5Y and HELA cells, we found that exon E8a was included only in neuronal cells while FOX1 was sufficient to promote exon E8b alternative splicing both in neuronal and in non-neuronal cells.

In order to understand which was exactly the intronic portion included into the Minigene 2750 mature transcript, we cut the two bands from the agarose gel (Fig. 1.8.5) to perform a sequencing PCR reaction using forward and reverse FN1 primers. We identified the new 77-bp exon at the end of the exon E8a downstream intron, nearby the 2 FOX1 binding sites that regulate its inclusion (Fig.1.8.8).



GTCTTCCAGCTTGATGACCTGCTTTGCAATCTGTTTGCTTCATCCACATCAATGAGGAGCCAGACCTTCTGGTCCTTCATTTACTCATAG
ACTCTCTTAAGAGCTTTGTACATTCAATTCTGAACTTATTGGTGCCATAAAATATTATTACTGTGCAAAAGAATAAAAGTGCTAGTGTGTT
CTTATTAGTAGCCATGAATTTCATTTCTCTTTTCCCTCTTCCCACCTG**AG**ATTCATCTTTGAGGGGAAGCCAGATACCTTGGGCTTCTTA
TTTTCCCATCCTGCCCTTATTCTTGGAAGGTCCTCAG**GT**TTGTTGACATGGTTTTTGTTCTCTGGGACTTTTCTAGTGTGTTGTT**TGCAT**
GTCAGCCAGCTCAACTTAA**TGCATG**TCAGCTCAGTGCCACTTGTGCTCCATGTGACTTCAGTTAGCACCAGTGTAAATCTTGTGGCTGT
TAATAGAGGTGGAGTCCCAGCTCTGACGGCAGTAGTGGGAGAAGAAGAGGAAATTGTCAGGGCCCAAACATTTCATATCTCATTTAAAT
GGATCTGACCACTTCATAGAAAAGAAAATTTGAATCATTGTGCTCAAAGGTTTCCCCTTTGAGATACATATTTTGACCTTTATAAACTATA
GAGCATTTGTGTTTTCATGGATGGTTATGTTTAGTTTTATAAGGACCAACTCTGCTATACTTCGGATTTTCAG

5'.....ACCTG**AG**ATTCATCTT**TGA**GGGGAAGCCAGATACCTTGGGCTTCTT
ATTTTCCCATCCTGCCCTTATTCTTGGAAGGTCCTCAG**GT**TTGTTGACA.....3'

Fig. 1.8.8 Exon E8b nucleotide sequence

Sequencing reaction let us to define the 77-bp exon E8b nucleotide sequence, which is highlighted in green. Exon E8b is nearby the 2 FOX1 binding sites (in red), at the end of the exon E8a downstream intron. In bolt are indicated donor and acceptor splicing sites that flank the exon E8b. TGA is the premature STOP codon found inside the exon E8b.

Fig. 1.8.8 highlights two important features of the exon E8b: the first one is that donor (5' end of the downstream intron) and acceptor (3' end of the upstream intron) splicing sites were GT and AG respectively, that are exactly the almost invariant sequences that usually identify an exon. The second important feature was the identification of the premature STOP codon " TGA" at the 5' end of the alternative exon. This immediately led us to think about a degradation of all the exon E8b-containing transcripts by the surveillance cellular pathway of the Non-sense Mediated mRna Decay (NMD), that usually prevents the generation of truncated proteins.

However the identification of the exon E8b inside Minigene 2750 transcripts is not suffcient to prove its existence in the endogenous LSD1 mature RNA and to speculate about its function. Thus we performed a PCR reaction on endogenous cDNA in order to identify the exon E8b. We designed forward primer on the exon E8b and the reverse one on the exon 10 of LSD1, in order to clearly discriminate among possible PCR products that include introns in addition to the exons E8b, E9 and E10 (Fig. 1.8.9).

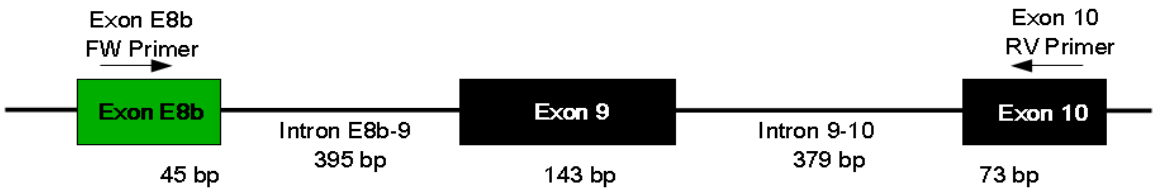


Fig. 1.8.9 PCR strategy used to search for endogenous exon E8b containing transcripts

We designed forward primer on the exon E8b and the reverse one on the exon 10. Base pair length of each exon and intron is reported. 45-bp is the distance of the forward primer from the end of the exon E8b (in green). 73-bp is the distance of the reverse primer from the beginning of the exon 10.

We performed PCR reactions for at least 50 cycles since the total amount of exon E8b-containing transcripts was supposed to be very low because of its degradation by NMD. We first used as reaction template cDNA obtained from Hela and SH-SY5Y cells, overexpressing FOX1 splicing factor. From this PCR we didn't find any endogenous exon E8b-containing amplicon. Also the treatment of the same FOX1 over-expressing cells with Puromycin (300 ug/ml), which blocks protein

synthesis and degradation, did not give any positive result (Fig.1.8.10 B). This means that neuronal and non-neuronal cell lines do not endogenously express exon E8b.

We first identified exon E8b-containing transcripts in human brain tissues (neuronal tissue of Rett and Epileptic human patients) (Fig. 1.8.10 A). PCR reaction on these cDNAs, performed after DNase treatment, gave two different amplicons. By sequencing them we found that the lower band on the agarose gel (261-bp) originated from the amplification of only the exons between forward and reverse primers, while the upper band included also the 395-bp intron between exon E8b and exon E9 (Fig. 1.8.10 A).

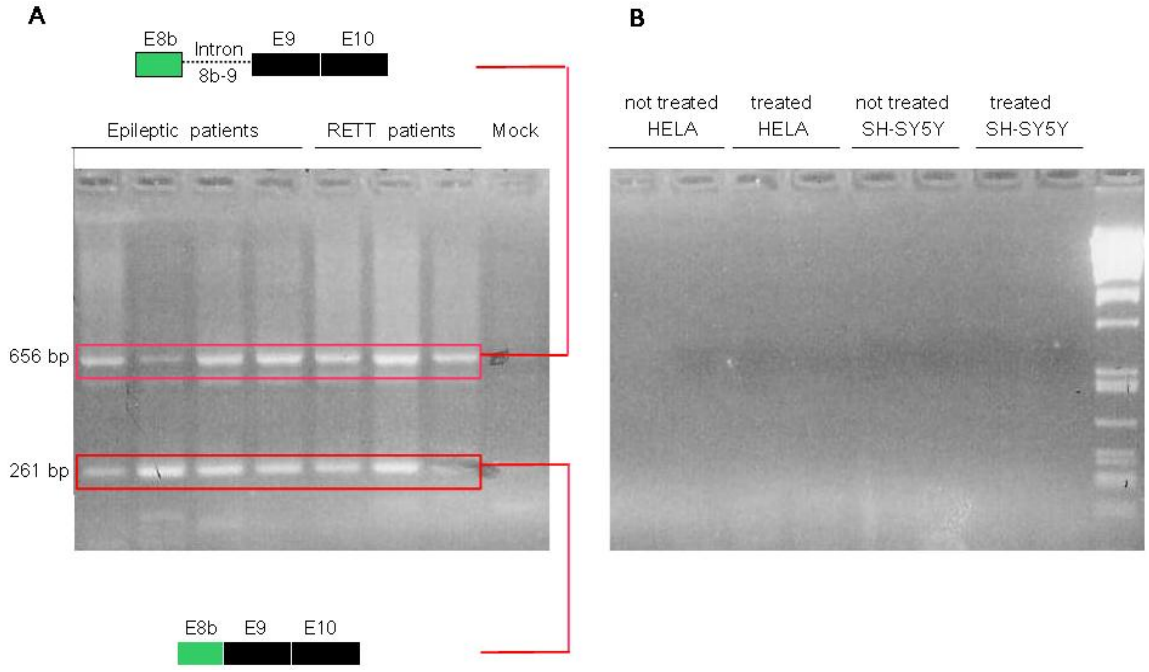


Fig. 1.8.10 Endogenous E8b containing transcripts

Fig. 1.8.10 A Two are the endogenous populations of exon E8b-containing transcripts in Epileptic and RETT human patients. One of them is represented by the 261-bp amplicon including only the exons E8b, E9 and E10; the other one is represented by the 656-bp PCR product coming from the inclusion of the intron E8b-9 in addition to the exons above. Fig. 1.8.10 B Exon E8b is not expressed in HELA and SH-SY5Y cell lines, treated or not with Puromycin.

This finding could not be the result of a DNA contamination because inside the LSD1 mature transcripts we always found the same E8b-9 intron. This means that in human brain there are not only LSD1 endogenous transcripts including the exon E8b but also mature transcripts with exon E8b and its downstream E8b-E9 intron. The function of these two new LSD1 splicing variants is not clear at the moment, since we know that the presence of the premature STOP codon inside the exon E8b causes RNA degradation before protein synthesis.

However we looked for their expression performing the same PCR on a multiple cDNA panel of human tissues (Ambion).

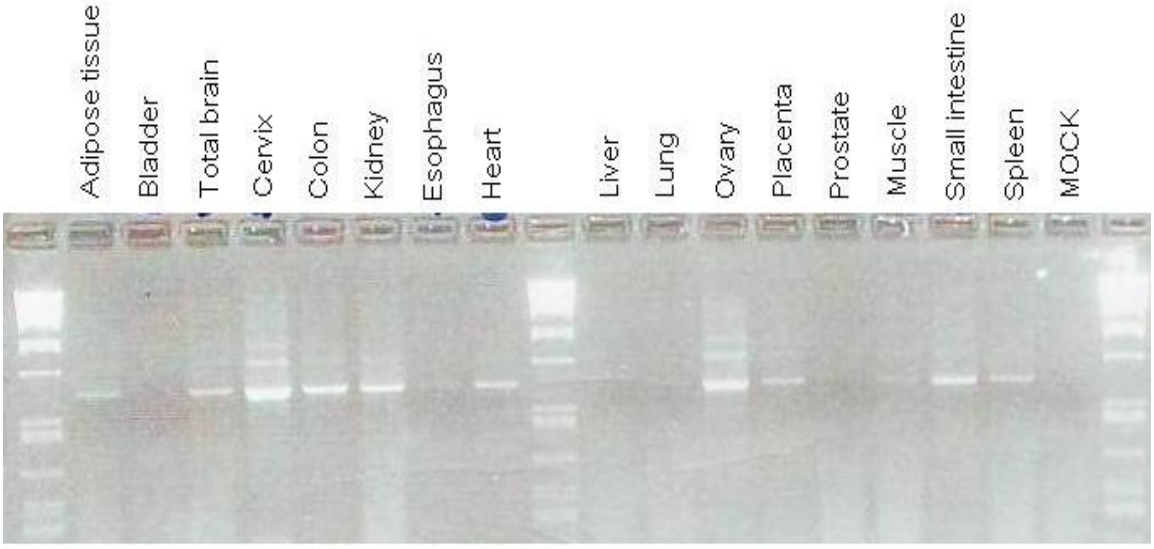


Fig. 1.8.11 Exon E8b-containing transcripts in multiple human tissues

PCR on human cDNA panel (Ambion), performed with forward primer on the exon E8b and reverse primer on the exon E10 of LSD1, shows that the only PCR amplicons detectable on the agarose gel correspond to the 656-bp LSD1 transcript including exon E8b and its downstream E8b-E9 intron.

From the PCR above (Fig. 1.8.11) we found that the endogenous expression of the exon E8b-containing transcripts is not restricted to the brain but occurs in different

human tissue with unrelated embryonic origin. Furthermore the bp-length of the only amplicon that we obtained corresponded to a mature transcript including the exon E8b but also the E8b-E9 intron. We did not find any PCR product deriving from the sole exons inclusion.

By UCSC genome browser we then performed a multiple alignment of the human exon E8b sequence with the same genomic portion of different mammals to find out its conservation degree. Fig. 1.8.12 shows only the 5' end of the exon E8b alignement and its intronic acceptor splicing site AG (inside the red square). The alignment clearly shows a very low level of conservation of the exon E8b among mammals, starting from the AG acceptor site which is not present in any other species with the exception of Rhesus. On the contrary the premature STOP codon “TGA” is conserved among mammals but most probably the absence of the acceptor splicing site causes the exon E8b not to be recognized by the cellular splicing machinery.

This finding is in agreement with the lack of any detectable exon E8b-including transcript in mouse brain tissues (data not shown).

Fig. 1.8.12 Multiple alignment of the human exon E8b

UCSC alignement of the 5' terminal end of the human exon E8b and its intronic acceptor site (red box) with the same genomic region of different mammals. Horizontal traits are the nucleotide gaps.

Taken together all these data show the existence of a new LSD1 alternative exon, that we called “exon E8b”, whose expression is restricted to human tissues. Such exon is not not neurospecifc, and its inclusion into LSD1 mature transcript is strongly positively regulated by FOX1, which is necessary and sufficient modulate exon E8b alternative splicing. The presence of a premature STOP codon inside the exon E8b causes the degradation of all the exon E8b-containing transcripts by the NMD, thus at the moment we can't say anything about its function but we think that it may be involved in regulating LSD1 expression level.

2. GENERATION OF EXON E8a-limited LSD1 KO MICE

To deeply investigate the biological implication of neurospecific LSD1 isoform i*n vivo,* in collaboration with Dr. F. Altruda and Dr. E. Turco of the “Centre for in vivo functional genomics” in Turin, we generated exon E8a-limited-LSD1 knock-out mice to abolish the sole neuronal LSD1 expression. Using homologous recombination in mouse ES cells, we replaced exon E8a of mouse nLSD1 with a *Neo* selection cassette, flanked by two LoxP sites. The picture below shows the overall strategy that we used to generate exon E8a KO mice.

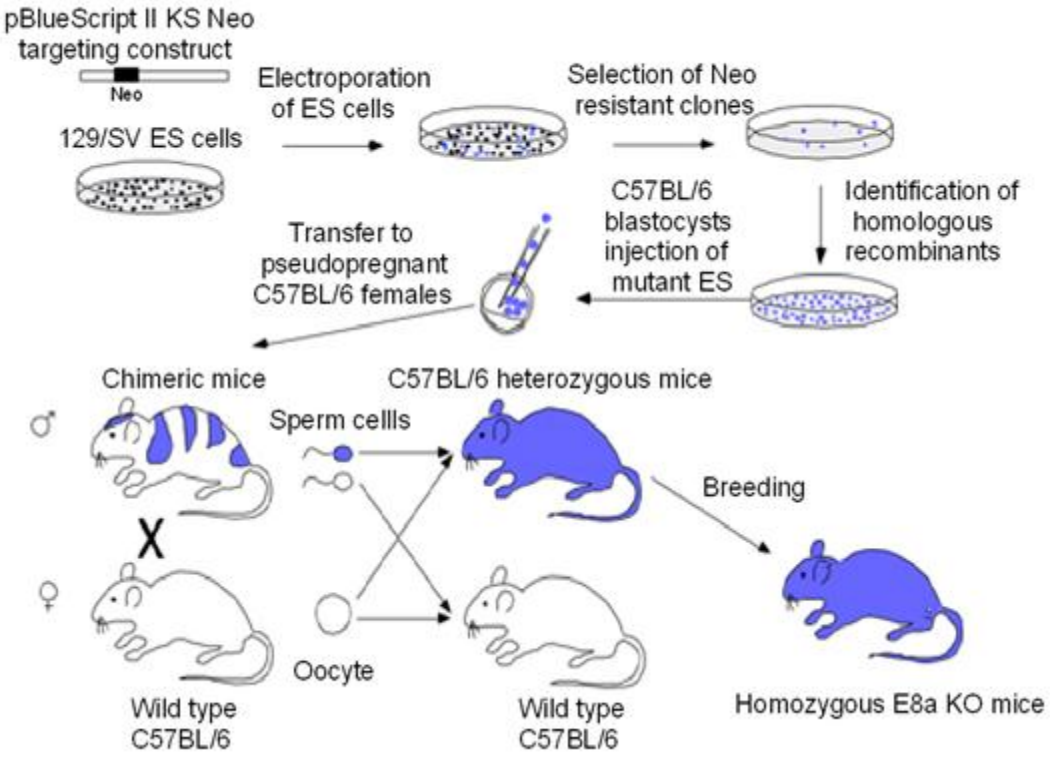


Fig. 2.0.1 Generation of LSD1-exon 8a Knockout mouse model.

Our genetic modification was inserted in a targeting vector, which carried Neo resistant selection marker and homologous regions. The homologous regions defined the specific region of genomic integration. Targeting vector was then electroporated into 129/SV mouse embryonic stem (ES) cells. Targeted ES cells were selected by antibiotics and expanded. ES cells were then injected into an early mouse embryo and the embryo implanted into the uterus of a C57BL/6 pseudopregnant mouse. Chimeric mice, which developed from the mixture of modified ES cells and cells of the host embryo, had the fur of two colours. These chimeric mice were mated with C57BL/6 wild type mice to obtain mice that derived from the modified ES cells. Such mice were hetorozygous for the deletion. By the breeding of male and female heterozygous animals we obtained homozygous LSD1-E8a KO mice.

2.1 GENERATION OF pBlueScript II KS TARGETTING VECTOR

A convenient method for subcloning the 20 Kb genomic fragment of LSD1 surrounding neurospecific exon E8a, from the BAC clone into pBlueScript II KS vector is represented by homologous recombination via a process known as *gap repair*. To generate our gap-repaired plasmid we made use of 2 long homology

arms that significantly increased the frequency of subcloning and decreased unwanted recombination products. To this aim we designed two couple of PCR primers (primers A and B, C and D) and we used them to amplify two 1000 bp regions of the LSD1-8a BAC clone, flanking the 20 Kb fragment that have to be subcloned. The PCR products were purified using spin columns and digested with *BamHI-EcoR*I and *EcoR*I-*Sal*I restriction enzymes. Restriction sites for these enzymes were included in the amplification primers to allow directional cloning of the PCR products into pBluescript II KS vector. The digested fragments were then ligated to a *BamHI-* and *SalI*-cut pBlueScript II KS vector cointaining a *TK* (*MC1TK*) gene for negative selection in ES cells. The retrieval pBlueScript vector was subsequently linearized with *EcoR*I to create a double-strand break for gap repairing. When we electroporated linear pBlueScript II KS plasmid into electro-competent cells, which contained *LSD1-8a* BAC clone, we found that we were able to generate several Amp^r colonies in a single electroporation experiment. Some of these Amp^r colonies were background colonies (derived from self-ligation of the linearized gap repair plasmid or from uncut DNA) but the others contained gap-repaired plasmids with the expected genomic inserts (Fig. 2.1.1).

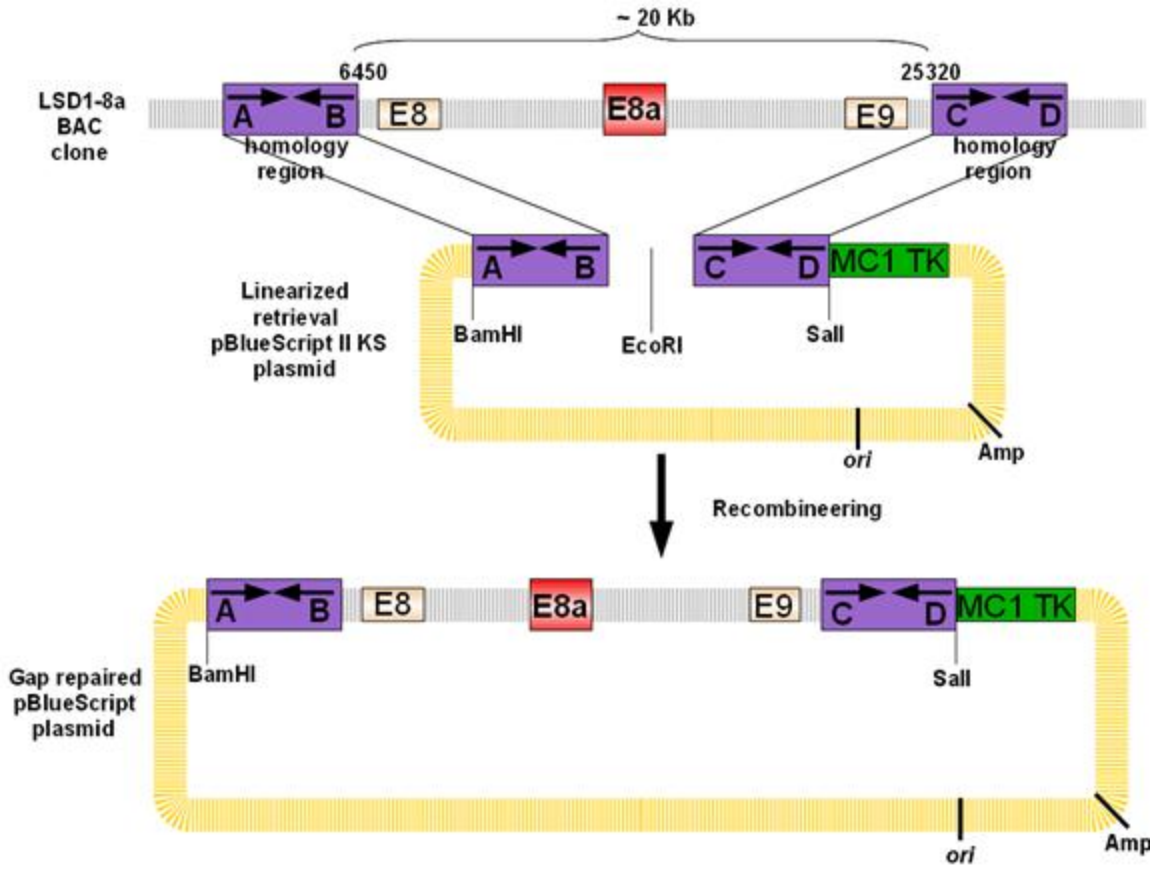


Fig. 2.1.1 Subcloning a DNA fragment from a BAC clone into pBluescript II KS vector by using gap repair with short homology arms.

The two homology arms (purple) used for subcloning have been PCR-amplified from LSD1 BAC clone DNA. They were obtained using primers A and B, and primers C and D. Subsequently they were cloned into an MC1TK-containing plasmid to generate retrieval plasmid. The retrieval plasmid was linearized with EcoRI to create a DNA double-strand break for gap repair. The linearized pBluescript vector containing the two homology arms was then transformed into recombination-competent cells that carried the LSD1 BAC clone. Gap-repaired plasmids were selected by their Ampicillin resistance.

2.2 TARGETTING Neo CASSETTE AND LoxP SITES INTO pBlueScript II KS PLASMID

The next step in generating exon E8a knock-out targeting construct was the introduction of *NeoR* gene flanked by 2 *lox*P sites into the subcloned pBlueScript II KS plasmid in order to replace the exon E8a. This was accomplished by introducing

a floxed Neomycin resistance (*Neo*) cassette (PL452) via homologous recombination. The floxed *Neo* gene in PL452 was expressed by a hybrid PGK–EM7 promoter that allowed efficient *Neo* expression both in mammalian cells (PGK promoter) and in bacterial cells (EM7 promoter). To introduce the floxed *Neo cassette* at the correct location, we generated two 1Kb arms homologous to the targeting site. These homology arms, as described above, were generated by PCR amplification of E8a flanking regions, inside upstream and downstream introns, using LSD1-8a BAC clone as template. PCR primers were engineered to contain *NotI* and *EcoRI* (primers E and F) or *BamHI* and *SalI* (primers G and H) restriction sites for the directional cloning of the homology arms. Following PCR amplification, the products were purified, restriction digested, and ligated to the floxed *Neo* cassette, excised from PL452 with *Eco*RI and *Bam*HI, and to a pEGFP vector linearized by *Not*I and *Sal*I. After the purification of the *Neo* cassette together with flanking homology arms, obtained by *Not*I and *Sal*I digestion of the pEGFP plasmid, we co-electroporated the cassette along with the gap-repaired subcloned pBlueScript II KS plasmid into EL350 cells. Transformants were selected by their kanamycin resistance, conferred in prokaryotes by *Neo gene (Fig. 2.2.1).* The obtained exon E8a knock-out targeting vector was subsequently linearized and electroporated into mouse 129/SV ES cells. Transformants were selected for their G418 and ganciclovir (Ganc) resistance and picked for Southern analysis.

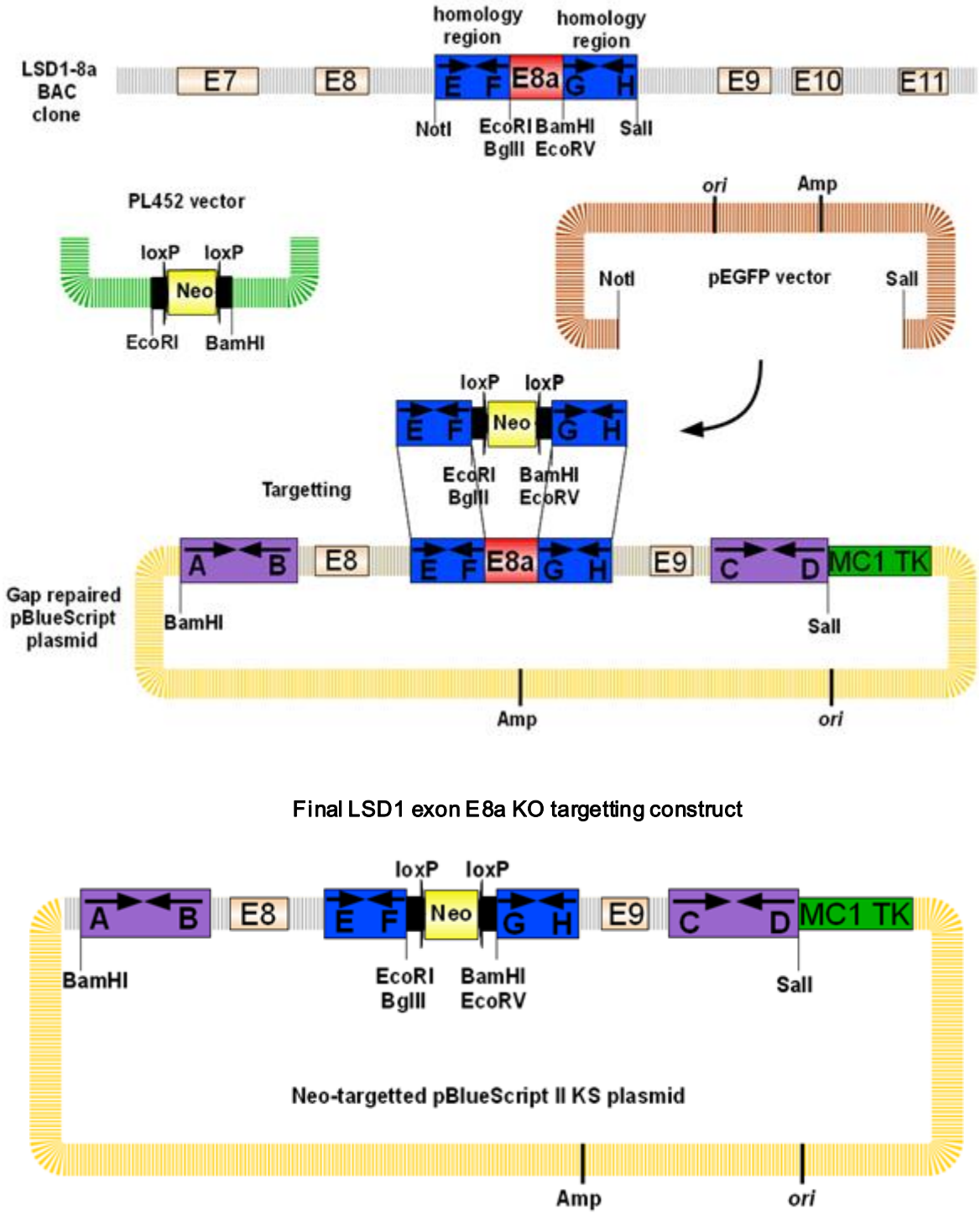


Fig. 2.2.1 An improved procedure for constructing LSD1 exon E8a KO-targeting vector

pEGF targeting vector was constructed by ligating together the two PCR products generated by amplification of LSD1 BAC DNA with primers E-F and G-H (blue), and the floxed Neo selection cassette (black arrow: loxP site) deriving from PL452 vector. The targeting cassette was excised by NotI and SalI digestion. The gap-repaired pBlueScript plasmid and the excised targeting cassette were co-transformed into recombination-competent EL350 cells.

2.3 BREEDING SCHEME FOR ES CELL CHIMERAS

From “Centre for *in vivo* functional genomics” we obtained 4 male and 1 female chimeras.

Most of 129SV ES cells were male derived and so the sex ratio of chimeras was skewed in favour of males. Furthermore female chimeras do not generally transmit the ES cell genotype because XY ES cells rarely undergo meiosis in ovary, therefore it is not worth test-breeding them. Despite this we bred all the 5 mice with wild type C57BL/6 animals. The goal of test-breeding ES cell chimeras was the identification of mice that were able to transmit the ES cell genotype through their germ line. This was assessed by the coat colour of their pups. Indeed, ES cells derived from 129/SV substrain, produced agouti fur. When these ES cells are microinjected into C57BL/6 (black fur) blastocysts the resulting mice will have patches of agouti (brown) and patches of black fur. However it is not unusual to get chimeras that are 90%-100% agouti, because it seems that 129 cells may grow faster than C57BL/6 cells in the developing mouse embryo. Secondly, the agouti pigment is a secreted protein, so the brown coat color may spread to cover areas that are derived from black cells. ES cell-mouse chimeras with high coat color contribution from the ES cells are likely to transmit through the germline more quickly than low coat color contribution chimeras. Thus, among the offspring all

pure black mice (non-agouti) were from the C57BL/6 component of the chimera and all yellowish-brown one (agouti) were from the ES cell component (Fig.2.3.1).

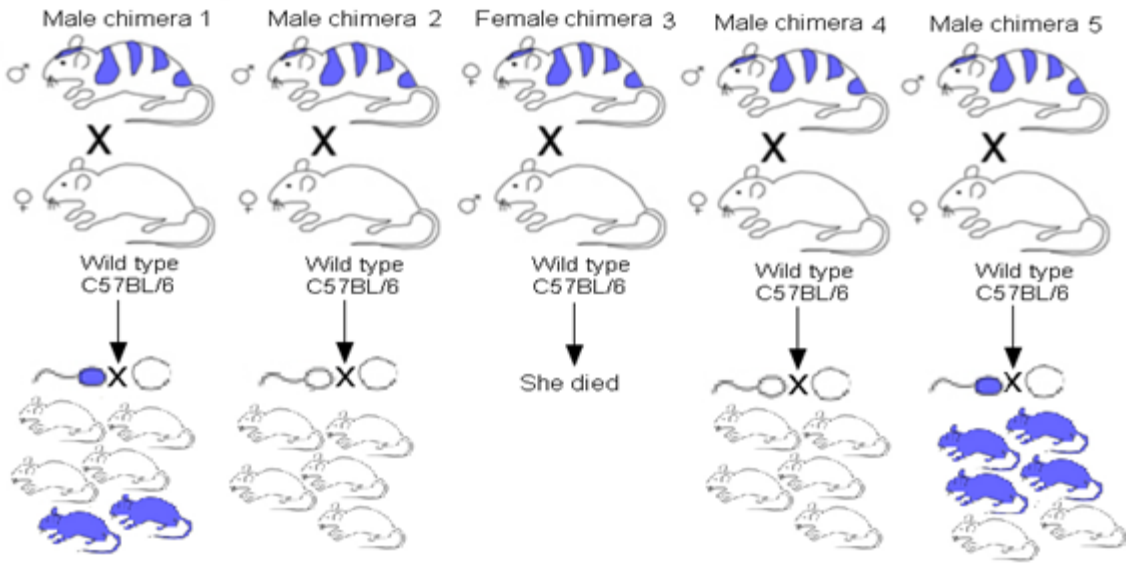


Fig. 2.3.1 Chimeric mice breeding.

We bred 4 male and 1 female chimeras (blue and white animals) with C57BL/6 willd type mice (white animals) but only male chimera 1 and 5 gave rise agouti pups (blue in the picture), since their 129/SV mutated ES cells reached their reproductive apparatus. Here, during meiosis their gametes received or mutated or wild type allele, for this reason agouti pups had also to be genotyped.

Female chimera died before giving pups so we could verify the germline trasmission of the four male chimeras only. Among them, two animals did not transmitted the ES cell genotype and gave rise to a completely black offspring (chimera 2 and 4), on the contrary the other two chimeras (1 and 5) gave together 82 brown (agouti) pups, because the 129 ES cells have contributed to the chimera's germline. All 82 brown pups were then genotyped to find out which ones received the targeted allele: since in ES cells only one of the two alleles were mutated in the

exon E8a locus of LSD1, only about 50% of the agouti mice should have inherited a mutated allele and become heterozygous mice. If we exclude from the count the animals that died before being genotyped, in the first-generation offspring (F1) we found 14 wild type and 10 eterozygous mice deriving from chimera 1, and 27 wild type and 23 eterozygous mice from chimera 5. (Fig.2.3.2)

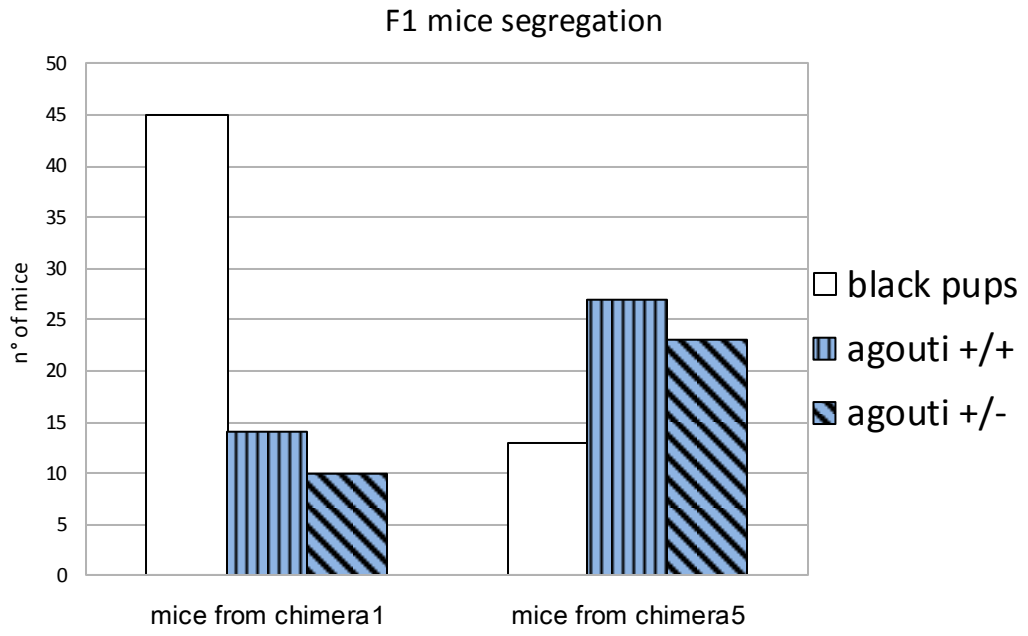


Fig. 2.3.2 F1 mice segregation

White bars represent black pups from chimera 1 (45 pups) and 5 (13 pups) that were not genotyped because of their fur color. On the contrary, blue bars depict all the 82 agouti pups, that were genotyped and classified in wild type (vertical stripes) and eterozygous (diagonal stripes).

In order to generate exon E8a Knock-out mice we mated 11 heterozygous F_1 males with 10 heterozygous F_1 females and we obtained 292 F_2 pups. 63 of them were knock-out, so the ratio was in accordance with the Mendelian low of segregation.

2.4 MICE GENOTYPING

To identify homologous recombination events between the exon E8a of LSD1 and the Neomycin cassette, we first performed a PCR analysis of genomic DNA taken from finger biopsies. Results were then confirmed by RT-PCR on a representative litter of adult mice cerebellum (C) and cortex (X), tested for the sole exon E8a expression. We also performed an isoform relative quantification of all the four LSD1 splicing variants in representative adult mice's cerebellum (CB) using the Relative Quantity Fluorescent-PCR (Fig. 2.4.1).

The targeting vector used to generate E8a-limited-LSD1 knockout mice contained a 800bp Neo-resistance cassette in place of the exon E8a (Fig. 2.4.1 A and B). With the purpose of establish mice genotype by PCR, we designed a set of oligos that let us discriminate between wild type, heterozygous and knock-out animals in the same reaction tube (Fig. 2.4.1 C).

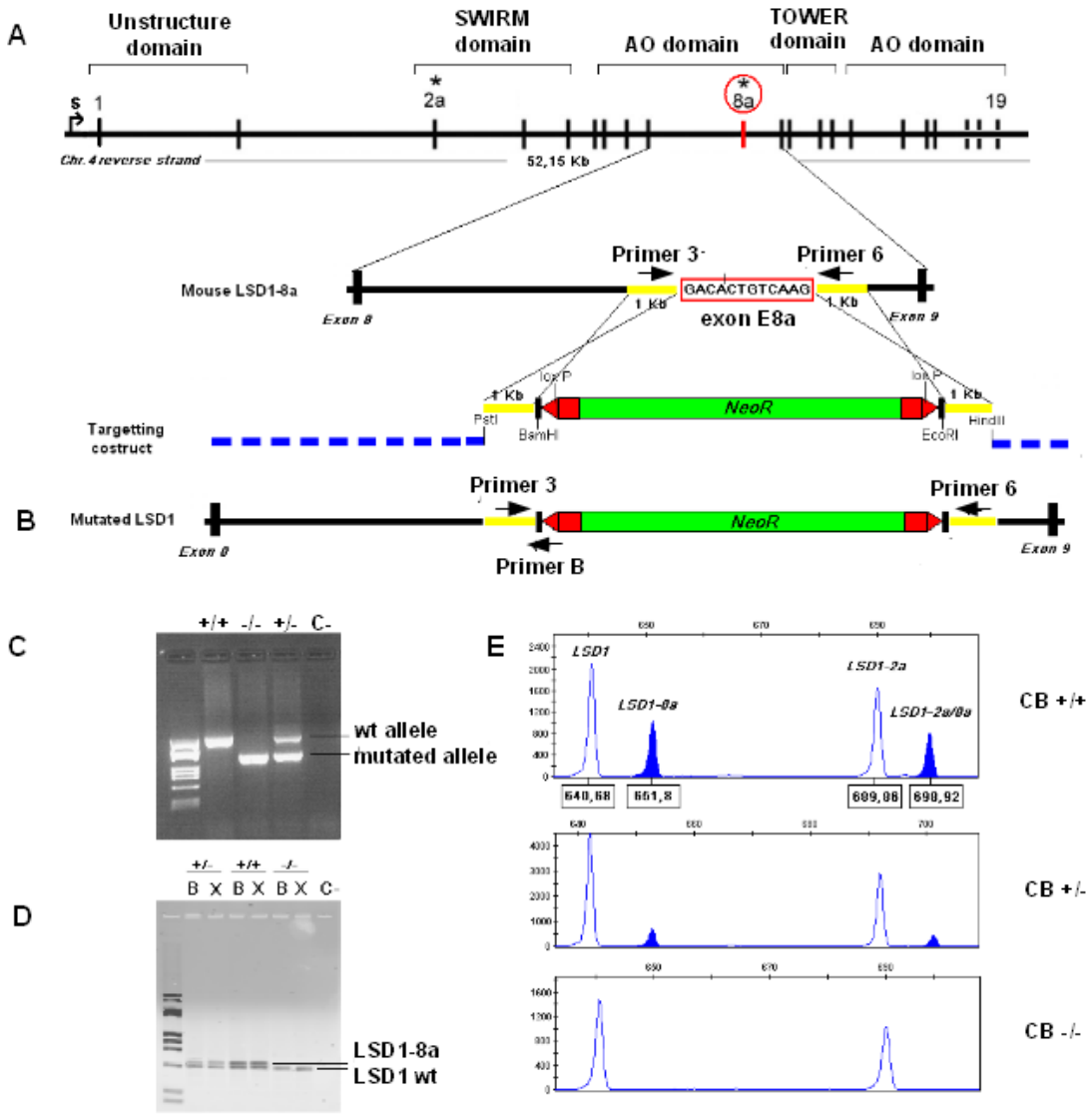


Fig. 2.4.1 Strategy for specific inactivation of the exon E8a and its detection

(A) Schematic representation of the mouse LSD1-8a gene and of the gene-targetting with Neo cassette. S indicates the transcriptional initiation site for LSD1-8a mRNA. Exons 8 and 9 are two coding exons of wild type LSD1 gene, whereas E8a is the alternative neurospecific exon. Yellow regions indicate 1 Kb-long homologous sequences used for construct the targeting vector. Primers 3, 6 and B are the oligos for routine genotyping of mice. LoxP sites are two 13 bp-long inverted repeats required for Cre-mediated recombination. (B) Structure of the mouse LSD1-8a locus in the mutated allele. Neo-cassette, flanked by 2 loxP sites, substitutes only the exon E8a. (C) PCR performed with primers 3,6 and B on a representative litter of LSD1 mice, detecting wild type and mutated alleles. (D) A representative litter of adult mice's cerebellum (C) and cortex (X), tested for Exon 8a expression by RT-PCR.The cDNA were amplified with primers including only E8a. (E) Isoform-relative quantification of LSD1 splicing variants in representative adult mice's cerebellum (Cb). rqf-PCR was performed on cDNA obtained from total RNA of the indicated samples. Amplicons were quantified by related fluorescence units (RFU) by GeneMapper software. E8a-including LSD1 isoforms are shown in blue.

P3 and P6 primers hybridized inside the 5' and 3' homology regions (yellow in the picture) flanking the Neo cassette, while PB primer sat exactly on the junction between 5' loxP site (red region) and the 5' homology region. Wild type DNA gave only one amplicon of 579 bp, deriving uniquely from P3-P6 PCR that efficiently amplified this region when this was not distrupted by Neo cassette. Also knock-out mice exhibited only one PCR product of 380 bp but generated by P3-PB PCR, because the presence of the Neo cassette allowed primer PB to anneal but at the same time prevented the low functioning P3-P6 PCR. On the contrary, heterozygous mice showed both the two amplicons of 579 and 380 bp, because of the contemporary presence of a wild-type and a mutated allele.

We confirmed PCR results performing a semi-quantitative RT-PCR (Fig. 2.4.1 D) with a couple of primers annealing on the exon 8 (hmFAMex8 FW) and on the exon 9 of LSD1 (mLSD1ex9 RV). cDNA was obtained by retro-transcription of the same amount of RNA deriving from cerebullum and cortex of representative wild type, heterozygous and knock-out litters. Our purpose was to demostrate the absence of exon E8a-containing isoform in the tissues of knock-out mice. As expected, wild type mouse showed two amplicons, since in CNS alternative splicing generates both ubiquitous isoform and the exon E8a-containing one. On the contrary, knock-out animal gave only one amplicon corrisponding to the sole ubiquitous LSD1 isoform, because of the exon E8a replacing with the unspliceable Neo-cassette on both the alleles. As control we tested also an heterozygous mouse

that had about half the amount of the exon E8a containing amplicon compared to wild type one.

Since RT-PCR performed around the exon E8a gave us a clear but limited information, we further supported our evidences by rqf-PCR with a couple of fluorescinated primers, the forward one on exon 2 (mFAMex2 FW) and the second one on exon 9 (mLSD1ex9 RV), that amplified in the same reaction all the four splicing variants of LSD1. Thanks to the qrf-PCR we were able to evaluate each PCR amplicon as single peak on the electropherogram and to compare different peaks' height in each sample to have a full-scale picture of what was the amount of the alternative splicing products (Fig. 2.4.1 E). In figure 2.4.1E cerebellar amount of the 4 different LSD1 splicing isoforms is shown for each genotype and, as expected, WT mice presented all the four isoforms but KO mice were totally lacking in LSD1-8a as well as in LSD1-2a8a variants (blue picks). Heterozygous mice confirmed the positive exon E8a replacing with an intermediate amount of the above-mentioned transcripts.

Furthermore, by qRT-PCR performed with a couple of primers that amplified all the LSD1 transcripts we demonstrated that total LSD1 amount was the same in the three different mice genotypes (Fig. 2.4.2).

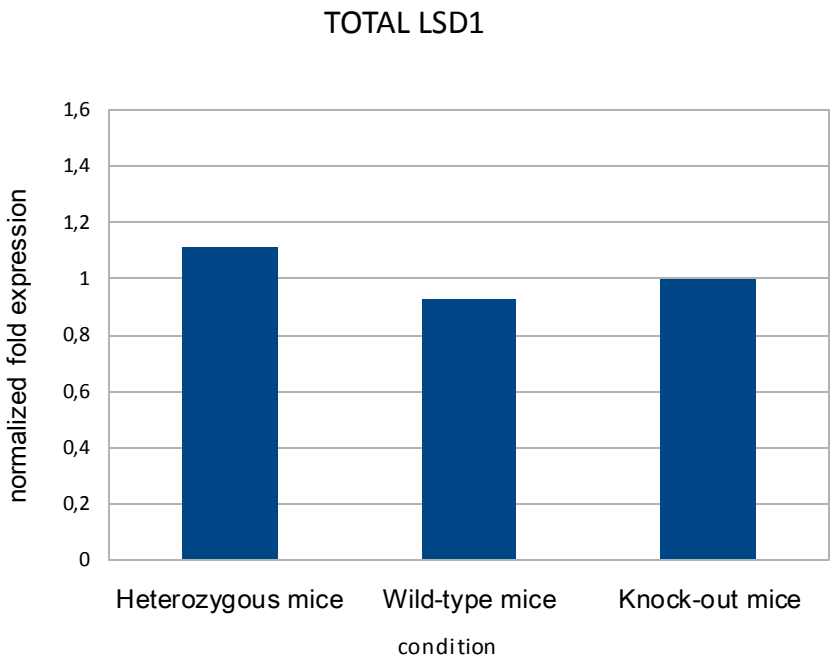


Fig.2.4.2 Level of total LSD1 expression in the three different mice genotypes.

qRT-PCRs were performed using B-actin as housekeeping gene. Ct values obtained for LSD1 were normalized over the housekeeping gene to correct for sample variations. A Student's t test (I Stat t I ≥ T α/2) was applied to LSD1 normalized expression level by comparing heterozygous and wild-type mice LSD1 expression to the knock-out one. *p<0.05; **p<0.01; ***p<0.001. Differences in the LSD1 level of expression are not statistically significative.

We then confirmed this evidence by immunostaining a whole brain section with an anti-panLSD1 antibody (policlonal anti-LSD1 (AB 17721)) and we found that also LSD1 protein level did not change in knock-out mice compared to wild-type ones (data not shown).

Together these results suggest that the homologous recombination between exon E8a and Neo-cassette succesfully occured in our mice and that only the exon E8a containing isoforms were turned off. Moreover the presence of the Neo-cassette did not affect total LSD1 transcription efficiency, without leading to unexpected effects due to low LSD1 protein dosage.

2.5 BACKCROSS BREEDING TO MAKE A CONGENIC STRAIN

The exon E8a-limited-LSD1 KO mice that we obtained by breeding heterozygous F1 mice inevitably had a mixed genetic background, deriving from 129 ES cells strain and C57BL/6 strain of the blastocysts that we used. This represents a drawback because the effects of neuronal LSD1 knocking down are subjected to variation in phenotype depending from the background genotype on which they are examined. For this reason we transfered the E8a knock-out allele into a fully C57BL/6 genotype, in order to recreat a congenic strain. Our aim, indeed, was to obtain E8a-limited-LSD1 KO mice genetically identical to C57BL/6 ones, except for the chromosomal region surrounding the mutated alleles. We chose C57BL/6 as inbred strain since these mice are widely used in research, available in large numbers and their genome has been completely sequenced, assembled and notated. To make our congenic strain we backcrossed heterozygous E8a-limited-LSD1 mutants with C57BL/6 inbred mice and we selected heterozygous offspring from the first generation (N_1) to cross back to the inbred strain for the next generation (N_2): at each generation (N) the genetic diversity decreased and the mice contained more and more of the inbred genome. A mutant strain is consider congenic when it reaches 99,8% of identity with the inbred genome. At this point we simply bred heterozygous male and female backcrossed mice to obtain congenic E8a-limited-LSD1 KO animals for analysis (Fig. 2.5.1).

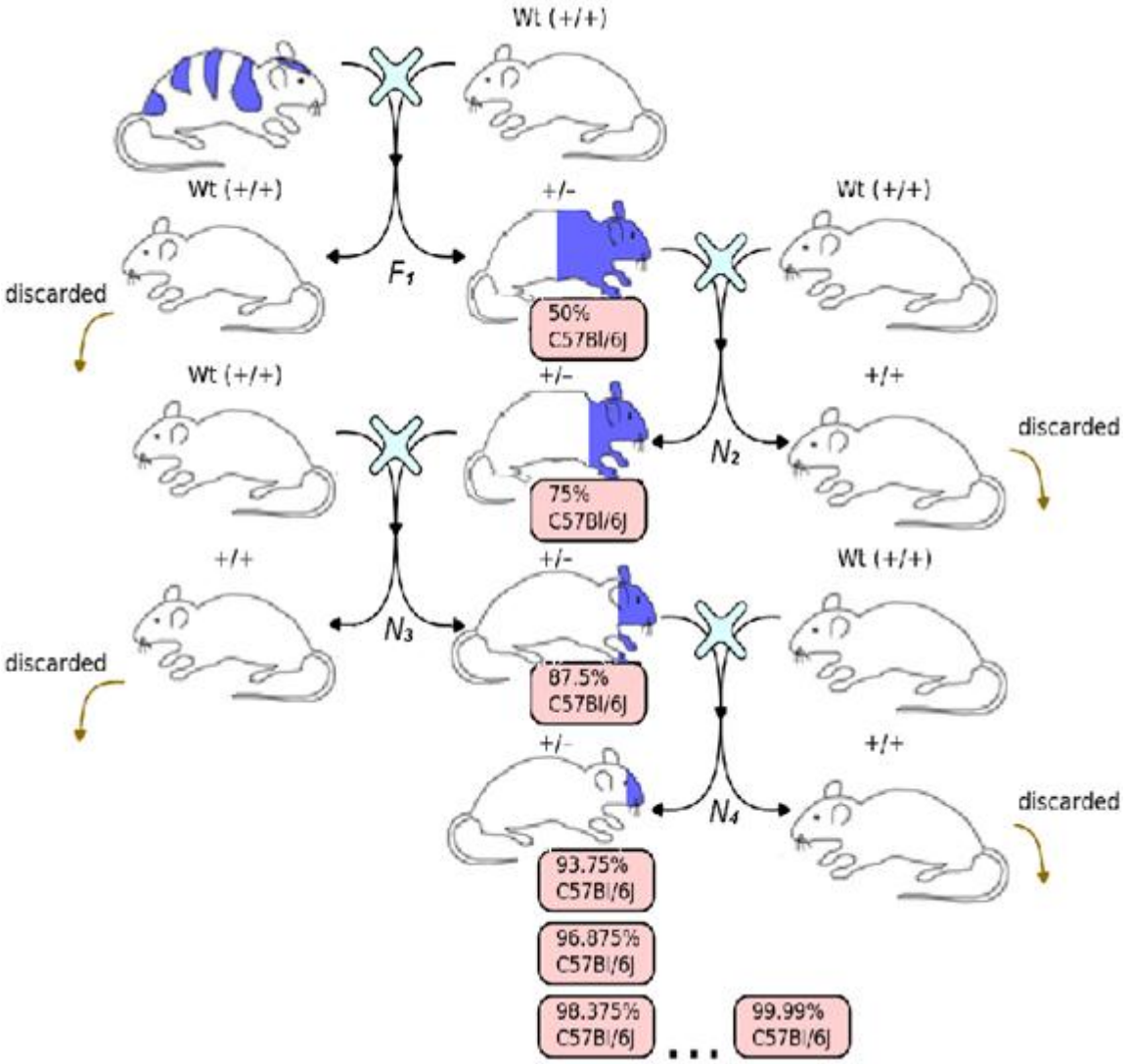


Fig. 2.5.1 Mice backcrossing

The picture shows the backcrossing of heterozygous mice, from a mixed genetic background to C57BL/6 congenic background. The exon E8a knocking-out was performed in 129/SV cells (blue) that were then implanted into the blastocyst of a C57BL/6 pseudopregnant female mouse (white). Thus, F1 heterozygous pups were C57BL/6 for 50% of their genome. To move our mutation in a congenic strain it was necessary to perform consecutive backcrossing with C57BL/6 wild type animals, in order to increase the percentage of C57BL/6 DNA that forms the genome of the offspring. When such percentage reached 98% we bred male and female heterozygous congenic pups to obtain backcrossed KO animals.

2.6 ACCELERATED BACKCROSS BREEDING

In order to speed up backcross breeding we used MAX-BAX (Marker-Assisted Accelerated Backcrossing) system. It consists in a 384 SNPs panel of C57BL/6 DNA markers (Charles River) that is able to dramatically reduce the time required to fully backcross a mouse line. This reduction is achieved by selecting animals with preferred genetic background as breeders for the next generation, resulting in a fully backcrossed animal in only five generations (Tab. 2.6.1).

Traditional Backcross		Speed Backcross	
Generation	Recipient Genome	Generation	Recipient Genome
F1	50.00%	F1	50.00%
N2	75.00%	N2	~80.00%
N3	87.50%	N3	~94.00%
N4	93.75%	N4	~99.00%
N5	96.88%	N5	~100.00%
N6	98.44%		
N7	99.22%		
N8	99.61%		
N9	99.81%		
N10	99.90%		

Tab. 2.6.1 Congenic strain generation strategies

Using traditional random backcrossing methods, it takes 10 generations (upwards of 2.5 years) to produce a congenic strain. Selectively breeding individuals containing more of the recipient genome from each generation allows accelerated congenic strain production. Indeed, the animals carrying the locus of interest with the highest percentage of recipient versus donor strain DNA are preferentially bred.

We performed MAX-BAX analysis on tail snips of 8 N_3-backcrossed heterozygous mice that showed an average recipient-strain's contribution percentage of 95,13%, with values ranging from 93,47% to 97,40% (Tab. 2.6.2), according to the mean

values reported in “Speed Backcross” table 2.6.1.

	N_3 mice analyzed							
	N3-1	N3-8	N3-10	N3-12	N3-13	N3-14	N3-15	N3-16
% of C57BL/6	94,01042	95,57292	97,39583	95,82245	94,66146	93,47258	94,14063	95,953

Tab. 2.6.2 MAX-BAX results

Table shows the percentage of C57BL/6 recipient genome present in each of the 8 mice analyzed. In yellow the 4 mice with a C57BL/6 percentage higher than 95% that we used to carry mice backcrossing on.

Thanks to these results we decided to continue accelerated backcrossing to N_5 generation using only the 4 mice with C57BL/6 contribution percentage higher than 95% (yellow in table 2.6.2). At the end of backcrossing we bred 6 heterozygous N_5 male with 13 N_5 female mice and so far we have obtained about 230 pups, with 1:1 ratio between male and female animals (Fig. 2.6.3 A). 29,5% of them were homozygous E8a-limited-LSD1 knock-out animals; the percentage is in according to the Mendelian frequency (Fig. 2.6.3 B). The great majority of them survived to adulthood and was fully fertile.

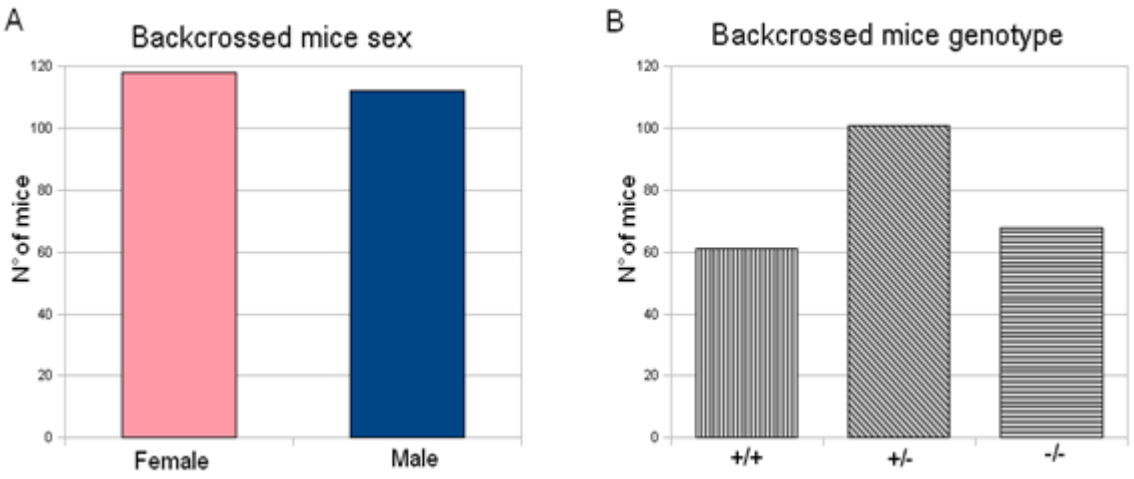


Fig. 2.6.3 Sex and genotype ratio in backcrossed mice

(A) Sex ratio 1:1 is respected in backcrossed pups, with 118 females (pink bar) and 112 male mice (blue bar). (B) Genotype segregation is accordance with Mendelian low, with 61 wildtype (vertical stripes), 101 heterozygous (diagonal stripes) and 68 knock-out mice (horizontal stripes)

3. CHARACTERIZATION OF E8a-limited LSD1 KO MICE

It's becoming increasingly clear that it is almost always impossible to predict the resultant phenotype deriving from a gene inactivation because it often differs from the reasearchers expectations. Indeed, our E8a-limited LSD1 knock-out KO mice do not have an obvious phenotype; on the contrary they rather seem to not have a peculiar phenotype, since they are not sicken, they do not die prenatally or right after birth, neither during the first days or weeks of life. The phenotype of a knock-out mouse is always a very critical aspect to point out. For this reason we approached this issue from different points of view. First of all we looked at the internal organs with a macroscopic and microscopic histophatological examinations of KO and WT animals, then we characterized them with brain-specific immunostainings and finally we evaluated their behavioural and biochemical response to a specific chemoconvulsant treatment.

3.1 NECROSCOPIC EXAMINATION

Mice examined:
2 WT female and 2 KO female mice
Developmental stage: 3 weeks (1 WT and 1 KO) and 3 months (1 WT and 1 KO)

To begin with knock-out and wild type mice characterization we decided carry out a complete necroscopic internal organs examination, in terms of size, shape, volume, reciprocal position contacts and histopathological evaluations. To this aim, in collaboration with Dr. Eugenio Scanziani of Mouse and Animal Pathology Lab. (Filarete foundation, Milan) we performed routine macroscopic analysis and microscopic histological studies of the complete mice, with major organs dissection, weight and fixation in paraffin, followed by a regular Hematoxylin and Eosin staining.

The following organs and tissues were fixed:

- representative samples from glossy detectable lesions
- CNS
- Thoracic cavity: heart, lungs, thymus, thyroid
- Parenchymatous organs of the abdominal cavity
- Hollows organs of the abdominal cavity
- Genital tract
- salivary glands and lymph nodes

The CNS, after a short fixation, was submitted to transversal sectioning by using a “mouse brain matrix”. Six slices of 2 mm of thickness were obtained. All samples were processed for paraffin blocks embedding. For histopathological analysis, 4 um-thick tissue sections were stained with Hematoxylin and Eosin.

Table below summarizes values obtained for the weight of the major internal organs of each mouse analyzed.

	weight (g)				
	brain	**liver**	**kidney (Dx)**	**kidney (Sx)**	**total body**
F 59	0,50	1,17	0,16	0,17	20,85
F 65	0,57	1,34	0,20	0,19	27,90
F 156	0,46	0,62	0,07	0,07	12,34
F 160	0,43	0,49	0,07	0,07	14,20

Tab. 3.1.1 Summary of the organs weight

No statistically significant differences existed in the total body weight of the age-matched knock-out and wild type mice as well as in the weight of all the isolated organs (Tab. 3.1.1). Macroscopic examination of the vital organs didn't reveal any abnormality. No relevant pathologies were found by necroscopy, except for a moderate mesenteric lymphadenomegaly in the 3 months-old WT animal. No inflammation, hemorrhage, fluid accumulation or any unfavourable changes were seen in the mutated animals tissue sections. Histopathologic analysis of the tissues showed that all the vital organs from the knock-out and control groups were normal in the main.

Medical report obtained from the necroscopic analysis is posted here:

Brain: the gross morphology of brain from *E8a* knockout animals do not differ from that of wild-type brain.

Heart, lungs, thymus, thyroid, and salivary glands: these tissues do not show any

anomalies.

Liver: rare and small foci of cellular of extramedullary hematopoiesis were noted in the liver of all the animals but the damage was considered very mild and non-specific, and could have been due to a great number of aetiological factors such as immunologic insult.

Kidneys: rare and minimal findings of necrosis of tubular epithelial cells was osserved in kidneys of 3 weeks old KO mouse. Small perivascular lympho/plasmacytic infiltrates with one tube atrophy of epitelial cells were noted in wild type mice.

Spleen, pancreas and lymph nodes: normal

Stomach: focal eosinophilic infiltrate was seen in lamina propria of almost all the mice

Small intestine: from moderate to severe, diffuse lympho/plasmacytic infiltrate in lamina propria with a high number of flagellate protozoa (most likely Giardia) in the intestinal lumen of all the mice, except for the 3 weeks-old knock-out mouse

Large intestine: from moderate to severe, diffuse lympho/plasmacytic infiltrate in lamina propria with a moderate number of pyriform protozoa (most likely Tritrichimonas) in all the mice. Large number of rod shape, elongated organisms within the lumen of intestinal crypts in 3 weeks-old mice.

Uterus: no pathological findings

Ovaries: some corpora lutea in 3 months old mice.

Taken together necroscopy and histopathological analysis showed no particular evidences in mutant mice compared to the controls, both in CNS and in the other organs/tissues. Only “spontaneous” pathological findings have been recorded and in this regard it was worth noting the intestinal parasites seen in all animals. In 3/4 cases protozoan flagellate were present and in 2/4 case pinworms, but these findings were unrelated from the exon E8a involvement.

3.2 IMMUNOHISTOCHEMICAL CHARACTERIZATION

Animals examined:
3 WT and 3 KO mice
Developmental stages: 6 weeks

Since we know from existing data of our laboratory [12] that LSD1-E8a expression is restricted to CNS but the gross brain morphology of *E8a* knock-out animals did not differ from that of wild-type ones, a closer inspection of distinct brain areas was necessary to look more carefully into possible alterations in the organization of mutant mice brain. Thus, in cooperation with prof. Silvia De Biasi of the Biomolecolar science and Biotechnology department (University of Milan) we performed immunohistochemical analysis using several neuronal markers.

We started from evaluating Calbindin-D28K (CB) and Calretinin (CR) distribution as

they are two calcium-binding proteins of great interest in neuroanatomy and neuropathology to characterize different subpopulations of inhibitory neurons in normal developing and adult nervous system. We immunostained cortex, hippocampus and cerebellum areas of 6 weeks-old knock-out mice and we immediately detected very deep differences compared to wild-type animals, in particular in Calbindin staining which was the most informative one. For all the markers and the brain areas considered, we obtained a remarkably consistent and reproducible labeling pattern in all the mice. In the following descriptions, indications regarding labeling intensities always referred to differences observed in the distribution of one marker between in the same section of mutant and control mice.

CALBINDIN IMMUNOSTAINING

Cerebellum

Wild-type mouse Knock-out mouse

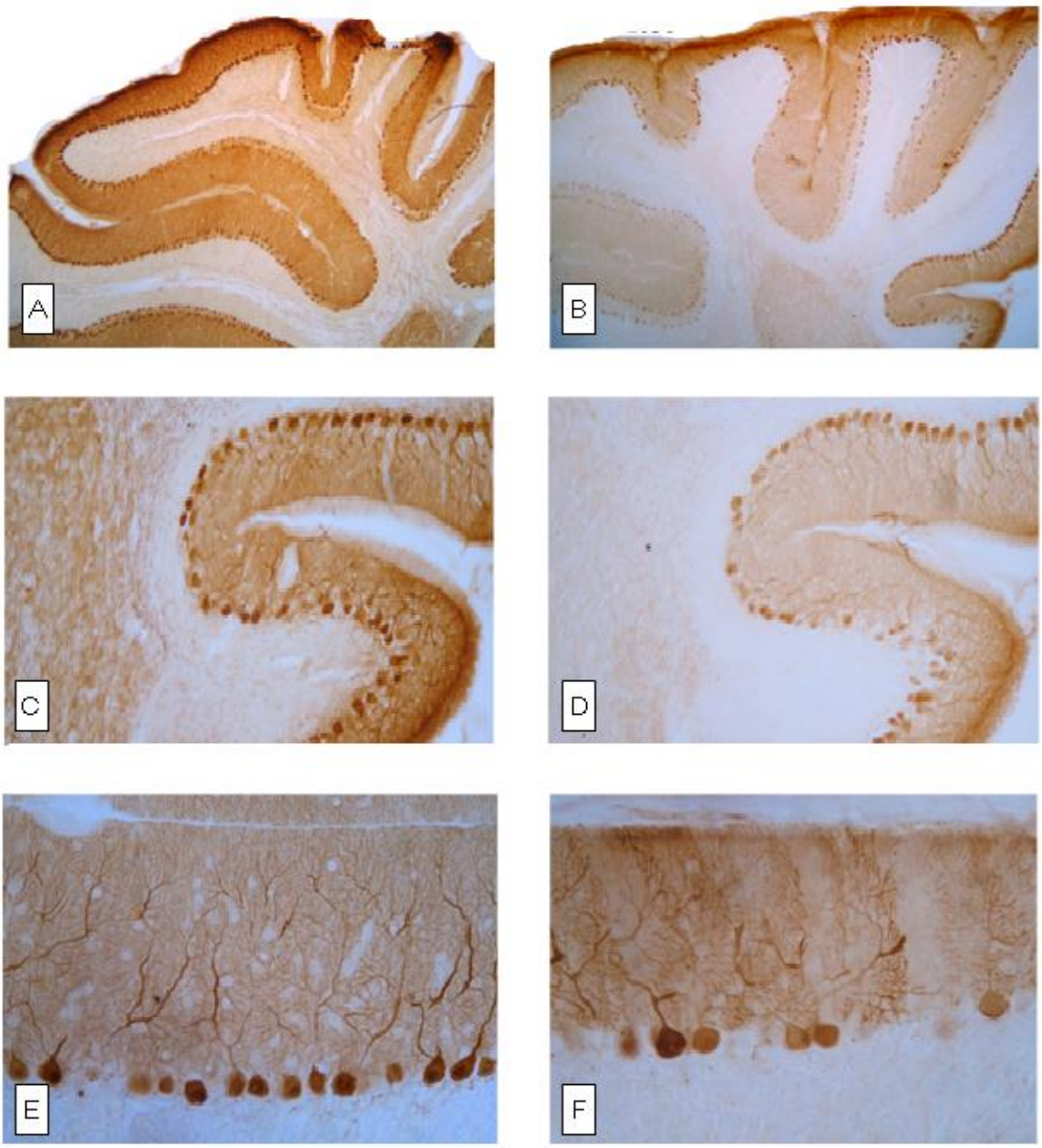


Fig. 3.2.1 Changes in calbindin D28k (CB) immunoreactivity in the cerebellum of E8a-limited-LSD1 knockout mice (B,D,F) compared with wild-type mice (A,C,E).

Parasaggital sections of the cerebellum are stained with antibodies against calbindin to detect Purkinje cells. Purkinje cells that normally constitute a monolayer (PCL) between the molecular layer and the granule cell layer (A,C,E) are highly disorganized in the nLSD1 -/- cerebellum (B,D,F)

In cerebellum, Purkinje cells express calbindin whereas basket, stellate and Golgi cells are not calbindin-D28K immunoreactive. These cells are, almost all, calretinin negative. On the contrary granule, Lugaro and unipolar brush cells present an opposite immunoreactive profile: they are all calretin positive but lacking in calbindin. The developmental pattern of apparence of these proteins seems to follow the maturation of neurons. Calbindin appears early, shortly after the cessation of mitosis when neurons become ready to start migration and differentiation.

Calbindin immunostaining revealed anatomy disruption of the Purkinje and molecular layers of the cerebellum in Knock-out mice (Fig. 3.2.1 B, D, F) compared to wild-type animals (Fig. 3.2.1 A, C, E). Such alteration was expressed as overall disorganization, changes in dendritic arborization, reduction in cell soma size and decrease in the numbers of Purkinje cells (Fig. 3.2.1 E-F). Furthermore we also observed a very clear defects in the misalignment of the Purkinje cells as well as a drastic reduction of the molecular layer staining in nLSD1 -/- mice (Fig. 3.2.1 C-D).

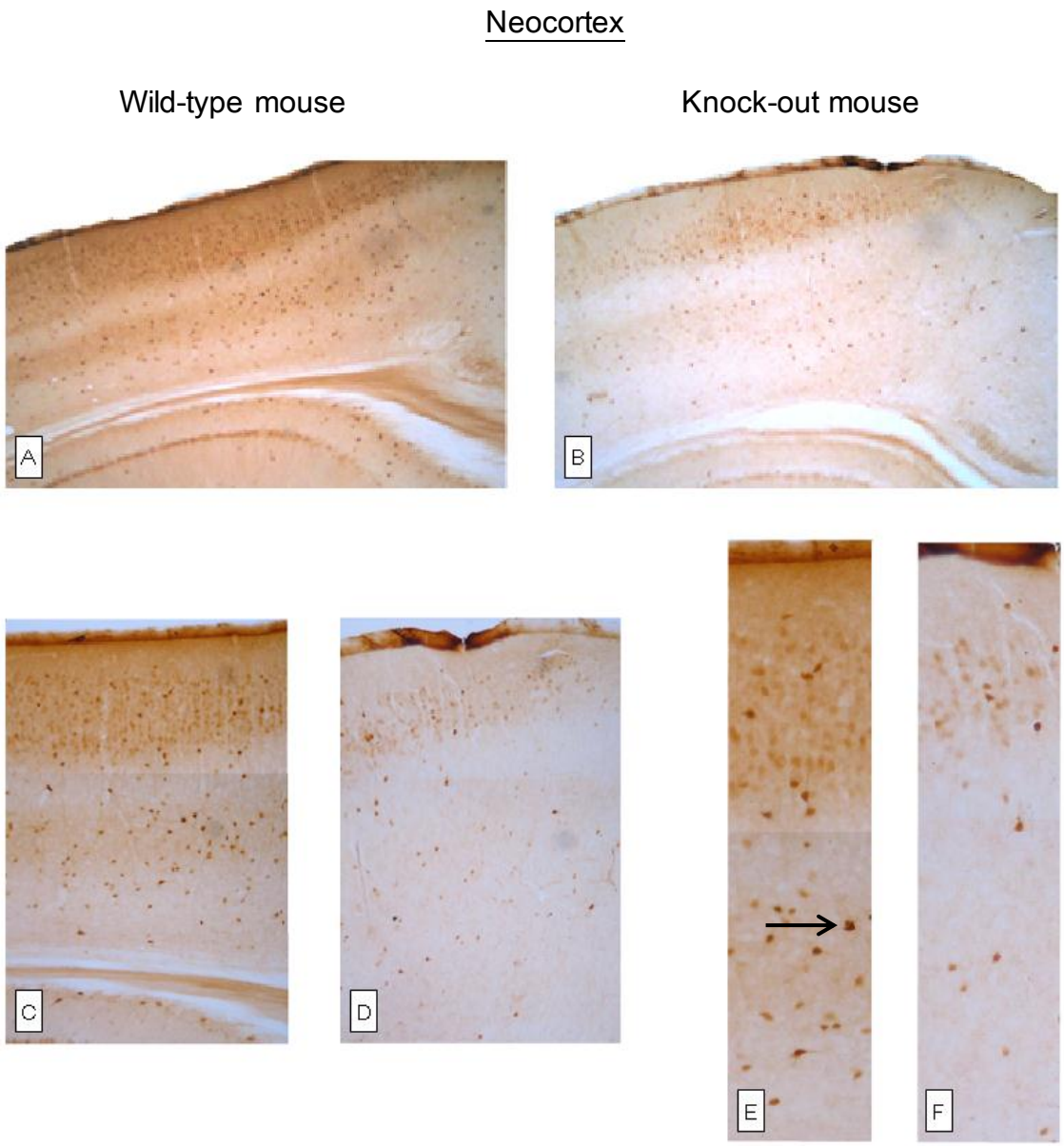


Fig. 3.2.2 Changes in calbindin D28k (CB) immunoreactivity in the cerebral cortex of E8a-limited-LSD1 knockout mice (B,D,F) compared with wild-type mice (A,C,E).

(C, D) are high power views of all layers of the cerebral cortex while (E, F) are high power views of layers I to V. The distribution pattern and morphology of CB-immunoreactive neurons are different in the control and nLSD1 knock-out mice. CB immunoreactivity in nLSD1 -/- mice is much lower than that of control mice (A, B) and the high power views show that the number of the CB-immunoreactive neurons in the nLSD1 -/- mice (D,F) is lower than those of the control mice (C,E).

In the neocortex there are basically two major classes of neurons: projecting neurons (~80% of neocortical neurons) and interneurons (~20%). Projecting neurons are represented by pyramidal cells which are located in all layers except for layer I and are excitatory (glutamatergic). Interneurons are subdivided into excitatory spiny stellate cells, located in the middle layers of the cortex, and inhibitory smooth non-pyramidal neurons, present in all layers. All or the majority of CB-immunoreactive cells in cortex are smooth non-pyramidal neurons.

Numerous CB-immunoreactive neurons were observed in the cerebral cortex of wild-type mice (Fig. 3.2.2 A,C,E), which were divided into two categories: neurons expressing high levels of CB in their neurites and cell bodies (black arrowheads in Fig. 3.2.2 E), and those expressing CB only in their cell bodies (light brown dots). The former neurons were distributed in all layers of the cerebral cortex except layer I (Fig. 3.2.2 A, C) and had long CB-immunoreactive neurites that branched many times (black arrowheads in Fig. 3.2.2 E). The former seemed to express more CB than the latter. The latter type of neurons were localized in layers II and III and they outnumbered the former (Fig. 3.2.2 A, C). These two kinds of CB-immunoreactive neurons were also observed in nLSD1 -/- mice, but their morphology and distribution pattern were different. The number of neurons expressing high CB levels in their cell bodies and neurites was reduced significantly in nLSD1-/- mice (Fig. 3.2.2 B, D, E) and each individual neuron of this type showed lower levels of CB immunoreactivity than that in control mice. On high power, the lengths of

CB-immunoreactive neurites were also reduced in nLSD1-/- mice and they had fewer branches (Fig. 3.2.2 F). As for the former type neurons, the number of neurons expressing CB only in the cell bodies and their CB immunoreactivity was reduced in nLSD1-/- mice (Fig. 3.2.2 D, F).

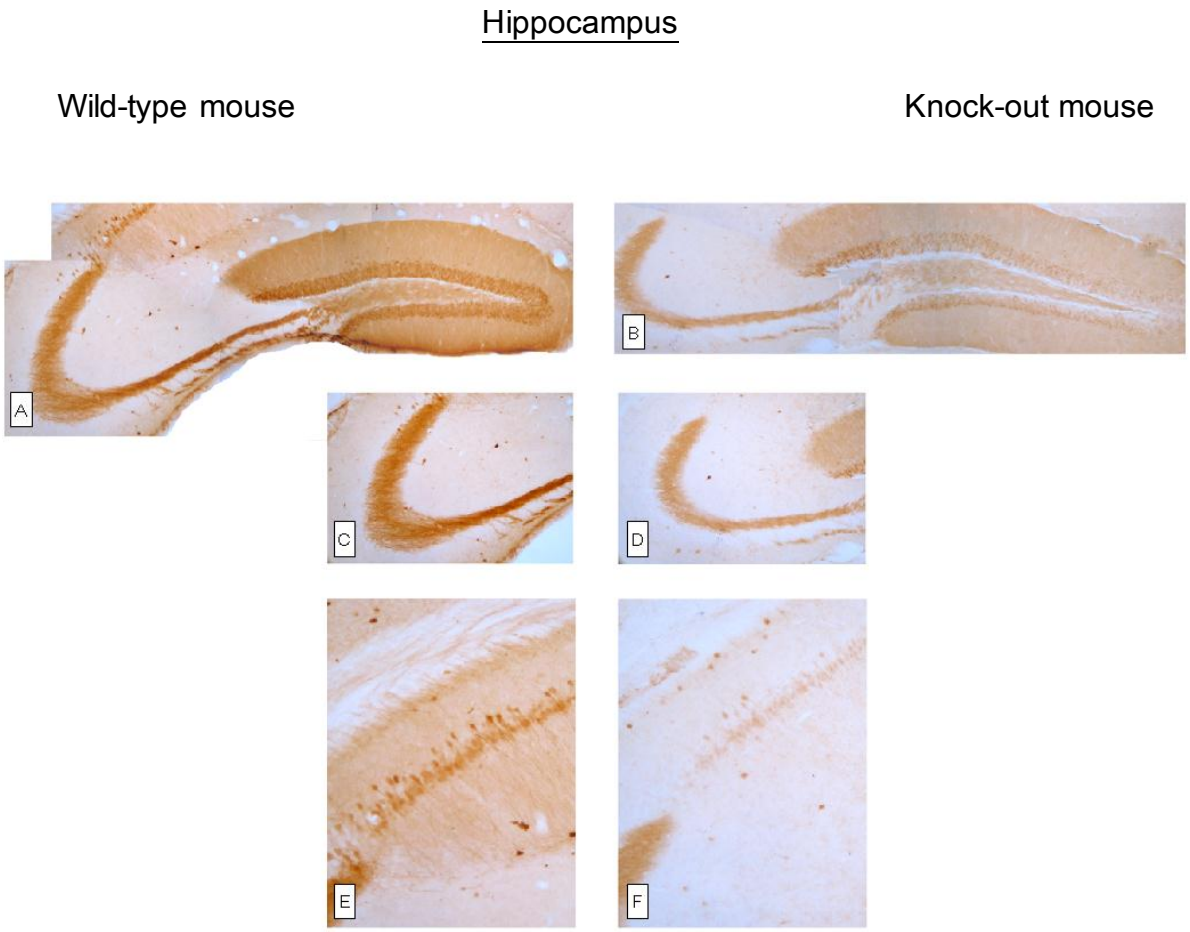


Fig. 3.2.3 Changes in calbindin D28k (CB) immunoreactivity in the hippocampus of E8a-limited-LSD1 knockout mice (B,D,F) compared with wild-type mice (A,C,E).

Decreased calbindin D28k (CB) expression in the hippocampal region of neuronal LSD1 knock-out mice (B, D, F). C, D are high power views of the CA3 of Ammon's horn; E and F are high power views of the CA2 of Ammon's horn. The overall distribution patterns of CB-immunoreactivity and the morphology of CB-immunoreactive neurons are very different in the control and nLSD1 -/- mice (A-F)

In the hippocampal formation there is a heterogeneous population of non-pyramidal neurons most of which expressing GABA. Largely non-overlapping subsets of these GABAergic interneurons express distinct calcium binding proteins (CBPs) and the distribution of these CBP-positive interneurons in the hippocampal subfields has been well characterized in rodents. Calbindin immunostaining marks granule cells of the dentate gyrus and their mossy fibers and axons. Furthermore calbindin fills in soma and dendrites of interneurons located at the border of stratum granulosum and rarely inside the hilus, but marks also and the superficial pyramidal cells of the CA1 region.

The hippocampal region of control mice showed a highly specific CB expression pattern. Granule cells in the dentate gyrus expressed high levels of CB in their cell bodies, and relatively high CB immunoreactivity was observed in the neuropil of the molecular layer and in the polymorph layer (Fig. 3.2.3 A). In Ammon's horn of the anterior hippocampal region, cell bodies of pyramidal cells in CA1 showed CB immunoreactivity, but those of pyramidal cells in CA3 did not (Fig. 3.2.3 C). Instead, the neuropil of the stratum lucidum in CA3 expressed high levels of CB (Fig. 3.2.3 C). As observed in the cerebral cortex, neurons and neuropil expressing CB in the hippocampal region of control mice showed differential CB expression pattern in the hippocampal region of nLSD1-/- mice. The CB expression levels of these neurons and neuropil was lower in the nLSD1-/- mice than that in control mice (Fig. 3.2.3 B). The number of CB-immunoreactive neurons other than pyramidal cells in

CA2 was significantly lower and each neuron showed less CB immunoreactivity in nLSD1-/- mice than that in controls (Fig. 3.2.3 E, F).

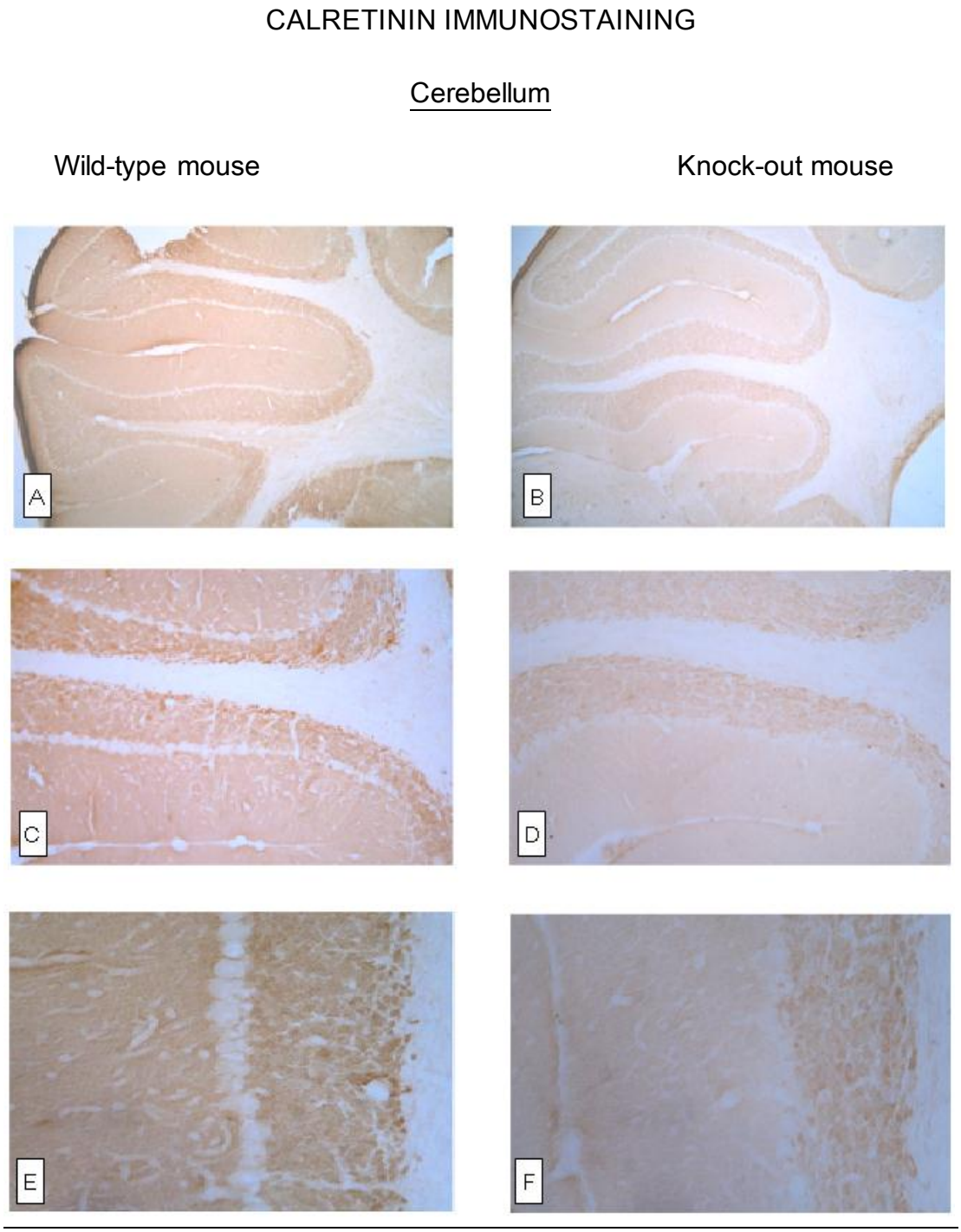


Fig. 3.2.4 Calretinin (CR) immunoreactivity in the cerebellum of E8a-limited-LSD1 knockout mice (B,D,F) compared with wild-type mice (A,C, E).

Sections of the cerebellum are stained with antibodies against calretinin to detect granule, Lugaro and unipolar brush cells. These cells that normally constitute the molecular layer and the granule cell layer are very well-organized in the nLSD1 -/- cerebellum (B,D,F) and in wild type one (A,C,E)

Calretinin-positive cells in the cerebellar cortex are interneurons, in particular they are known as Lugaro, Golgi and unipolar brush cells. In both wild-type and knock-out mice the localization of CR expression was very similar (Fig. 3.2.4 A-F), indeed both the molecular and granular layers showed positive immunoreactivity (Fig. 3.2.4 A, B). Only at high power views Calretinin-positive interneurons seemed to be slightly lower immunoreactive in nLSD1 knock-out mice (Fig. 3.2.4 C, E) compared to wild-type (Fig. 3.2.4 D, F) but the difference is not significative.

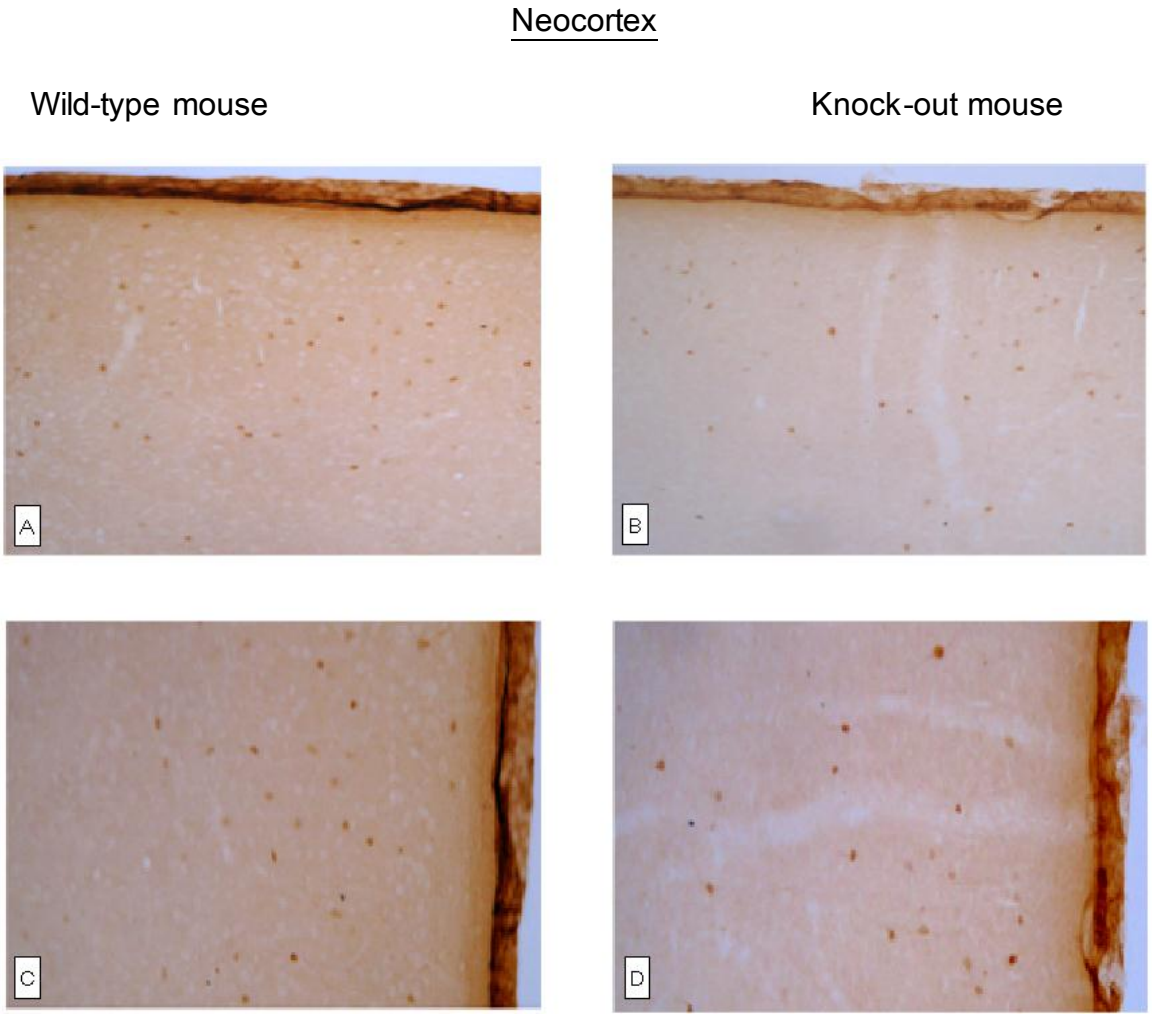


Fig. 3.2.5 Calretinin (CR) immunoreactivity in the cerebral cortex of E8a-limited-LSD1 knockout mice (B) compared with wild-type mice (A).

The distribution pattern and morphology of CR-immunoreactive neurons are te same in the control and nLSD1 knock-out mice. CR immunoreactivity in nLSD1 -/- mice is similar to that of control mice (A, B) and also the number of the CR-immunoreactive neurons in the nLSD1 -/- mice (B) is comparable to that of the control mice (A).

CR+ neurons in neocortex are concentrated predominantly in cortical layers II and III, both in rodents and primates. The density of CR+ neurons decreases with increasing depth in neocortex and therefore they are quite rare in infragranular layers, when compared to the supragranular ones. Neocortex of nLSD1 mutants (Fig. 3.2.5 B) mice revealed at this magnification a slightly lower frequency of calretinin-positive neurons compared to wild-type one (Fig. 3.2.5 A), togheter with a general very mild tissue disorganization (Fig. 3.2.5 A).

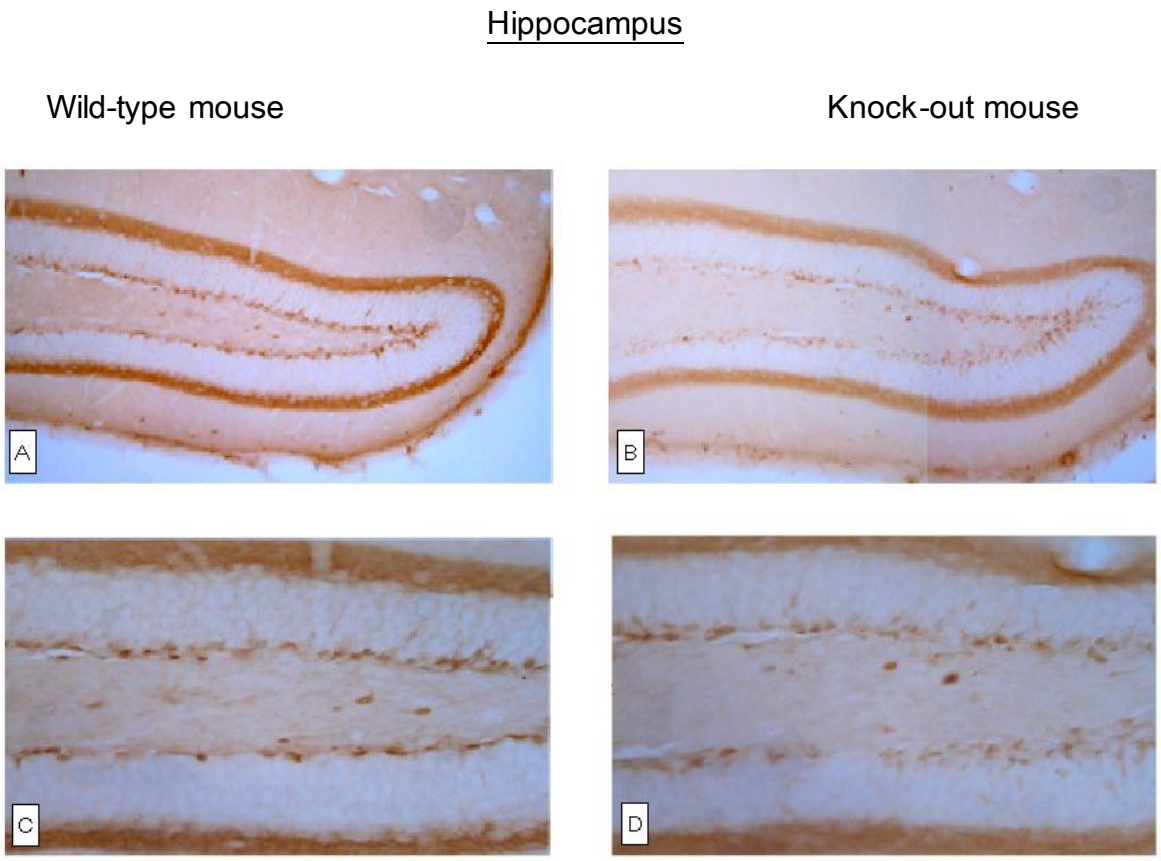


Fig.3.2.6 Calretinin (CR) immunoreactivity in the hippocampus of E8a-limited-LSD1 knockout mice (B,D) compared with wild-type mice (A,C).

Calretinin (CR) expression in the hippocampal dentate gyrus of neuronal LSD1 knock-out mice (B, D, F). C, D are high power views of the dentate gyrus. The overall distribution patterns of CR-immunoreactivity and the morphology of CR-immunoreactive neurons are quite similar in the control and nLSD1 -/- mice (A-D)

Calretinin is expressed mainly in the soma and the proximal and distal dendrites of interneurons that are specialized to innervate either principal cell dendrites or other interneurons in all layers and subregions of the hippocampus. Calretinin immunostains with the highest density the stratum pyramidale of the CA1 region and the projections of the mossy fiber into the inner molecular layer.

In our restricted views of the dentate gyrus (Fig. 3.2.6 A-D) there was a little

reduction of the calretinin immunoreactivity of the supragranular band in nLSD1 -/- mouse (Fig. 3.2.6 B) compared to wild-type mouse (Fig. 3.2.6 A), but such evidence was no more clear in the higher power views (Fig. 3.2.6 C-D) where we rather found in mutated mice a small decreasing in the number of immature interneurons compared to wild-type ones.

VGLUTs IMMUNOSTAINING

Glutamate is the predominant excitatory neurotransmitter in the central nervous system. It is stored in synaptic vesicles and released upon stimulation. The homeostasis of glutamatergic system is maintained by a set of transporters present in plasma membrane and in the membrane of synaptic vesicles. The family of vesicular glutamate transporters in mammals is comprised of three highly homologous proteins: VGLUT-1, VGLUT-2 and VGLUT-3. The expression of particular VGLUTs is largely complementary with limited overlap and so far they are the most specific markers for neurons that use glutamate as neurotransmitter. VGLUTs are developmentally regulated and determine distinct populations of glutamatergic neurons.

VGLUT-1 is localized mainly in neocortex (layers I-III), enthorinal and piriform cortex, hippocampus, amygdala and subiculum. VGLUT-2 can be observed in

olfactory bulb, layer IV of the cerebral cortex, granular layer of the dentate gyrus, thalamus, hypothalamus and in the brain stem cells. In cerebellum, VGLUT-1 is expressed in parallel fibers and VGLUT-2 in climbing fibers and in Purkinje cell dendrites. VGLUT-1 is present at synapses with low release probability, which are known to exhibit long term potentiation (LTP), whereas VGLUT-2 is expressed in synapses that exhibit high release probability and long term depression (LTD). The third vescicular glutamate transporter VGLUT-3 is localized in a very limited numbers of glutamatergic neurons in multiple brain regions and so we didn't perform any immunostaining against this transporter to characterize our mutated mice.

VGLUT-1 IMMUNOSTAINING

Cerebellum

Wild-type mouse Knock-out mouse

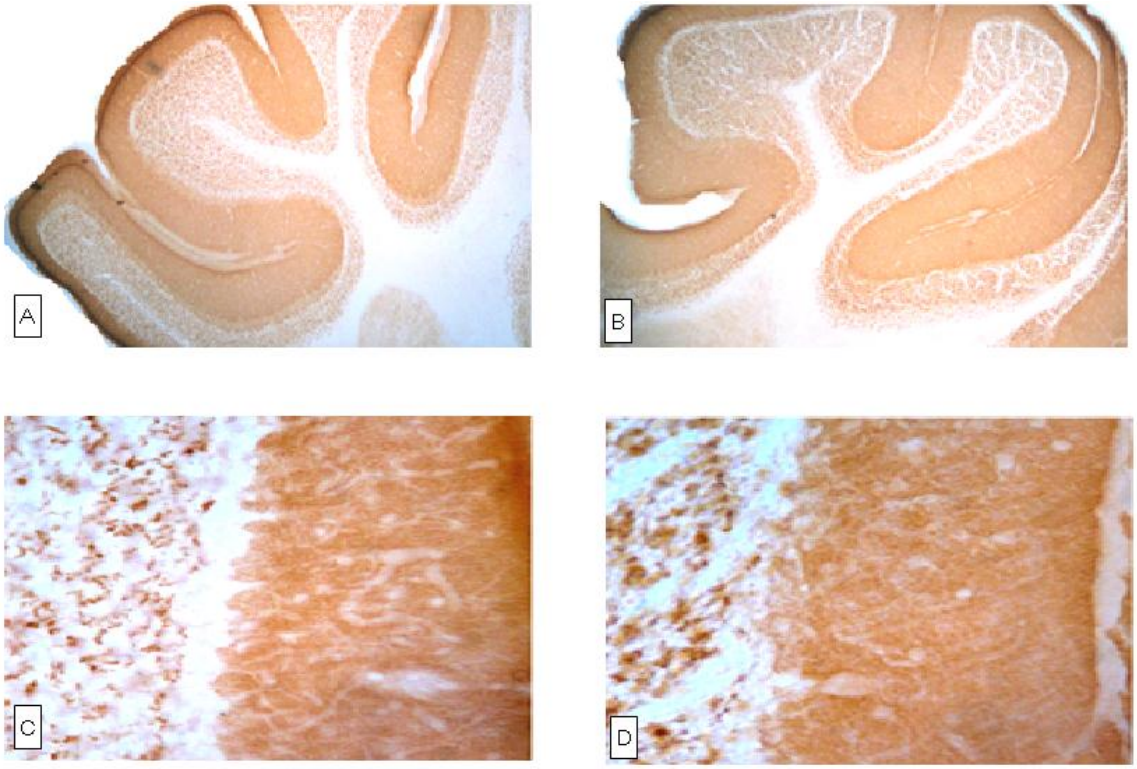


Fig. 3.2.7 Vesicular Glutamate Transporter 1 (VGLUT1) immunoreactivity in the cerebellum of E8a-limited-LSD1 knockout mice (B,D) compared with wild-type mice (A,C).

Sagittal sections of cerebellum stained with antibody against VGLUT1 to mark distribution of parallel fiber terminals to Purkinje cells and of mossy fiber terminals. These cells normally constitute the molecular layer and the granular layer, respectively. Parallel fibers are very well-organized in the nLSD1 -/- cerebellum (B,D) and in wild type one (A,C), while a mild general disorganization is visible in the granular layer of nLSD1 -/- mice (B,D) compared to wild type (A,C).

In mouse cerebellar cortex there are three kinds of excitatory nerve terminals: axon terminals of climbing fibers, parallel fibers and mossy fibers. The parallel fibers that are derived from cerebellar granule cells are well known to use L-glutamate for their excitatory transmission and recently also mossy fibers have been supposed to

employ L-glutamate as transmitter. These two kind of excitatory cells are immunostained by anti-VGLUT1 antibody. While no differences were seen in parallel fibers to Purkinje cells distribution between nLSD1 knock-out (Fig. 3.2.7 B) and wild-type (Fig. 3.2.7 A) mice, at high power views a general mild disorganization was found in the granular layer of knock-out mice (Fig. 3.2.7 D) compared to control mice (Fig. 3.2.7 C). Also Purkinje-cells monolayer, located between the molecular and the granular one, seemed more irregular and thinner in knock-out mice (Fig. 3.2.7 D) than in wild-type ones (Fig. 3.2.7 C).

Hippocampus

Wild-type mouse

Knock-out mouse

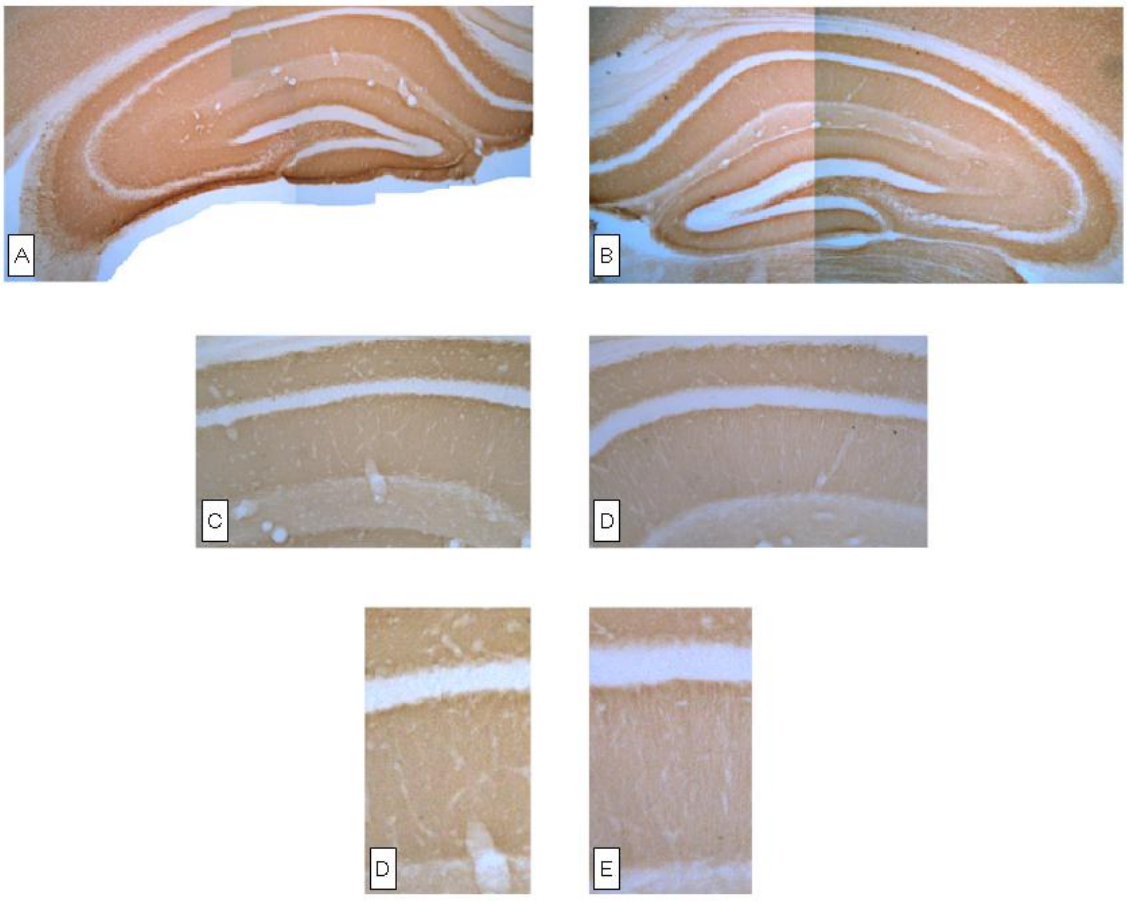


Fig. 3.2.8 Vesicular Glutamate Transporter *1* (VGLUT1) immunoreactivity in the hippocampus of E8a-limited-LSD1 knockout mice (B,D,E) compared with wild-type mice (A,C,D).

Sagittal sections of hippocampus stained with antibody against VGLUT1 to mark all the hippocampal areas except for pyramidal cells and granule cells. In nLSD1 -/- cerebellum (B,D,E) and in wild type one (A,C,D) no differences were observed in VGLUT1 immunoreactivity.

In the hippocampus, VGLUT1 is the dominant transporter subtype in adult animals, whereas VGLUT2 is transiently expressed during early postnatal development. VGLUT1 immunoreactivity is normally detectable in all hippocampal regions except for pyramidal cell layers and for granule cell layer of the dentate gyrus, exactly as we found in all the stainings (Fig.3.2.8 A-F). No differences were observed in VGLUT1 immunoreactivity of nLSD1 knock-out mice hippocampus compared to wild-type one (Fig.3.2.8 A-F).

VGLUT-2 IMMUNOSTAINING

Cerebellum

Wild-type mouse knock-out mouse

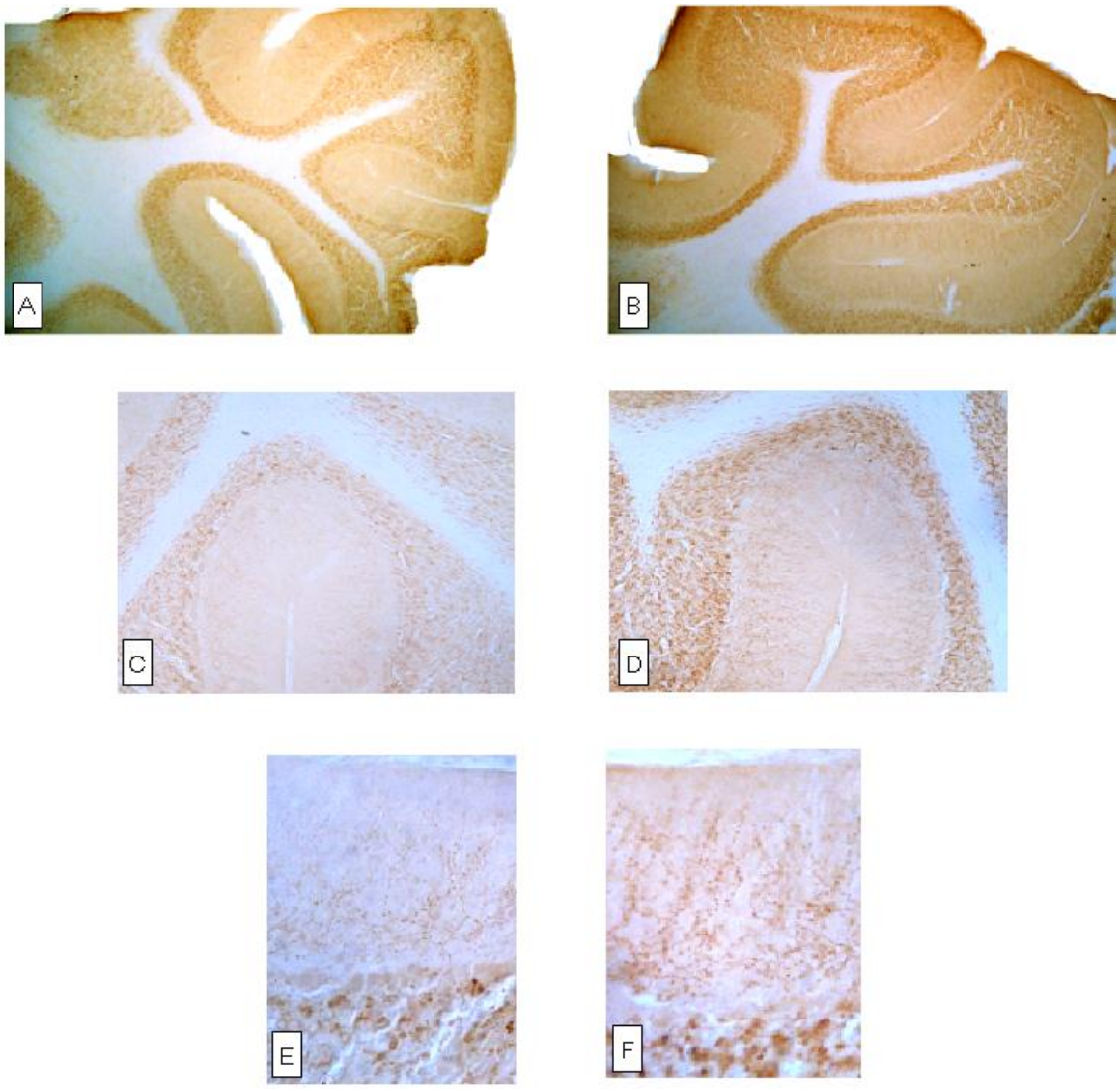


Fig. 3.2.9 *Vesicular Glutamate Transporter 2* (VGLUT2) immunoreactivity in the cerebellum of E8a-limited LSD1 knock-out mice (B,D,F) compared with wild-type mice (A,C,E).

Sagittal sections of cerebellum stained with antibody against VGLUT2 to mark distribution of dimbing fibers and of mossy fiber terminals. These cells normally constitute the molecular layer and the granular layer, respectively. A general increasing in VGLUT2 staining was clear in knock-out mice (Fig. 3.X B, D, F) compared to wild-type ones (Fig. 3.X A, C, E). At high power view an increasing also in the number of climbing fibers and mossy fiber terminals were seen in the nLSD1 -/- cerebellum (F) compared to wild type one (E).

In cerebellar cortex VGLUT2 expression is strong in mossy fibers located in the granule cell layer and in climbing fiber terminals in the molecular layer. Our knock-out and wild-type mice showed a different pattern of staining, with knock-out animals (Fig. 3.2.9 B, D, F) that were much more immunostained than wild-type ones (Fig. 3.2.9 A, C, E), both at the level of molecular and granular zone. In particular, Fig. 3.2.9 E and F highlighted a strong increasing in the number of climbing fibers in knock-out mice, as well as a stronger VGLUT2 affinity for mossy fibers terminals of the granule layer.

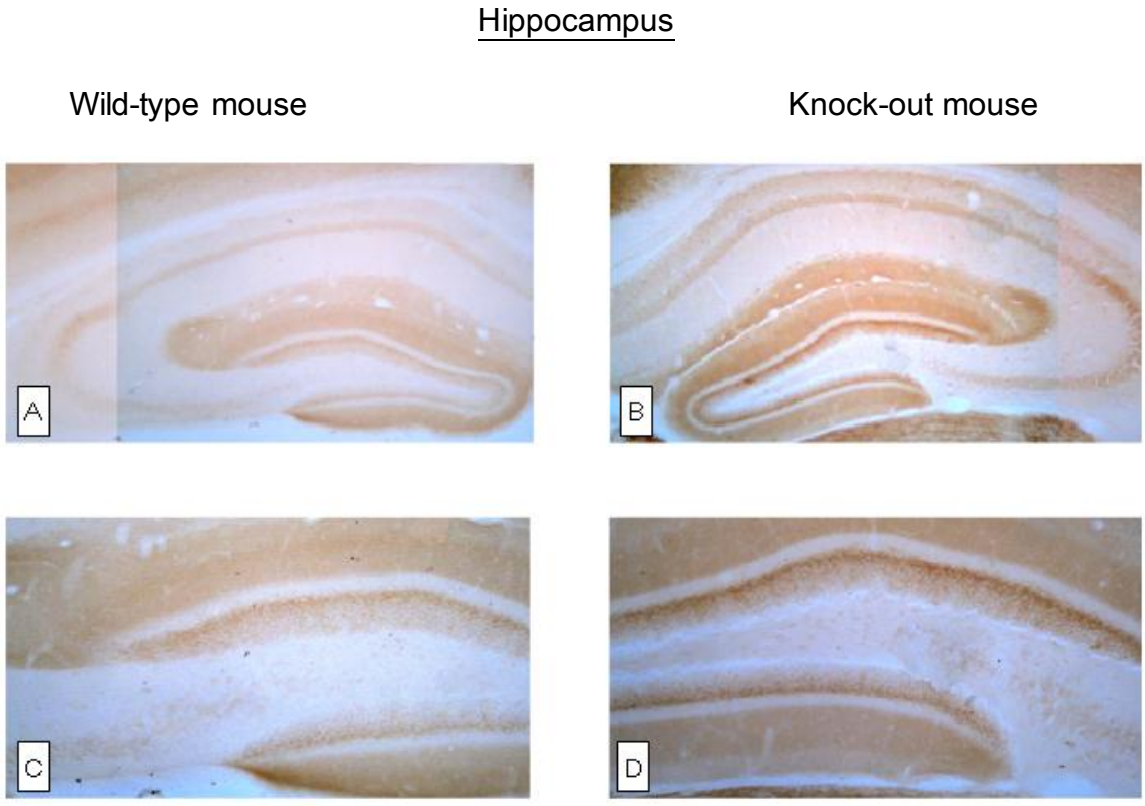


Fig. 3.2.10 Vesicular Glutamate Transporter 1 (VGLUT2) immunoreactivity in the hippocampus of E8a-limited-LSD1 knockout mice (B,D) compared with wild-type mice (A,C).

An overall increasing in VGLUT2 staining was clear in all the hippocampal regions of knock-out mice (Fig. 3.2.10 B, D) compared to wild-type ones (Fig. 3.2.10 A, C). At high power views it was also more evident that VGLUT2 immunoreactivity in the dentate gyrus was higher in knock-out mice (D) than in control mice (C)

VGLUT2 immunoreactivity is routinely detected in nerve terminals in the pyramidal cell layer of the hippocampus, in the granule cell layer and in the outer two-thirds of the molecular layer of the dentate gyrus.

A general increasing in VGLUT2 staining was observed in knock-out mice hippocampus compared to wild-typre one (Fig. 3.2.10 A and B). In particular, at high power view we saw much more immunoreactivity at the level of granule cell layer of the dentate gyrus of mutated mice (Fig. 3.2.10 D).

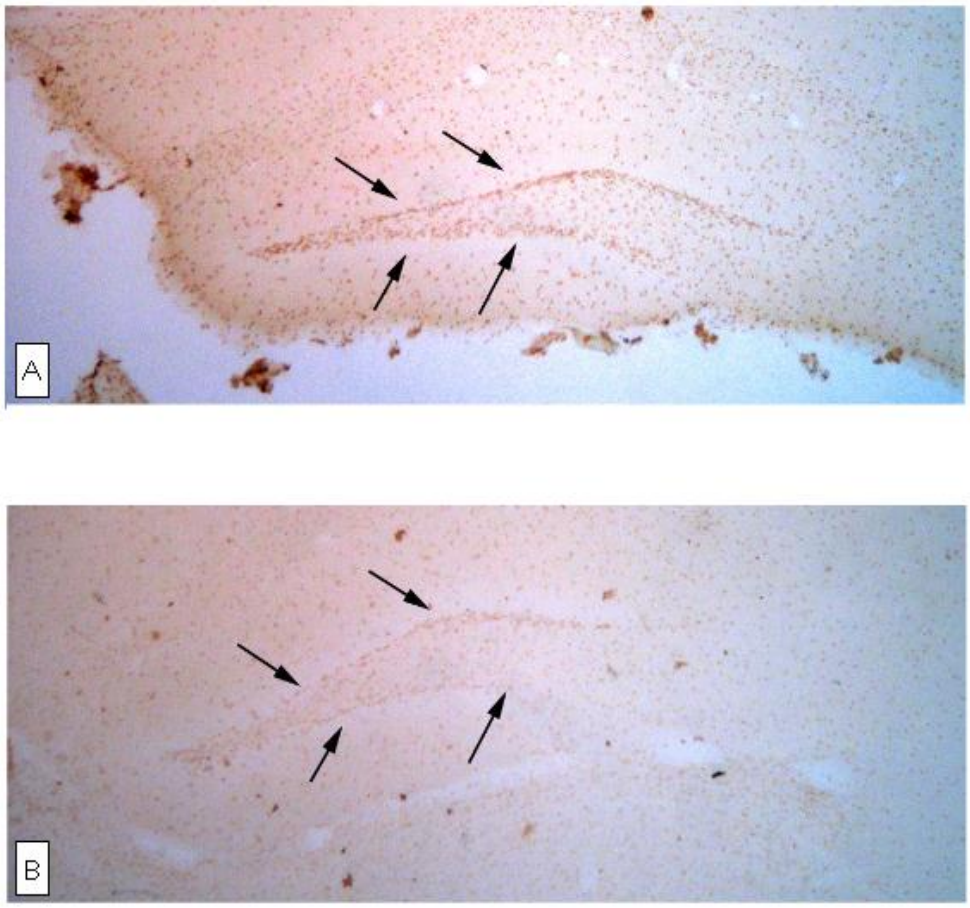
SOX-2 IMMUNOSTAINING

Since we were interested in understanding if nLSD1 could be involved in memory and learning pathways, we performed an immunostaining of the subgranular zone (SGZ) of the dentate gyrus with an anti-SOX2 antibody. We know that the dentate gyrus has an important role in learning and memory, since the subgranular zone is one of the two major sites in the hippocampus, along with the subventricular zone (SVZ), where adult neurogenesis occurs, and adult neurogenesis may play an important role in the acquisition of new memories. The subgranular zone is characterized by several types of cells, but the most prominent are neural stem cells (NSCs) in various stages of development. However, in addition to NSCs, there are also astrocytes, endothelial cells, blood vessels and other components, which form a microenvironment that supports the NSCs and regulates their proliferation,

migration, and differentiation. SGZ generates only one type of differentiated cells: granule cells, the primary excitatory neurons in the dentate gyrus which are thought to contribute to cognitive functions such as memory and learning. The progression from neural stem cell to granule cell in the SGZ can be described by tracing different biomarkers. Sox2 is a biomarker associated with neural stem cells and in the SGZ it is expressed by radial glia cells.

Cerebellum

Wild-type mouse



Knock-out mouse

Fig. 3.2.11 Sox-2 immunoreactivity in the hippocampus of E8a-limited-LSD1 knockout mice (B) compared with wild-type mice (A).

The subgranular zone (SGZ) (black narrows) of the dentate gyrus is marked by the stem cell marker Sox-2. The number of Radial glia cells (indicated by the arrows) that normally constitute SGZ is lower in nLSD1 -/- hippocampus (B) compared to the wild-type one (A).

Radial glial cells are a subset of astrocytes which are typically thought as non-neuronal support cells, indeed they are astrocytic in their morphology. However, unlike most astrocytes, they also act as neurogenic progenitors; in fact,

they are widely considered to be the neural stem cells that give rise to subsequent neuronal precursor cells. Radial glial cells often divide asymmetrically, producing one new stem cell and one neuronal precursor cell per division. Both the number and the immunoreactivity of the Radial glia cells to Sox-2 marker were lower in nLSD1 -/- knock-out (Fig. 3.2.11 B) compared to wild-type mice (Fig. 3.2.11 A).

3.3 PILOCARPINE TREATMENT RESPONSIVENESS

Animals examined:
35 KO and 35 WT mice
Developmental stage: 5 weeks

There is a reciprocal strong relationship between neuronal activation, neurogenesis in the SGZ and learning and memories. On the one hand, high rates of neuronal activation induce neurogenesis that may increase memory abilities. On the other hand, learning and memories have a positive effect on cell survival and induce cell proliferation through increased synaptic activity and neurotransmitter release.
Since our goal is to understand if nLSD1 could be involved in memory and learning pathway we performed the well characterized lithium-pilocarpine treatment to induce neuronal activation in wild-type and mutants mice. This treatment is commonly referred to as a model of Human Mesial Temporal Lobe Epilepsy (MTLE) but seizures in epilepsy are widely thought to develop from highly synchronized

neuronal activation.

Pilocarpine treatment highlighted very different neuronal activation responsiveness in wild-type and knock-out mice, evaluated as number of seizures observed (Fig. 3.3.1) and time of latency to seizures (Fig. 3.3.2)

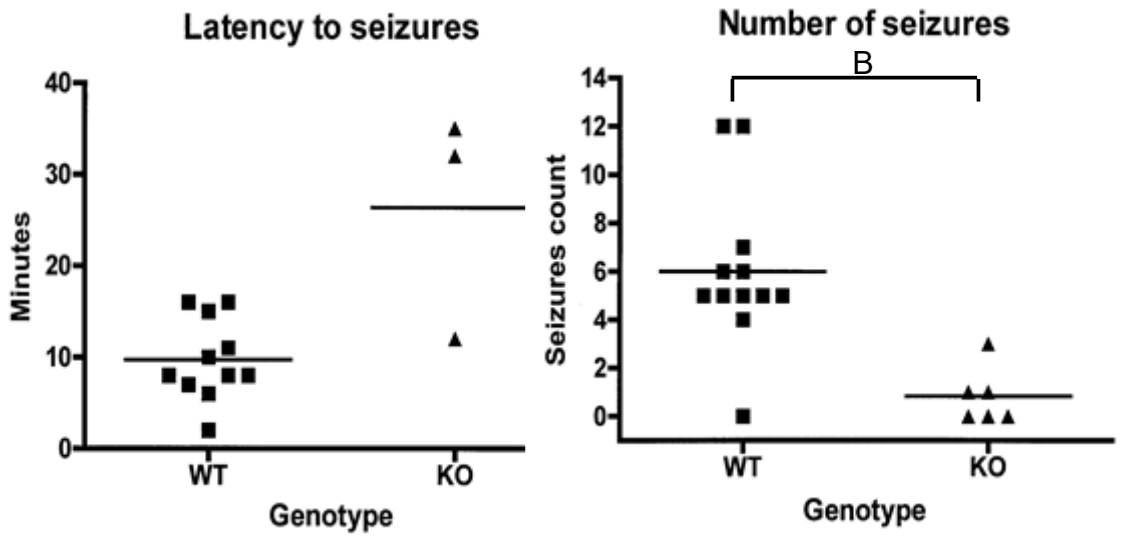


Fig. 3.3.1 Time of latency to seizures Fig.3.3.2 Number of seizures experience

Black squares symbolize wild-type mice while black triangles are for knock-out mice; horizontal lines are the mean values obtained in each analysis

After pilocarpine-induced neuronal stroke with the same pilocarpine dosage described in Results-chapter 1, in nLSD1 knock-out mice we observed an extension of the latent period in which usually seizures do not occur, compared to wild-type mice. Indeed, such time of latency in wild-type animals was about 10 min. and then almost all of the animals experienced tonic-clonic seizures; on the contrary knock-out mice started to display the effects of the treatment only about 25 min. after injection and only a small number of them had tonic clonic convulsion

(Fig. 3.3.1). Also the number of experienced seizures was significantly different between the two mice populations (Fig. 3.3.2): wild-type animals displayed 6 seizure events on the average while mutated mice had about only 1 convulsion in the same period of treatment.

Studying pilocarpine-evoked neuronal activation in nLSD1 knock-out mouse model is interesting also from the biochemical point of view. Indeed, unpublished data from our laboratory revealed a strong physical interaction between LSD1 protein and the Serum Response Factor (SRF), which is a positive transcription factor that binds to Serum Response Element (SRE) in the promoter region of target genes. We know from literature that SRF regulates the activity of many immediate-early genes; among them there are c-Fos and Egr-1, whose induction in the brain is strongly associated with neuronal activity.

C-fos is an immediate-early gene that undergoes dramatic and transient expression in response to a diversity of stimuli but a striking activation in neurons is obtained by depolarization *in vitro* and by seizures *in vivo*. Also egr-1 encodes for a transcription factor protein which has a distinct pattern of expression in the brain, where it has been demonstrated to be elicited by seizure activity.

Since we found that LSD1 takes part to SRF transcriptional complex, we hypothesized that LSD1 might have a role in regulating its transcriptional activity at the level of above-mentioned immediate-early genes, which have in their promoter region more than one SRF Responsive Elements (SRE).

That being so, we evaluated by qRT-PCR the expression level of c-Fos and Egr-1 gene on RNA taken from hippocampal region of wild-type and knock-out mice, 7 hours after seizures onset.

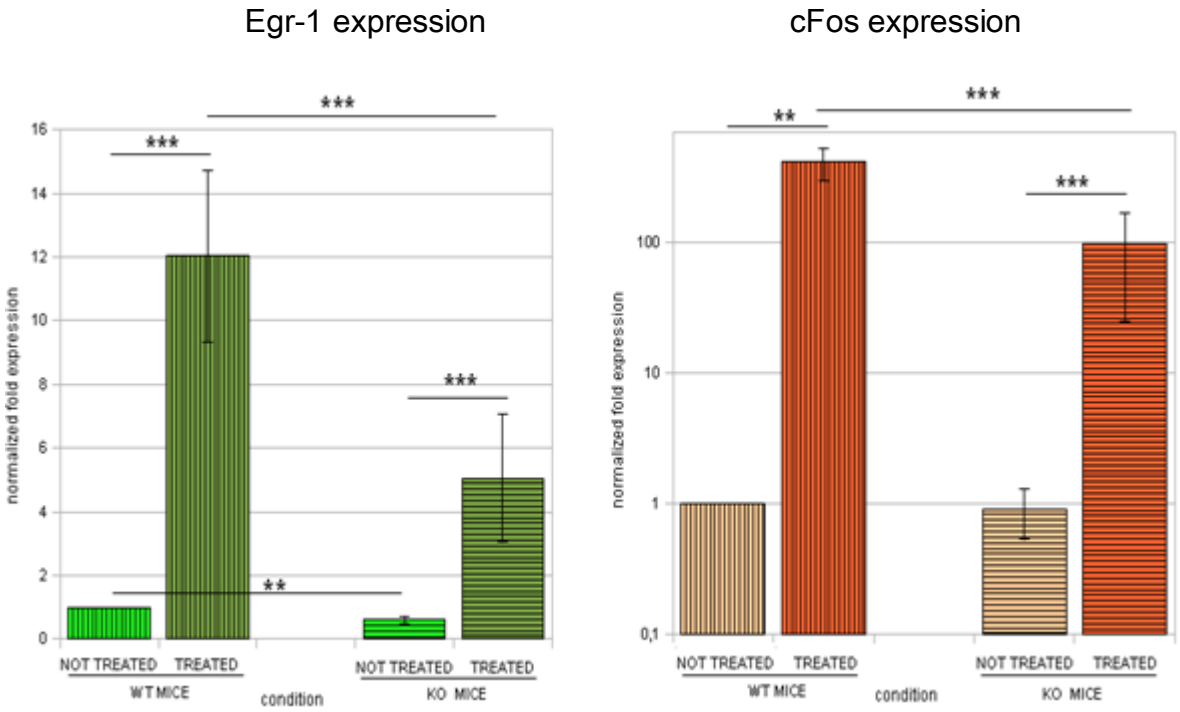


Fig.3.3.3 Level of hippocampal Egr1 and cFos expression in treated and not treated KO mice compared to WT

qRT-PCRs were performed using RPSA as housekeeping gene. Ct values obtained for Egr-1 and cFos were normalized over the housekeeping gene to correct for sample variations in RT-PCR efficiency (dCt). In the charts, the level of expression of the two genes is presented as *fold* change relative to the resting WT mice condition. **are referred to p value<0,01 *** are referred to p value<0,001

As we expected, the expression level of the two immediate early gene significatively changed after the Pilocarpine treatment, with a burst that proved their strong responsivness to seizures. However farmacologically-induced neuronal activation did not had the same effect on both the two mice populations. Wild-type mice, indeed, experienced a massive increasing in Egr-1 and cFos expression

compared to knock-out mice, who on the contrary showed a lower and restrained burst. Such significative difference in their responsiveness may be explained by their different genotype: knock-out mice in fact are totally lacking in the E8a-containing isoform of LSD1, which is known to repress the transcription less than LSD1 ubiquitary isoform. Thus knock-out mice have much more repressive potential compared to wild type mice and this was clearly highlighted by this experiment.

Knock-out mice characterization is still in progress and these results come from preliminary experiments that have to be performed more than one time to be significative. For this reason no discussion will be presented about these data.

DISCUSSION

Alternative splicing is considered one of the most powerful biological mechanisms that leads to diversification of gene function without a corresponding increase in gene number, giving rise to evolutionary complexity especially for mammalian species. Notably, most of the splicing events occur in CNS-related transcripts and participate in fundamental nervous processes, from cell-related ones, such as axon guidance and synapse formation, to higher cognitive functions, including learning and memory [12].

Although LSD1 and its epigenetic role in the central nervous system have been extensively investigated, LSD1 alternative splicing regulation as well as neurospecific LSD1 function *in vivo* has not been elucidated yet.

LSD1 is the first identified Lysine specific demethylase that removes methyl groups from mono- or di-methylated Histone 3 Lys4 (H3K4). It is known to be part of a repressing complex containing also the proteins CoREST and HDAC1/2. As LSD1 is known to be a corepressor of the transcriptional silencer REST, its general function is to inhibit the expression of neuronal genes in non-neuronal cells. LSD1 gene contains two alternative exons: exon E2a and exon E8a. The inclusion in the mature transcripts of the alternative exon E8a is restricted to the nervous system and the LSD1-E8a isoform is characterized by a less neuronal gene repressing action. The exon E8a is only 12-bp long and it codes for 4 aa that form a protruding loop in the LSD1 catalytic domain. Moreover exon E8a contains a Threonine residue that can be phosphorylated [12], and that is required to induce neuronal

maturation and neurite outgrowth and to abrogates neuronal LSD1-8a repressive activity [13]. All these observations suggest an important regulatory role played by the exon E8a inclusion into neuronal LSD1 protein. Thus, it is relevant to understand what are the main *cis*- and *trans*-acting factors involved in exon E8a retaining inside the neuronal LSD1 transcripts.

Here I will point out the main evidences that I obtained from my PhD project about neurospecific LSD1 splicing regulation.

- EXON E8a INCLUSION IS POSITIVELY REGULATED BY NOVA1 and nPTB

We began our investigation of alternative splicing regulation by analyzing the intronic regions flanking the exon E8a, where we identified an extremely conserved 800-bp sequence surrounding the alternative exon. We then cloned this sequence in a Minigene reporter system to study exon E8a inclusion in different cellular context. The Minigene vector is a good reporter model that let us to examine exon E8a inclusion in transcripts exclusively deriving from the reporter gene. Thanks to the high-resolution power of capillary electrophoresis, the assay allowed us to precisely detect exon E8a inclusion in the mature Minigene transcript and to quantitatively measure its percentages. With other techniques, like the polyacrylamide gel electrophoresis, the detection and quantitative analysis would have been more difficult and less precise and reproducible.

We observed exon E8a inclusion selectively in neuronal-like cells. This indicates that within this region are present the *cis*-acting elements necessary to positively drive splicing inclusion in the Minigene transcripts in the neuronal-like cells SH-SY5Y. Interestingly, this cell line, likewise all the other neuroblastoma cell lines we analyzed, do not express endogenous LSD1-8a transcripts. This apparent incongruity suggests that although in this cellular context are present several neuro-restricted *trans*-acting factors positively regulating exon E8a inclusion, within the analyzed genomic region are missing the negative regulatory elements that would recapitulate the endogenous splicing.

Indeed, neuronal splicing is modulated by a subset of neuro-specific *trans*-acting factors whose function have been deeply elucidated during the last few years. Among these factors we started investigating the role of NOVA1 and nPTB. NOVA1 was found by R. Darnell group to be directly bound to an RNA containing the exon E8a nucleotide sequence [67] while nPTB was demonstrated to be part of the same NOVA1-multiprotein complex [69]. The bioinformatics analysis of the cis-acting elements recognized by these two factors indicated the presence of several consensus sequences inside the highly conserved region cloned into Minigene 800. Another important evidence suggesting the implication of both NOVA1 and nPTB in exon E8a alternative splicing is their developmental expression profile in rat brain cortex that closely overlaps with LSD1-8a.

Using the Minigene 800 we were able to demonstrate a positive role of both

NOVA1 and nPTB in exon E8a splicing modulation. In particular in neuronal cells, the overexpression of NOVA1 factor was able to induce a dose-dependent increase in the exon E8a inclusion and in the same cellular context its down-regulation using an anti-NOVA1 ShRNA decreased exon E8a usage. Similarly, we found a positive effect mediated by nPTB over-expression in neuronal cells.

In both cases, NOVA1 and nPTB alone are not able to induce exon E8a usage in non-neuronal cells, indicating that they are both necessary but not sufficient to mediate exon E8a alternative splicing. Their inability in promoting exon E8a inclusion in non-neuronal cells revealed that this splicing event needs other/s neurospecific factor/s present in SH-SY5Y but not in HELA cells. We are now evaluating if a combinatorial effect between NOVA1 and nPTB is required to promote exon E8a inclusion in HELA cells and to increase it in neuronal cells, both transfected with the Minigene 800 vector.

Furthermore, the positive role of NOVA1 in modulating exon E8a alternative splicing was supported by the identification of a pathological condition where NOVA1 and exon E8a were co-regulated. Indeed, in a mouse model of pharmacologically induced Mesial Temporal Lobe Epilepsy, we observed that upon Pilocarpine treatment both the percentage of exon E8a inclusion and the level of NOVA1 expression were reduced. This discovery is important because on the one hand it proves the involvement of the neuronal LSD1 isoform in one of the most

frequent neurological disorder in humans, opening the way to a possible treatment of this pathology, and on the other hand it corroborates also *in vivo* the positive role of NOVA1 in promoting exon E8a inclusion into mature LSD1 transcripts.

Using the same splicing assay, we further tested the possible role of another important neurospecific splicing factor, Sam-68. Bio-informatic analysis using SpliceAid2 prediction tool, indicated Sam-68 as one of the most promising exon E8a splicing regulator, with a score even higher than that found for NOVA1. Both in neuronal and non-neuronal cells we did not see any positive effect on exon E8a alternative splicing, even upon treatment with TPA, a Phorbol ester which is known to induce Sam-68 phosphorylation by activating Ras-signalling pathway. All these experiments allowed us to conclude that exon E8a alternative splicing is not regulated by Sam-68.

4. LSD1 RNA SECONDARY STRUCTURE PLAYS A PIVOTAL ROLE IN MODULATING NEUROSPECIFIC EXON E8a ALTERNATIVE SPLICING

The identification of a downstream sequence complementary to the exon E8a, shed a new light on exon E8a splicing regulation. Pre-mRNAs are typically depicted in a linear fashion although a good portion of them exists in a double-stranded conformation originating higher-order structures. Often, such structures play important roles in regulating RNA metabolism. This is what happens to the neuronal LSD1 pre-mRNAs, where the exon E8a and its complementary and

inverted sequence, separated by 300 nucleotides, promote the generation of a stable stem structure that keeps the exon E8a sequence and the splicing acceptor and donor sites embedded inside it. Since the stem structure is thermodynamically very stable, it could interfere and negatively modulate splice-sites recognition. In principle, RNA secondary structures are able to inhibit spliceosome assembly because the recognition of a splice site, or enhancers or silencers typically depends upon interactions between *trans-acting* splicing factors and single-stranded *cis-acting* elements of the pre-mRNA.

To test the hypotheses that the complementary/inverted downstream sequence negatively regulated exon E8a inclusion, we generated a deleted Minigene 800, missing the 12-bp core sequence of the complementary/inverted region. mFold browser predicted that this deletion would have strongly affected the secondary mRNA structure, “unmasking” exon E8a containing sequence. Indeed, when we transfected deleted Minigene 800 in neuronal cells we observed a strong increasing in the exon E8a inclusion. This result let us to conclude that the deletion perturbed the pre-mRNA secondary structure surrounding the exon E8a, allowing positive trans-acting spicing regulators to easily reach their binding sites.

This effect was not observed in non-neuronal cells where the removal of the complementary sequence did not exert any positive effect. This result indicated that:

- in non-neuronal cells it is not possible to have exon E8a inclusion because,

although the deleted Minigene 800 pre-mRNA is unfolded, these cells are lacking in the neurospecific splicing factor/s that control/s exon E8a alternative splicing;

- in neuronal cells, on the contrary, such neurospecific *trans-acting* factor (or more than one) is expressed but its function is inhibited by the inclusion of exon E8a in the double strand structure.

Intriguingly, the overexpression of NOVA1 or nPTB with the deleted Minigene 800 did not further increase exon E8a inclusion, as expected, suggesting that their positive effect might be exerted by inhibiting or contrasting exon E8a trapping into the double strand RNA structure. Thus, we postulated that the small but significant exon E8a inclusion that we observed over-expressing NOVA1 or nPTB with Minigene 800, was induced by their ability to partially “unlock” Minigene 800 stem structure, allowing the other neurospecific positive splicing regulators to partially play their role.

In the future it will be interesting to discover this neurospecific trans-acting factor (or complex) that is sufficient to strongly drive exon E8a inclusion. This might clear the way to find new possible involvement of LSD1 in physiological or pathological condition with a deregulated expression of such splicing regulator.

- LSD1 HAS A NEW ALTERNATIVE EXON, THAT WE CALLED “EXON E8b”, WHOSE INCLUSION IS PROMOTED BY FOX1

The discovery of this new alternative exon occurred serendipitously when we evaluated the effect of FOX1 on neurospecific LSD1 alternative splicing. Using an extended version of the Minigene 800, named Minigene 2750, including the full length downstream intron with the 2 FOX1 binding sites at the 3' end, we found that FOX1 was completely unable to modulate exon E8a alternative splicing but it had an important role in promoting the “exonization” of a 77-bp downstream intronic sequence. Importantly, exon E8b inclusion occurred almost at the same level in neuronal and non-neuronal cells, thus we can say that FOX1 is sufficient *per se* to regulate exon E8b alternative splicing. Indeed, exon E8b inclusion was not affected by NOVA1 nor by nPTB. This 77-bp “exonized” fragment had all the features of an exon, such as canonical acceptor and donor splicing sites, and is was proximal to *cis*-acting FOX1 binding sequences that regulated its inclusion. The exon E8b sequence analysis revealed that, if included in LSD1 mature transcript containing exon E8 or exon E8a, it introduced a premature STOP codon. This would predict that all the endogenous exon E8b-containing mRNA should be rapidly degraded by the Non-sense mRNA Mediated Decay (NMD). Such degradation does not occur for hybrid exon E8b-including Minigene transcripts, since minigene-derived transcript are not translated. As expected for transcripts regulated my NMD, their identification *in vivo* was very difficult. Indeed, to detect them we had to perform PCRs with more than 50 cycles. We found exon E8b-including transcripts only in human tissues and in particular in brain tissues

from RETT and Epileptic patients. Analyzing these RNA we noticed the presence of another mRNA population presenting a longer exon E8b that included the entire downstream intron fused to exon E9. This second LSD1 splicing variant arose in many different human tissues when we screened an Ambion human RNA panel, although only upon more than 50 PCR cycles. We conclude that in human tissues there are two new LSD1 splicing variants: one containing the sole exon E8b and the other one with the exon E8b and its downstream intron. Importantly, we found exon E8b only in human tissues and not in mouse or rat. This is in agreement with the fact that the acceptor splice site is present only in Human and in Rhesus and the exon E8b nucleotide sequence is not highly conserved. Thus, so far about the exon E8b we have discovered that:

3. it exists not only inside the Minigene 2750 transcripts but also in the endogenous LSD1 mRNAs;
4. it is present only in human LSD1 transcripts;
5. its expression is not restricted to central nervous system;
6. its splicing is tightly regulated by FOX1;
7. when present in LSD1 transcripts, it induces *in vivo* degradation of the mRNA by the Non-sense Mediated mRNA Decay (NMD);
8. it is included, although not preferentially, together with the exon E8a in a small percentage of LSD1 splicing transcripts;

Its very low endogenous expression level, makes it difficult to study its functional

role. At the moment we can only say that its inclusion into LSD1 mature transcripts could be a tool at cells disposal to finely tune LSD1 transcripts amount, providing a new human-restricted level of regulation of LSD1 expression. In support of this hypothesis we know that there are pathological conditions in which LSD1 levels are deregulated. For instance, while in the majority of tumors LSD1 is upregulated, in some cases it is downregulated. This is the case of breast carcinomas, where LSD1 reduction enhances the metastatic spread of the tumor [79]. LSD1 was also found to be progressively downregulated during human embryonic stem cells differentiation, indicating that LSD1 may be involved in maintaining pluripotency [80]. All these findings are consistent with a possible role of the exon E8b in the negative regulation of LSD1 expression. Furthermore exon E8b was found in Minigene 2750 reporter system to be modulated by FOX1, an important splicing factor involved in a lot of phenotypes exhibiting mental retardation, epilepsy and autism. Thus in the future we will investigate pathological conditions characterized by the concomitant deregulation of FOX1 and LSD1.

In conclusion, we can say that the discovery of the human-restricted exon E8b may be a very important step forward for the LSD1 characterization. Indeed, it is interesting that an extremely conserved enzyme like LSD1 has a neurospecific splicing isoform present only in mammals and a newly discovered alternative exon expressed only in human transcripts. This means that the exons E8a and E8b could have given their small but significative contribution to phylogenetic

evolution from yeast, where LSD1 initially appeared, to human. Now all we have to do is understand their function.

BIBLIOGRAPHY

1. Meaney MJ, Ferguson-Smith AC. Epigenetic regulation of the neural transcriptome: the meaning of the marks. *Nat Neurosci. 2010; Nov;13(11):1313-8.*

2. Borrelli E., Nestler E.J., Allis C.D., Sassone-Corsi P. Decoding the epigenetic language of neuronal plasticity. *Neuron. 2008; Dec 26;60(6):961-74.*

3. Sweatt JD. Experience-dependent epigenetic modifications in the central nervous system. *Biol. Psychiatry. 2009 Feb 1;65(3):191-7.*

4. Mansuy IM. and Mohanna S. Epigenetics and the human brain: where nature meets nurture. *Cerebrum, May 2011.*

5. Bernstein E. and Allis C.D. RNA meets chromatin *Genes Dev. 2005; Jul 15;19(14):1635-55.*

6. Maze I., Noh K.M., Allis C.D. Histone Regulation in the CNS: Basic Principles of Epigenetic Plasticity. *Neuropsychopharmacology. 2013; Jan;38(1):3-22.*

7. Lewin B. Genes IX. *Jones and Bartlett, 2008.*

8. Ramakrishnan V. Histone structure and the organization of the nucleosome. *Annu Rev Biophys Biomol Struct. 1997; 26:83-112.*

9. Hamon M.A., Cossart P. Histone Modifications and Chromatin Remodeling during Bacterial Infections. *Cell Host Microbe 2008; Aug 14;4(2):100-9.*

10. Vasquero A. et. al., The Constantly Changing Face of Chromatin. *Sci Aging Knowledge Environ. 2003; Apr 9;2003(14)*

11. Cosgrove M.S. et. al., Regulated Nucleosome Mobility and the Histone Code. *Nature Structural Molecular Biology 2004; 11: 1037-1043.*

12. Zibetti C., Adamo A., Binda C., Forneris F., Toffolo E., Verpelli C., Ginelli E., Mattevi A., Sala C., Battaglioli E. Alternative splicing of the histone demethylase LSD1/KDM1 contributes to the modulation of neurite morphogenesis in the mammalian nervous system. *J Neurosci. 2010 Feb 17;30(7):2521-32.*

13. Toffolo E., Paganini L., Frusconi F. Binda C.,Tortorici M., Verpelli C., Tedeschi G., Maggioli E., Sala C., Mattevi A. and Battaglioli E. Loss of transcriptional repressive activity of neuronal LSD1/KDM1A by neuro-specific phosphorylation modulates morphogenesis in cortical neurons. *In press.*

14. Forneris F., Battaglioli E., Mattevi A., Binda C. New roles of flavoproteins in molecular cell biology: Histone demethylase LSD1 and chromatin. *FEBS Journal 2009; 276 4304-4312.*

15. Strahl B.D. and Allis C.D. The language of covalent histone modifications. *Nature 2000; Vol. 403.*

16. Simmons D., Epigenetic Influences and Disease. *Nature Education 2008.*

17. Fischle W., Wang Y., Allis C.D. Histone and chromatin cross-talk. *Curr Opin Cell Biol. 2003; Apr;15(2):172-83.*

18. Sims R.J, Reinberg D. Is there a code embedded in proteins that is based on post-translational modifications? *Nat Rev Mol Cell Biol. 2008; Oct;9(10):815-20.*

19. Webby CJ, Wolf A, Gromak N, Dreger M, Kramer H, Kessler B, Nielsen ML, Schmitz C, Butler DS, Yates JR , Delahunty CM, Hahn P, Lengeling A, Mann M, Proudfoot NJ, Schofield CJ, Böttger A. Jmjd6 catalyses lysyl-hydroxylation of U2AF65, a protein associated with RNA splicing. *Science 2009; 325(5936):90-3.*

20. Shilatifard A. Molecular Implementation and Physiological Roles for Histone H3 Lysine 4 (H3K4) Methylation. *Curr Opin Cell Biol. 2008; Jun;20(3):341-8.*

21. Culhane J.C., Cole P.A. *LSD1 and the chemistry of histone demethylation. Curr Opin Chem Biol. 2007; Oct;11(5):561-8.*

22. Mosammaparast N. and Shi Y. Reversal of histone methylation: biochemical and molecular mechanisms of histone demethylases. *Annu. Rev. Biochem. 2010; 79:155-79.*

23. Foster C.T., Dovey O.M., Lezina L., Luo J.L., Gant T.W., Barlev N., Bradley A., Cowley S.M. Lysine Specific Demethylase 1 regulates embryonic transcriptome and CoREST stability. *Mol Cell Biol. 2010; Oct;30(20):4851-63.*

24. Lee M.G., Wynder C., Cooch N., Shiekhattar R. An essential role for CoREST in nucleosomal

histone 3lysine 4 demethylation. *Nature 2005; Sep 15;437(7057):432-5.*

25. Forneris F, Binda C, Adamo A, Battaglioli E, Mattevi A. Structural basis of LSD1-CoREST selectivity in histone H3 recognition. *J Biol Chem. 2007; Jul 13;282(28):20070-4.*

26. Sun G., Alzayady K., Stewart R., Ye P., Yang S., Li W., Shi Y. Histone demethylase LSD1 regulates neural stem cell proliferation. *Mol Cell Biol. 2010; Apr;30(8):1997-2005.*

27. Fuentes P., Cánovas J., Berndt F.A., Noctor S.C., Kukuljan M. CoREST/LSD1 control the development of pyramidal cortical neurons. *Cereb Cortex. 2012; Jun;22(6):1431-41.*

28. Neelamegam R., Ricq E.L., Malvaez M., Patnaik D., Norton S., Carlin S.M., Hill I.T., Wood M.A., Haggarty S.J., Hooker J.M. Brain-penetrant LSD1 inhibitors can block memory consolidation. *ACS Chem Neurosci. 2012; Feb 15;3(2):120-128.*

29. Black D.L. Mechanisms of alternative pre-messenger RNA splicing. *Annu Rev Biochem. 2003;72:291-336.*

30. Cartegni L., Wang J., Zhu Z., Zhang M.Q., Krainer A.R. ESEfinder: A web resource to identify exonic splicing enhancers. *Nucleic Acids Res. 2003; Jul 1;31(13):3568-71.*

31. Wang Z., Burge C.B. Splicing regulation: from a parts list of regulatory elements to an integrated splicing code. *RNA 2008; May;14(5):802-13.*

32. Norris A.D., Calarco J.A. Emerging Roles of Alternative pre-mRNA Splicing Regulation in Neuronal Development and Function. Front Neurosci. *2012;6:122.*

33. Witten J.T. and Ule J. Understanding splicing regulation through RNA splicing maps. *Trends in Genetics 2010; Vol. 27, No. 3.*

34. Jensen KB, Musunuru K, Lewis HA, Burley SK, Darnell RB. The tetranucleotide UCAY directs the specific recognition of RNA by the Nova K-homology 3 domain. *Proc Natl Acad Sci U S A. 2000 May 23;97(11):5740-5.*

35. Ule J, Ule A, Spencer J, Williams A, Hu JS, Cline M, Wang H, Clark T, Fraser C, Ruggiu M, Zeeberg BR, Kane D, Weinstein JN, Blume J, Darnell RB.

36. Licatalosi D.D. et al. HITS-CLIP yields genome-wide insights into brain alternative RNA processing. *Nature. 2008 Nov 27;456(7221):464-9.*

37. Zhang C, Frias MA, Mele A, Ruggiu M, Eom T, Marney CB, Wang H, Licatalosi DD, Fak JJ, Darnell RB. Integrative modeling defines the Nova splicing-regulatory network and its combinatorial controls. *Science. 2010 Jul 23;329(5990):439-43. doi: 10.1126/science.1191150. Epub 2010 Jun 17.*

38. Ule J, Stefani G, Mele A, Ruggiu M, Wang X, Taneri B, Gaasterland T, Blencowe BJ, Darnell RB. An RNA map predicting Nova-dependent splicing regulation. *Nature. 2006 Nov 30;444(7119):580-6. Epub 2006 Oct 25.*

39. Witten J.T., Ule J. Understanding splicing regulation through RNA splicing maps. *Trends Genet. 2011; Mar;27(3):89-97.*

40. Buckanovich, R. J., Posner, J. B., and Darnell, R. B. Nova, the paraneoplastic Ri antigen, is homologous to an RNA-binding protein and is specifically expressed in the developing motor system. *Neuron 1993; Vol. 11, 657-72.*

41. Huang C.S., Shi S.H., Ule J., Ruggiu M., Barker L.A., Darnell R.B., Jan Y.N., Jan L.Y. Common Molecular Pathways Mediate Long-Term Potentiation of Synaptic Excitation and Slow Synaptic Inhibition. *Cell. 2005; Oct 7;123(1):105-18.*

42. Buckanovich R.J., Darnell R.B. The neuronal RNA binding protein Nova-1 recognizes specific RNA targets in vitro and in vivo. *Mol.Cell.Biol. 1997;17:3194-3201.*

43. Calarco JA, Zhen M, Blencowe BJ. Networking in a global world: Establishing functional connections between neural splicing regulators and their target transcripts. *RNA* 2011; *May; 17(5): 775-791.*

44. Ratti A. et al. Post-transcriptional Regulation of Neuro-oncological Ventral Antigen 1 by the Neuronal RNA-binding Proteins ELAV. *J.Biol.Chem.; 2008 Vol. 283, 12: 7531-7541.*

45. Racca C. et al. The neuronal splicing factor Nova co-localizes with target RNAs in the dendrite. *Frontiers in Neural Circuits 2010; Mar 3;4:5.*

46. Calarco JA, Superina S, O'Hanlon D, Gabut M, Raj B, Pan Q, Skalska U, Clarke L, Gelinas D, van der Kooy D, Zhen M, Ciruna B, Blencowe BJ. Regulation of vertebrate nervous system alternative splicing and development by an SR-related protein. *Cell. 2009 Sep 4;138(5):898-910.*

47. Raj B, O'Hanlon D, Vessey JP, Pan Q, Ray D, Buckley NJ, Miller FD, Blencowe BJ. Cross-regulation between an alternative splicing activator and a transcription repressor controls neurogenesis. *Mol Cell. 2011 Sep 2;43(5):843-50. doi: 10.1016/j.molcel.2011.08.014.*

48. Paquette AJ, Perez SE, Anderson DJ. Constitutive expression of the neuron-restrictive silencer factor (NRSF)/REST in differentiating neurons disrupts neuronal gene expression and causes axon pathfinding errors in vivo. *Proc Natl Acad Sci U S A. 2000 Oct 24;97(22):12318-23.*

49. Boutz PL, Stoilov P, Li Q, Lin CH, Chawla G, Ostrow K, Shiue L, Ares M Jr, Black DL. A post-transcriptional regulatory switch in polypyrimidine tract-binding proteins reprograms alternative splicing in developing neurons. *Genes Dev.2007 Jul 1;21(13):1636-52.*

50. Spellman R, Rideau A, Matlin A, Gooding C, Robinson F, McGlincy N, Grellscheid SN, Southby J, Wollerton M, Smith CW. Regulation of alternative splicing by PTB and associated factors. *Biochem Soc Trans. 2005 Jun;33 (Pt 3):457-60.*

51. Makeyev EV, Zhang J, Carrasco MA, Maniatis T. The MicroRNA miR-124 promotes neuronal differentiation by triggering brain-specific alternative pre-mRNA splicing. *Mol Cell. 2007 Aug 3;27(3):435-48.*

52. Zheng S, Gray EE, Chawla G, Porse BT, O'Dell TJ, Black DL. PSD-95 is post-transcriptionally repressed during early neural development by PTBP1 and PTBP2. *Nat Neurosci. 2012 Jan 15;15(3):381-8* .

53. Norris A.D. and Calarco J.A. Emerging roles of alternative pre-mRNA splicing regulation in neuronal development and function. *Front Neurosci. 2012; 6:122.*

54. Gehman LT, Stoilov P, Maguire J, Damianov A, Lin CH, Shiue L, Ares M Jr, Mody I, Black DL.. The splicing regulator Rbfox1 (A2BP1) controls neuronal excitation in the mammalian brain. *Nat Genet. 2011 May 29;43(7):706-11.*

55. Gehman LT, Meera P, Stoilov P, Shiue L, O'Brien JE, Meisler MH, Ares M Jr, Otis TS, Black DL. The splicing regulator Rbfox2 is required for both cerebellar development and mature motor function. *Genes Dev. 2012 Mar 1;26(5):445-60. doi: 10.1101/gad.182477.111. Epub 2012 Feb 22.*

56. Chawla G., Lin C.H., Han A., Shiue L., Ares M. Jr, Black D.L. Sam68 regulates a set of alternatively spliced exons during neurogenesis. *Mol Cell Biol. 2009; Jan;29(1):201-13.*

57. Matter N., Herrlich P., Konig H. Signal-dependent regulation of splicing via phosphorylation of Sam68. *Nature 2002; 420:691-695.*

58. Paronetto M.P., Achsel T., Massiello A., Chalfant C.E., Sette C. The RNA-binding protein Sam68 modulates the alternative splicing of Bcl-x. *J.Cell Biol. 2007; 176:929-939.*

59. Richard S. et al. Ablation of the Sam68 RNA binding protein protects mice from age-related bone loss. *PLoS Genet. 2005; Dec;1(6):e74.*

60. Di Fruscio M., Chen T., Richard S. Characterization of Sam68-like mammalian proteins SLM-1 and SLM-2: SLM-1 is a Src substrate during mitosis. *Proc Natl Acad Sci U S A. 1999; Mar 16;96(6):2710-5.*

61. Buratti E. and Baralle F.E. Influence of RNA Secondary Structure on the Pre-mRNA Splicing Process. *Mol Cell Biol. 2004; Dec;24(24):10505-14.*

62. Zheng S, Chen Y, Donahue CP, Wolfe MS, Varani G. Structural basis for stabilization of the tau pre-mRNA splicing regulatory element by novantrone (mitoxantrone). *Chem Biol. 2009 May 29;16(5):557-66.*

63. Jacobs E, Mills J.D., Janitz M. The role of RNA structure in post -transcriptional regulation of gene expression. *J. Genet. Genomics. 2012; Oct 20;39(10):535-43.*

64. Hu X.L., Cheng X., Cai L., Tan G.H., Xu L., Feng X.Y., Lu T.J., Xiong H., Fei J., Xiong Z.Q. Conditional deletion of NRSF in forebrain neurons accelerates epileptogenesis in the kindling model. *Cereb Cortex 2011; 21:2158-2165.*

65. Pitkänen A. Therapeutic Approaches to Epileptogenesis - Hope on the Horizon *Epilepsia 2010; Jul, 51 Suppl 3:2-17.*

66. Sorek R., Ast G. Intronic sequences flanking alternatively spliced exons are conserved between human and mouse. *Genome Res. 2003; Jul;13(7):1631-7.*

67. Zhang C. et al. Integrative Modeling Defines the Nova Splicing-Regulatory Network and Its Combinatorial Controls. *Science 2010; Jul 23;329(5990):439-43.*

68. Witten J.T. and Ule J. Understanding splicing regulation through RNA splicing maps. *Trends Genet. 2011 Mar;27(3):89-97.*

69. Keppetipola N., Sharma S., Li Q., Black D.L. Neuronal regulation of pre-mRNA splicing by polypyrimidine tract binding proteins, PTBP1 and PTBP2. *Crit Rev Biochem Mol Biol. 2012; Jul-Aug;47(4):360-78.*

70. Underwood J.G., Boutz P.L., Dougherty J.D., Stoilov P., Black D.L. Homologues of the *Caenorhabditis elegans* Fox-1 Protein Are Neuronal Splicing Regulators in Mammals *Mol Cell Biol. 2005 Nov;25(22):10005-16.*

71. Polydorides A.D., Okano H.J., Yang Y.Y.L., Stefani G., Darnell R.B. A brain-enriched polypyrimidine tract-binding protein antagonizes the ability of Nova to regulate neuron-specific alternative splicing. *Proc. Natl. Acad. Sci. USA. 2000; Jun 6;97(12):6350-5.*

72. Abranches E. et al. Neural differentiation of embryonic stem cells *in vitro*: a road map to neurogenesis in the embryo. *PLoS One. 2009 Jul 21;4(7):e6286.*

73. Robinson F, Jackson RJ, Smith CW. Expression of human nPTB is limited by extreme suboptimal codon content. *PLoS One. 2008 Mar 12;3(3):e1801.*

74. Iijima T., Wu K., Witte H., Hanno-Iijima Y., Glatter T., Richard S., Scheiffele P. SAM68 regulates neuronal activity-dependent alternative splicing of neurexin-1. *Cell. 2011; Dec 23;147(7):1601-14.*

75. Grange J., Belly A., Dupas S., Trembleau A., Sadoul R., Goldberg Y. Specific interaction between Sam68 and neuronal mRNAs: implication for the activity-dependent biosynthesis of elongation factor eEF1A. *J Neurosci Res. 2009; Jan;87(1):12-25.*

76. Matter N., Herrlich P., König H. Signal-dependent regulation of splicing via phosphorylation of Sam68. *Nature; 420, 691–695.*

77. Jelen N, Ule J, Zivin M Cholinergic regulation of striatal Nova mRNAs. *Neuroscience. 2010 Aug 25;169(2):619-27.*

78. Bhalla K, Phillips HA, Crawford J, McKenzie OL, Mulley JC, Eyre H, Gardner AE, Kremmidiotis G, Callen DF. The de novo chromosome 16 translocations of two patients with abnormal phenotypes (mental retardation and epilepsy) disrupt the A2BP1 gene. *J Hum Genet. 2004;49(6):308-11. Epub 2004 May 18.*

79. Wang Y, Zhang H, Chen Y, Sun Y, Yang F, Yu W, Liang J, Sun L, Yang X, Shi L, Li R, Li Y, Zhang Y, Li Q, Yi X, Shang Y. LSD1 is a subunit of the NuRD complex and targets the metastasis programs in breast cancer. *Cell. 2009 Aug 21;138(4):660-72.*

80. Adamo A, Sesé B, Boue S, Castaño J, Paramonov I, Barrero MJ, Izpisua Belmonte JC. LSD1 regulates the balance between self-renewal and differentiation in human embryonic stem cells. *Nat Cell Biol. 2011 Jun;13(6):652-9. Epub 2011 May 22.*