PhD PROGRAM IN TRANSLATIONAL
AND MOLECULAR MEDICINE
DIMET

University of Milano-Bicocca
School of Medicine and Faculty of Science

Mitochondrial stress deregulates the expression of Brahma, a chromatin-remodeling factor that controls transcription and splicing of genes involved in axon growth and guidance

Coordinator: Prof. Andrea Biondi
Tutor: Prof. Silvia Barabino

Dr. Gabriele Alessandro Fontana
Matr. No. 042788

XXIV Cycle
Academic Year 2010-2011

Né la pioggia che corrode,
né il vento impetuoso
potranno abbatterti,
né l'interminabile corso degli anni
e la fuga del tempo.
Non morirò del tutto, anzi, una gran parte di me
eviterà la morte
(Orazio, III, 30)

- Alla mia cara Nonna Enue -

Table of Contents

Chapter 1

General Introduction

1.1 Two mechanisms that regulate gene expression: alternative promoters usage and alternative splicing

The complete sequencing of human genome has revealed that it contains a smaller than expected number of genes (Lander et al., 2001). On the other hand, one of the most remarkable observations stemming from the comparison of the different genomes belonging to different species is that the number of protein-coding genes in a given organism does not correlate with its cellular complexity. These considerations have renewed the interest in the mechanisms that are able to increase the coding potential of the cell, which are the same mechanisms that have played a major role in the diversification of gene functions throughout evolution (Blencowe et al., 2006). In this paragraph, I will discuss the two major mechanisms that, acting together, generate multiples transcripts of the same gene: the usage of alternative promoters and the alternative splicing process.

Alternative promoters

The most recent evidences suggest that 30-50% of human genes have multiple alternative promoters, indicating the prevalence of this regulatory mechanism (Davuluri et al., 2006). The alternative promoter is defined as an alternative regulatory region from which transcripts of a gene originate. The existence of multiple transcripts, produced by a single gene but that differ in their 5' termini, reflects the presence of such a regulatory sequence. The alternative promoter usage results in proteins with different N-terminal regions (Landry et al., 2003). From the functional point of view, alternative promoters play critical roles in the regulation of the expression of the different isoforms of the gene during physiological and pathological states. It has been demonstrated that genes that are involved in the differentiation process have alternative promoters, and that the activities of these regulatory regions differ during different differentiation and developmental stages. (Davuluri et al., 2006). For example, the *Ly49* gene, controlling the expression of a family of Natural Killer cells surface proteins, has two alternative promoters: the first one is preferentially active during embryonic and post-natal stages, while the other is active in the adult organism (Landry et al.,2003). Moreover, it has been shown that alternative promoters are present in genes which encode multiple protein isoforms which are expressed in different tissues. For example, the *CYP19A1* gene posses 10 alternative promoters which control the trascription of the gene in a tissue-specific fashion (Davuluri et al., 2006).

Alternative splicing

Alternative splicing (AS) is a regulatory mechanism that affects nearly 95% of human genes (Witten et al., 2011), and it is defined as the production of multiples mature mRNAs that vary in composition because of the usage of different splice sites during the pre-mRNA maturation process.

Basing on the most recent genome-wide approaches and sistematic ESTs analyses, it has been shown that basically six different types of alternative splicing participate to the increasing of the proteome complexity (Figure 1) (Blencowe et al., 2006). The most common type of AS involves cassette exons (Figure 1A). AS of cassette exons, which accounts for the 30% of all the AS events, involves alternative exons, flanked by introns, which are either included or excluded from the mature transcript. The second most frequent AS event is the alternative selection of 5' or 3' splice sites (Figure 1B-C), which together account for nearly the 25% of AS events. It has been proposed that the subtle change introduced in the coding sequences by the alternative selection of 5' or 3' splice sites may be biologically relevant. Concerning this topic, it has been shown that alternative selection of 5' or 3' splice sites could arise from stochastic binding of the spliceosome components to less active or cryptic splicing sites contained inside the coding region of the exons. Finally, the remaining 45% of the total AS events are consituted by intron retention, mutually exclusive alternative exons and alternative last exon/polyadenylation site events (Figure 1 D-F).

The regulation of alternative splicing

From the molecular point of view, AS is a strictly regulated process that takes advantage of different mechanisms to finely tune gene expression.

The well established presence of exonic/intronic splicing enhancers (ESEs and ISEs) and exonic/intronic splicing silencers (ESSs and ISSs) sequences provides the first “layer” of AS regulation. As a matter of fact, these *cis*-acting sequences, which are present in the pre-mRNA, play important roles in the regulation of the inclusion or the exclusion of the exon in which are located (in the case of ESEs and ESSs, respectively) or of the downstrem exon (in the case of ISEs and ISSs, respectively). These *cis*-acting sequences regulate the alternative splicing process by binding with *trans*-acting factors. In particular, the splicing enhancers are recognized by protein containing the RS domain, such as Serine-Arginine (SR)-rich proteins or SR-related proteins. These proteins act as exon inclusion enhancers by promoting the stable assembly of the U1 and U2 snRNPs to the 5’ splice site and the branch site, respectively, resulting in exon inclusion. On the other hand, the splicing silencers sequences are recognized by heterogeneous nuclear ribonucloproteins (hnRNP). These proteins act as inhibitors of exon inclusion by competing or disrupting the binding of the SR proteins, resulting in exon skipping (Blencowe et al.,2006).

Another level of AS regulation is provided by the connections between splicing and transcription. As I will discuss in more

details in the paragraph 1.3, it has been shown that splicing could occur cotranscriptionally, and this evidence has suggested novel and intriguing mechanisms that could couple transcription, AS and the chromatin environment. In the frame of “cotranscriptionality”, it has been demonstrated that transcription factors that act at the transcription start sites have an impact in AS events (Nogués et al.,2002). Several reports have demonstrated that the processivity of the RNA Polymerase II greatly alters the splicing outcome: as a matter of fact, chemical or genetic inhibitions of RNA Polymerase II processivity cause the inclusion of alternative exons in the mature transcript (de la Mata et al.,2003; Batsché et al.,2006; Kornblihtt, 2006), establishing a direct link between transcription and splicing. Moreover, it has also been shown that chromatin structure has an impact on AS patterns, because histone post-translational modifications directly act on the polymerase processivity (Nogués et al.,2002; Hnilicova et al., 2011; Gunderson et al.,2011).

Finally, the presence of long nuclear-retained regulatory and noncoding RNAs, such as the MALAT1 RNA, regulate the AS process, acting on the nuclear localization and on the phosphorylation of SR proteins and other splicing regulators (Anko et al., 2010).

The alternative splicing of 3' terminal exons and the usage of alternative polyadenylation sites

About half of mammalian genes generate alternative mRNAs which differ on their 3' regions. Such differences could be related to the AS of different 3' terminal exons or to the usage of different polyadenylation sites. These two distinct events, which sometimes are grouped together as a single event, differ in their dependence on splicing and on their final outcomes. The choice of alternative 3' terminal (or last) exons, which is also called "3' exon switching", is splicing-dependent, and it generates two mRNAs which differ in their 3' terminal regions. In turn, these two mRNAs would generate two proteins with different C-terminal regions or different 3' untranslated regions (3' UTRs). On the other hand, the usage of different polyadenylation sites (PAS), which is also called "tandem UTRs", rely on the choice of different PAS of the same 3' terminal exon. This event is splicing-independent, and in turn generates two mRNAs with the same exons composition, but with different 3' UTRs. This difference does not have an impact on the protein sequence, but rather on the mRNA stability, because 3' UTRs are often targets of micro-RNAs (Zlotorynski et al.,2008) and/or 3' UTRs-binding proteins (Li et al., 2009).

Even if the 3' exon switching and the tandem UTRs events differ in their dependence on the splicing process, they are both dependent on transcription, and in particular on RNA polymerase II processivity. As a matter of fact, the RNA polymerase II is subjected to a "terminal exon pausing", which

corresponds to a change from its “fast” elongating state to a slow processive state. This pause, which is often observed 250 base pairs before the PAS, is functional to delay the 3’ pre-mRNA processing and to provide time for the contranscriptional splicing to occur. It has been suggested that, during this “pause”, the splicing decisions regarding the 3’ last exons are made, and that the 3’ pre-mRNA processing takes place, chosing one of the possible tandem UTRs present in the transcript (Oesterreich et al.,2011). From the molecular point of view, this terminal exon pause is dictated by the components of the 3’ mRNA processing. As a matter of fact, it has been demonstrated that the CPSF (cleavage polyadenylation specificty factor) complex, which binds to the classic AAUAAA hexamer of the PAS, interacts with the body of the RNA polymerase II and slows down its processivity (Nag et al.,2007).

1.2 The epigenetic regulation of gene expression

The complete sequencing of human genome (officially started in 1990 and fullly completed in 2006, when the sequence of the long human chromosome 1 was finally published) has given to us a potent tool to gain insight into the mechanisms which control gene expression (Lander et al.,2001). But despite all the celebrations associated with this important scientific conquest,

our ability to describe the genome function remains quite limited. It has been demonstrated that we cannot simply describe the coordinated activities of the thousands genes present in the cell on the basis of the DNA sequences, such as promoters, enhancers and splicing sites. It has became a common idea that the primary sequence of the DNA is only the starting point for understanding how genetic program is read (Bernstein et al.,2007). It is also true that, during these last years, the researchers have made great efforts to shed some light on the complex mechanisms that ensure the correct functioning of gene activities in different cell types and developmental stages. In particular, current researches are trying to understand how genomes are organized in the nucleus, the basic principles of nuclear architecture and how these features are linked to regulation of gene expression. Taken all together, the results of these researches opened up a new and exciting field of interest, know as the epigenetic regulation of gene expression. Basically, the word “epigenetic” indicates all the mechanisms of gene expression regulation that are “over” the common “genetic” features present in the DNA sequence. From the practical point of view, an epigenetic feature is constitued by any heritable influence in the gene activity that is not accompained by a change in nucleotide sequence (El-Osta et al.,, 2000; Bernstein et al., 2007). So many different cellular activities are involved in the establishment of epigenetic tracts, but basically all the epigenetic mechanisms that control gene expression fall in

three groups: DNA methylation, histones post-translational modifications and nucleosome remodeling (Cremer et al., 2001; Bernstein et al., 2007). In this paragraph, I will discuss the basic epigenetic mechanisms that control gene expression, with a particular focus on the mechanisms of nucleosome remodeling. But initially, it is mandatory to discuss how the DNA is organized in the nucleus, and how this particular structural organization allows the epigenetic mechanisms to take place. As a matter of fact, the organization of the DNA in the complex structure known as chromatin controls the activities of the enzymes that are involved in the generation of the various epigenetic marks, and vice versa (Allemand et al. 2008). The interplay between chromatin and "epigenetic" enzymes is an exciting and very recent field of research, which constantly reminds us both the beauty and the complexity of biological systems.

The eukariotyc genome is organized in chromatin

The human genome, which contains all the genetic informations requested to ensure the correct cell functions, is composed of about $3{,}3 \times 10^9$ base pairs. If strecthed out completely, it would represent more than 2 meters of DNA. On the other hand, the spherical structure known as the cell nucleus contains two copies of this DNA, and it does not even reach a diameter of 6 µm (Craig et al., 2010). In order to obtain this extraordinary grade of compaction, human cells, like all other eukaryotic cells, organize the DNA in chromatin. Chromatin is a complex

structure composed by DNA and architectural proteins, which interact and forms contacts between each others to form a packaged, high-ordered macromolecular structure. The basic building block of chromatin is the nucleosome, which consists of 147 base pairs of duplex DNA wrapped around a protein multi-subunit complex called histone octamer. The histone octamer is composed of two copies of each of the four coventional histones proteins: H2A, H2B, H3 and H4 (Khorasanizadeth et al.,2004). The positively charged residues present in the histones contact the phosphate backbone of the DNA every 10,4 base pairs, so that the 147 bases stretches of the DNA wrapped around the histone octamer make nearly 14 contacts. Even if a single contact is very weak, the presence of 14 contacts per nucleosome and the positional effect make the nucleosome-DNA interaction very stable (Clapier et al., 2009). The distance between nucleosomes varies between organisms and species, but is usually from 10 to 50 base pairs. The spacing between neighbouring nucleosomes is important for proper gene expression regulation, mostly at the *cis*-acting control regions, such as promoters and enhancers (Khorasanizade et al., 2004).

The eukaryotic cells contain also some histones variant, which are less abundant and which are incorporated in the nuclesomes and/or nucleosome-related particles. However, the incorporation of these variants is restricted to small regions of chromatin and it is functional to “mark” some peculiar sites. For

example, histone variant H2A.Z is incorporated in the nucleosomes localized in close proximity of transcription start sites (Clapier et al.,2009). The variant histone H1 and H5 join the nucleosomes, thus promoting the high-ordered chromatin structures that form the “chromatosome”, defined as the very stable asimmetric chromatin unit constituted by the compaction of neighbouring nucleosomes linked together by these histone variant. Respect to H5, the H1 variant seems to be more enriched in the heterochromatin than in euchromatin, leading to the higher compaction and “closed” aspect of these particular chromatin territories (Hargreaves et al.,2011).

DNA methylation

DNA methylation is a DNA modification that occurs both in plants and animals, and consists in the biochemical addition of a methyl group to the cytosine C5 in a cytosine-phosphate-guanine (CpG) dinucleotide via a methyltransferase enzyme. In mammalian cells, the DNA methylation occurs almost exclusively in the context of CpG dinucleotides. Non-CpG methylation (for example, methylation in the CNG and CNN trinucleotides) has been observed in the cells of the early mouse embryo and in the embryonic stem cells, but it decreases dramatically in the somatic and differentiated tissues, where it becomes very rare (Bernstein et al.,2007). Although CpG dinucleotides are very infrequent in mammalian genomes, they tend to cluster in discrete regions (of 100-200 nucleotides), termed “CpG islands”, that are characterized by

and high C+G content (at least 50% of Cs and Gs) and that constitues nearly the 1% of the human genome. These clusters of the CG dinucleotide tend to localize upstream to the transcription start site. Due to this peculiar localization, finding a CpG island inside a genome sequence is a good evidence that indicates the presence of a trascription start site (Gardiner-Garden et al.,1987). The CpG islands are important elements that regulate eukaryotic transcription. Genome-wide studies revealed that unmethylated CpG islands are prominent in undifferentiated cell and in the embryo. This methylation-free state is associated with actively transcribed genes. During differentiation, some CpG islands acquire the methylation mark in a tissue and cell-specific fashion, and this methylated state is usually associated with silencing of the downstream gene. For these reasons, it has been proposed that different patterns of CpG methylation regulate the differentiation-dependent and the tissue-specific expression of one given gene (Bird et al., 2002).

Histone post-translational modifications

All the four histone types that are included in the nucleosome have an amino-terminal region consisting in 25-40 residues that protrude beyond the nucleosome surface. This region, called “tail”, is the target of a wide variety of post-translation modifications (PTMs), such as methylation, acetylation, phosphorylation and ubiquitination (Bernstein et al.,2007). In particular, it has been demonstrated that lysine acetylation (controlled by histone acetylases and deacetylases) and lysine

or arginine methylation (catalyzed by methylases and demethylases) play important roles in the epigenetic regulation of gene expression. As a matter of fact, the whole-genome mapping of the histones modifications has revealed that different histone PTMs map in different regions of the body of the gene, creating an “histone code” that epigenetically control the activity of the gene. Particular histone PTMs, such as dimethylation of histone 3 at lysine 27 (H3K27me2) and trimethylation of histone 3 at lysine 4 (H3K4me3) are enriched in exons present at the 5’ region of the genes; some others, like the trymethylation of histone 3 at lysine 36 (H3K36me3), are enriched in the internal exons (Spies et al.,2009). From the functional point of view, it has been demonstrated that a close cross-talk between the histone PTMs and the cotranscriptional splicing occur. As a matter of fact, H3K36me3 marks the alternative exons which are included in the transcripts, and some other PTMs, such as H3K79me3, control the elongation rate of the RNA Polymerase II (Spies et al.,2009), which in turn is linked to the inclusion of alternative exons (Luco et al., 2011). Taken all together, these observation reveal the high complexity of the epigenetic control exerted by the histone code. In contrast to the DNA methylation, that is the only DNA chemical modification so far identified, histones have at least 100 different PTMs, and the number is still growing. The understanding of the “code” that lies beneath the different histone PTMs is a new and fresh field of research that forces the researchers to develop novel techniques to identify the

connections between the different histone PTMs and the regulation of gene expression.

The chromatin remodeling complexes

In order to enable the dynamic access to the chromatin packaged DNA and to finely regulate the nucleosome composition in a specific chromosomal region, the eukaryotic cells have evolved a class of enzymes know as chromatin remodeling complexes (Clapier et al., 2009). All these remodelers share some basic properties that make them able to allow nucleosome engagement, selection and remodeling.

The first peculiar feature that defines this class of enzymes is their enzymatic activity. As a matter of fact, all the chromatin remodeling enzymes are characterized by two abilities: first of all, they are able to alter or disrupt the histone-DNA contact. Second, they can move the nucleosomes along the DNA, exchange the positions of the nucleosomes or remove them completely from the chromatin. These activities are aimed to regulate the accessibility of the different regions of the chromosome to the proteins that need to access the DNA during the physiological cellular processes (Hargreaves et al., 2011). All these activities are driven by the energy derived from the hydrolysis of the ATP. This characteristic leads to the second common features shared by all the chromatin remodeling complexes: the presence of an ATPase subunit.

The ATPase subunit present in the chromatin remodeling enzymes is the “motor” that allows them to alter the chromatin structure. This subunit is present in all the different chromatin remodeling complexes, is evolutionally conserved between all the proteins that belong to this family, and also conserved in different species. Despite the high grade of homology of the core ATPase subunits present in the different remodelers, the proteins are not genetically redundant *in vivo,* and they show different functions respect to different chromosome region, indicating that they have specialized functions. The specificity of the different ATPase and/or chromatin remodeling complexes is not due to exclusive expression patterns (for example, tissue-specific expression), because co-expression of different ATPase subunits, as well as their deletion, exhibits very different phenotypes (Eisen et al., 1995). In the Figure 3 are highlighted the different domains that are contained in the various ATPase subunits. As I will mention in the next part of this paragraph, these differences are the basis to distinguish and classify the different ATPase subunits.

The last peculiar feature that joins all the members of the chromatin remodeling family is their multi-subunits structure. As a matter of fact, these complexes function as “macromolecular machines”, in which each subunit is devoted to one specific activity. The different remodelers could contain from 2 to 15 subunits (depending on organism and the complex family), generating a multi-subunits, multi-enzymatic complex that could

easily reach the molecular weight of 2 MDa. One example of subunit that has been already mentioned is the common ATPase subunit which is peculiar of all the chromatin remodelers. Another example is represented by monomeric nuclear actin and actin-related proteins (ARPs), which are both constitutive subunits that directly bind to the ATPase of all the chromatin remodeling complexes (Hargreaves et al., 2011). Other subunits of these enzymes could be either constitutive or alternative, and their incorporation in the complex could vary throughout several differentiation stages or different cell lines (Yoo et al., 2009). The presence or the absence of the alternative subunits could alter the affinity for the nucleosome (independently from the DNA itself), could make the remodeler recognize particular histone modifications, could modify the affinity between the enzyme and other DNA or chromatin interactors and, finally, could modulate and finely tune the constitutive ATPase activity (Clapier et al., 2009). The question that could arise is why evolution has selected such a huge macromolecular complexes in which each subunit is endowed with a single specific enzymatic activity, rather than relying on the activities of the single and isolated subunits. The answers that could reply to this question are basically two. In the first scenario, it has been proposed that the presence in close proximity of all the subunits at the same time and in the same chromatin region helps coupling the different biochemical reactions that together lead to chromatin remodeling. The single reactions could be very “slow” in term of reaction rates, but with

the enzymatic coupling of the different subunits they can proceed very fast and be able to respond to the cell needs in a shorter time period (Hargreaves et al., 2011). The second hypothesis that tries to explain the presence of such huge multi-enzymatic complexes relies on the opportunity to easily change the chromatin remodeling reaction outcome by exchanging few alternative subunits. In this scenario, the presence of the different subunits is functional to a combinatorial assembly of the complexes, which in turn explains also the specificity and activity of the different remodeling complexes. This latter hypothesis is sustained by the evidence that the different alternative subunits of the chromatin remodeling complexes are encoded by gene families (Aigner et al.,2007; Lessard et al., 2007).

The mechanisms and the aftermath of the activity of the chromatin remodeling complexes

The enzymatic activity of the chromatin remodelers lead to different outcomes, depending on the combinatorial assembly of the complex, on their interactions with other proteins and on the structure of the chromatin.

Basically, the remodeling complexes activities could be classified in two main categories, which have been both verified *in vitro* and *in vivo*: “site exposure” or “alteration of the nucleosomes composition” (Figure 2).

Site exposure: this activity takes place when the chromatin remodeling enzyme uses the energy derived from ATP hydrolysis to remodel the chromatin in a specific chromosome locus. During this process, the nucleosomes could undergo to repositioning (the selected nucleosome slides along the DNA, and it is repositioned in a nearby place), ejection (the remodeler "extracts" the selected nucleosome from the chromatin) or unwrapping (the enzyme maintains the previous position of the selected nucleosome but disrupts locally the contacts between the DNA and the nucleosome). The site exposure activity exerted by the remodelers is functional to expose a specific site in the DNA and to make it interact with other proteins, such as DNA or chromatin binding proteins.

Alteration of the nucleosome composition: this activity takes place when the histone content present in the nucleosome is modified, and can lead to histones octamer replacements (for example, exchange of the dimer H2A-H2B with a dimer containing one histone variant) or to dimer ejection.

Recently, some structural and biochemical studies have been conducted to gain insight into the molecular mechanism that allows the chromatin remodeling complexes to couple the ATP hydrolysis to the chromatin remodeling activities such as site exposure or alterations in the nucleosome composition. These studies demonstrated that the DNA present in the nucleosome is very stably associated with the histone octamer, throught

electrostatic interactions that make the unwrapping reaction energetically unfavorable. Nevertheless, the chromatin remodeling complexes are able to disrupt the contacts between the DNA and the histone octamer. But how does this reaction take place? An early hypothesis postulated the model of the "twist diffusion", in which the remodelers are able to induce a propagation of the DNA wrapped around the histone, making the nucleosome "slide" along the DNA. However, this model was then rejected, on the basis of the observations that great impediments to DNA twisting produced no defects in nucleosome sliding (Aoyagi et al., 2002). A more realistic model is the "loop recapture" model, which postulates that the remodeler uses the energy of the ATP hydrolysis to generate a loop of DNA that interacts with neighbouring linker DNA. This loop then moves along the DNA, making the nucleosome slide. This loop seems to be generated by the translocase/helicase activity that is present in the ATPase subunit of the chromatin remodeling complex. This model is strongly sustained by the observation that the ATPase subunits is able to bind to a specific location in the nucleosome, which is characterized by weak DNA-histones contacts. This binding is followed by ATP hydrolysis and activation of the translocase activity (Saha et al., 2005).

The four families of the chromatin remodeling complexes

The different chromatin remodelers display an ATPase subunit, endowed with an evolutionally conserved ATPase domain.

However, the ATPase subunits of the remodelers differ in the other domains composition. The differencies in the domain composition of the ATPase subunits constitues a common way to group all the chromatin remodeling complexes in four distinct families: the SWI/SNF, the ISWI, the CHD and the INO80 families (Clapier et al.,2009).

As shown in Figure 3, all the remodelers contain an ATPase subunit characterized by an ATPase domain which is splitted in to units: a DExx-box and a Helicase, spaced by a linker (Tang et al., 2010). The other domains which are adjacent to ATPase domain differ from one to another in the four families of chromatin remodelers. I will now discuss the peculiar features that distinguish the ISWI, CHD and INO80 families, and next I will focus specifically on the SWI/SNF family.

ISWI family remodelers: the ISWI ("Imitation SWItch") remodelers contain 2 to 4 subunits. They were initially identified in an *in vitro* screening aimed to test the nucleosome remodeling activity of *Drosophila* embryo extracts (Tsukiyama et al., 1995). The ISWI catalytic subunit, in addition to the ATPase domain, contains a SANT and a SLIDE domain, which together form a nucleosome recognition motif that binds to unmodified histone tails (Clapier et al., 2009). The ISWI ATPase subunit forms at least three different chromatin remodeling complexes, termed ACF, CHRAC and NURF, which were initially identified in *Drosophila*. The three homologous

human ISWI complexes differ for the number of subunits incorporated, and, depending on that, for their relative enzymatic activity. Additional subunits incorporated in the ISWI remodelers comprise DNA-binding/ histone fold domains (CHRAC 15-17), bromodomains (BPTF and ACF1) and DNA-binding domain (HMG1(Y)). The resulting enzymatic activity of the ISWI complexes ranges from chromatin assembly, to chromatin remodeling, to the maintenance of the euchromatin and of the heterochromatin, and finally to the assistance to the RNA Polymerase to facilitate or repress the transcription of target genes. The different activities depend from the attendant subunits (Clapier et al., 2009; Hargreaves et al., 2011).

CHD family remodelers: the CHD ("Chromodomain, Helicase, DNA binding") complexes contain from 1 to 10 subunits.They were firstly purified from *Xenopus leavis* (Marfella et al., 2007). The ATPase incorporated in the CHD chromatin remodeling complexes contains two characteristic chromodomain, arranged in tandem in the N-terminal region (Figure 3). This catalytic subunit is monomeric in the lower eukaryotes, but forms large complexes in vertebrates. Attendant subunits comprise DNA-binding and SANT domains-containing proteins, histone deacetylases (HDAC 1/2) and methyl-CpG binding proteins (MBD). The human CHD family contains nine members, named CHD1-9, which differ for the ATPase subunit incorporated. As a matter of fact, in human there are nine different CHD ATPase subunits, which are further categorized in three sub-groups

basing on their domain composition: CHD 1-2 ATPases contain a C-terminal DNA-binding domain, CHD3-4 ATPases lack the DNA-binding domain but have two N-terminal PHD fingers, while CDH5-9 have additional domains (Hargreaves et al., 2011). CHD family members are mostly studied for their roles in differentiation and in the regulation of genome stability (Marfella et al.,2007).

INO80 family remodelers: the INO80 ("INOsitol requiring 80") family contains more than 10 subunits, and they were firstly purified from *Saccharomyces cerevisiae* in a screening aimed to identify the regulators of phosphlipid biosynthesis (Ebbert et al., 1999). The human orthologs of the yeast complex contains the Ino80, SRCAP and SWR1 ATPases. The defining feature of the ATPase subunit incorporated in the INO80 remodelers is the presence of a "split" ATPase domain. The two sub-domains of the ATPase are divided by a long spacer region, to which the helicase-related Rvb1/2 and ARP proteins bind. The INO80 family members have different activities in human cells ranging from promotion of transcription to DNA repair. The SWR1 member is unique in its ability to restructure the nucleosome: as a matter of fact, it can replace the canonical H2A-H2B dimer with the variant dimer H2A.Z-H2B dimer, specifically in the chromosome regions that mark transcription start sites (Clapier et al., 2009).

The SWI/SNF chromatin remodeling complexes

The SWI/SNF (mating type SWItch or SWItching defective/Sucrose NonFermenting) chromatin remodeling complexes were initially described in *Saccharomyces cerevisiae* as positive regulators of *HO* and *SUC2* genes (Lander et al., 1992). Shortly after the identification of this complex in yeast, the homolog complexes present in fruit fly and human were identified. As in the *S.cerevisiae* protein, these complexes contains the ATPase subunit, which is homologous to the yeast SWI2/SNF2 protein, as well as other attendant subunits, which are homologous to the yeast SNF5, SWI3 and SWP73 subunits (Muchardt et al., 1999). Together, these four proteins form the core of the eukaryotic SWI/SNF chromatin remodeling complex. This consideration is based on two evidencies: the first one comes from the experimental observation that these four subunits are able to reconstitute the chromatin remodeling activity *in vitro* (Phelan et al., 1999); the second one comes from the observation that homologs of these four proteins have been identified in other species, such as *C. elegans*, *Xenopus*, chick and mouse, and also in *Arabidopsis*, further indicating that SWI/SNF function is very well conserved throughout evolution. On the other hand, some other proteins which are contained in the human SWI/SNF complexes are not present in other species. For example, BAF57, a protein contained in the human BAF complexes, has a counterpart in *Drosophila* but is not present in yeasts (Papoulas et al., 1998). Table 1 recapitulates the homologous subunits that constitute

the SWI/SNF chromatin remodeling complex in the different species. As reported by many papers, the human SWI/SNF complex can also be named BAF, which is the acronym of Brahma Associated Factors, by the name of one of the two alternative ATPases present in the complex, Brahma. As a consequence, the protein that are incorporated in the human complex are named BAF, but, alternatively, they could also retain the yeast homologs name or could be identified by the name SMARC (SWI/SNF-related, Actin containing, Regulators of Chromatin).

Depending on the different eukaryotic species, the SWI/SNF complex could contain from 8 to 14 subunits, generating a huge macromolecular complex. For example, a typical human BAF complex is composed by five yeast ortholog subunits (one of the two alternative ATPase subunits Brm or Brg1, BAF155/170, BAF60a/b/c, Baf53a/b and BAF47), plus several unique, human-restricted subunits (such as BAF57, BAF250a/b, BAF200, BAF45a/b/c/d) and monomeric nuclear actin. This huge “macromolecular machine” reaches the molecular weight of 2 MDa, which is larger than the calculated molecular weights of the known subunits, indicating that several additional interactors have yet to be identified (Hargreaves et al., 2011). It is long known that, following biochemical purification of the SWI/SNF-BAF complexes, it is possible to obtain several fractions, each one representing a single particular complex. Having identified the four constitutive subunits, the

combinatorial assembly of the others attendant subunits generates a huge amount of possible combinations, each one corresponding to one remodeler complex that may or may not be present in a particular cell of a tissue or in specific differentiation stage. The combinatorial generation of the different complexes reflects a specialization in the tasks accomplished by the different SWI/SNF-BAF multi-subunits enzymes (Wang et al., 1996).

One peculiar mechanism of catalysis distinguishes the SWI/SNF remodelers from the other three families of chromatin remodeling complexes. As a matter of fact, in addition to the mechanisms of action already discussed above, it has been observed that the SWI/SNF chromatin remodeling complexes can use another mechanism to promote the ejection of nucleosomes. It has been recently observed that SWI/SNF remodelers use the energy of ATP hydrolysis to catalyze a two-step reaction whose outcome consists in the eviction of one entire nucleosome. In the first reaction, the SWI/SNF complex induces the exit of the histone dimer H2A-H2B from one nucleosome, while in the second reaction substitutes the lost histone dimer with one identical dimer from the neighbouring nucleosome. The nucleosome which loses the dimer now lacks the H2A-H2B, and needs to get a new dimer that can be obtained from the next neighbouring nuclesome, as so the two reaction start again. The coupling of these two reactions creates a “wave” of chromatin remodeling, whose aftermath

consists in the localized displacement and in the movement of nucleosomes in a particular region of the chromosome (Dechassa et al., 2010). This observation indicates that, at least *in vitro*, SWI/SNF is able to catalyze the eviction of one nucleosome by exploiting the presence of another nucleosome localized in close proximity. This neighbouring nucleosome is requested for the catalysis to increase the processivity of the wave of displacement. When the “wave” of remodeling reaches a DNA region in which no more nucleosomes are present, the reaction stops, leaving a nucleosome with the H3 and H4 dimers lacking the H2A-H2B dimer. The H3 and H4 dimers are ejected from the DNA, so that the final outcome is the shifting of the nucleosomes in one direction and the ejection of one single nucleosome (Liu et al., 2010).

The two alternative ATPase subunits which are incorporated in the human SWI/SNF-BAF complexes are named Brahma (Brm, hBrm, SNF2α) and Brg1 (Brahma-Related Gene 1, hBrg1 SNF2β). These two proteins are 75% identical (Kadam et al., 2003) and elute in the same fraction upon purification by conventional chromatography (Muchardt et al., 1999). In order to exhaustively review the main features of the two ATPase, I will firstly describe their shared characteristics, and then I will highlight the differences that make each enzyme unique respect to its couterpart.

Brahma and Brg1, the two ATPase subunits of the human SWI/SNF-BAF chromatin remodeling complex, share common features

The evolutional divergence between the two ATPases Brm and Brg1 is present starting from mouse (mBrm and mBrg1). Other eukaryotic species, such as yeasts and *Drosophila*, have only one SWI/SNF ATPase subunit, and this observation, together with the high level of sequence homology, indicates that the two ATPase genes derive from the duplication of one single gene (Muchardt et al., 1999; Clapier et al., 2009; Hargreaves et al., 2011). The two human genes, *SMARCA2* and *SMARCA4* code for two proteins of about 200 kDa, respectively encoding Brm and Brg1. The two genes are mostly co-expressed in many cells, even if it is also important to note that some differences in the relative expression of the two genes have been observed during physiological and pathological processes. A comparison of the two protein indicates an high grade of homology, as reported in the schematic structure of the two protein (Figure 4). The two proteins share a common ATPase domain, which includes a Helicase-SANT (HSA) domain. In addition, they also share a common E7 sequence, a protein module that interacts with Retinoblastoma protein Rb (Kadam et al., 2003). The peculiarity of both these two enzymes consists in the presence of a C-terminal Bromodomain. Firstly identified in the fruit fly homolog (Lander et al., 1992), this domain is unique to the SWI/SNF family of ATPase subunits. In the other families of remodelers, the Bromodomain is not present in the ATPase

subunits, but is provided by attendant proteins. The bromodomain allows the interactions with the acetylated histone tails, an interaction important for the recrutiment of the SWI/SNF-BAF complexes in specific regions of the chromatin. It has also been demonstrated that the bromodomain can interact with specific acetylated histone residues, to promote localized gene activation (Clapier et al., 2009), and that bromodomain deletions inactivate the chromatin-targeting activity exerted by this protein module (Winston et al., 1999). Moreover, it has been demonstrated that the bromodomain can also interact with regulators of histone acetylation, such as HATs and HDACs, thus creating an auto-regulatory regulatory loop (Hargreaves et al., 2011).

Brahma and Brg1 are functionally and genetically not redundant

Even if Brm and Brg share a high grade of homology in the gene and in the polypetide sequences, the differences between them are more marked than their similarities. The experimental evidencies so far produced by different approaches sustain the hypothesis that Brm and Brg1 are not functionally and genetically redundant (Flowers et al., 2009). I will now highlight the main differencies and peculiarities of the two ATPases.

1) Brm and Brg1 interacts with different co-activators

Brm and Brg1 share an high grade of protein sequence homology (nearly 75%), but the two proteins differ in specific

regions of the polypeptide chain. The N-terminal region of Brg1, specifically the residues 1-282 (which contain the Proline/Glycine-rich domain), show a high grade of divergence from the corresponding region of Brm. Moreover, Brg1 contains an insertion of 33 aminoacids which spans part of the C-terminal E7 and the Lysine-Arginine rich regions. These divergences in the protein sequence are localized in regions which are specifically devoted to the protein-protein interactions, and it has been demonstrated that Brm and Brg1 interact with different proteins thanks to this domain differences. As a matter of fact, the N-terminal region of Brg1 allows to this protein to interact with zinc finger-containing proteins (such as transcription factors of the KLF and GATA families), while the same region allows the interactions between Brm and ankyrin repeats-containing proteins (such as proteins of the Notch pathways). These different interactions are functional to localize both the Brm- and Brg1-containing BAF complexes and the correct interactors on specific promoter and/or chromosome regions (Kadam et al., 2003).

2) Brm and Brg1 have different ATPase activities

From the biochemical point of view, Brm and Brg1 are incorporated in the SWI/SNF-BAF complexes in a mutually exclusive fashion and in 1:1 stoichiometry (one ATPase subunit per single complex). The purification of the complexes which includes Brm or Brg1 lead to the isolation of three different complexes: one contains Brm and the other two contain Brg1.

The three SWI/SNF-BAF complexes are endowed with ATPase-depedent chromatin remodeling activity and are able to increase the accessibility of the DNA to nucleases in an *in vitro* assay. However, the three complexes differ not only in subunit composition but also in the ATPase activity. The two Brg1-containing complexes (which differ in the composition of attendant subunits) show high ATPase activity, while Brm-containing complexes show a lower (a five fold decrease) enzymatic activity (Sif et al., 2001).

3) Brm and Brg1 play antagonistic roles during the development and the differentiation process

The difference in the activity of the two ATPases, which is mirrored by the difference in the interactors that they could bind, is strongly sustained by the observation of the phenotypes of the respective knock-out animal models. The two knock-out mouse models were generated by two independent groups at the end of the 90s, using a gene targeting approach. The knock out of mBrg1 in mice causes a very severe phenotype: the mBrg1$^{-/-}$ die during the peri-implantation stage (E3.5-E5.5), because of an impairment in the formation of the inner cell mass and the trophectoderm. The lethality in this particular stage is earlier than any other mammalian transcriptional regulator mutant so far reported. The mBrg1$^{+/-}$ survive, but these mice are predisposed to exencephaly and are prone to a wide variety of differentiated epithelial tumors. (Bultman et al., 2000). In these mice, there's no evidence of upregulation of

mBrm, which may be one possible way adopted by cells (Strobeck et al., 2002) to overcome the loss of part of mBrg1. On the other hand, the knock-out of mBrm has a quite mild phenotype. The mBrm$^{-/-}$ mice develop normally, with no defects in the embryonic development and no visible impairment in the adult life. They are heavier than normal littermates, because of an increased cellular proliferation due to a impairment on the retinoblastoma pathway. From the molecular point of view, the cells of the mBrm$^{-/-}$ mice upregulate the expression of mBrg1, to compensate mBrm loss (Reyes et al., 1998).

Taken together, these data suggest a scenario in which Brm and Brg1 play a role in different stages of development. As the loss of mBrg1 causes a very severe phenotype in the homozygous mice, it seems that Brg1 is requested in the earlier stages of development. On the other hand, loss of mBrm does not impair the development process, even if homozygous mice suffer of an increase in cellular proliferation. This hypothesis is strongly sustained by another very recent research (Flowers et al., 2009), which uses an *in vitro* differentiation system to study the different aftermath of the depletion of Brm and Brg1. Flowers and co-workers demonstrated that Brm-depleted cells, once stimulated to differentiate, show a very fast differentiation process respect to the Brg1-depleted cells. This observation has been verified also from the molecular point of view, by checking the upregulation of differentiation markers in the differentiated Brm-depleted cells and the presence of Brm- and Brg1-containing BAF complexes in the promoters of the

monitored markers. For what concerns the occupacy of Brm and Brg1 at the level of the promoters of the differentiation markers, the athours state that these promoters are targets of both Brm- and Brg1-BAF containing complexes. Brg1 complexes are present at the promoter level well before the induction of differentiation, an observation which is in line with the role of Brg1 during the earlier stage of development. Brm complexes become associated with the promoters in later stages, and exert a repressive role over Brg1-induced transcriptional activation. Once the cells start to differentiate and Brm is depleted, the Brg1-induced activation of transcription is no more inhibited, and the cells differentiate faster. These observations open up a new and exciting scenario, in which Brm and Brg, at the very beginning of the differentiation process, form a network in which they have antagonistic roles (Flowers et al., 2009). This hypothesis have been verified with other approaches, and all of them highlights that Brg1 is enriched during earlier stages of development and in undifferentiated cells, and, as the differentiation procedes, becomes downregulated in the differentiated cells. On the other hands, Brm is enriched in differentiated cells, but very low levels are detectable in the undifferentiated cells or cells from early stages of development (Muchardt et al.,1999). The levels of Brm and Brg1 are also regulated in an opposite way during the cell cycle. As a matter of fact, at the G2/M transition, both proteins are phosphorylated (Muchardt et al., 1996). This phosphorylation leads to Brm proteolytic degradation during

mitosis, while Brg1 remains stable. This process favours the formation of Brg1-containg BAF complexes during the early G1 phase (Sif et al., 1998).

1.3 An example of the crosstalk between the genetic and the epigenetic regulation of gene expression: the cotranscriptional pre-mRNA splicing

One of the most astonishing and interesting features of the biological systems consists in the perfectly coordinated regulation between the different processes that act together to sustain the correct functioning of the cell. After the discovery of the epigenetic mechanisms of gene expression regulation, it has became clear that these processes should work together with the “classic” mechanisms of gene expression regulation to properly orchestrate the thousands of activities that ensure cell viability. This observation has induced the researchers to try to understand the possible connections and the networks generated by the close crosstalk between the different processes.

It is quite impossible to exaustively describe all the connections made by the different processes of gene expression regulation, mostly because of their high grade of complexity and for their links with other cell activities which are not mentioned in this

thesis. In this paragraph, I will focus on one single network of connections between the different layers of gene expression regulation events: the cotranscriptional pre-mRNA splicing. The “cotranscriptionality” is one of the most well known examples of crosstalk between different gene expression regulation processes, because it contains different aspects of RNA biosynthesis and processing.

The general features of cotranscriptional pre-mRNA splicing

The concept of “cotranscriptionality” usually indicates a pre-mRNA splicing event in which introns are removed from the nascent pre-RNA transcript while transcription is still taking place, and RNA is still tethered to the DNA by the transcriptional machinery (Allemand et al., 2008). The concept that some events of pre-mRNA splicing can occur cotranscriptionally is quite recent. The first observation of such a splicing event dates back in the late 80s, when Beyer and colleagues described the cotranscriptional splicing in *Drosophila* embryos (Beyer et al., 1988). Later, cotranscriptional splicing has been demonstrated in mammals, in particular for one of the longest human genes, the human dystrophyn gene. For this specific gene, cotranscriptional splicing is mandatory, because if its splicing occurs independently from transcription, it would take 16 hours to be completed (Tennyson et al., 1995). This observation leads to the first feature of cotranscriptionality: as a general rule,

cotranscriptional splicing is mainly required for the long mammalian genes, to “boost” their splicing mechanisms (Luco et al., 2011).

As the different steps of cotranscriptional mRNA biogenesis occur at the same time (during transcription) and place (the actively transcribed gene), these processes become “coupled”. The coupling is the second, and most distinctive, feature of cotranscriptional splicing (Perales et al., 2009). For example, it has been observed that splicing factors are recruited on their sites in a cotranscriptional way, meaning that they are positioned on an actively transcribed gene at the same time that the RNA polymerase II (RNA Pol II) is still transcribing the gene (Luco et al., 2011). This close collaboration between different mechanisms indicates that the RNA Pol II is the key player of cotranscriptionality, because it is able to couple the transcription and the splicing reactions in a proper way. It has been demonstrated that, among all the RNA polymerases, only RNA Pol II is able to efficiently promote cotranscriptional splicing, mainly because of the distinctive features of its carboxy-terminal domain (CTD) (Bird et al., 2004). The importance of RNA Pol II in the regulation of the cotranscriptional pre-mRNA splicing is also sustained by other evidencies. As a matter of fact, it has been shown that this enzyme directly interacts *in vivo* with different splicing factors, such as SR proteins and U1snRNP (Das et al., 2006) and that deletion of the CTD affects splicing of different transcripts (Luco et al., 2011). As I

will describe in the next part of this paragraph, some interactions made by the CTD directly determine a shift in the enzymatic activity of the RNA Pol II, thus promoting cotranscriptional splicing events.

The third distinctive feature of cotranscriptional pre-mRNA splicing is its close collaboration with the epigenetic marks that are present in the chromosome. As a matter of fact, DNA methylation, histone modifications, and chromatin structure have all been linked to the regulation of cotranscriptional pre-mRNA splicing (Allemand et al., 2008). For example, it has been demonstrated that DNA methylation is nonrandomly distributed along the genome, and the different patterns of methylation not only mark regulatory regions, but also exons (Luco et al., 2011). On the other hand, histone post-translational modifications (PTMs) are emerging as important regulators of cotranscriptional splicing. Different histone PTMs have been mapped in the genome, creating an “histone code” which is functional to mark exons, which are enriched in the H3K36me3, H3K4me3 and H3K27me2 histone PTMs, and introns, which are instead enriched in the H3K39me3 histone PTM (Sims et al., 2007; Schwartz et al.,2009; Luco et al., 2010). Other researchers have demonstrated that also histone acetylation is a determinant of cotranscriptional splicing (Hnlicova et al.,2011; Gunderson et al., 2011). It has been proposed that different histone PTMs recruit distinct interactors, thus promoting or inhibiting the transcriptional and the splicing

activities (Luco et al., 2011). Finally, chromatin structure plays an important role in creating a scaffold on which the different actors of the cotranscriptional pre-mRNA splicing are co-localized and interact. As a matter of facts, the presence of a nucleosome upstream or downstream of a splicing acceptor site may affect the inclusion of one alternative exon, and nuclesome occupancy, that could change cotranscriptionally, generally enhances exon definition (Allemand et al., 2008; Tilgner et al., 2010).

Very recently it has also been demonstrated that there is another actor in the cotranscriptional regulation of pre-mRNA splicing: the splicing process itself. Kim and coworkers (Kim et al., 2011) demonstrated that there is a close correlation between cotranscriptional splice site selection and chromatin marks. In particular, they have demonstrated that the deletions of specific splicing sites in a reporter gene induce a change in the relative distribution of the methylations of the lysine 36 of histone 3 (H3K36), creating a histone methylation code which is very important for the crosstalk between chromatin and splicing. Similarly, inhibition of splicing causes a comparable repositioning of H3K36me3 along the same gene. Taken together, these observations let us hypothesize that a novel and unexpected mechanism of cotranscriptional splicing regulation should be added to the list, creating a new “layer” and greatly increasing the complexity of the entire network.

The mechanisms that regulate the cotranscriptional pre-mRNA splicing

As I have discussed in the previous part of this paragraph, the most important feature of cotranscriptionality is the crosstalk between the RNA Pol II, the splicing machinery and the epigenetic marks. These three components are all present at the same time and in the same place in the chromosome. The interplay between these "actors" generates a very complex network that researches have very recently tried to clarify. Essentially, two mechanisms have been proposed to describe the cotranscriptional pre-mRNA splicing: the kinetic coupling model and the recruitment model. Even if these two mechanisms focus on different events that take place during the cotranscriptional splicing and they stress the importance of different aspects of contranscriptionality, they are not mutually exclusive, but rather co-existent (Luco et al., 2011).

<u>The kinetic coupling model</u>

One of the already well established evidencies in the relatively new field of cotranscriptionality is the oustanding role played by the RNA Pol II. The CTD of the RNA Pol II can be seen as a "scaffold" to which different interactors bind, and in which take place the majority of the different events that regulate the contranscriptional splicing. The human RNA Pol II CTD is built of 52 repetitions of the heptad $YS_2PTS_5PS_7$. Among all the residues contained in one single heptad, the serines 2, 5 and 7 are the targets of phosphorylation events that are able to

regulate the enzymatic activity of the RNA Pol II during the different steps of transcription (Allemand et al., 2009). As a matter of facts, the phosphorylation events at the level of the different serine residues are able to induce a series of conformational changes that in turn regulate the processivity of the RNA Pol II, and create a “CTD code” that is “read” by different interactors. Upon transcription initiation, the polymerase is localized in the promoter-proximal regions and it contains the heptads phosphorylated on serine 5 (S5), a pattern of phosphorylation associated with slow processive enzymatic activity (Batsché et al., 2006). After the “promoter escape” phase, the phosphorylation in tre residue S5 decreases, and concomitantly the phosphorylation of serine 2 (S2) increases, catalyzed by the pTEFb kinase complex. At this stage, the polymerase is fully processive, and enters in the productive elongation phase of transcription. It has been proposed that the RNA Pol II phosphorylated at S2 residues corresponds to a very highly processive enzyme, which rapidly “reads” the gene and elongates the nascent transcript. As demonstrated by reporter minigene assays (de la Mata et al., 2003), this highly processive enzyme is able to identify only the strong 5’ alternative splice sites, while it “skips” the weak 5’ splice sites. In this scenario, the elongation rate is the main feature that dictates which exon is cotranscriptionally included in the nascent transcripts. This hypothesis is strongly sustained by the evidence that chemical or the genetic inhibition of RNA Pol II processivity promotes the inclusion of exons with weak 5’ splice

sites (de la Mata et al., 2003). In accordance with these results, it has been demonstrated that the S5 slow processive variant of RNA Pol II is detectable not only in the promoter proximal regions, but also in the coding regions of various genes, and in particular in the chromosome regions containing alternative exons (Batsché et al., 2006). Taken together, these results suggest a kinetic coupling model of transcription and cotranscriptional pre-mRNA splicing. This model suggests that the processivity of the RNA Pol II, which is in turn modulated by different phosphorylation events in the serines present in CTD heptads is the main factor that regulates the splicing outcome. In this case, the enzymes that modify the CTD code (such as pTEFb, Cdk7 and Cdk9) and the RNA Pol II intrinsic enzymatic activity are coupled, and the changes in the RNA Pol II processivity dictate the cotranscriptional inclusion of a specific alternative exon (Kornblihtt, 2006).

The recruitment model

The kinetic model of pre-mRNA cotranscriptional splicing suggests that the relative processivity of the RNA Pol II is the principal cause that dictates the inclusion of specific exons. Artificial manipulations of the CTD phosphorylation state have demonstrated that this is really the case. But how the CTD phosphorylation state changes during an *in vivo* physiological cotranscriptional splicing event? How can the changes in the phosphorylation patterns of the CTD be triggered in a specific exon, and not in others? Some researches suggest that the

specific events that allow to the kinetic coupling to take place are determined by the interaction of the CTD with different interactors and by the chromatin state. An interesting observation that moves in this way is that the splicing outcome could be altered by the interactions between the CTD with S5 phosphorylated and some elongation factors, such as SF2/ASF2. This interaction takes place at the level of the promoter, before the RNA Pol II is able to leave the promoter and to enter in the productive elongation phase. These interactions are directly dictated by the promoter structure, because the same reporter minigene under the control of different promoters gives rise to differentially spliced transcripts (Luco et al., 2011). Another observation that strongly sustains this recruitment model is that potent chemical modulators of chromatin state, such as TSA, are able to modulate exon inclusion, demonstranting that interaction of the transcription machinery with different histone codes is able to completely alter the splicing outcomes (Nogués et al., 2002; Hlincova et al., 2011). It has also been reported that human STAGA and other deacetylase complexes physically interact with U2 snRNP proteins (Martinez et al., 2001), and that the histone methyltransferase CARM1 interacts with U1snRNP proteins (Luco et al., 2011), which are both components of the splicesome. All these results suggest the intriguing hypothesis that the protein interactions cotranscriptionally made by the RNA Pol II and the cotranscriptional recruitment of different co-factors dictate the splicing outcome.

Brm as a regulator of cotranscriptional alternative splicing

So far, I have mentioned that both the processivity of the RNA Pol II and the presence of different co-factors are able to alter the cotranscriptional pre-mRNA splicing events. Now I will discuss one specific example that involves both these two activities and that clearly demonstrates that the kinetic model and the recruitment model could co-regulate the same splicing event. In 2006 it came out a very pioneering work, made by Christian Muchardt group (Batshcé et al., 2006). This oustanding research paper demonstrates that a specifc subset of human SWI/SNF-BAF chromatin remodeling complexes, in particular the ones containing Brm as the ATPase subunit, are able to modulate various cotranscriptional splicing events. In particular, Batsché and coworkers demonstrated that overexpression of Brm in human cells in culture caused the inclusion of alternative exons in the E-cadherin, BIM, Cyclin D1 and CD44 transcripts. It is noteworthy that all these events are inclusions of cassette exons which harbor weak 5' splice sites. As further proved by the silencing of different SWI/SNF-BAF subunits, the Brm-dependent alternative splicing events take place only when Brm (and not Brg1) is complexed into the chromatin remodeling complex. Mutations that inactivate the ATPase activity of Brm demonstrated that its effects on the alternative splicing are independent from its chromatin remodeling activity. They have demonstrated that Brm directly interacts with U1snRNP and with U5snRNP, which are both components of the splicesome involved in 5' splice-site

recognition. Moreover, Brm also interacts with Sam68, an exon inclusion enhancer. In turn, Sam68 is able to directly bind to the nascent transcripts to promote cotranscriptional exon inclusion. Finally, Batsché and coworkers proposed a molecular mechanism to explain how Brm can modulate alternative splicing. They proposed that the BAF/Brm-Sam68-U5 complex directly interacts with the RNA Pol II CTD, as the polymerase leaves the promoter. In the frame of this interaction, Sam68 constitutes a “bridge” between the huge BAF macromolecular complex and the nascent transcripts. The Brm-containing SWI/SNF complex then interacts with the RNA Pol II CTD, and the polymerase slows down its processivity by changing its phosphorylation state from S2 to S5 phosporylation. It is not clear if Brm directly induces the S5 phoshorylation or if Brm, being part of the huge SWI/SNF-BAF macromolecular complex, simply acts as a “roadblock” to the polymerase processivity. What is known is that the RNA Pol II CTD-S5 fraction is enriched in the genomic regions that contain the included alternative exons. It is possible that a yet unidentified protein, able to change the phosphorylation state of the CTD, directly interacts with Brm. In the frame of this molecular mechanism, Brm regulates the splicing outcomes by slowing down the polymerase processivity and favours the usage of weak splice sites by delaying the synthesis of downstream splice sites, thus facilitating the recognition of suboptimal exons.

The work published by Batsché clearly established a role for Brm in the regulation of some cotranscriptional splicing events

and demonstrated that both RNA Pol II processivity and recruitment of different cofactors are important for the splicing outcomes (Luco et al., 2011).

Another very recent paper published by the group of Neus Visa (Waldholm et al., 2011) moves in the same direction and establishes a new role for Brm in the choice of alternative 3' terminal exons and alternative polyadenylation (polyA) sites. Waldholm and coworkers used *Drosophila melanogaster* S2 cells to study the role of Brm in the cotranscriptional alternative polyadenylation choice. These cells are very useful to study the putative roles of Brm in these events, because *D.melanogaster* posses only Brm as the ATPase subunit of the SWI/SNF (BAP) complex. The paper takes advantage of some data collected from S2 cells in which Brm was depleted. The researchers noticed that Brm depletion in S2 cells, as well as the depletion of other BAP subunits, impairs both alternative promoter choice and alternative splicing events of two classes of genes: the genes involved in the dorso-ventral patterning and the genes involved in eye development, respectively. The paper then focuses on the splicing events, and in particular in the alternative last exon/polyA choice. The researchers validate three independent alternative polyadenylation events, specifically the events of the *CG388A*, *lola* and *mdg4* genes. Even if the results obtained from these three genes are not concordant (in two case, Brm depletion induces the choice of proximal exon/polyadenylation sites, while in the other promotes the choice of the distal exon), this paper represents

an important step in the understanding of the activity exerted by Brm in the regulation of alternative last exon choice.

1.4 An epigenetic view on Amyotrophic Lateral Sclerosis

Amyotrophic Lateral Sclerosis (ALS) is one of the most common adult-onset, progressive neurodegenerative disorders characterized by the selective loss of upper an lower motor neurons of the cerebral cortex, brainstem and spinal chord. The typical age of onset is between 50 and 60 years, and the disease is fatal within 2-5 years of onset, because it leads to a progressive muscular atrophy that eventually evolves in paralysis. The etiology of ALS is complex, and it combines both environmental and genetic causes. Although 90-95% of ALS cases are sporadic (sALS), the remaining 5-10% of the cases are familial and have genetic causes (fALS) (Rosen et al., 1993). From the molecular point of view, mutations in many genes have been to linked fALS, and the common features that connect most of them are their role in the RNA processing pathway and in the epigenetic regulation of gene expression.
I will begin this chapter focusing on SOD1, the most known protein that has been linked to fALS, and then I will review the most recent discoveries that connect the impairments of the epigenetic regulation of gene expression to the onset of ALS pathology.

The most common cause of human fALS: the mutation G93A in the SOD1 gene

Mutations in the SuperOxyde Dismutase 1 (*SOD1*) gene and their connection to fALS were firstly described in the 90s (Rosen et al., 1993). More recently, it has been proposed that only a small fraction of the ALS cases (nearly 15% of the fALS cases and 1-2% of all ALS cases) has *SOD1* mutations as the major cause of neurodegeneration (Dion et al., 2009).

SOD1 gene codes for an abundant, ubiquitously expressed cytosolic enzyme named SuperOxyde Dismutase1. This enzyme functions as a Cu/Zn-dependent homodimer and converts toxic superoxyde free radicals to molecular oxygen and hydrogen peroxide, preventing the generation of reactive oxygen species (ROS) (Rosen et al., 1993). The protein is composed of 153 aminoacids, and more than 125 different aminoacid changes have been linked to ALS. One of the most common human mutation present in the SOD1 enzyme and linked to ALS is the substitution of the Glycine in position 93 with an Alanine (G93A) It has been proposed that the pathological effect of this mutation does not cause a loss of function, but rather a gain of function, by which the protein acquires some toxic properties. The major proposed mechanisms of SOD1 G93A-dependent toxicity range from protein aggregation, to defect in axonal transport and to impairment in mitochondrial function (Rothstein et al., 2009). However, some recent reports tried to go deeper in the understanding of the SOD1 G93A-dependent pathological

mechanism. These reports took advantage of the generation of transgenic mice that express the human SOD1 G93A exclusively either in neurons or astrocytes. In both cases, no neurodegeneration was observed. This observation was further analyzed using chimeric mice that are characterized by the presence of a mixture of normal cells and cells epxressing the mutant SOD1. These experiments demonstrated that motor neurons expressing the mutated SOD1 survive longer without degeneration when surrounded by glial cells that express the “normal” SOD1. Moreover, the chimeric mice show signs of neuroinflammation and astrogliosis, which are two features present in the ALS patients. These observations suggest that the SOD1 G93A-dependent ALS is not cell-autonomous, and that also the behavior of surronding glial cells is important for the disease progression. The most recent view on SOD1 G93A-dependent ALS depicts a scenario in which the mutated gene seems to be more an initiator of the pathology than the central player (Lobsiger et al., 2007).

An epigenetic view on ALS

Currently, the molecular and translational medicine is discovering the increasing importance of the epigenetic factors in the onset of multi-factorial human pathologies. As a matter of fact, the epigenetic regulation of gene expression provides functional links between the environment and the alterations in gene expression that may lead to disease phenotypes. In the case of ALS, these studies are still ongoing, but some

interesting observations have already been made. First of all, it has been proposed that a number of epigenetic mechanisms (from nucleosomes repositiong to changes in the histone codes) are involved in the silencing of genes which are critical for the survival of the motor neurons (Dion et al., 2009). Moreover, the major role of RNA regulators of the proteins involved in ALS is now been reconsidered taking advantage of the new discoveries in the epigenetic field.

Recently, it has been reported a new and unexpected function for the SOD1 G93A protein. As a matter of fact, it has been demonstrated that the mutated SOD1 competes with the Hu-antigen R (also known as HuR or ELAVL1) for the binding to adenine/uridine rich sequences present in the 3' UTR region of vascular endothelial factor (VEGF), a gene which has been previously linked to the onset of ALS in mice and humans. In this scenario, SOD1 G93A would compete with a regulator of mRNA stability, causing the premature degradation of the mRNA encoding the VEGF protein (Li et al., 2009).

The discovery in 2009 of the mutations in the FUS and in the TDP43 proteins (Kwiatkowski et al., 2009; Vance et al., 2009) generates a promising advance in the identification of proteins implicated in the epigenetic component of ALS molecular pathogenesis. FUS and TDP43 are evolutionally conserved nuclear proteins, distinct but highly similar. They encompass two C-terminal RNA recognition motifs, by which they binds to hnRNPs, and a DNA binding domain. These proteins shuttle between the cytoplasm and the nucleus. When mutations are

present in their RNA binding motifs, they undergo cytoplasmatic retention and recruitment to toxic aggregates present in the cytoplasm of the motor neurons.

Another gene whose impared activity in ALS and neurodegeneration cases is now being reinterpreted is Elongator protein 3 (ELP3). Low expression and mutations in the ELP3 gene have been linked to increased susceptibility to sporadic ALS. It has been recently observed that ELP3 has a role in the epigenetic regulation of gene expression. First of all, ELP3 is component of the RNA Polymerase II elongator complex, which assists the polymerase during the productive elongation phase of transcription. Moreover, the elongator complex is also involved in the acetylation of H3 and H4 histone tails, an important process that regulates chromatin accessibility during transcription. Moreover, it has been discovered that ELP3 is directly linked to the post-transcriptional processing of the tRNA, especially in the conversion of uridine present in the tRNA wobble position (Blitterswijk et al., 2010).

Another protein mutated in ALS cases and that is now been linked to the epigenetic component of this neurodegenerative pathology is senataxin. Previosly, mutations in the *SETX* gene have been linked to the emergence of juvenile ALS. STX encodes the protein senataxin, an RNA/DNA helicase that resides both in the nucleus and the cytoplasm and that exhibits a high grade of homology with other helicases such as RENT1 ans Sen1p. Mutations in the catalytic domain of senataxin are involved in the genome-wide alterations of the RNA polymerase

II distribution, and are a cause of premature termination of transcription (Blitterswijk et al., 2010).
Taken al together, these data open up a brand new scenario in which impairments of the epigenetic gene expression regulation may be the cause of the emergence of both sporadic and familial ALS.

1.5 Scope of the thesis

My PhD project was aimed to investigate the putative roles played by the human protein Brahma (Brm, SNF2α) in neuronal cells in culture exposed to mitochondrial stress.

Chapter 1: General introduction
This chapter underlines the most recent topics regarding the epigenetic gene expression regulation, the “cotranscriptional” splicing mechanism and their possible links with Amyotropic Lateral Sclerosis (ALS) pathology.

Chapter 2: The chromatin-remodeling factor Brahma modulate sthe choice of alternative 3’ terminal exons
In chapter 2, I report the results that I have obtained regarding the molecular mechanism by which Brm differentially modulates the choice of 3’ terminal exons in a class of transcripts encoding proteins invoved in axon growth and guidance. My results point towards a novel Brm-dependent mechanism of cotranscriptional

alternative splicing regulation, which takes place at the level of the proximal last exon.

Chapter 3: Identification of Brm short, a class of isoforms of the chromatin-remodeling factor Brahma

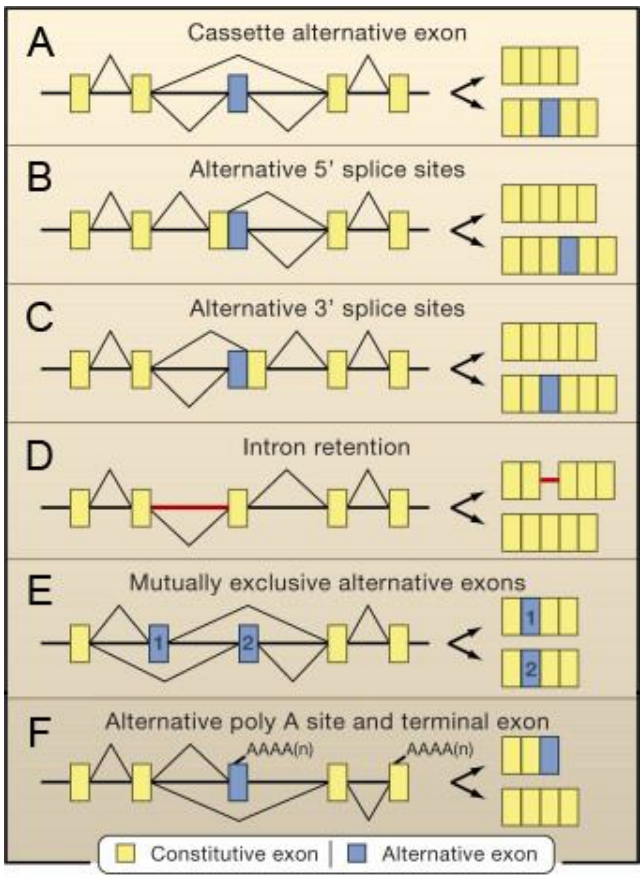
In chapter 3, I report the identification and the characterization of a class of “short” isoforms of Brahma (BrmS) and of the *SMARCA2* alternative promoter that controls their transcription.

Chapter 4: Conclusions and future perspectives

The last chapter summarizes the results obtained and underlines the possible future perspectives, focusing the attention on the putative translational applications of my research.

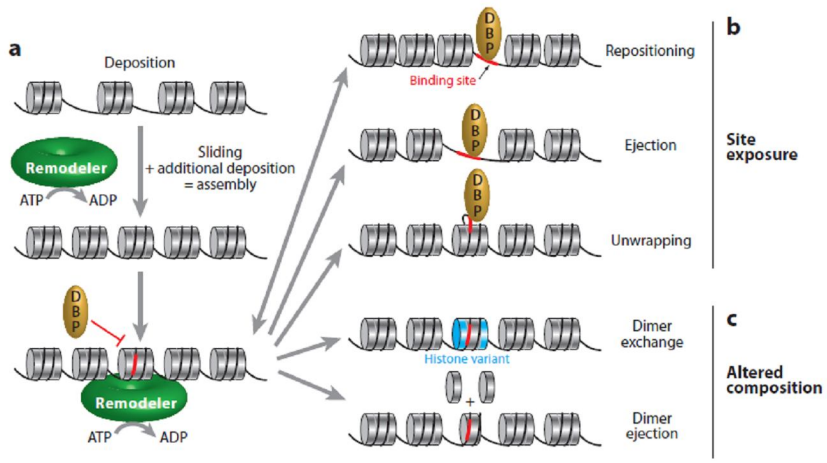
1.6 Figures and Tables

Figure 1: **The major alternative splicing patterns**



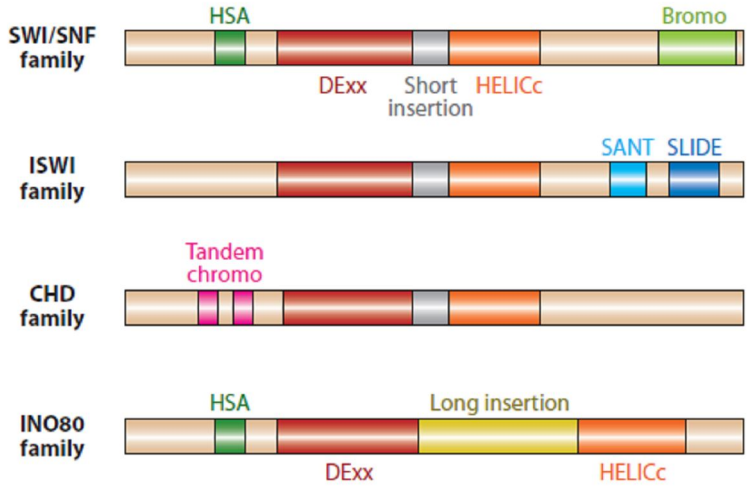
The figure (modified from Blencowe, 2006) depicts the different types of alternative splicing that are responsible for the generation of functionally distinct transcripts. Yellow boxes indicate the constitutive exons, while blue boxes indicate alternative exons.

Figure 2: **The mechanisms and the aftermath of chromatin remodeling**



The figure (from Clapier et al., 2009) describes how a chromatin remodeler (green) uses the energy derived from ATP hydrolysis to move already deposited histone octamers, generating room for additional deposition. The deposition of nucleosomes may "hide" the binding site for a specific DNA-binding protein (DBP) (a). Remodeler action results in two different events. The first is "site exposure" (b), in which a binding site (red) for a DBP, initially occluded by the histone octamer, becomes accessible by nucleosomal sliding (repositioning), nucleosomal eviction (ejection) or localized unwrapping. The second cathegory is "altered composition" (c), in which the nucleosome content is modified by dimer replacement (for example, the exchange of H2A-H2B with histone variants, in blue) or through dimer ejection.

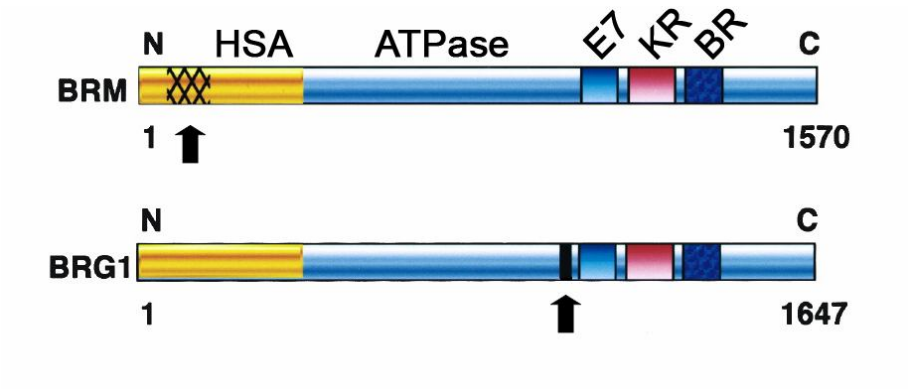
Figure 3: **The ATPases subunits which define the four families of eukaryotic chromatin remodeling complexes**



The four eukaryotic chromatin remodeling families contain an ATPase subunit characterized by an ATPase domain that is split in two parts: DExx (red) and HELICc (orange). What distinguishes each family are the unique domains residing within, or adjacent to, the ATPase domain. Remodelers of the SWI/SNF, ISWI, and CHD families each have a distinctive short insertion (gray) within the ATPase domain, whereas remodelers of INO80 family contain a long insertion (yellow). Each family is further defined by distinct combinations of flanking domains, such as Bromodomains (light green), helicase-SANT (HSA) domain (dark green), SANT-SLIDE module (blue) and chromodomains (magenta).

The figure is from Clapier et al., 2006.

Figure 4: **Schematic representation of human Brm and Brg1 proteins**



The figure (modified from Kadam et al., 2003) depicts a schematic representation of human Brahma (Brm, SNF2α, encoded by the *SMARCA2* gene) and Brahma-Related Gene 1 (Brg1, SNF2β, encoded by the *SMARCA4* gene) proteins, the two mutually exclusive ATPases that could be present in the human SWI/SNF (or Brahma-Associated Factors, BAF) chromatin remodeling complexes. The figure highlights the domain composition of the two proteins, which share the 75% of homology. HSA = Helicase-Sant domain, E7= domain of interaction with Rb protein, K/R= Lys/Arg rich domain, BR= bromodomain. The black arrows indicate the regions where the aminoacid sequences of the two proteins diverge and is nonhomologous.

Table 1: **The subunits of the SWI/SNF chromatin remodeling complexes in different species**

Subunits	Yeasts (SWI/SNF)	Drosophila (BAP)	Human (BAF)
ATPase	Snf2/Sw2	Brahma	hBrm (SmarcA2)
			hBrg1 (SmarcA4)
Constitutive subunits	Snf5	Snr1	hBAF155 (SmarcC1)
	Swi3	moira	hBAF170 (SmarCC2)
	Swp73	BAP60	hBAF60a/b/c (SmarcD1/2/3)
		BAP111	
Attendant subunits	ARP7	BAP55	β-Actin
	ARP9	BAP47	hBAF53a/b (Actl6a/b)
	Swi1		hBAF250a/b (Arid1a/b)
			hBAF45a/b/c (SmarCC1/2/3)
			hBAF47 (SmarcCB1)
			hBAF57 (SmarcE1)
			hBAF200 (Arid2)

The table indicates the subunits of the SWI/SNF chromatin remodeling complex in yeasts, *Drosophila melanogaster* and humans. Note that not all the subunits are evolutionally conserved. The name of the complex varies in the three analyzed species: in yeasts it is called SWI/SNF (SWItching defective/Sucrose NonFermenting), in fruit fly BAP (Brahma-Associated Proteins) and in humans BAF (Brahma-Associated Factors). For the human subunits are reported the name as “BAF” or as “SMARC” (SWI/SNF-related, Actin containing, Regulators of Chromatin).

1.7 References

Aigner S., Denli A.M., Gage F.H. *A novel model for an older remodeler: the BAF swap in neurogenesis*. Neuron, 2007; **55**:171-3.

Allemand E., Batsché E., Muchardt C*. Splicing, transcription, and chromatin: a ménage à trois*. Curr. Op. Gen. Dev., 2008; **18**:145-151.

Anko M., Neugebauer K.M. *Long noncoding RNAs add another layer to pre-mRNA splicing regulation*. Mol. Cell, 2010; **39**:833-834

Aoyagi S., Hayes J.J. *hSWI/SNF-catalyzed nucleosome sliding does not occur solely via a twist-diffusion mechanism*. Mol Cell. Bio., 2002; **12**:683-690.

Batsché E., Yaniv M., Muchardt C. *The human SWI / SNF subunit Brm is a regulator of alternative splicing*. Nat. Struct. & Mol. Biol., 2006;**13**:22-29.

Bernstein B.E., Meissner A., Lander E.S. *The mammalian epigenome*. Cell, 2007; **128**:669-681

Beyer A.L., Osheim Y.N. *Splice site selection, rate of splicing, and alternative splicing of nascent transcripts*. Genes Dev., 1988; **2**:754-765.

Bird A. *DNA methylation patterns and epigenetic memory*. Genes Dev., 2002. **16**:6-21.

Bird G., Zorio D.A., Bentley D.L. *RNA polymerase II carboxy-terminal domain phosphorylation is required for cotranscriptional pre-mRNA splicing and -3' end formation.* Mol Cell. Biol., 2004; **24**:8963-8969.

Blencowe B.J. *Alternative Splicing: new insights from global analyses*. Cell, 2006; **126**:37-47.

Blitterswijk M.V., Landers J.E. *RNA processing pathways in amyotrophic lateral sclerosis*. Neurogenetics, 2010; **11**:275-290.

Bultman S., Gebuhr T., Yee D., La Mantia C., Nicholson J., Gilliam A., Randazzo F., Metzger D., Chambon P., Crabtree G., Magnuson T. *A Brg1 null mutation in the mouse reveals functional differences among mammalian SWI/SNF complexes*., 2000;**6**:1287-1295.

Clapier C.R., Cairns B.R. *The biology of chromatin remodeling complexes*. Ann. Rev. Biochem., 2009; **78**:273-304.

Corona D.F., Tamkun J.W. *Multiple rolefor ISWI in transcription, chromosome organization and DNA replication*. Biochim. Biophys. Acta, 2004; **1677**:113-119.

Craig N., Cohen-Fix O., Green R., Greider C., Storz G., Wolberger C. *Molecular Biology – Principles of genome function*. 2010. Oxford University Press, New York.

Cremer T., Cremer C. *Chromosome territories, nuclear architecture and gene regulation in mammalian cells*. Genetics, 2001; **2**:292-301.

Das R., Yu J., Zhang Z., Gygi M.P., Krainer A.R., Gygi S.P., Reed R. *SR proteins function in coupling RNAP II transcription to pre-mRNA splicing*. Mol. Cell., 2006; **26**:867-881.

Davuluri R.V., Suzuki Y., Sugano S., Plass C., Huang T.H. *The functional consequences of alternative promoter use in mammalian genomes.* Trends Gen., 2008; **24**:167-177.

de la Mata M., Alonso C. R., Kadener S., Fededa J.P., Blaustein M., Pelisch F., Cramer P., Bentley D., Kornblihtt A.R. *A slow RNA polymeras II affects alternative splicing in vivo*. Mol. Cell, 2003; **12**:525-532.

Dechassa M.L., Sabri A., Pondugula S., Kassabov S.R., Chatterjee N., Kladde M.P., Bartholomew B. *SWI/SNF has*

intrinsic nucleosome disassemblly activity that is dependent on adjacent nucleosomes. Mol. Cell., 2010: **38**:590-602.

Dion P.A., Daoud H., Rouleau G.A. *Genetics of motor neurons disorders: new insights into pathogenic mechanisms*. Nature, 2009; **10**:769-781.

Ebbert R., Birkmann A., Schuller H.J. *The product of the SNF2/SWI2 paralogue INO80 of Saccharomyces cerevisiae required for the efficient expression of various yeast structural genes is a part of a high-molecular-weight protein complex*. Mol.Microbiol., 1999; **32**:741-751.

Eisen J.A., Sweder K.S., Hanawalt P.C. *Evolution of the SNF2 family of proteins: subfamilies with distinct sequences and functions*. Nucleic Acid Res., 1995; **23**:2715-2723.

Flowers S., Nagl N.G., Beck G.R., Moran E. Antagonistic roles for BRM and BRG1 SWI/SNF complexes in differentiation. J. Biol. Chem., 2009;**284**:10067-10075.

Gardiner-Garden M., Frommer M. *CpG islands in vertebrate genome*. J. Mol. Biol., 1987. **196**:261-282.

Gunderson F.Q., Merkhofer E.C., Johnson T.L. *Dynamic histone acetylation is critical for cotranscriptional spliceosome*

assembly and spliceosomal rearrangements. PNAS, 2011; **5**:2004-2009.

Hargreaves D.C., Crabtree R.C. *ATP-dependent chromatin remodeling: genetics, genomics and mechanisms*. Cell. Res., 2011; **21**:396-420.

Hnilicova J., Hozeifi S., Duskova E., Icha J., Tomankova T., Stanek D. *Histone deacetylase activity modulates alternative splicing*. PLOS One. 2011;**6**:1-11.

Kadam S., Emerson B.M. *Transcriptional specificity of human SWI/SNF BRM chromatin remodeling complexes*. Mol Cell, 2003; **11**:377-389.

Khorasanizadeth S. *The nucleosome: from genomic organization to genomic regulation*. Cell, 2004; **116**:259-272.

Kim S., Hyunmin K., Fong N., Erickson B., Bentley D.L. *Pre-mRNA splicing is a determinant of histone H3K36 methylation*. PNAS, 2011; **108**:13564-13569.

Kornblihtt A.R. *Chromatin, transcript elongation and alternative splicing*. Nat. Struct. Mol. Biol., 2006;**13**:7-9.

Kwiatkowski T. J., Bosco D. A., LeClerc A. L., Tamrazian E., et al. *Mutations in the FUS/TLS gene on chromosome 16 cause*

familial amyotrophic lateral sclerosis. Science, 2009; **323**:1205-1208.

Lander E.S., Linton L.M., Birren B., Nusbaum C., Zody M.C., Baldwin J., Devon K., Dewar K., Doyle K., et al. *Initial sequencing and analysis of the human genome*. Nature, 2001; **409**:860-921

Landry J.R., Mager D.L., Wilhelm B.T. *Complex controls : the role of alternative promoters in mammalian genomes*. Trends in Gen.,2003;**19**:640-648.

Laurent B. C., Yang X., Carlson M. *An essential Saccharomyces cerevisiae gene homologous to SNF2 encodes a helicase-related protein in a new family.* Mol. Cell. Biol., 1992; **12**:1893-1902.

Lessard J., Wu J.I., Ranish J.A., Wan M., Winslow M.M., Staahl B.T., Wu H., Aerbersold R., Graef I.A., Crabtree G.R. *An essential switch in subunit composition of a chromatin remodeling complex during neural development.* Neuron, 2007; **55**:201-215.

Li X., Lu L., Bush D.J., Zhang X., Zheng L., Suswam E.A., King P.H. *Mutant copper-zinc soperoxyde dismutase associated with amyotropic lateral sclerosis binds to adenine/uridine-rich*

stability elements in the vascular endothelial growth factor 3'-untranslated regions. J.Neurochem., 2009; **108**:1032-1044.

Liu N., Hayes J.J. *When push comes to shove: SWI/SNF uses a nucleosome to get rid of a nucleosome*. Mol. Cell, 2010; **38**:484-486

Lobsiger C.S., Cleveland D.W. *Glial cells as intrinsic components of non-cell-autonomous neurodegenerative disease*. Nat. Neurosc. 2007;**10**:1355-1360.

Luco R.F., Pan Q., Tominaga K., Blencowe B., Pereira-Smith O., Misteli T. *Regulation of alternative splicing by histone modifications*. Science, 2010; **327**:996-1000.

Luco R.F., Allo M., Schor I.E., Kornblihtt A.R., Misteli T. *Epigenetics in alternative pre-mRNA splicing*. Cell, 2011; **144**:16-26.

Marfella C.G., Imbalzano A.N. *The Chd family of chromatin remodeling complexes*. Mutat. Res., 2007;**618**:30-40.

Martinez E., Pahlan V.B., Tjernberg A., Lymar E.S., Gamper A.M., Kundu T.K., Chait B.T., Roeder R.G. *Human STAGA complex is a chromatin-acetylating transcription coactivator that interacts with pre-mRNA splicing and DNA damage-binding factors in vivo*. Mol. Cell Biol., 2001;**21**:6782-6795.

Muchardt C., Yaniv M. *ATP-dependent chromatin remodelling: SWI / SNF and Co. are on the job*. J. Mol. Biol., 1999; **293**:187-198

Muchardt C., Reyes J., Bourachot B., Legouy E., Yaniv M. *The hbrm and BRG-1 proteins , components of the human SNF / SWI complex , are phosphorylated and excluded from the condensed chromosomes during mitosis*. EMBO Journal. 1996;**15**:3394-3402.

Nag A., Narsinh K., Martinson H.G. *The poly(A)-dependent transcriptional pause is mediated by CPSF acting on the body of the polymerase*. Nat. Struct. Mol.Biol., 2007;14:662-669.

Nogués G., Kadener S., Cramer P., Bentley D., Kornblihtt A.R. *Transcriptional activators differ in their abilities to control alternative splicing*. J. Bio. Chem., 2002;**277**:43110 -43114.

Oesterreich F.C., Bieberstein N., Neugebauer K.M. *Pause locally, splice globally*. Trends Cell. Biol., 2011; **21**:328-335.

Papoulas O., Beek S. J., Moseley S. L., McCallum C. M., Sarte M., Shearn A., Tamkun J. W. The Drosophila trithorax group proteins BRM, ASH1 and ASH2 are subunits *of distinct protein complexes*. Development, 1998;**125**:3955-3966.

Perales R., Bentley D. *"Cotranscriptionality": the transcription elongation complex as a nexus for nuclear transactions*. Cell, 2009; **36**:178-191.

Phelan M. L., Sif S., Narlikar G. J., Kingston R. E. *Reconstitution of a core chromatin remodeling complex from SWI/SNF subunits*. Mol. Cell., 1999; **3**:247-253.

Reyes J.C., Barra J., Muchardt C., Camus A., Babinet C., Yaniv M. *Altered control of cellular proliferation in the absence of mammalian brahma (SNF2α)*. EMBO J., 1998;**17**:6979-6991.

Rosen D.R., Siddique T., Patterson D., Figlewicz D.A., Sapp P., Hentati A., Donaldson D., Goto J., O'Regan J.P., *et al*. *Mutations in Cu/Zn superoxyde dismutase gene are associated with familial amyotropic lateral sclerosis*. Nature, 1993; **362**:59-63.

Rothstein J.D. *Current hyphoteses for the underlying biology of amyotropic lateral sclerosis*. Ann. Neurol., 2009; **65**:3-9.

Saha A. Wittmeyer J., Cairns B.R. *Chromatin remodeling through directional DNA translocation from an internal nucleosomal site*. Nature Struct. & Mol. Biol, 2005;**12**:747-755.

Sif S., Saurin A.J., Imbalzano A.N., Kingston R.E. *Purification and characterization of mSin3A-containing Brg1 and hBrm*

chromatin remodeling complexes. Genes Dev., 2001; 15:603-618.

Sif S., Stukenberg P. T., Kirschner M. W., Kingston R. E. *Mitotic inactivation of a human SWI/SNF chromatin remodeling complex*. Genes Dev., 1998; **12**:2842-2851.

Sims R.J., Millhouse S., Chen C., Lewis B.A., Erdjument-Bromage H., Tempst P., Manley J.L., Reinberg D*. Recognition of trimethylated histone H3 lysine 4 facilitates the recruitment of transcription postinitiation factors and pre-mRNA splicing*. Mol. Cell, 2007; **28:**665-676.

Spies N., Nielsen C.B., Padgett R.A., Burge C.B. *Biased Chromatin Signatures around Polyadenylation Sites and Exons*. Mol.Cell.,2009;**36**:245-254.

Strobeck M.W., Reisman D.N., Gunawardena R.W., Betz B. L., Angus S.P., Knudsen K.E., Kowalik T.F., Weissman B.E., Knudsen E.S. *Compensation of BRG-1 function by Brm.* Biochemistry, 2002;**277**:4782-4789.

Schwartz S., Meshorer E., Ast G. *Chromatin organization marks exon-intron structure*. Nat Struct. Mol. Biol., 2009; **9**:990-995.

Tang L., Nogales E., Ciferri C. *Structure and Function of SWI/SNF Chromatin Remodeling Complexes and Mechanistic*

Implications for Transcription. Prog Biophys Mol Biol.; **102**:122-128.

Tennyson C.N., Klamut H.J, Worton R.G. *The human dystrophin gene requires 16 hours to be transcribed and is cotranscriptionally spliced*. Nat. Genet., 1995; **9**:184-190

Tilgner H., Guigò R. *From chromatin to splicing*. Epigenetics, 2010; **5**:180-184.

Tsukiyama T., Daniel C., Tamkun J., Wu C. *ISWI, a member of the SWI2/SNF2 ATPase family, encodes the 140 kDa subunit of the nucleosome remodeling factor*. Cell,1995;**83**:1021-1026.

Vance C., Rogelj B., Hortobagyi T., De Vos K., Lumi Nishimura A,,et al. *Mutations in FUS, an RNA processing protein, cause familial amyotropic lateral sclerosis type 6*. Science, 2009; **323**:1208-1211.

Waldholm J., Wang Z., Brodin D., Tyagi A., Yu S., Theopold U., Ostlund Farrants A.K., Visa N.. *SWI / SNF regulates the alternative processing of a specific subset of pre-mRNAs in Drosophila melanogaster*. BMC Mol. Biol., 2011;**12**:46-58.

Wang W., Xue Y., Zhou S., Kuo A., Cairns B.R., Crabtree G.R. *Diversity and specialization of the mammalian SWI/SNF complexes*. Genes Dev., 1996; **10**:2117-2130.

Winston F., Allis C.D. *The bromodomain: a chromatin-targeting module?* Nature, 1999; **6**:601-604

Witten J.T., Ule J. *Understanding splicing regulation through RNA splicing maps*. Trends in Gen., 2011;**27**:89-97.

Wolffe, A. *Chromatin, Structure and Function*, 1998. Academic Press, San Diego.

Yoo A.S., Crabtree E.G. *ATP-dependent chromatin remodeling in neural development*. Curr. Opin. Neurobiol., 2009; **19**:120-126.

Zlotorynski E., Agami R. *A PASport to cellular proliferation*. Cell, 2008; **134**:208-209.

Chapter 2

Manuscript in preparation

The chromatin-remodeling factor Brahma regulates the inclusion of alternative 3' terminal exons

Gabriele Alessandro Fontana[1], Aurora Rigamonti[1], Natalia Jurdziak[1], Silvia Carolina Lenzken[1], and Silvia M.L. Barabino[1*]

[1]Department of Biotechnology and Biosciences, University of Milano-Bicocca, Piazza della Scienza, 2, Milan, Italy

*Correspondence to: Silvia M.L. Barabino, Department of Biotechnology and Biosciences, University of Milano-Bicocca, Piazza della Scienza, 2; 20126 Milan, Italy; Phone: +39-02-6448 3352; Fax: +39-02-6448 3569; Email: silvia.barabino@unimib.it

Keywords: SMARCA2/Brahma – cotranscriptionality – 3' terminal exon – Bard1- Cstf

2.1 Abstract

The human protein Brahma, encoded by the *SMARCA2* gene, is one of the two mutually exclusive ATPase subunits of the mammalian SWI/SNF-BAF chromatin-remodelling complex. Brm-containing BAF complexes are enriched in neurons, where they play crucial roles in the regulation of genes involved in neuronal differentiation. Moreover, it has been reported that Brm associates with components of the spliceosome to regulate the inclusion of alternative internal exons.

While investigating with splicing-sensitive microarrays the gene expression changes triggered by mitochondrial stress, we found that Brm is strongly downregulated in SH-SY5Y human neuroblastoma cells overexpressing the SOD1 (G93A) protein, one of the genetic causes of Amyotrophic Lateral Sclerosis (ALS). We found that this downregulation is due to a mitochondrial stress-induced impairment in the *SMARCA2* promoter activity.

Among the genes deregulated at the splicing level by SOD1 (G93A) expression, we identified several targets that are regulated by alternative 3' terminal exon usage in a Brm-dependent manner. Specifically, we found that Brm promotes the skipping of the proximal terminal exon in five out of six genes that were analyzed. In order to define the molecular mechanism that allow to Brm to modulate the choice of alternative 3' terminal exons, we used one of these genes, *RPRD1A*, as a model. We found that Brm inhibits the choice of

the proximal *RPRD1A* last exon by directly localizing in its genomic region. In turn, the presence of Brm modulates the processivity of RNA Polymerase II, preventing the “terminal exon pausing” event. On the other hand, through a direct interaction with Bard1, Brm recruits the Bard1-Cstf complex on the *RPRD1A* proximal last exon, a complex known to inhibit the 3’ end processing of the pre-mRNA. These observations let us hypothesize an inhibitory role for Brm, which is exerted both at the level of the cotranscriptional choice of the proximal last exon and at the level of the 3’ end pre-mRNA processing.

2.2 Introduction

Alternative splicing (AS) of the pre-mRNA affects nearly 95% of human genes (Witten et al., 2011), and it results in the production of multiples mature mRNAs that vary in exon composition. Alterations in the AS process have been linked to the a vast number of pathologies, ranging from cancer (David et al., 2010) to neurodegeneration processes (Mills et al., 2011). AS is controlled at various levels, from *cis*-acting sequences (exonic/intronic splicing enhancers, ESEs or ISEs, and exonic/intronic splicing silencers, ESSs or ISSs) to *trans*-acting factors (splicing enhancers, such as the SR proteins, and splicing silencers, such as hnRNP ribonucleoproteins). The concerted effects of these regulatory elements have an impact

on the inclusion of one single alternative exon, resulting in the variation on the exon composition of the mRNA. From various recent reports, it has became clear that AS can occur cotranscriptionally (Kornblihtt, 2006), and that AS regulation strongly relies on promoter structure, on transcriptional activators (Nogués et al., 2002), on chromatin structure (Allemand et al., 2008) and, finally, on epigenetic marks (Luco et al., 2011). All these “layers” of regulation play crucial roles in the control of the RNA Polymerase II (RNA Pol II) processivity (Perales et al., 2009): as a matter of fact, the presence of the slow processive form of the polymerase (which is associated to the phosphorylation of serine 5 of the CTD hepteads, pSer5, in contrast to pSer2, a modification linked to the fast processive enzyme) has been linked to exon inclusion (de la Mata et al., 2003).

A recent report demonstrated that Brahma (Brm, SNF2α, encoded by *SMARCA2* gene), one of the two alternative ATPase subunits of the human SWI/SNF-BAF complex, regulates alternative splicing of internal cassette exons in various mRNAs (Batsché et al., 2006). This activity is exerted through an interaction with Sam68, and exon inclusion enhancer, and components of the spliceosome. These molecular interactions are functional in creating a “roadblock” to the polymerase, which in turn slows down its processivity and facilitates the recruitment of the splicing machinery to variant exons with sub-optimal splice sites.

The modulation of the elongation rate of the RNA Pol II is not only a determinant of the AS of internal exons, but also important for the definition of the 3' terminal (or last) exons. From a high-density tilling microarray global survey of *S.cerevisiae* nascent RNAs emerges that the RNA Pol II "pauses" not only at the very early steps of transcription, but also at the level of the last exons (Oesterreich et al., 2010). Specifically, the polymerase slows down its processivity (shifting from pSer2 to pSer5) roughly 250 bp upsytream to the polyadenylation (poly(A)) signal of the last exon. This "terminal exon pausing" triggers the inclusion of these exons in the mRNAs, and is typical of yeasts gene endowed with short (less than 750 bp) last exons. In higher eukaryotes, a pause in the RNA Pol II processivity has been demonstrated at the level of poly(A) sites (Nag et al., 2007), but the presence of such "terminal exon pausing" has never been demonstrated, for example, in human cells.

The modulation of alternative last exons inclusion in the final transcripts is emerging as a crucial mechanism to enhance the genome coding potential. As a matter of fact, the presence of a different 3' terminal exon, as well as a different 3'UTR, can strongly modulate the outcome of transcription (Campigli di Giammartino et al., 2011; Proudfott, 2011).

Here, we show that human Brm, one of the two alternative ATPase subunits of the SWI/SNF-BAF chromatin remodeling complex, is downregulated in two neuronal paradigms of mitochondrial stress. This downregulation is specific for Brm,

and not for Brg1, and it is triggered by a mitochondrial stress-induced impairment in the *SMARCA2* gene promoter activity. We also report that chronic mitochondrial stress (overexpression of the SOD1 (G93A) protein, the most common SOD1 mutation found of the genetic cases of familial Amyotropic Lateral Sclerosis, Shi et al., 2010) induces alterations in the usage of the last exons of a number of genes encoding factors involved in axon growth and guidance. Our analysis shows that Brm regulates the last exon of five out of six genes belonging to this class. In particular, Brm induces the inclusion of the more promoter-distal last exon in all these genes, strongly suggesting that a common molecular mechanism underlies these AS events. Exploiting *RPRD1A*, one of these genes, as a model, we found that Brm localizes at the level of its proximal 3'terminal exon. In the same genomic location, Brm interacts with the the Bard1-Cstf complex, a complex known to be involved in repression of the 3' end processing of the pre-mRNA (Kleiman et al., 2001). This processing repression could explain the preferential exclusion of the proximal exon. On the other hand, the loss of localization of Brm at the level of the proximal last exon correlates with an accumulation of "slow processive" RNA Pol II (providing the first evidence that the "terminal exon pausing" is present also in higher eukaryotes) and with the preferential inclusion of the proximal 3' terminal exon.

Our results suggest an inhibitory role for Brm in the inclusion of the proximal last exon, a function which is exerted through an interaction with the Bard1-Cstf complex.

2.3 Results

Brm is downregulated by mitochondrial stress

Previous results obtained in our laboratory (Lenzken et al., 2011) indicated that mitochondrial stress strongly alters the expression and the alternative splicing patterns of genes controlling neuritogenesis in two models of neurodegeneration: an acute stress model, namely the human SH-SY5Y neuroblastoma cells treated with Paraquat (PQ, N,N'-dimethyl-4,4'-bipyridinium dichloride, a chemical that inhibits mitochondrial complex 1) (Maracchioni et al., 2007), and chronic stress model, namely the same cell line overexpressing the SOD1 protein carrying the G93A mutation, one of the genetic causes of Amyotropic Lateral Sclerosis (ALS) (Shi et al., 2010). The exon-sensitive microarray analysis carried out on these two paradigms of mitochondrial stress revealed that deregulations in gene expression are accompained by profound alterations in the patterns of alternative splicing. Interestingly, one of the most downregulated genes in both paradigms appears to be *SMARCA2*, which encodes the human protein Brahma (Brm, SNF2α), one of the two alternative ATPase subunits present in the human SWI/SNF-BAF complex. To validate the *SMARCA2* microarray data both at the mRNA and

at the protein level, we carried out quantitative real time PCRs (qPCRs) and western blots, respectively. The extent of the *SMARCA2* downregulation expression reported by the microarray is in line with the qPCR validation carried out on both mitochondrial stress models (Figure 1A). The validation at the protein level confirmed that Brm is downregulated in both mitochondrial stress models (Figure 1B and C); in particular, the overexpression of the mutated SOD1 (G93A) protein (Figure 1C) caused a more pronounced downregulation of Brm respect to the cells treated with PQ (Figure 1B). In order to establish if the mitochondrial stress-induced downregulation is specific for this particular ATPase subunit, we monitored the protein levels of Brg1 (Brahma-related gene 1, SNF2β), the second mutually exclusive ATPase that could be present in the SWI/SNF-BAF complexes. Brg1 is not downregulated by acute (Figure 1B) or chronic (Figure 1C) mitochondrial stress, suggesting that Brm downregulation in response to mitochondrial stress is specific.

SMARCA2 **promoter activity is impaired by mitochondrial stress**

In order to investigate if *SMARCA2* gene downregulation could be caused by a transcriptional impairment triggered by mitochondrial stress, we cloned different fragments of its putative regulatory region upstream of the luciferase gene in a promoterless vector, and tested their activities in the SH-SY5Y cells. We started to analyze a region of 3400 bp, ranging from base -3344 to base +57 of the *SMARCA2* gene, respect to the

transcription start site (Figure 2A). The bionformatic analysis of this region identified peaks of evolutional conservation, a CpG island and a putative DNAse hypersensible region in the -749/+57 stretch, all elements that indicates the presence of a promoter sequence (Figure 2B). Our analysis did not identify any functional TATA box, but instead a high GC-content (especially in the Exon 1-proximal region), in the analyzed sequence, indicating that the promoter putatively present in the genomic region of interest may belong to the class of GC-promoters (*data not shown*). The UCSC Genome Browser, basing on submitted ChIP-Seq data obtained from different cell lines, identified in the -749/+57 region peaks of enrichment of trimethylation of histone 3 lysine 4 (H3K4me3), an histone mark usually associated to regulatory regions located in close proximity to transcription start sites (Kolasinska-Zwierz et al., 2009). Subsequent restriction endonuclease digestions allowed us to map in the -3344/-146 sequence the region which is responsive to mitochondrial stress: as a matter of fact, the luciferase activities of the constructs containing this region decrease in response to exposure to both acute and chronic mitochondrial stress (Figure 2C). All the luciferase constructs analyzed, except the one containing the -76/+57 region, show high promoter activity in neuronal cells and low activity in other cell lines (*data not shown*), indicating that these sequences may be endowed with neuron-specific promoter activity.

These results indicate that mitochondrial stress impairs *SMARCA2* transcription, causing the downregulation of Brm expression.

Brm modulates the choice of alternative 3' terminal exons in a subset of genes involved in axon growth and guidance

A previous report indicates that Brm is involved in the regulation of the cotranscriptional splicing of alternative cassette exons. In particular, it has been demonstrated that Brm overexpression favors the inclusion of alternative internal exons in the E-cadherin, BIM, Cyclin D1 and CD44 mRNAs, and that this activity is exerted through an interaction with the exon inclusion enhancer Sam68 and components of the spliceosome (Batsché et al., 2006). Our exon-sensitive microarray results indicate that the mitochondrial stress caused by the mutated SOD1 protein triggers the alteration of 405 alternative splicing events (ASEs), 35 of which affect the choice of alternative 3' terminal exons and alternative polyadenylation sites (*data not shown*). Having assumed that alternative splicing of 3' terminal exons is one of the major mechanisms that could alter the protein composition (Zlotorynski et al., 2008), we decided to asses if Brm could regulate this particular type of splicing event, and if SOD1 (G93A)-induced Brm downregulation could be the cause of the observed alterations in the last exon choice. To answer these questions, we reconstitued, by stably transfection, the expression of Brm in the SH-SY5Y SOD1 (G93A) cells; concomitantly, using short hairpin RNAs, we decreased Brm

expression in the SH-SY5Y SOD1 cells. Having considered that axon retraction is one of the first hallmarks of neurodegeneration (Schmidt et al., 2009), we initially focused our attention on genes involved or related to axon growth and guidance. Using a retro-transcription PCR (RT-PCR) approach, we firstly validated the alterations in the 3' terminal exon choice of six genes present in the SOD1 (G93A) cells (Table 1). In all these genes, chronic mitochondrial stress induces a common shift in the splicing patterns, resulting in the increase in the preferential choice of their proximal 3' terminal exons. We found that five of them respond to Brm overexpression and silencing. As an example, we show in Figure 3 the results obtained from the analysis of *RPRD1A* and *SLC6A15* genes. As the other genes examined, both *RPRD1A* and *SLC6A15* encode two transcripts, one terminating at a proximal exon (exon 8 for *RPRD1A* and exon 5 for *SLC6A15*) and another terminating at a distal exon (exon 9/10 for *RPRD1A* and exon 12 for *SLC6A15*). The rescuing of Brm expression in the SOD1 (G93A) cells caused an increase in the choice of the distal last exon in the *RPRD1A* and *SLC6A15* transcripts. Concomitantly, the depletion of Brm in the SOD1 cells induced the choice of the proximal last exons in the same transcripts, a result consistent to what is observed in the SOD1 (G93A) cells, where Brm expression is low. Overexpression and/or silencing of Brg1 did not impact the last exon choice in the same transcripts, indicating that the effects observed upon the modulation of Brm expression are specific.

Taken together, these results suggest that Brm modulates the choice of alternative 3' terminal exons in a class of transcripts encoding factors involved in axon growth and guidance. In five genes out of six, Brm induces the choice of the distal last exons, indicating that Brm could regulate the alternative 3' terminal exon splicing through a common mechanism.

Brm localization and modified CTD phosphorylation at the level of *RPRD1A* proximal last exon modulate its cotranscriptional splicing

A recent report demonstrate that brm, the *Drosophila melanogaster* homolog of Brm, is involved in the pre-mRNA processing of a a particular subset of transcripts encoding proteins involved in differentiation and development, both *in vitro* and *in vivo* (Waldholm et al., 2011). However, the mechanism by which Brm exerts this activity remains poorly understood. To investigate the molecular mechanism by which Brm modulates the choice of alternative 3' last exons and/or pre-mRNA processing, we firstly wondered wheter Brm directly localizes in the *RPRD1A* (or *p15RS*) gene, one of the genes identified in our screening and that we used as a model. We carried out Chromatin Immunoprecipitation (ChIP) assays, focusing our attention in the 3' regions of both constitutive (exon 7) and alternative (exon 8, 9, and 10) exons, and in intronic (intron 5 and 7) regions (Figure 4A). In the SH-SY5Y SOD1 cells, where Brm is expressed at physiological levels, we found a peak of Brm enrichment in intron 5, intron 7 and in the

alternative 3' terminal exon 8 (Figure 4B). Interestingly, the *RPRD1A* transcripts produced by these cells preferentially terminate at the distal exon 9/10 (Figure 3), suggesting that in these cells the choice of the "first come" proximal last exon 8 is inhibited and choice of the distal last exons 9/10 is favoured. In the SH-SY5Y SOD1 (G93A) cells, Brm displays a low localization in all the monitored regions of *RPRD1A* gene (Figure 4B), which is probably due to the strong downregulation of Brm expression present in these cells. In opposition to what happens in the SOD1 cells, The *RPRD1A* transcripts present in the SOD1 (G93A) cells preferentially terminate at exon 8, the "first come" last exon (Figure 3).

Recently, it has been demonstrated that RNA polymerase II (RNA Pol II) processivity plays an important role in the regulation of mRNA biogenesis (Oesterreich et al., 2011). The accumulation of the RNA Pol II phosphorylated at the serine 5 (pSer5) contained in heptads of the carboxy-terminal domain (CTD), a modification linked to slow processivity, is a general feature of promoter-proximal sites (Morris et al., 2005). However, some reports demonstrated that the RNA PolI-pSer5 also accumulates in some genomic regions containing alternative exons (Batsché et al., 2006) and that the slow processive RNA Pol II enhances exon inclusion *in vivo* (de la Mata et al., 2003). On the other hand, it has been shown that, in yeasts, the RNA Pol II slows down its processivity in proximity of the last exons (Oesterreich et al., 2010). The evidence of this "terminal exon pausing" in higher eukaryotes is still missing,

even if a pause in the RNA Pol II processivity has been linked to poly(A) processing (Nag et al., 2007).

These observations proned us to investigate if the slow processing RNA Pol II is present in the *RPRD1A* genomic region of interest. We carried out ChIP assays using an antibody that specifically recognizes the pSer5 modification of the polymerase (Batsché et al. 2006). In the SOD1 cells, the pSer5 RNA Pol II is present at the level of the 3' end of exon 10 (Figure 4C), the last exon preferentially included in the *RPRD1A* transcripts produced by these cells. On the other hand, in the SOD1 (G93A) cells, the peak of enrichment of the slow processive polymerase is shifted on exon7, intron 7 and the 3' end of exon 8 (Figure 4C). This result is in line with the observation that, in the SOD1 G93A cells, the *RPRD1A* transcripts preferentially terminate at exon 8.

Taken together, these results indicate that a change in the RNA Pol II processivity is involved in the cotranscriptional inclusion of the proximal, "first come" *RPRD1A* 3' terminal exon, providing an evidence that the "terminal exon pausing" is present also in higher eukaryotes. However, when Brm localizes on exon 8, the enrichment of the slow processive RNA Pol II is shifted on exon 9/10, resulting in the preferential inclusion of these last exons. These observations suggest that Brm has an inhibitory role on the production of transcripts terminating at the proximal last exon, and concomitantly favours the inclusion of the distal terminal exon. Brm may exert this activity by directly inhibiting

the contranscriptional splicing and/or the pre-mRNA processing machinery.

A Bard1/Cstf complex, involved in the inhibition of the 3' end processing, localizes at the level of *RPRD1A* exon 8

Our results suggest that Brm localizes at the level of the *RPRD1A* proximal last exon and inhibits its inclusion. This observation is consistent with a repression of the cotranscriptional splicing and/or the 3' end processing of the pre-mRNA, two activities which depend on the phosphorylation state of the RNA Pol II CTD (McCracken et al., 1997). 3' end pre-mRNA processing is carried out by the cleavage-polyadenylation specificity factor (Cpsf) and by the cleavage stimulation factor (Cstf) complexes, which respectively interact with the AAUAAA hexamer and with the G/U rich region contained in the core polyadenylation (poly(A)) signal present in the pre-mRNA. These two complexes cooperatively interact with each others (Takagaki et al., 2000) and with the RNA Pol II CTD (McCracken et al., 1997) to enhance the pre-mRNA 3' maturation. The cross-talk between the cleavage/poly(A) factors and the CTD of the RNA Pol II that takes place at the level of the last exon promotes the transcriptional termination. This observation constitutes the basis to the "torpedo model" of transcriptional termination, which states that transcriptional termination is a "kinetic competition" between the elongating polymerase and the exonuclease Xrn2, which degrades the nascent RNA. As the exonuclease catches up with the RNA Pol

II, the polymerase detaches from the DNA and transcription ends, resulting in the production of a pre-mRNA which contains the more proximal 3' terminal exon already processed (Proudfoot, 2011). It has been also shown that 3' end processing is negatively regulated by protein interactions, and that the interaction between the Cstf50 subunit of the Cstf complex and the BRCA1-associated RING domain 1 (Bard1) protein represses the 3' processing (Kleiman et al., 2001). To asses if the Cstf-Bard1 complex may localize in the region of interest and modulate the production of the two *RPRD1A* alternative-terminating variants in the SOD1 cells, we carried out ChIP assays. In the SOD1 cells, Bard1 accumulates at exon 8, the proximal 3' terminal exon which is preferentially excluded in these cells (Figure 5A). In the same cells, Cstf64, one of the constitutive subunits of the Cstf complex (Proudfoot, 2011), peaks in the same genomic region (Figure 5B). On the other hand, in the SOD1 (G93A) cells, Bard1 (Figure 5A) and Cstf64 (Figure 5B) fail to localize at the level of exon 8, and both are "shifted" to more promoter-proximal regions. The different genomic localization of Cstf64 in the two cell lines does not depend from a SOD1 (G93A)-induced alteration of the expression levels of this protein. This is is true also for the other Cstf and Cpsf subunits (*data not shown*). These observations are consistent with a Bard1-dependent inhibition of 3' processing of the *RPRD1A* proximal last exon. In order to further sustain the presence of the Bard1-Cstf50 interaction at the level of exon 8, we carried out a ChIP-ReChIP assay, using

an anti-Bard1 antibody as the first chromatin immunoprecipitating antibody, and then run a second ChIP on this Bard1-immunoprecipitated fraction using an anti-Cstf50 antibody. These experiments demonstrated that, in the SOD1 cells, Bard1 interacts with Cstf50 at the level of *RPRD1A* exon 8 (Figure 5C), creating a complex previously identified as a negative regulator of 3' end processing (Kleiman et al., 2001).
Taken together, these results suggest that, in the SOD1 cells, a Bard1-Cstf complex localizes at the level of *RPRD1A* terminal exon 8, inhibits the 3' pre-mRNA processing, and concomitantly favours the preferential inclusion of the downstream alternative last exon. This is not true in the SOD1 (G93A) cells, where Bard1 and the Cstf complex fail to localize at the level of the same exon. In this case, the 3' end processing is not inhibited, and the pre-mRNA containing exon 8 as the last exon is preferentially produced.

Brm recruits Bard1 at the level of *RPRD1A* exon 8

Brm contains several protein domains, each one endowed with a specific role in the catalytic activity and in the protein-protein interactions. In particular, the N-terminal domain of Brm interacts with ankyrin-repeats containing proteins. This interaction is specific for Brm, as Brg1 did not display this preference (Kadam et al., 2003). Having assumed that Bard1 contains an ankyrin-repeats domain (Fox et al., 2008) and that both co-localize in the *RPRD1A* exon 8, we assessed if Brm and Bard1 interact. We carried out ChIP-ReChIP experiments,

using an anti-Brm antibody for the first chromatin immunoprecipitation, and an anti-Bard1 antibody for the second ChIP. We found that, in the SOD1 cells, Brm and Bard1 interact at the level of *RPRD1A* exon 8 (Figure 6). Interestingly, at the level of this exon, Bard1 is also interacting with the Cstf complex via the Cstf50 subunits, generating a macromolecular complex known to inhibit the 3' end pre-mRNA processing. This result is in line with the observation that the inclusion of exon 8 is preferentially inhibited in the SOD1 cells. On the other hand, in the SOD1 (G93A) cells, where Brm levels are low and this protein does not localize in exon 8, no enrichment in the ChIP-ReChIP fraction is observed (Figure 5B), indicating that the inhibitory Bard1-Cstf complex is not recruited in this exon. This observation is consistent with the preferential inclusion of exon 8 in the *RPRD1A* transcripts produced by the SOD1 (G93A) cells.

Taken together, these results suggest that Brm localizes in the *RPRD1A* exon 8 and recruits the Bard1-Cstf complex. This complex inhibits the cotranscriptional processing of the pre-mRNA at the level of this terminal exon. This, in turn, would favour the production of the transcripts terminating at the distal last exon. When Brm is not localized in the "first come" proximal exon 8, as in the SOD1 (G93A) cells, the Bard1-Cstf inhibitory complex is not recruited and the 3' end processing takes place, resulting in the preferential choice of this last exon.

2.4 Discussion

Considering that roughly 95% of human genes are thought to encode two or more splicing isoforms (Witten al., 2011), alternative pre-mRNA splicing (AS) is considered a major source of diversity in the human proteome. In addition to changing internal “cassette” exons, AS can affect the 3’ end of the mRNA, by combining alternative splicing with alternative polyadenylation (APA) sites. Recent reports have demonstrated that APA is more widespread and complex than previously anticipated, and that nearly 50% of human genes encode multiple transcripts derived from APA (Tian et al., 2005). Two general classes of APA are generally recognized. The first is the “coding region APA” (CR-APA, known also as “3’ exon switching”), which is splicing-dependent, involves APA sites located in different exons and produces different protein isoforms. The second class is UTR-APA (or “tandem UTRs”), which involves APA sites located in the same exon, is splicing-independent and results in the production of transcripts encoding the same protein but with different lengths (Campigli di Gianmartino et al., 2011).

Many different studies have provided evidence that AS, as other processing events required for the synthesis of the mature mRNA, is coupled to transcription (Luco et al., 2011). The mechanisms proposed to explain how AS can be cotranscriptional involve both the differential recruitment of factors on the transcribing polymerase and the kinetic coupling

between transcription and splicing. (Allemand et al., 2008). In the kinetic coupling scenario, slowing down of the polymerase processivity favours the use of weak splice sites by delaying the synthesis of downstream splice sites, thus facilitating the recognition of suboptimal exons (Kornblihtt, 2006). Further evidence has been provided by the report that the catalytic subunit Brahma (Brm, Snf2a) of the chromatin remodelling complex SWI/SNF-BAF modulates AS by affecting RNA Pol II elongation rate (Batsché et al., 2006).

We show here that Brm is strongly downregulated in SH-SY5Y human neuroblastoma cells overexpressing the mutated SOD1 (G93A) protein, one of the genetic causes of familiar ALS. The downregulation of Brm is present both at the mRNA level (Figure 1A) and at the protein level (Figure 1C), but the extent of the protein deregulation is more pronounced respect to the one registered at the mRNA level, suggesting that other post-translational mechanisms may contribute to the regulation of Brm expression. Interestingly, some reports indicate Brm transcripts as targets for microRNA (miRNAs) regulation (Sakurai et al., 2011), and that epigenetic alterations in Brm expression are present in a wide variety of tumors (Gramling et al., 2011). Concerning the extent of Brm donwregulation in the two mitochondrial stress paradigms examined, we can observe that the deregulation is more pronounced in the chronic (SOD1 (G93A), Figure 1C) paradigm respect to the acute stress (PQ, Figure 1B) model. This could be due to the different natures of

the stresses, as well as to adaptation mechanisms triggered by stable expression of the mutated SOD1 protein. No alterations in Brg1 expression have been highlighted, indicating that the effects evoked by mitochondrial stress are specific for Brm.

We demonstrated that Brm donwregulation is triggered by a mitochondrial stress-induced transcriptional impairment, and that in the -3344/-146 regulatory region is present the “mitochondrial stress-responsive” sequence (Figure 2). In this region, which is endowed with “classical” regulatory sequence features (CpG islands and evolutional conservation, Figure 2B) the bionformatic prediction identified several putative binding sites for transcription factors, such as Notch family beta helix-loop-helix E-Box-binding proteins, SOX2 and LIM factors (Figure 2C). Interestingly, some of them result to be involved in the transcriptional regulation of genes expressed in neurons. This observation is consistent with the strong promoter activity displayed by the analyzed fragments of *SMARCA2* promoter in neuronal versus non-neuronal cells (*data not shown*). In the future, it will be of major interest to identify the transcription factor(s) involved in Brm transcriptional regulation, and to identify the one(s) responsible for the deregulated expression. Our results suggest also a novel transcriptional mechanism, other than miRNA-mediated silencing (Sakurai et al., 2011) by which Brm expression could be altered by “pathological” conditions.

Using the SOD1 (G93A) cells as a model of stable Brm depletion, we were able to identify several genes, deregulated by mitochondrial stress at the level of the choice of their alternative last exons, which respond to Brm overexpression and silencing (Table 1). All these genes fall in the CR-APA cathegory, because they have two different terminal exons endowed with distinct APA sites. We report that Brm commonly modulates the choice of their 3' terminal exons, always favouring the usage of the more distal termination (Figure 3B and C). Interestingly, some of these genes encode proteins involved in axon growth and guidance, a process which is impaired in neurodegeneration and in ALS pathogenesis (Schmidt et al., 2009). Among the selected targets, we characterized *RPRD1A* gene as a model. We found that Brm physiologically localizes in the genomic region containing the proximal last exon (*RPRD1A* exon 8) (Figure 4B). In the SOD1 (G93A), Brm fails to localize in the same region, maybe because of its low levels of expression. When Brm does not localize in the proximal terminal exon, we concomitantly registered an accumulation of slow processive RNA Pol II (pSer5, Figure 4C). This observation demonstrates for the first time the presence of "terminal exon pausing" in higher eukaryotes, a feature of the polymerase which has been previuosly described in yeasts (Oesterreich et al., 2010), and that consists in a decrease in the polymerase processivity well before to the APA site. The presence of slow processive RNA Pol II is consistent with proximal exon inclusion (de la Mata et

al., 2003) and correlates with the “first come, first served” AS theory (Beyer et al. 1988).

Our ChIP-ReChIP data (Figure 5 and 6) suggest that Brm, localizing at the level of the proximal last exon, recruits the Bard1-Cstf complex, a complex involved in the inhibition of 3’ pre-mRNA processing (Kleiman et al., 2001). This complex is possibly formed thought an interaction between the N-terminal region of Brm (which has been known to contain ankyrin repeats-binding domains, Kadam et al., 2003) and the ankyrin repeats domain of Bard1 (Fox et al., 2008). In the future, it will be interesting to map the regions of interactions present in the two proteins. When this macromolecular complex is formed, the proximal exon is preferentially not included in the final transcripts, and the distal termination (exon 9/10) is included in the mRNA.

In conclusion, we propose a molecular mechanism in which Brm modulates the choice of the last exon by acting as a coupling factor that links cotranscriptional splicing and 3’ end processing (Figure 7). In this scenario, Brm containing SWI/SNF-BAF complexes (which can be recruited onto specific terminal exons by a particular chromatin environment, as suggested by Allemand et al., 2009) are able to control the cotrascriptional inclusion of terminal exons by an inhibition of the 3’ end processing machinery, which in turn is due to a Brm-Bard1/Cstf interaction. Concomitantly, the RNA Pol II does not

change its processivity rate, and this effect is probably mediated by Brm itself. Even if we do not have any direct proof of this effect, the observation that Brm loss of localization in *RPRD1A* exon 8 is accompained by an accumulation of the RNA Pol II phosphorylated at serine 5 (Figure 4B and C), let us hypothesize this intriguing possibility. The Brm-mediated inhibition of the proximal 3' processing results in the preferential inclusion of the *RPRD1A* distal last exon, indicating that the polymerase has transcribed the locus and has reached the alternative 3' terminal exon. Following the "torpedo model" (which sees the termination of transcription as a "kinetic competition" between the polymerase processivity and the exonucleolitic degradation of the transcripts mediated by Xrn2) (Proudfoot, 2011), the kinetic competition is won by the polymerase. On the other hand, if Brm is not present, the pSer5 polymerase accumulates at exon 8, losing the kinetic competition with Xrn2. In this case, the exonucleolitic enzyme reaches the polymerase, and detaches it from the template, resulting in "short" transcript which includes the first come exon 8.

The previous results regarding Brm as a regulator of alternative splicing (Batsché et al., 2006) described this ATPase as an internal exon inclusion enhancer. Instead, our results demonstrated that Brm acts as an inhibitor of exon inclusion. However, our results, rather to be in contrast to the previous literature, add a new activity to Brm, which is specifically related

to the inclusion of the 3' terminal exons, a mechanism which is differently regulated respect to AS of cassette exons.

2.5 Matherials and Methods

Bionformatic analysis of human SMARCA2 regulatory region

The bionformatic analysis of the human *SMARCA2* putative regulatory region was performed using these prediction softwares: UCSC Genome Browser (for evolutional conservation, H3K4me3 enrichment and DNAse hypersensitivity analysis databases; http://genome.ucsc.edu/); EMBOSSCpGplot (for the CpG island analysis; http://www.ebi.ac.uk/Tools/emboss/cpgplot/index.html). For the putative transcription factors binding sites analysis, we used: TESS software (www.cbil.upenn.edu/cgi-bin/tess),MatInspector (www.genomatix.de/online_help/help_matinspector),and TFSearch (www.cbrc.jp/research/db/TFSEARCH.html).

Plasmids construction

The -3344/+56 region of human *SMARCA2* putative regulatory sequences (+1 denotes the transcription start site as reported by Ensembl Genome Browser) was amplified by PCR using the Phusion Hot Start High-Fidelity DNA Polymerase (Finnzymes) according to manufacturer's instructions. The human genomic DNA extracted from HeLa cells was used as the template in the

PCR reaction. The sequences of the oligos used in this PCR reaction were: hSMARCA2 promoter forward: 5'-GCAGTGAGCCAAGATCCCGCCA-3', hSMARCA2 promoter reverse: 5'- TCGCGAGGAGTGTGCTGGCTGA-3'. The resulting 3400 bp PCR product was purified and cloned into the pGEM®-T Easy vector system (Promega). The insert was verified by nucleotide sequencing. The *SMARCA2* promoter -3344/+57 region was then subcloned into the Sma I – Xho I sites of the pGL2 Basic vector (Promega), a promoterless vector which allows the cloning of putative promoter sequences upstream of the *Firefly* luciferase gene. In order to obtain shorter promoter sequences, subsequent 5' deletions were operated using restriction endonuclease digestions (Figure 2C); all the restriction enzymes and corresponding buffers were from New England Biolabs.

The cDNAs encoding the human Brm (Brahma, Genbank accession number NM_003070.3) and Brg1 (Brahma-Related Gene 1, Genbank accession number NM_003072.3) were a kind gift from Beverly Emerson (Kadam et al., 2003). Both cDNAs encode a N-terminal flag-tagged protein. These cDNAs were subcloned into the Sma I – Sal I sites of the pAD5-CMV-Wpre-PGK-Puro expression vector harboring a puromycin selectable marker, which was a kind gift from Sjaak Philipsen (Gutiérrez et al., 2007). These vectors were used for the overexpression experiments.

The interfering short-hairpin RNAs (shRNAis) targeting human *SMARCA2* and *SMARCA4* RNAs, were designed using the

SiDesign Center (Dharmacon). Primers were (capitol letters indicate the shRNA regions): shRNAi_*SMARCA2*_forward:5'-gatccccGCAGGAAACCGAAGAGAAAttcaagagaTTTCTCTTCGGTTTCCTGCtttttggaaa-3',shRNAi_*SMARCA2*_reverse: 5'-agcttttccaaaaaGCAGGAAACCGAAGAGAAAtctcttgaaTTTCTCTTCGGTTTCCTGCggg-3'; shRNAi_*SMARCA4*_forward: 5—gatccccGCTCAGAAGAAGAGGAAGAttcaagagaTCTTCCTCTTCTTCTGAGCtttttggaaa-3'; shRNAi_SMARCA4_reverse: 5'-agcttttccaaaaaGCTCAGAAGAAGAGGAAGAtctcttgaaTCTTCCTCTTCTTCTGAGCggg-3'. The shRNAis primers were aligned, phosphorylated, cut with Bgl II and Hind III, and subcloned into the pSuperPuro vector (Oligoengine). Following plasmid purification, the plasmids were verified by nucleotide sequencing. The pSuperPuro vector targeting the β2 T-Cell Receptor Beta, used as an unrelated control shRNAi, was a kind gift from Marc David Ruepp.

Cell cultures, transfections, and drug treatments

Human neuroblastoma SH-SY5Y cells untransfected or stably transfected with cDNAs coding for wild type SOD1 or the mutant SOD1(G93A) were maintained in Dulbecco's modified Eagle's Medium (Euroclone) supplemented with antibiotics (100 U/mL streptomycin and 100 µg/mL penicillin), 2,5 mM L-Glutamine and 10% Foetal Bovine Serum (all from EuroClone) at 37°C with 5% CO_2. SOD1-stably transfected cells were also maintained in the presence of 400 µg/ml Geneticin (G418 sulphate, Euroclone, prepared as 40 mg/ml stock solution a in

water). Cells were fed every 2–3 days and passed once a week. Paraquat (PQ, N,N'-dimethyl-4,4'-bipyridinium dichloride, Sigma, prepared as a 100 mM stock solution in H_2O) treatment was performed for 18 hours at a final concentration of 750 µM.

For the luciferase assay experiments, 1,5 x 10^5 SH-SY5Y cells per well were seeded in 24 multiwell, and the next day plasmids (a total of 0,75 µg) were co-transfected using Polyethylenimine (PEI, Sigma, 100 mM in H_20 pH 7.00) according to the manufacturer's instruction. For the luciferase assays experiments carried out in the presence of PQ, the drug was added to the medium 3 hours after the transfection. Transfected cells were maintained for 24 hours before lysis with Passive Lysis Buffer 1x (Promega).

For the overexpression and silencing experiments, 5 x 10^6 SH-SY5Y cells were plated on a 10 cm plate, and the next day plasmids (a total of 20 µg) were transfected using Polyethylenimine according to the manufacturer's instruction. The transfected cells were incubated for further 24 hours before the addition of 1µg/ml puromycin (Sigma, prepared as a 1mg/ml stock solution in water). In order to obtain a population of cells which retains the expression of the transfected vectors, the cells were maintained in constant Puromycin (1gµ/ml, Sigma) selection.

RNA extraction and reverse-transcription

SH-SY5Y cells were seeded on 10 cm plates, and 24 hours after plating the RNAs were extracted using TRIzol Reagent

(Invitrogen) and subsequently purified using silica membrane spin columns from the RNeasy Mini Kit (Qiagen). RNA quantity and purity were assessed using a NanoDrop® instrument (Thermo Fisher Scientific Inc.). 2 µg of total RNA were reverese-transcribed using the random hexamers-based High Capacity cDNA Reverse-Transcription Kit (Applied Biosystems), according to manufacturer's instructions. In order to inhibit RNAse activity, RNAsin Plus reagent (Promega) was added to the reverse-transcription reactions.

Quantitative Real-Time PCR assays (qPCRs)

In order to validate the microarray data regarding *SMARCA2* gene level downregulation, and to obtain data regarding the relative fold enrichment in ChIP experiments, Quantitative Real-Time PCR (qPCRs) amplifications were performed in a final volume of 25 µl with SYBR® Green qPCR master mix (Applied Biosystems), 1 µl cDNA diluted (1:50), and 0.2 µM of each primer. The primers used for human SMARCA2 gene were: forward: 5'- AAACCTGTAGTGAGCGATT-3', reverse: 5'-TCATCATCCGTCCCACTT-3'. Amplifications were performed in triplicate using an ABI PRISM 7500 real time system (Applied Biosystem). Only for gene expression experiments, normalization of cDNA loading was obtained running all samples in parallel using human *GAPDH* as housekeeping gene. The primers for *GAPDH* were: forward: 5'-ACGGATTTGGTCGTATTGGG-3', reverse: 5'-TGATTTTGGAGGGATCTCGC-3'. The amplification protocol

was as follows: an initial denaturation and activation step at 95°C for 10 min, followed by 40 cycles of 95°C for 15 s and 60°C for 1 min. After the amplification phase, a dissociation step as carried out at 95°C for 15 s, 60°C for 1 min, and 95°C for 15 s. Normalization of the target amplification over the normalizer was performed using Excel software.

Antibodies

The antibodies used for immunoblotting were: rabbit polyclonal anti SMARCA2 / BRM - ChIP Grade (ab15597, Abcam, 1:1000 dilution, also used for ChIP), rabbit polyclonal anti BRG1-ChIP grade (ab4081, Abcam, 1:1000 dilution), mouse monoclonal anti-β-actin (ab8226, Abcam, 1:5000 dilution).

Additional antibodies used for Chromatin Immunoprecipitation (ChIP) were: mouse monoclonal RNA Polymerase II CTD Phospho-Ser5 (H14, MMS-134R, Covance), rabbit polyclonal anti-Cstf64 (H300, C-28201, Santa Cruz), rabbit polyclonal anti-Bard1 (A300-263A, Bethyl Laboratories Inc.), rabbit polyclonal anti-Cstf50 (A301-251A, Bethyl Laboratories Inc.). As a control for ChIP, non-immune rabbit IgGs (PP64, Millipore) were used. As a control for ChIP-ReChIP experiments, a mouse monoclonal IgM that recognizes phoshoepitopes of SR proteins was used (a kind gift from Karla Neugebauer).

Protein extracts and immunoblotting

Cells on 10 cm plate were washed once in PBS1x (Euroclone) and the lysed in 1 mL of cold Lysis Buffer (Tris HCl 50 mM pH

7.5, NaCl 150 mM, 1% NP40, 5 mM EGTA, 5 mM EDTA) with protease inhibitors (Roche). The samples were then centrifuged at 15,000 rcf for 15 min at 4°C, and the supernatants were collected. An aliquot of the cell lysate was used for protein analysis with the Bradford kit (Bio-Rad) for protein quantification.

Proteins were separated in 6% to 10% SDS-polyacrylamide gels (classic Laemli conditions) and transferred to nitrocellulose membranes (Whatman GmbH), in Transfer Buffer (25 mM Tris, 192 mM Glicine, 20% Methanol), and trasfer was carried out overnight 15 Volts, at 4°C. Membranes were blocked using 5% non fat dried milk in PBST (0.1% (v/v) Tween 20 in 1x PBS) for 1 hour at room temperature and incubated with a primary antibody diluted in PBST supplemented with 5% non fat dried milk. Membranes were immunoblotted with antibodies of interest for 1 to 2 hours. The antibodies used are describe above in the “Antibodies” section. After washing 3 times with PBST, membranes were incubated with peroxidase-conjugated secondary antibody anti-mouse IgG (GE Healthcare, 1:10000 dilution) or anti-rabbit IgG (Pierce, 1:10000 dilution), in PBST with 5% non fat dried milk for 45 minutes at room temperature. After washing as above, the chemio-luminescent signals developed by ECL reagents (Millipore) were detected using films Amersham Hyperfilm ECL (GE Healthcare). Quantifications of the signals were performed using ImageJ software, normalizing each band intensity with the average grey

value and then with the corresponding housekeeping band intensity, thus obtaining relative intensity values.

Luciferase reporter assays

1,5 x 10^5 human neuroblastoma SH-SY5Y cells, untransfected or stably overexpressing the SOD1 WT or the SOD1 (G93A) variant, were seeded in 24 mutiwells as described above. 24 hour after the seeding, cells were co-trasfected with a total of 0,75 µg of plasmid DNA, composed by the *Firefly* luciferase vector, a constant amount of the *Renilla* expression vector pRL-TK (50 ng/well, Promega), and a variable amount of the empty pGL2 Basic vector. In this manner, all the vectors containing different stretches of the putative *SMARCA2* promoter sequences were transfected as equimolar amounts. As a positive control, pGL2-Promoter vector (Promega, a vector in which luciferase expression is driven by SV40 promoter) was transfected in parallel wells. The experiments were carried out in duplicate. Luciferase expression was maintained for 24 hours, and then cells were lysed in 100 µg/well of the Passive Lysis Buffer 1x according to manufacturer's instructions. To obtain a complete lysis, the cells were subjected to a freeze-thaw cycle. 10 µl of the lysates were subjected to the luciferase assays, which was carried out using the Dual-Luciferase Reporter Assay System (Promega) and a Berthold luminometer (Berthold Inc.). The relative luminescence units (RLUs) were obtained normalizing the *Firefly* luciferase readings to the corresponding *Renilla* luciferase readings. In the graphs, the

RLUs of the SH-SY5Y untreated or SOD1 WT overexpressing cells are set to 1, while the RLUs of the Paraquat-treated or SOD1 G93A overexpressing cells were expressed as a fraction of the corresponding values.

Validation of the alternative termination events by retro-transcription PCR (RT-PCR)

Validation of the alternative termination splicing changes predicted by microarray experiments (Lenzken et al., 2011) was performed by retro-transcription PCR (RT-PCR) analysis, using a three gene-specific primers able to discriminate between the two different alternative termination events. The target genes are listed in Table 1. cDNA synthesis was performed as described above. For primers sequences, see Table 2. Assay conditions were optimized for each gene with respect to primer annealing temperatures, primer concentration, and MgCl2 concentrations. The number of amplification cycles used for each reaction was determined to ensure that transcript amplification was within a linear range (22 to 36 cycles). PCR products were separated by electrophoresis on 2% agarose gels. The amplified PCR products were extracted from gel, cloned in th pGEM T-Easy vector system and sequenced to ensure the identity of the PCR products. Quantification was performed with Bioanalyzer 2100 (Agilent Technologies). Statistical analysis was performed with Prism GraphPad Instant Software (GraphPad Software Inc.).

Chromatin Immunoprecipitation (ChIP) and ChIP-ReChIP

10^7 SH-SY5Y cells overexpressing the wild-type or mutated SOD1 or G93A protein were cross-linked with 1% Formaldehyde (Molecular Biology Reagent, Sigma) for 10 minutes at room temperature. After quenching with 125 mM Glycine (Sigma), cells were washed twice with cold PBS1x and then collected by screaping. Pellets were lysed in 750 µl of FA lysis buffer (50 mM HEPES-KOH pH 7.5, 140 mM NaCl, 1 mM EDTA pH 8, 1% Triton X-100, 0,1% Sodium Deoxycholate, 0,1 % SDS) supplemented with fresh protease inhibitors for 1 hours on ice, and then sonicated with a Branson 250 sonifier (Branson Inc.) to shear chromatin to a final average size of 400 base pairs. After checking the average chromatin fragments size, aliquots corresponding to $5x10^6$ flash-frozen in liquid nitrogen, and maintained at -80°C. For each immunoprecipitation, an aliquot of input, corresponding to the 1% of the total genomic DNA present in the reaction, was taken out. Immunoprecipitation was carried out using the Chromatin IP Assay Kit (Millipore), following manufacturer's instructions. Briefly, immunoprecipitation was carried out overnight, using 5 µg of the indicated antibodies for $2x10^6$ cells. As a control, a parallel sample, in which immunoprecipitation was carried out with non-immune isotype antibodies were performed. Beads were washed extensively, and bound matherial was eluted with 500 µl of Eluition Buffer (1% SDS, 100 mM $NaHCO_3$) for 1 hour at room temperature. Crosslink was reversed at 65°C overnight, with the addition of 250 mM NaCl and 2 µg/ml Proteinase K

(Sigma). De-crosslinked DNA was isolated by phenol-chloroform method and the precipitated, by the addition of 300 mM NaAcetate and 1ml of Ethanol, supplemented with 5 µg of tRNA (Sigma) as a carrier. After centrifugation, pellets were washed with a 70% Ethanol solution and spinned again. Pelleted DNA was resuspended in 100 µl DNAse-free water. 3 µl of the sample were subjected to Quantitative Real Time PCRs, with the conditions described above. For primers sequences Table 3. Primers were designed in the human *RPRD1A* genomic locus. Data were normalized by the Fold Enrichment Method, calculating signals over IgGs background, as follows: fold enrichment = $2^{-(\Delta Ct\ antibody-\ \Delta Ct\ IgG)}$.

For double-ChIP (ChIP-ReChIP) experiments, eleuates were diluted in the ChIP Diluition buffer (Millipore) to reduce the SDS concentration to 0,1% (w/v). Then, the second ChIP was performed.

2.6 Figures and Tables

Figure 1: ***SMARCA2*/Brm is downregulated by mitochondrial stress**

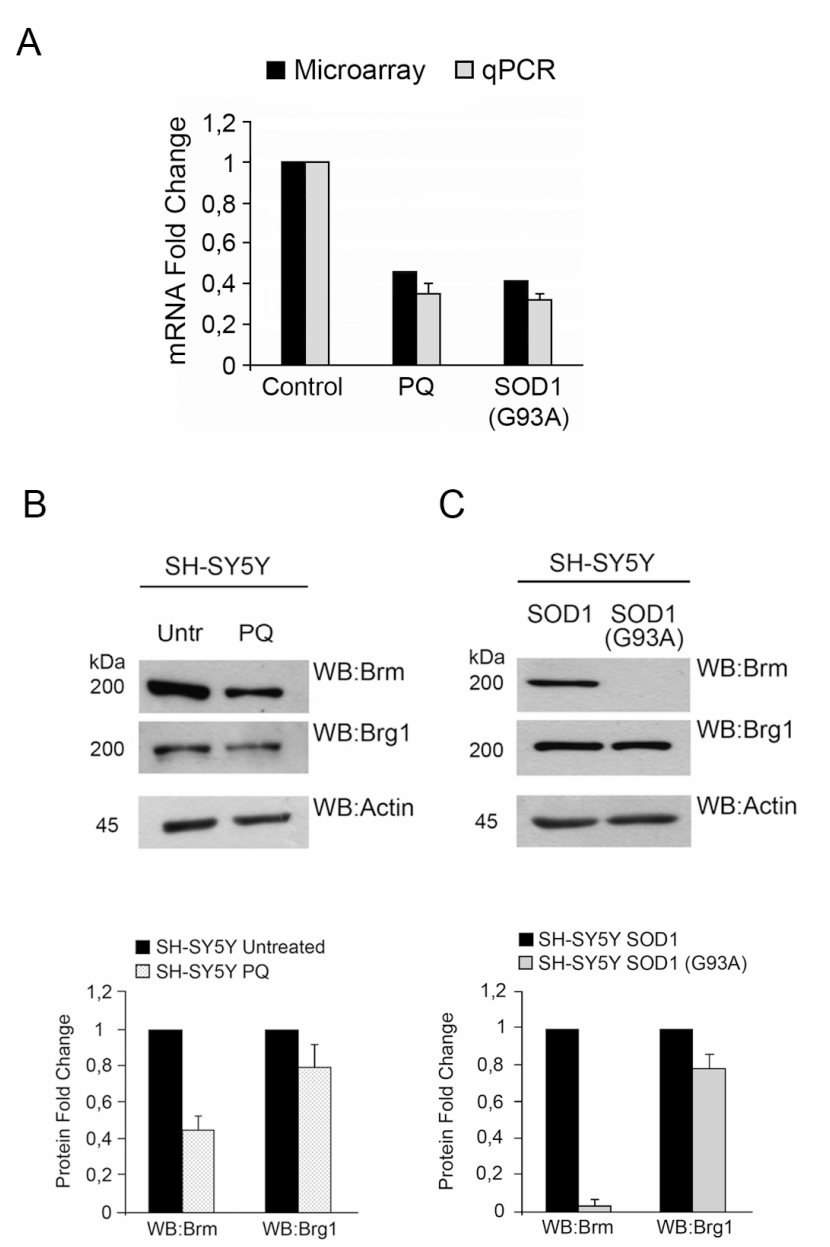


Fig. 1A: Quantitative real time PCRs (qPCRs) were carried out to validate the microarray data obtained from the neuronal acute stress paradigm (SH-SY5Y cells treated with Paraquat, PQ) and the chronic paradigm (the same cell line overexpressing the mutated SOD1 protein). The *SMARCA2* mRNA levels of the control cells (untreated cells and cells overexpressing the SOD1 wild-type protein) were set at 1, while the relative mRNA levels of the PQ-treated or SOD1 (G93A) overexpressing cells were expressed as fractions. The black bars represent the data collected by the microarray (n=5), while the grey bars represent the data collected by qPRCs (n=3). The experimental variability is expressed as standard deviation.

Fig. 1B: Western blots were carried out to validate the microarray data. The figure shows the results obtained from the SH-SY5Y cells treated with PQ. *Upper panel*: a representative gel. *Lower panel*: densitometry quantification (n=3). The protein levels of the untreated were set at 1, while the relative protein levels of the PQ-treated cells were expressed as fractions.

Fig. 1C: Western blots were carried out to validate the microarray data. The figure shows the results obtained from the SH-SY5Y cells overepressing the mutated (G93A) SOD1 protein. *Upper panel*: a representative gel. *Lower panel*: densitometry quantification (n=3). The protein levels of the SOD1 cells were set at 1, while the relative protein levels of the SOD1 (G93A) cells were expressed as fractions.

Figure 2: ***SMARCA2*** **promoter activity is impaired by mitochondrial stress**

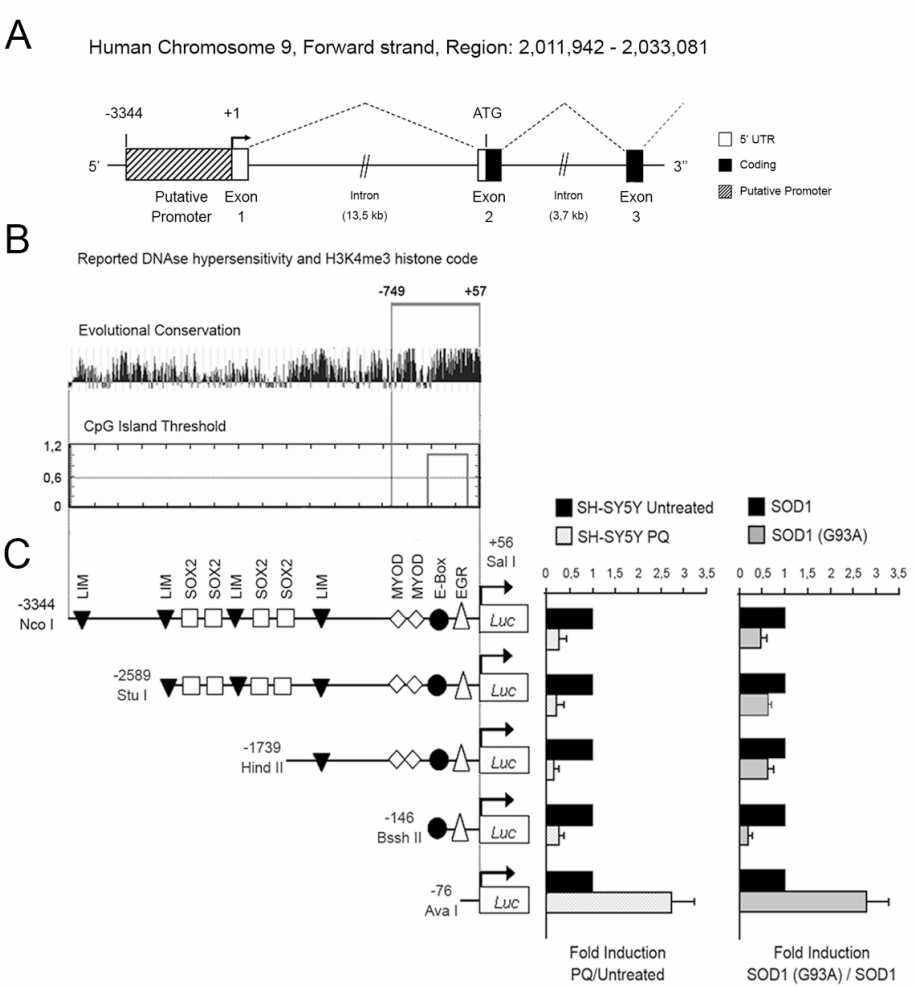


Fig 2A: Schematic representation of the region of interest of the human chromosome 9. White boxes indicates the 5'UTR regions, black boxes indicates the coding regions, while introns are represented by lines. The putative *SMARCA2* regulatory region examined in this paper is indicated by the striped box.

Fig. 2B: Results of the bioinformatic study carried out on the -3344/+56 region located at the 5' side of human *SMARCA2* gene. The evolutional conservation was evaluated using UCSC Genome Browser, the presence of the Dnase hypersensitive region and H3K4me3 occupancy were evaluated using Ensembl Genome Browser, while the CpG islands content was evaluated using Emboss CpGPlot.

Fig. 2C: Results of the luciferase (*luc*) assays *Left*: schematic representations of the luciferase constructs used in this study. The numbers indicate the relative distance from the transcription start site (+1). The restriction sites for the enzymes used for the cloning are indicated. The panel shows also the most relevant predicted binding sites for transcription factors (evaluated using TFSearch, MatInspector and TESS). *Right*: results obtained from the PQ (n=3) and the SOD1 (G93A) (n=3) mitochondrial stress paradigms. The relative *luc* units (RLUs) of the controls are set to 1, while the RLUs obtained from the PQ-treated and the SOD1 (G93A) overexpressing cellsls exposed to mitochondrial stresses are expressed as fractions. The experimental variability is expressed as standard deviation.

Figure 3: **Brm modulates the choice of the alternative 3' terminal exons of the *RPRD1A* and *SLC6A15* genes**

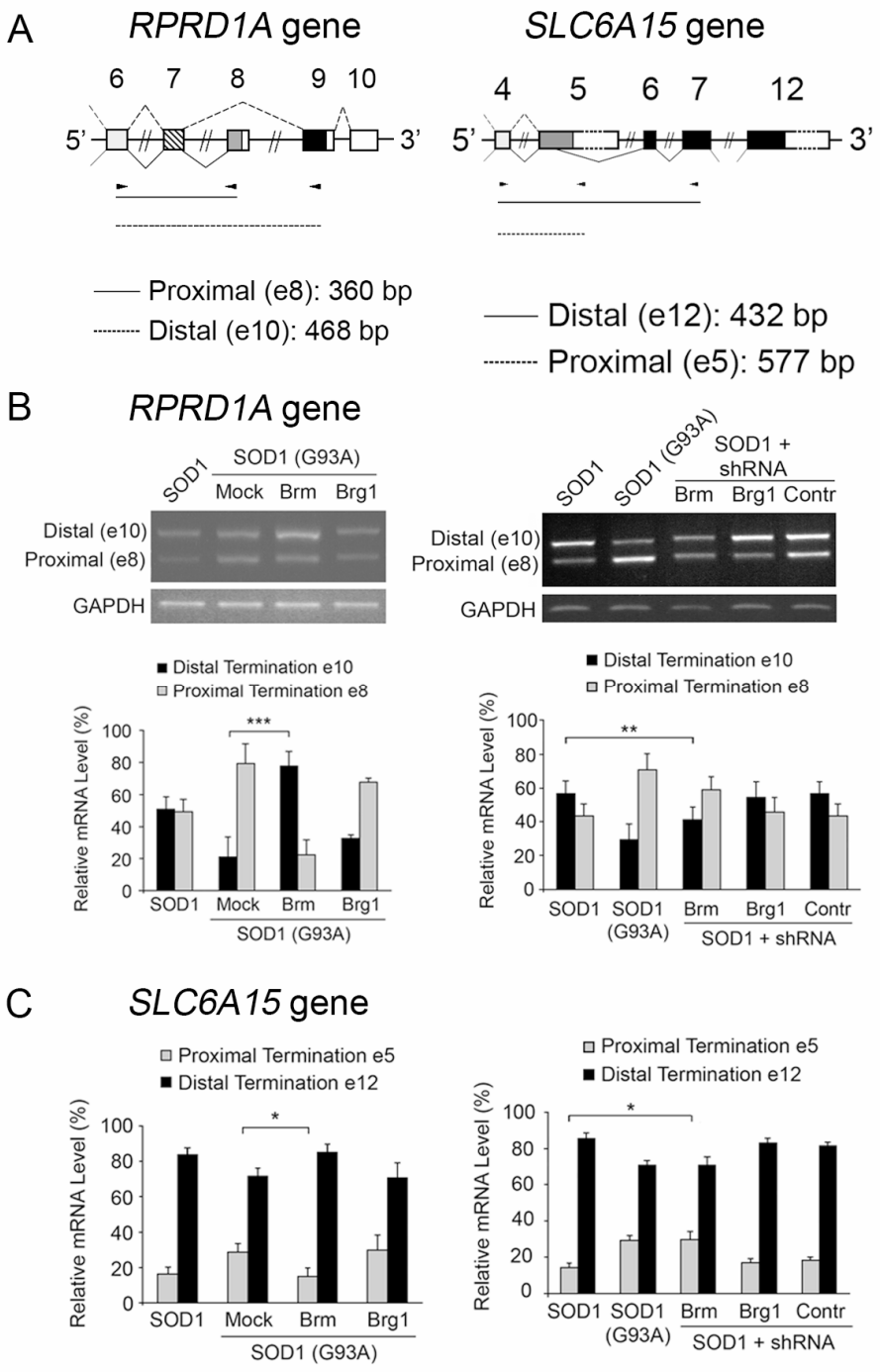


Fig. 3A: Schematic representation of the human *RPRD1A* and *SLC6A15* genes. The coding exons are represented as black or grey boxes, while the 3' UTRs are represented by white boxes. The lines represent the two different patterns of splicing. The the localization of the primers (arrows) used for the retro-transcription PCRs (RT-PCRs), and the relative PCR products are indicated.

Fig 3B: Results of the RT-PCRs carried out on *RPRD1A* gene. *Left*: results of Brm rescuing in the SOD1 (G93A) cells. *Right*: results of the shRNA-mediated silencing of Brm in the SOD1 cells. *Lower panels*: quantifications of six independent experiments. The experimental variability is expressed as standard deviations T-test: *= $p < 0,05$, ***= $p < 0,001$. *Upper panels*: representative gels. GAPDH was used as a loading control.

Fig 3C Results of the RT-PCRs carried out on *SLC6A15* gene. *Left*: results of Brm rescuing in the SOD1 (G93A) cells. *Right*: results of the shRNA-mediated silencing of Brm in the SOD1 cells. The graphics show the results of the quantifications of six independent experiments. The experimental variability is expressed as standard deviations. T-test: *= $p < 0,05$, **= $p < 0,01$.

Figure 4: **Brm and RNA Pol II pSer5 localization at the level of *RPRD1A* gene**

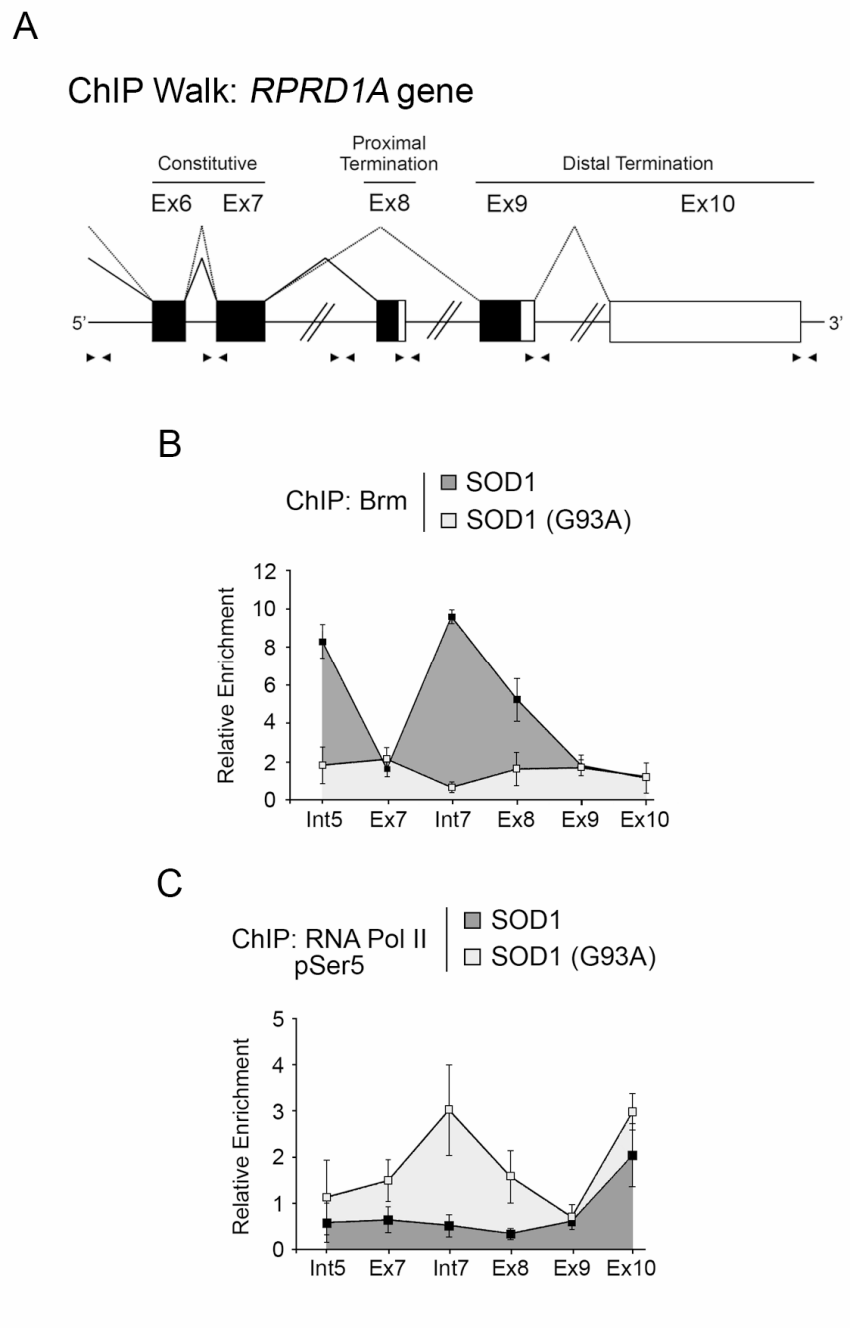


Fig. 4A: Schematic representation of region of interst present of the human *RPRD1A* gene. The coding exons are represented as black boxes, while the 3' UTRs are represented by white boxes. Exon 6 and 7 are constitutive. Exon 8 is the promoter-proximal termination, while Exon9-10 represent the more promoter-distal termination. The two possible patterns of splcing are represented by a continuous line (termination at proximal Exon 8) or a dashed line (termination at the distal Exons 9-10). The localization of the primers (arrows) used for the Chromatin ImmunoPrecipitation (ChIP) experiments are indicated.

Fig 4B: Results of the ChIP-walking experiments carried out on *RPRD1A* gene using an anti-Brm antibody. The results are represented as relative (over the IgG) enrichment, and derive from three independent experiments. The experimental variability is expressed as SEM.

Fig 4C: Results of the ChIP-walking experiments carried out on *SLC6A15* gene using an antibody which specifically recognizes the phosphorylated serine 5 (pSer5) present in the RNA Pol II CTD, a modification linked to the slow processivity of this enzyme. The results are represented as relative (over the IgG) enrichment, and derive from three independent experiments. The experimental variability is expressed as SEM.

Figure 5: **Bard1 and Cstf complex interact at the level of exon 8 of the *RPRD1A* gene**

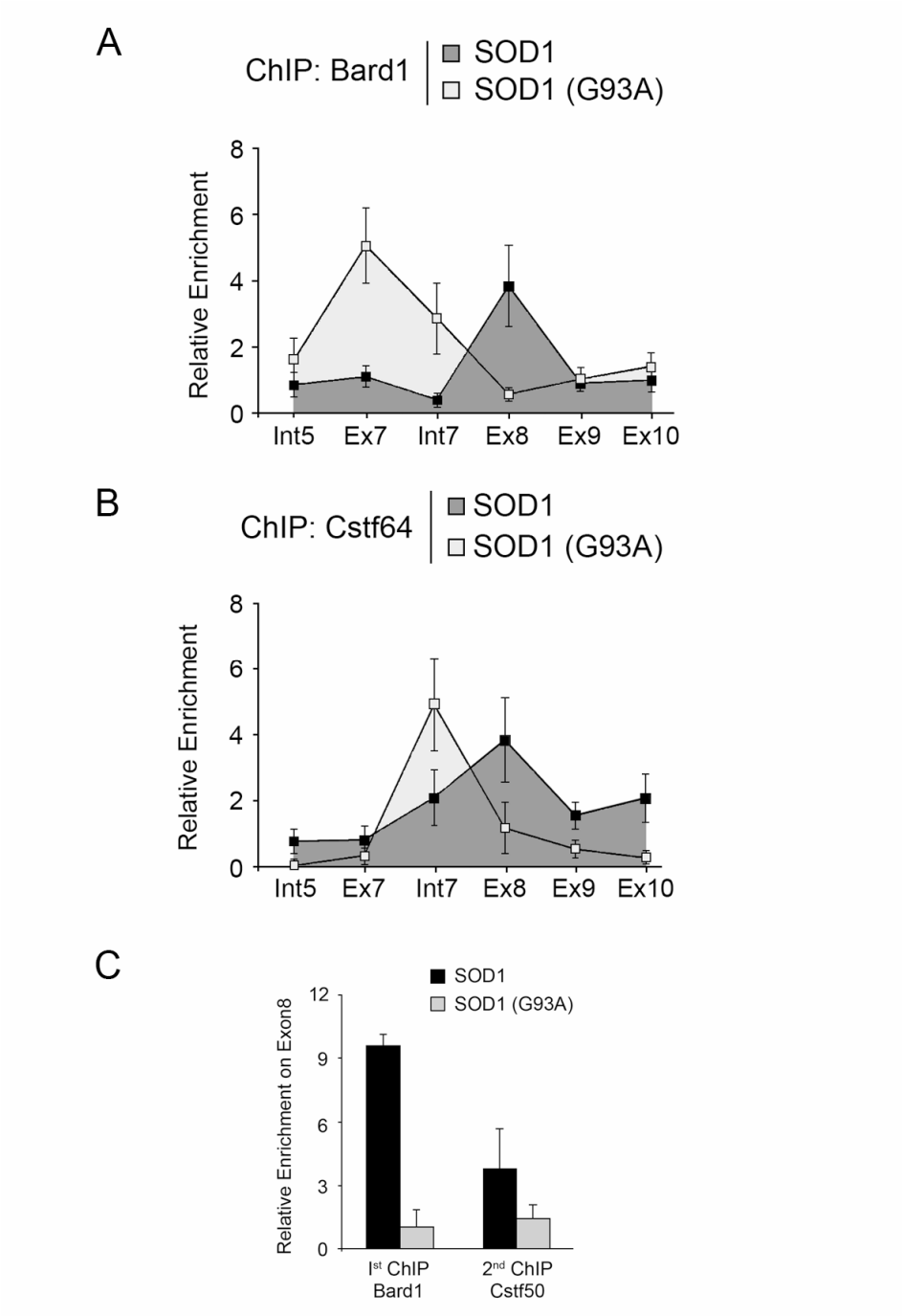
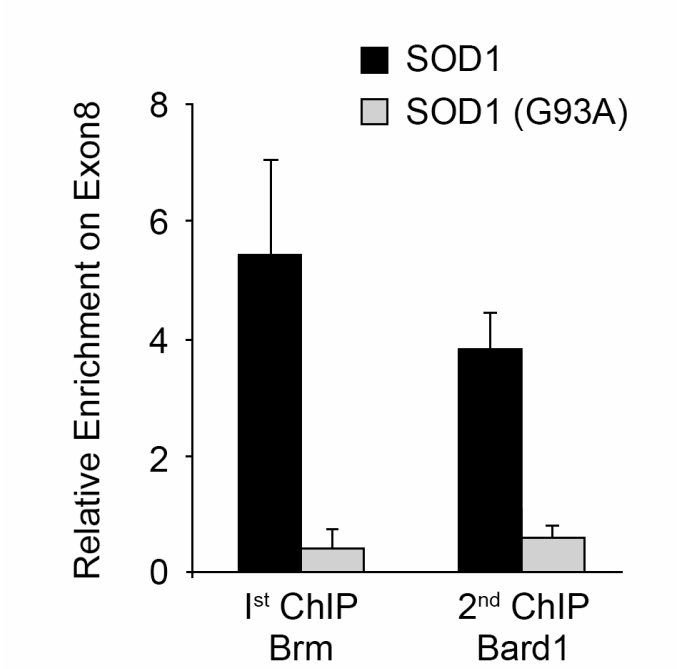


Fig 5A: Results of the ChIP-walking experiments carried out on *RPRD1A* gene using an anti-Bard1 antibody. The results are represented as relative (over the IgG) enrichment, and derive from three independent experiments. The experimental variability is expressed as SEM.

Fig 5B: Results of the ChIP-walking experiments carried out on *RPRD1A* gene using an anti-Cstf64 antibody. The results are represented as relative (over the IgG) enrichment, and derive from three independent experiments. The experimental variability is expressed as SEM.

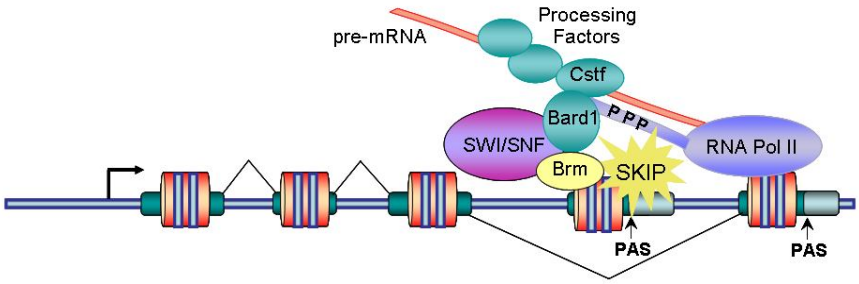
Fig 5C: Results of the ChIP-ReChIP experiments. The first immunoprecipitation (ChIP) was performed using an anti-Bard1 antiboby. Then, a second ChIP (ReChIP) was carried out, starting from the Bard1-immunoprecipitated fraction, using an anti-Cstf50 antibody. The presence of the Bard1/Cstf50 complex was then evaluated by qPCR, focusing in the region containing *RPRD1A* Exon 8. *Left*: results of the first anti-Bard1 ChIP. *Right*: results of the anti-Cstf50 ReChIP. The results are shown as relative (over the IgGs for the first ChIP, over the IgMs for the second ReChIP) enrichment, and derive from two independent experiments. The experimental variability is expressed as standard deviation.

Figure 6: **Brm and Bard1 interact at the level of *RPRD1A* Exon 8**



Results of the ChIP-ReChIP experiments. The first immunoprecipitation (ChIP) was performed using an anti-Brm antiboby, while the ReChIP was carried out using an anti-Bard1 antibody. The presence of the Brm/Bard1 complex was then evaluated by qPCR, focusing in the region containing *RPRD1A* Exon 8. *Left*: results of the first anti-Brm ChIP. *Right*: results of the anti-Bard1 ReChIP. The results are shown as relative (over the IgGs for the first ChIP, over the IgMs for the second ReChIP) enrichment, and derive from two independent experiments. The experimental variability is expressed as standard deviation.

Figure 7: **A model for the Brm-dependent regulation of the last exon alternative splicing**



While the RNA Pol II-dependent transcription takes place, Brm-containing SWI/SNF-BAF complexes localize at the level of the more promoter-proximal last exon. The presence of this complex can in turn modulate the processivity of the RNA Pol II, inhibiting the transition to the slow processive (pSer5) state linked to the “terminal exon pausing”. In the meantime, Brm recruits the Bard1-Cstf complex, which inhibits the 3’ end processing of the pre-mRNA. The combination of these two events (and, possibly, the concominant presence of other processing co-factors) results in the inhibition of the cotranscriptional inclusion of the proximal last exon. Thus, the more promoter-distal exon will be preferentially included in the mRNA products.

Table 1: **The genes with SOD1 (G93A)-altered alternative terminations examined in this study**

Gene	Proposed functions	Responsive to Brm re-expression/depletion?
RPRD1A (p15RS)	Modulator of WNT/β-Catenin pathway (Wu et al, 2010)	Yes
	Regulator of RNA Pol II transcription (Ni et al., 2011)	
SLC6A15	Brain-restricted, sodium-coupled aminoacid transporter (Takanaga et al., 2005)	Yes
IGSF1	Cell-to-cell interaction (Mazzarella et al, 1998)	Yes
STRADA	Regulation of neuronal polarity (Kim et al., 2010)	Yes
SETD3	H3K36 methyltransferase (Kim et al, 2011)	Yes
NFX1	Interactor of mSin3A and HDACs (Xu et al., 2008	No

Table 2: **Oligonucleotide primers used for the RT-PCRs assays**

Application	Primer Name	Primer sequence (from 5' to 3')
RPRD1A **RT-PCR**	*RPRD1A* RT-PCR FW ex6	CTGGGATCTCCAAGTGAACC
	RPRD1A RT-PCR Rev ex8	GTAGATGTCTCCCGCAAAGG
	RPRD1A RT-PCR Rev ex9	CGTTAGAAGATACGCCCATGT
SLC6A15 **RT-PCR**	*SLC6A15* Common FW ex4	GGGATCAGTGTCCTTTGGTG
	SLC6A15 Proximal Rev ex5	CCACTTTCCCCAATTTCCAT
	SLC6A15 Distal Rev ex7	AGCTTGAAAAGGCAATGACA

IGSF1 **RT-PCR**	*IGSF1* Common FW ex12	GGTGCCCTTACTGAGTCCAA
	IGSF1 Proximal Rev ex13	ATTGGCTCCCATACATCTGC
	IGSF1 Distal Rev ex14	TCACCCAGATTTTCAGGACA
SETD3 **RT-PCR**	*SETD3* Common FW ex7	TGCCAACAAACTACCCTTGA
	SETD3 Proximal Rev ex8	TGAGGCTGGAATTATGGCTTA
	SETD3 Distal Rev ex9	GTAACAGGACAGGGCAGGAG
STRADA **RT-PCR**	*STRADA* Common ex11	GCCATGTCCCCTTTAAGGAT
	STRADA Proximal ex12	CCGAAATCCTGCCACTTATG

	STRADA Distal ex13-14	ACGTCGCTTGATCTGCTTG
NFX1 **RT-PCR**	*NFX1* Common FW ex 11	GAGCTTCCATGTACCAGTCTCA
	NFX1 Proximal Rev ex16	TCTGGGTAGACGCCCAGTAG
	NFX1 Distal Rev ex17	GGGTGACCACAGTCAGCTCT

Table 3: **Oligonucleotide primers used for qPCR analysis of the *RPRD1A* ChIP-walking experiments**

Localization	Primer Name	Primer sequence (from 5' to 3')
RPRD1A Intron 5	FW	AGTGATTGCAACTGGGTTCC
	Rev	GCCAAAAGGGTAAACAGCAA
RPRD1A Exon 7	FW	TTTGGGAGTTGGGATGAGAG
	Rev	CATCTATTTCTGCCGCCAAT
RPRD1A Intron 7	FW	TGTGCCAAAAATGGTGCTTA
	Rev	GAAGTGGGCAACCTCTTCAA
RPRD1A Exon 8	FW	TGGACATGGGCGTATCTTCT
	Rev	AAACAGTGACAAATGACCATCA
RPRD1A Exon 9	FW	ACAGGAGTACAAGCGCAAGC
	Rev	GTGACATTGGGCAATCGAG
RPRD1A Exon 10	FW	GGCAATAGCACATGGGAAGA
	Rev	ACTTTGCTTCCCTCCCAGTC

2.7 References

Allemand E., Batsché E., Muchardt C*. Splicing, transcription, and chromatin: a ménage à trois*. Curr. Op. Gen. Dev., 2008; **18**:145-151.

Beyer A.L., Osheim Y.N. *Splice site selection, rate of splicing and alternative splicing on nascent transcripts*. Gen. Dev., 1988; **2**:754-765.

Batsché E., Yaniv M., Muchardt C. *The human SWI / SNF subunit Brm is a regulator of alternative splicing*. Nat. Struct. & Mol. Biol., 2006;**13**:22-29.

Campigli di Gianmartino D., Nishida K., Manley J.L. *Mechanisms and consequences of polyadenylation*. Mol. Cell, 2011; **43**:853-866.

David C.J., Manley J.L. *Alternative pre-mRNA splicing regulation in cancer: pathways and programs unhinged*. Genes Dev, 2010; **24**:2343-2364.

de la Mata M., Alonso C. R., Kadener S., Fededa J.P., Blaustein M., Pelisch F., Cramer P., Bentley D., Kornblihtt A.R. *A slow RNA polymeras II affects alternative splicing in vivo*. Mol. Cell, 2003; **12**:525-532.

Fox D., Le Trong I., Rajagopal P., Brzovic P.S., E. Stenkamp R.E., Klevi R.E. *Crystal structure of the BARD1 ankyrin repeat domain and its functional consequences*. J. Biol. Chem., 2008; **283**:21179-21186.

Gramling S., Rogers C., Liu G. Reisman D. *Pharmacologic reversal of epigenetic silencing of the anticancer protein BRM: a novel targeted treatment strategy*. Oncogene, 2011; **30**:3289-3294.

Gutiérrez L., Nikolic T., van Dijk T. B., Hammad H., Vos N., Wilart M. Grosveld F., Philipsen S., Lambrecht B.N. *Gata1 regulates dendritic-cell development ans survival.* Blood, 2007; **110**:1933-1941.

Kadam S., Emerson B.M. *Transcriptional specificity of human SWI/SNF BRM chromatin remodeling complexes*. Mol Cell, 2003; **11**:377-389.

Kim J.S., Hung W., Narbonne P., Roy R., Zhen M. *C . elegans STRADα and SAD cooperatively regulate neuronal polarity and synaptic organization*. Development, 2010;**102**:93-102.

Kim D.K., Kim K.K., Kim J.K., Seo S.G. *Characterization of a novel histone H3K36 methyltransferase setd3 in zebrafish.* Biosci. Biotechnol. Biochem, 2011;**75**:289-294.

Kleiman F.E., Manley J.L. *The BARD1-Cstf-50 interaction links mRNA 3' end formation to DNA damage and tumor suppression*. Cell, 2001; **104**:743-753.

Kornblihtt A.R. *Chromatin, transcript elongation and alternative splicing*. Nat. Struct. Mol. Biol., 2006;**13**:7-9.

Kolasinska-Zwierz P., Down T., Latorre I., Liu T., Liu X.S., Ahringer J. *Differential chromatin marking of introns and expressed exons by H3K36me3*. Nat. Gen., 2009; **41**:376-381.

Lenzken S.C., Romeo V., Zolezzi F., Cordero F., Lamorte G., Bonanno D., Biancolini D., Cozzolino M., Pesaresi M.G., Maracchioni A., Snges R., Achsel T., Carrì M.T., Calogero R.A., Barabino S.M.L. *Mutant SOD1 and mitochondrial damage alter expression and splicing of genes controlling neuritogenesis in models of neurodegeneration*. Hum. Mut., 2011; **32**:168-182.

Luco R.F., Pan Q., Tominaga K., Blencowe B., Pereira-Smith O., Misteli T. *Regulation of alternative splicing by histone modifications*. Science, 2010; **327**:996-1000.

Luco R.F., Allo M., Schor I.E., Kornblihtt A.R., Misteli T. *Epigenetics in alternative pre-mRNA splicing*. Cell, 2011; **144**:16-26.

Maracchioni A., Totaro A., Angelini D.F., Di Penta A., Bernardi G., Carri M.T., Achsel T. *Mitochondrial damage modulates alternative splicing in neuronal cells: implications for neurodegeneration*. J Neurochem., 207;**100**:142–153

Mazzarella R., Pengue G., Jones J., Jones C., Schlessinger D. *Cloning and expression of an Immunoglobulin Superfamily Gene (IGSF1) in Xq25*. Mol. Microbiol., 1998;**162**:157-162.

McCracken S., Fong N., Yankulov K., Ballantyne S., Pan G., Greenblatt J., Patterson S.D., Wickens M., Bentley D.L. *The C-terminal domain of RNA polymerase II couples mRNA processing to transcription*. Nature, 1997; **385**:357-361.

Mills J.D., Janitz M. *Aternative splicing of mRNA in the molecular pathology of neurodegenerative diseases*. Neurobiol. Aging, 2011;epub ahed of print.

Morris D.P., Michelotti G.A., Schwinn D.A. *Evidence that phosphorylation of the RNA polymerase II carboxy-terminal repeats is similar in yeasts and humans.* J. Biol. Chem., 2005; **280**:31368-31377

Nag A., Narsinh K., Martinson H.G. *The poly(A)-dependent transcriptional pause is mediated by CPSF acting on the body of the polymerase*. Nat. Struct. Mol.Biol., 2007;**14**:662-669.

Ni Z., Olsen J.B., Guo X., Zhong G., Ruan E. D., Marcin E., Young P., Guo H., Li J., Moffat J., Emili A., Greenblat J.F. *Control of the RNA polymerase II phosphorylation state in promoter regions by CTD interaction domain-containing proteins RPRD1A and RPRD1B*. Transcription, 2011; **25**:237-242.

Nogués G., Kadener S., Cramer P., Bentley D., Kornblihtt A.R. *Transcriptional activators differ in their abilities to control alternative splicing*. J. Bio. Chem., 2002;**277**:43110 -43114.

Oesterreich F.C., Preibisch S., Neugebauer K.M. *Global analysis of nascent RNA reveals transcriptional pausing in terminal exons.* Mol Cell, 2010; **40**:571-581.

Oesterreich F.C., Bieberstein N., Neugebauer K.M. *Pause locally, splice globally*. Trends Cell. Biol., 2011; **21**:328-335

Perales R., Bentley D. *"Cotranscriptionality": the transcription elongation complex as a nexus for nuclear transactions*. Cell, 2009; **36**:178-191.

Proudfoot N.J. *Ending the message: poly(A) signals then and now.* Genes & Dev., 2011; **25**:1770-1782.

Sakurai K., Furukawa C., Haraguchi T., Inada K. et al. *MicroRNAs miR-199a-5p and -3p target the Brm subunit of SWI/SNF to generate a double-negative feedback loop in a variety of human cancers*. Canc. Res., 2011; **75**:1680-1689.

Schmidt E.R., Pasterkamp R.J., van den Berg L.H. *Axon guidance proteins: novel therapeutic targets for ALS ?* Progr. Neurobiol., 2009;**88**:286-301.

Shi P., Gal J., Kwinter D.M., Liu X., Zhu H. 2010. *Mitochondrial dysfunction in amyotrophic lateral sclerosis*. Biochem Biophys Acta, 2010;**1802**:45–51

Takagaki Y., Manley J.L. *Complex protein interactions within the human polyadenylation machinery identify a novel component*. Mol. Cell, 2000; **20**:1515-1525.

Takanaga H., Mackenzie B., Peng J., Hediger M.A. *Characterization of a branched-chain amino-acid transporter SBAT1 (SLC6A15) that is expressed in human brain*. Biochem. Biophys.Res. Comm., 2005;**337**:892-900.

Tian B., Hu J., Zhang H., Lutz C.S*. A large-scale analysis of mRNA polyadenylation of human and mouse genes*. Nucleic Ac. Res., 2005; **33**:201-212.

Waldholm J., Wang Z., Brodin D., Tyagi A., Yu S., Theopold U., Ostlund Farrants A.K., Visa N.. *SWI / SNF regulates the alternative processing of a specific subset of pre-mRNAs in Drosophila melanogaster*. BMC Mol. Biol., 2011;**12**:46-58.

Witten J.T., Ule J. *Understanding splicing regulation through RNA splicing maps*. Trends in Gen., 2011;**27**:89-97.

Wu Y., Zhang Y., Zhang H., Ren F., Liu H., Zhai Y., Jia R., Yu J., Chang Z.. *p15RS attenuates Wnt/Beta-Catenin signaling by disrupting Beta-Catenin/TCF4 interaction*. J. Biol. Chem., 2010; **285**:34621-34631.

Xu M., Luo W., Elzi D.J., Grandori C., Galloway D.A. *NFX1 interacts with mSin3A/Histone Deacetylase to repress hTERT transcription in keratinocytes*. Mol. Cell. Biol., 2008; **28**:4819-4828.

Zlotorynski E., Agami R. *A PASport to cellular proliferation*. Cell, 2008; **134**:208-209.

Chapter 3

Manuscript in preparation

Identification of Brahma short, a class of isoforms of the chromatin-remodeling factor Brahma

Gabriele Alessandro Fontana[1], Aurora Rigamonti[1], and Silvia M.L. Barabino[1*]

[1]Department of Biotechnology and Biosciences, University of Milano-Bicocca, Piazza della Scienza, 2, Italy

*Correspondence to: Silvia M.L. Barabino, Department of Biotechnology and Biosciences, University of Milano-Bicocca, Piazza della Scienza, 2; 20126 Milan, Italy; Phone: +39-02-6448 3352; Fax: +39-02-6448 3569;

Email: silvia.barabino@unimib.it

Keywords: SMARCA2/Brahma – alternative promoter – isoforms – mitochondrial stress- bromodomain

3.1 Abstract

The human protein Brahma (Brm), encoded by the *SMARCA2* gene, is one of the two alternative ATPase subunits contained in the SWI/SNF chromatin remodeling complex. We previously reported that *SMARCA2* is downregulated in two paradigms of neuronal mitochondrial stress.

In order to further characterize the human *SMARCA2* gene, we carried out bioinformatic researches and we noticed that this gene could potentially encode a number of shorter transcripts which could be generated by an alternative promoter usage. From our analysis, it resulted that the protein isoforms encoded by these “short” transcripts would lack the N-terminal region, which contains important catalityc domains. We called these isoforms “BrmS” (Brm Short), in contrast to the full length Brm FL isoforms.

Brm FL and BrmS share the C-terminal bromodomain, while they differ in the N-terminal region. In particular, the N-terminal domain of BrmS does not contain the ATPase/helicase domain which is present in Brm and instead contains a region of 33 aminoacids, which is unique to the short isoforms.

The relative expression of the mRNAs which encodes the two isoforms are characterized by an opposite expression pattern upon mitochondrial stress exposure. More specifically, the transcpits encoding Brm FL are downregulated by acute and chronic mitochondrial stress, while the transcripts encoding BrmS are upregulated by the same conditions. These data were

further validated at the protein level in the SOD1 (G93A) cells, using an antibody which recognizes a C-terminal epitope present in both isoforms.

We also mapped and cloned the putative regulatory region which controls the expression of BrmS. By luciferase assays, we verified that the putative promoter of BrmS is endowed with regulatory elements which are preferentially active in neurons and that respond to mitochondrial stress.

By immunofluorescence, we were able to confirm that BrmS, as well as Brm, localize in the nucleus.

Taken all together, these data suggest that human *SMARCA2* gene relies on alternative promoter usage to encode two distinct classes of transcripts. Both *SMARCA2* promoters are endowed with “mitochondrial stress responsiveness” and with elements that are preferentially active in neuronal cell lines. The two proteins localize in the nucleus, suggesting the intriguing hypothesis of possible functional connections between the two isoforms.

3.2 Introduction

One interesting observation derived from the sequencing of the genomes of different species, and especially the sequencing of human genome, is that a major “gap” is present between the total number of genes and the number of proteins that may be present in the cell. Specifically, the number of genes is strikingly

lower than the number of the possible final proteins that could originate from such coding sequences (Kimura et al., 2006). It has become clear that regulatory mechanisms such as alterntive splicing (AS) and alternative promoters (APs) play important roles in increasing the diversity and complexity of the genome function.

APs are defined as genomic regions, endowed with regulatory activities, from which alternative transcripts of one given gene originate (Landry et al., 2003). During the last years, researchers have made great efforts to identify such regulatory sequences. The direct comparation of the inventories of protein-coding transcripts (transcriptome) and genome sequences has opened a brand new field of research, in which the "promoterome" (all the promoters and APs present in one given genome) is the cental topic to unravel the complexity of gene expression. From these studies emerged that more than half of human genes are endowed with multiple transcription start sites and multiple APs (Kimura el atl., 2006). These reports have identified APs that are able to control the expression of multiple isoforms in many genes, from BDNF (Timmusk et al., 1999), to the subunits of GABA receptors (Steiger et al., 2004), as well as to the gene econding the p53 protein (Buordon et al., 2005). Interestingly, the presence of such a transcriptional regulatory mechanism is typical of genes which display different isoforms which are strictly regulated in a tissue-specific fashion or whose expression is modulated during differentiation and development (Landry et al., 2003; Kimura et al. 2006).

Here, we report the identification of an alternative promoter, present in the human *SMARCA2* gene which controls the transcription of a class of isoforms of the chromatin remodeling factor Brahma (Brm, SNF2α). Brm is one of the two mutually ATPase subunits present in the SWI/SNF-BAF complex, a multi-enzymatic complex which epigenetically controls the transcription of genes involved in differentiation and development (Muchardt et a., 1999). The AP of *SMARCA2* drives the expression of “short” isoforms of Brm, that we called “BrmS”, in contrast to the full legth Brm isoforms (Brm FL). Both isoforms display a nuclear localization. The regulatory region which controls the expression of BrmS isoforms, as well as the one which controls the expression of the Brm FL isoforms (Fontana et al., manuscript in preparation), respond to mitochondrial stress, an observation which is consistent with the data collected from mRNA and protein analysis. Moreover, the *SMARCA2* AP contains putative silencer elements and sequences which are preferentially active in neuronal cell lines.

3.3 Results

Bionformatic analysis of the transcripts encoded by the human *SMARCA2* gene

The human *SMARCA2* gene encodes Brm, one of the two alternative ATPase subunits contained in the SWI/SNF-BAF chromatin remodeling complex. (Hargreaves et al., 2011). Brm contains two N-terminal domains involved in protein-protein

interaction, a central Helicase/ATPase domain and a C-terminal bromodomain. Acting together, these domains confer to Brm the capability to exert its cataylic activity, namely the ATP hydrolysis, while it is contained in the SWI/SNF-BAF complex (Clapier et al., 2009).

Taking advantage of the Ensembl Genome Browser database, we examined the *SMARCA2* transcripts reported from various researchers and obtained using different experimental approaches. We found that human *SMARCA2* gene can potentially encode at least two classes of transcripts (Figure 1A). The first class is constituted by the transcripts (SMARCA2-203, 202, 003 and 004) encoding the full length Brm proteins (Brm FL). The protein prediction of these transcripts (Figure 1B) showed that they encode proteins of nearly 180 kDa containing all the previously known domains contained in the Brm protein, as well as four nuclear localization signals (NLS1-4). Among these NLS, NL2 (Bourachot et al., 1999), NLS 3 and 4 (Loe-Mie et al., 2010) are validated from experimental evidencies. On the other hand, NLS1, which resides in the HSA domain, only results from the bioinformatic prediction.

Even though the four full length transcripts exhibit different 5' first exons, the protein prediction did not highlight significant variations in the protein secondary structure and in the domain composition. This consideration is in line with the observation that the genomic region containing the alternative 5' first exons, as well as the *SMARCA2* promoter (Fontana et al., manuscript in preparation) contains a high GC content, which is especially

concentrated in the region near the transcription start site(s). The presence of a “GC promoter” (not TATA box) is usually associated with several multiple transcription start sites, thus generating different transcripts with different starting sequences but with the same coding potential (Salbaum et al., 1989; Deaton et a., 2011).

The second class is constituted by eight transcripts (SMARCA2-011, 201, 001, 014, 010, 012, 009 and 013), which are shorter than their counterpart (Figure 1A). As in the case of the transcripts encoding the Brm FL proteins, these transcripts show different transcription start sites. The protein prediction suggest that these transcripts could encode proteins of 35-40 kDa containing a C-terminal bromodomain and two NLS (Figure 1B). We called the proteins encoded by this class of transcripts “Brm short” (BrmS). The bromodomain of BrmS is common to the Brm FL proteins, as it is encoded by common 3’ exons. Instead, the N-terminal region of the two proteins is different. As a matter of fact, while Brm FL posses various catalytic domains, BrmS isoforms have a short 33 aminoacids region which is unique to these isoforms and that does not contain any known protein motif.

Taken together, these observations suggest that human *SMARCA2* gene could encode two class of proteins: the Brm FL proteins, which display all the domains already described in literature, and BrmS, which is characterized by a N-terminal unique region and by the bromodomain.

The expression of the human *SMARCA2* transcripts/Brm protein isoforms is impaired by mitochondrial stress

Recent evidencies collected in our laboratory (Fontana et al., manuscript in preparation) suggest that the expression of Brm FL isoforms is impaired in two paradigms of neuronal mitochondrial stress: an acute stress model, namely the SH-SY5Y human neuroblastoma cells treated with Paraquat (PQ, N,N'-dimethyl-4,4'-bipyridinium dichloride, a chemical that inhibits mitochondrial complex 1) (Maracchioni et al., 2007), and a chronic stress model, constituted by the same cell line overexpressing the SOD1 protein carrying the G93A mutation, one of the genetic causes of the onset of Amyotropic Lateral Sclerosis (ALS) (Shi et al., 2010). Mitochondrial stress, both acute and chronic, downregulates *SMARCA2* expression by a transcriptional impairment (Fontana et al., manuscript in preparation). Having assumed that Brm FL expression is deregulated upon exposure to mitochondrial stress, we wondered wheter BrmS expression is similalrly altered by the same treatments. To answer to this question, we monitored the expression of the transcripts encoding the two isoforms, as well as their relative protein levels, in the two models of stress. In order to monitor the expression of the two classes of transcripts, we set up a retro-transcription PCR (RT-PCR) approach. Using a three primers PCR, we were able to discriminate the relative expression of both groups of transcripts. From these experiments, we noticed that mitochondrial stress induces changes in the relative ratio of

SMARCA2 transcripts. As a matter of facts, PQ treatment and SOD1 (G93A) overexpression similarly induce a relative downregulation of the expression of the transcript encoding Brm FL, and a concomitant relative upregulation of the expression of the transcript encoding BrmS. This is true for the acute stress model (Figure 2A) as well as for the chronic stress model (Figure 2B).

We next examined the expression of the different Brm isoforms at the protein level, focusing our attention to the SOD1 (G93A) overexpressing cells (Figure 2C). We took advantage of one commercial antibody (anti-BRM/SNF2α KR-17) which specifically recognizes an epitope located in the C-terminal bromodomain, which is common to the two classes of isoforms. This antibody identifies the Brm FL isoforms as a 200 kDa band, while BrmS isoforms are represented by a 45 kDa band. From the western blot analysis, carried out using nuclear extracts from the SH-SY5Y cells, we noticed that the expression of Brm FL isoforms is strongly downregulated, a result which is in line with our previous evidences, collected using an antibody which recognizes an epitope located in the N-terminal region of Brm FL (Fontana et al., manuscript in preparation). On the contrary, the relative expression level of BrmS isoforms results upregulated.

Taken together, these results indicate that mitochondrial stress impairs the expression of both the *SMARCA2* transcripts and Brm isoforms. In particular, SOD1 (G93A) overexpression

induces a downregulation in Brm FL expression, and a concomitant relative upregulation of BrmS isoforms.

The genomic region located upstream of the first exon of the *SMARCA2* transcripts encoding BrmS is endowed with promoter features

The transcripts encoding BrmS have a different 5' sequences respect to the ones encoding Brm FL, and BrmS isoforms exhibits a different C-terminal region respect to Brm FL. These observations proned us to investigate the possibility that the transcripts encoding BrmS may derive from an alternative promoter usage. To address this question, we examined the genomic region located upstream to the transcription start site of the transcripts encoding the BrmS isoforms. The first two exons of the transcripts encoding BrmS are unique to this transcripts, and are located in a 34 kb intron (Figure 3A). Using various bioinformatic tools, we analyzed the -1232/+198 region (the numbers are relative to the BrmS transcription start site, indicated as +1). This region contains peaks of evolutional conservation, and five putative CpG islands (Figure 3B). The UCSC Genome Browser, basing on submitted ChIP-Seq data obtained from different cell lines, identified in the -912/+198 region peaks of enrichment of trimethylation of histone 3 lysine 4 (H3K4me3), an histone mark usually associated to regulatory regions located in close proximity to transcription start sites (Kolasinska-Zwierz et al., 2009). In the same region, this software also identified a DNAse hypersensible region, another

feature associated with transcription start sites. The analysis aimed to identify putative transcription factors binding sites revealed the presence of many possible consensus sequences for regulatory proteins (Figure 3C).

Taken together, the bioinformatic prediction results indicate that the transcripts encoding the BrmS isoforms may be encoded by an alternative promoter.

Characterization of the alternative promoter that controls the expression of the transcripts encoding Brms isoforms

In order to verify that the predicted sequence is endowed with promoter activity, we cloned the entire region (from base -1232 to +198) upstream of the luciferase gene in a promoterless vector. Following 5' deletions, we tested the luciferase (luc) activities of four different constructs (Figure 3C). Initially, we monitored the activities of the different constructs in the SH-SY5Y SOD1 cells and in HEK293 cells. In order to directly compare the relative luciferase units (RLUs) obtained from these different cell lines, we operate the normalization over the background (calculated basing on the empty luciferase vector) and over an unrelated promoter (the ubiquitously active SV40 promoter controlling the luciferase transcription inside the same vector backbone). We found that the vector containing the -1232/+198 region has a luc activity comparable to the background, in all the tested cell lines (Figure 4A). However, when the -1232/-850 region is deleted, the luc activity increases (Figure 4A, compare the results of the -1232/+198 and the

-850/+198 constructs). Interestingly, this is true for the neuroblastoma cell lines, but not for HEK293 cells, in which the luciferase activity of the -1232/+198 and -850/+198 constructs remain comparable. These observations suggest that in the –1232/-850 region may be present a silencer elements which is preferentially active in neuronal cells. The luc activities of the two other constructs (which respectively contain the -513/+918 and the -166/+198 regions) are comparable to the vector containing the -1232/+198 sequence, suggesting that the basic regulatory elements that control the activity of the human *SMARCA2* alternative promoter may reside in the minimal -166/+198 region.

In another set of experiments, we analyzed the response of the -850/+198, -513/+198 and -166/+198 regions to the overexpression of the SOD1 (G93A). From these experiments, we noticed that only the luc activity of the -850/+198 sequence changes in response to mitochondrial stress (Figure 4B). Specifically, we found that SOD1 (G93A) overexpression increases the luc activity of this sequence, while the activities of the other two regions tested are slightly inhibited from mitochondrial stress. These observations suggest that the “mitochondrial stress responsive” region of the *SMARCA2* alternative promoter is localized in the -850/-513 sequence (Figure 4B, compare the results obtained from the -850/+198 and the -513/+198 constructs).

Taken together, these results suggest that the genomic sequence of interest, contained in a intron of the human

SMARCA2 gene, contains elements which are able to drive the ectopic expression of the luciferase reporter gene. Moreover, the luc assays carried out on this regulatory region have revealed the presence of a putative silencer element (which is preferentially active in the SH-SY5Y cells respect to the HEK293 cells) and of a “mitochondrial stress responsive” sequence. The presence of a sequence that responds to mitochondrial stress with an increase in the promoter activity is consistent with the upregulation of BrmS expression upon SOD1 (G93A) overexpression (Figure 2).

Brm FL and BrmS localize in the cell nucleus

A very recent report (Yang et al., 2011) indicates that several isoforms of Brm, possibly derived from alternative splicing events, may be generated from the murine *SMARCA2* gene. However, their functions have not yet been evaluated. In order to begin to understand the putative functions of human BrmS isoforms, we cloned the cDNA of the SMARCA2-001 isoform and express it as a FLAG-tagged fusion protein in the SH-SY5Y neuroblastoma cells. Following immunoflorescence, we were able to study its sub-cellular localization. We found that this BrmS isoform localizes in the nucleus (Figure 5). This result is in line with the presence of NLS (specifically, NLS3 and NLS4) in this proteins (Figure 1B). We also monitored the localization of the Brm FL isoforms, both endogenous (using an antibody which specifically recognizes its N-terminal region) and exogenous (overexpressing one of the Brm FL isoforms as a

FLAG-tagged fusion protein and then monitoring its sub-cellular localization using an anti-FLAG anitbody). Also this isoform localizes in the nucleus, a result which is in line with its function in the chromatin remodeling activity (Figure 5)

Taken together, these results indicate that the analyzed BrmS isoform, as well as its Brm FL counterpart, localizes in the cell nucleus. These localizations may be due to the NLS present in both these proteins.

3.4 Discussion

A recent report (Pal et al., 2011) suggest that alternative promoter (AP) usage exceeds alternative splicing (AS) in the generation of transcriptome diversity in the cerebellum. Specifically, this work exploited tha data collected from a genome-wide survey and demonstrated that the genes which encode proteins preferentially expressed in neurons contains the highest number of AP, and that their transcriptional regulation plays important roles in the generation of the isoforms produced. Moreover, these genes encode proteins crucial for the differentiation of neuronal cells, and mostly of them are linked to the emergence of neurological diseases and medulloblastoma, one cancer that develops in the cerebellum.

The research of Pal and coworkers is a significant example that highlights some typical features of the APs. As a matter of fact, some previous reports already linked the presence of AP to the

regulation of gene expression during differentiation and development (Landry et al., 2003). Moreover, it has been demonstrated that AP usage is strikingly enriched in genes specifically expressed in neurons and linked to axon growth and guidance, as well as general neuronal functions. Some examples are provided from the multiple promoters present in the *BDNF* gene (Timmusk et al., 1999), from the generation of different isoforms of the GABA receptor subunits (Steiger et al., 2004) and from the APs-encoded isoforms of the murine *EFA6A* gene, which are endowed with distinct biological activities in the regulation of neuritogenesis (Sironi et al., 2009).
In the present work, we report the identification of an alternative promoter in the human *SMARCA2* gene, which encodes Brahma (Brm), one of the two alternative ATPase subunits of the SWI/SNF-BAF complex. By bioinformatic analysis (Figure 3), and subsequent experimental validation by luciferase reporter assays (Figure 4), we identified a regulatory region that controls the expression of a class of short Brm isoforms, that we called “BrmS”. These isofoms display a C-terminal bromodomain, which is common to the Brm full length (Brm FL) isoforms. However, BrmS differ from Brm FL in the N-terminal region: as a matter of facts, BrmS isoforms do not display the catalytic ATPase domain, but instead are endowed with a 33 aminoacids sequence which is unique to this class of isoforms (Figure 1). A previous report (Yang et al., 2011) suggested the presence of different murine Brm isoforms, that could potentially be generated by AS. Our present work further characterizes the

complexity of *SMARCA2* gene expression, experimentally validating the presence of an AP in this gene.

Inside the -1232/+198 sequence of the putative BrmS regulatory region, located upstream to the transcription start site of the transcripts, we identified different elements that can possibly regulate their expression. A silencer region, preferentially active in neuronal cell lines, may be located in the -1232/-850 sequence, while a “mitochondrial stress responsive” region may be present in the -850/+198 sequence (Figure 4). Interestingly, we previuosly reported that also the promoter that controls the expression of the transcripts encoding Brm FL isoforms contains a region which is sensitive to mitochondrial stress (Fontana et al., manuscript in preparation). However, the two regulatory regions respond differently to such a stressful stimuli: while the luciferase activity of promoter encoding the Brm FL transcripts is dowregulated by mitochondrial stress, the luciferase activity of the promoter encoding the BrmS transcripts is upregulated by the same treatments (Figure 4B). These observations are consistent with the data collected from our RT-PCRs and western blot analysis (Figure 2), that essentially show a relative downregulation of Brm FL transcripts/isoforms and a concomitant upregulation of BrmS transcripts/isoforms. Our previous report suggests that Brm FL may be involved in the regulation of AS of genes involved in axon growth and guidance, and that Brm FL downregulation may represent one of the causes to the deregulation of the expression of these genes in two models of mitochondrial

stress. The present work adds a new layer of complexity to this possible mechanism of regulation. As a matter of fact, we demonstrated that both Brm isoforms are nuclear (Figure 5). When mitochondrial stress occurs, the ratio between the two classes of Brm isoforms changes. This observation, and the consideration that both protein display a common C-terminal bromodomain, a domain of interaction with acetylated histones (Lavigne et al., 2009), may suggest a scenario in which BrmS isoforms act as "dominant negative" proteins. In particular, the increase in the BrmS abundance in the nucleus may saturate the sites of interaction between Brm FL and the acetylated histones. This intriguing hypothesis, that has yet to be verified, may constitue a regulatory loop that modulates Brm activity in case of mitochondrial stress.

3.5 Matherials and Methods

Bioinformatic analyses

The analysis of the human *SMARCA2* transcripts was carried out using the Ensembl Genome Browser database (http://www.ensembl.org/index.html).

The bionformatic analysis of the human *SMARCA2* putative alterative regulatory region was performed using these prediction softwares: UCSC Genome Browser (for evolutional conservation, H3K4me3 enrichment and DNAse hypersensitivity analysis databases; http://genome.ucsc.edu/);

EMBOSSCpGplot (for the CpG island analysis; http://www.ebi.ac.uk/Tools/emboss/cpgplot/index.html). For the putative transcription factors binding sites analysis, we used these softwares: TESS (www.cbil.upenn.edu/cgi-bin/tess), MatInspector (www.genomatix.de) and TFSearch (www.cbrc.jp/research/db/TFSEARCH.html).

The bioinformatic prediction of Brm FL and BrmS secondary domain structure and NLS localization was performed with PredictProtein software (http://www.predictprotein.org) and Expasy Prosite (http://prosite.expasy.org).

Plasmids construction

The -1232/+198 region of human *SMARCA2* putative alternative regulatory sequences (+1 denotes the transcription start site as reported by Ensembl Genome Browser) was amplified by PCR using the Phusion Hot Start High-Fidelity DNA Polymerase (Finnzymes) according to manufacturer's instructions. The human genomic DNA extracted from HeLa cells was used as the template in the PCR reaction The sequences of the oligos used in this PCR reaction were: hSMARCA2 alternative promoter forward: 5'-TTTTCTTGTTTGGGGGATCA-3', hSMARCA2 alternative promoter reverse: 5'-CCCCAAAACTTGCTACACAA-3'. The resulting 1430 bp PCR product was purified and cloned into the pGEM®-T Easy vector system (Promega). The insert was verified by nucleotide sequencing. The *SMARCA2* promoter -3344/+57 region was then subcloned into the Sma I – Sac I

sites of the pGL2 Basic vector (Promega), a promoterless vector which allows the cloning of putative promoter sequences upstream of the *Firefly* luciferase gene. In order to obtain shorter promoter sequences, subsequent 5' deletions were operated using restriction endonuclease digestions (Figure 3B); all the restriction enzymes and corresponding buffers were from New England Biolabs.

The cDNA encoding one of the human Brm FL isoforms (Brahma, Genbank accession number NM_003070.3, corresponding to Ensembl SMARCA2-202 transcript) was a kind gift from Beverly Emerson (Kadam et al., 2003). This cDNA encodes Brm as an N-terminal FLAg fusion protein, and it was subcloned into the Kpn– Xho I sites of pcDNA3 expression vector (Invitrogen).

The cDNA encoding one of the human BrmS isoforms (Ensembl SMARCA2-001) was obtained performing a PCR on the cDNA from SH-SY5Y cells. The PCR was performed using the Phusion High Fidelity polymerase (Finnymes) according to manufacturer's instructions. Primers were: BrmS FW: 5'-AAGAGACTAGCAGCTCGCTGC-3; BrmS Rev: 5'-TCACTCATCATCCGTCCCACTTC-3'. The resulting band was subcloned into pGEM®-T Easy vector system, and verified by sequencing. The cDNA encoding BrmS was then cloned into the EcoRI – EcoRV sites of the p3xFLAG-Myc-CMV-26 vector (Sigma), that allows to expressed this isoform as a N-terminal 3xFLAG-tagged protein.

Cell cultures, transfections, and drug treatments

Human neuroblastoma SH-SY5Y cells untransfected or stably transfected with cDNAs coding for wild type SOD1 or the mutant SOD1(G93A), as well as the human embryonic kidney HEK293 cells, were maintained in Dulbecco's modified Eagle's Medium (Euroclone) supplemented with antibiotics (100 U/mL streptomycin and 100 µg/mL penicillin), 2,5 mM L-Glutamine and 10% Foetal Bovine Serum (all from EuroClone) at 37°C with 5% CO_2. SOD1-stably transfected cells were also maintained in the presence of 400 µg/ml Geneticin (G418 sulphate, Euroclone, prepared as 40 mg/ml stock solution a in water). Cells were fed every 2–3 days and passed once a week.

The treatment with Paraquat (PQ, N,N'-dimethyl-4,4'-bipyridinium dichloride, Sigma, prepared as a 100 mM stock solution in H_2O) was performed for 18 hours at a final concentration of 750 µM.

For the luciferase assay experiments, 1,5 x 10^5 SH-SY5Y cells and 10^5 HEK293 cells were seeded in 24 multiwell, and the next day plasmids (a total of 0,75 µg) were co-transfected using Polyethylenimine (PEI, Sigma, 100 mM in H_20 pH 7.00) according to the manufacturer's instruction. Transfected cells were maintained for 24 hours before lysis with Passive Lysis Buffer 1x (Promega).

RNA extraction and reverse-transcription

SH-SY5Y SOD1 cells were seeded on 10 cm plates, and 24 hours after plating the RNAs were extracted using TRIzol Reagent (Invitrogen) and subsequently purified using silica membrane spin columns from the RNeasy Mini Kit (Qiagen). RNA quantity and purity were assessed using a NanoDrop® instrument (Thermo Fisher Scientific Inc.). 2 µg of total RNA were reverese-transcribed using the random hexamers-based High Capacity cDNA Reverse-Transcription Kit (Applied Biosystems), according to manufacturer's instructions. In order to inhibit RNAse activity, RNAsin Plus reagent (Promega) was added to the reverse-transcription reactions.

Retro-transcription PCR (RT-PCR)

In order to monitor the relative amount of the mRNAs encoding the human *SMARCA2* isoforms in the SH-SY5Y SOD1 cells, we used a three primers retro-transcription PCR (RT-PCR) approach. cDNA synthesis was performed as described above. The three gene-specific primers were: hSMARCA2 FL FW: 5'-ATCTTGGAGCATGAGGAGG-3', hSMARCA2 Short FW: 5'-GGGGTTTGCTTCTGTGATTT-3', and hSMARCA2 Rev common: 5'- GCGTTCATCTGCTTTGTCAG-3'. The RT-PCRs were perfomed using the polymerase GoTaq Flexi (Promega), following the manufacturer's instructions. We run the following PCR program: 95°C 30'' seconds, 60°C 30'', 72°C 40'', and the cycling was repeated for 35 cycles (we previously ensure that transcript amplification was within a linear range). PCR

products were: 510 bp (which represents the transcripts encoding the Brm FL isoforms), 350 bp and 230 bp (which represent the transcripts encoding the BrmS isoforms). The bands were separated by electrophoresis on 2% agarose gels. The amplified PCR products were extracted from gel, cloned in th pGEM T-Easy vector system and sequenced to ensure the identity of the PCR products. Quantification was performed with Bioanalyzer 2100 (Agilent Technologies). Statistical analysis was performed with Prism GraphPad Instant Software (GraphPad Software Inc.).

Antibodies

The antibodies used for immunoblotting were: rabbit polyclonal anti-hBRM/SNF2α (KR17) (h9787, Sigma, 1:1000 dilution; a commercial antibody which specifically recognizes an epitope, common to both Brm isoforms, located in the C-terminal bromodomain) and ant-Cpsf73 (ab72294, Abcam, 1:1000 dilution; an antibody used as a normalizer for nuclear extracts loading).

Nuclear protein extracts and immunoblotting

Cells on 10 cm plate were washed twice in cold PBS1x (Euroclone) and then screaped in cold PBS 1x. After mild centrifugation, the cell pellet was resuspended in 3 ml of cold Buffer A (10mM HEPES pH 7.9; 1,5mM $MgCl_2$; 10mM KCl; 0,5mM DTT; supplemented with Protease Inhibitor Complete, from Roche) and left 5 minutes on ice. The cell suspension was

then passed on a Douncer mortar to mechanically lyse the cell membrane. Folloqing 30-40 Douncer strokes, we verified at the microscope that the population of intact nuclei in the suspension is above 90%. The suspension is then collected and centrifuged at 228xg, for 5 minutes at 4°C. The surnatant fraction, which represents the cytosolic extract, is discarded. The pellet, which is constitued by the intact nuclei, is resuspended in 500 µl of S1 Buffer (0,25M saccharose; 10mM $MgCl_2$; supplemented with Protease Inhibitor Complete). Then, this nuclei suspension was layered on a S3 buffer cushion (0,88M saccharose; 0,5mM $MgCl_2$; supplemented with Protease Inhibitor Complete). Following centrifugation at 2800xg, 10 minutes at 4°C, we obtained the clean nuclear fraction. The surnatant was discarded and the nuclei pellet was lysed in RIPA 1x (50mM TRIS pH 7.5 ; 150mM NaCl; 1% NP40; 0,5% Na Deoxycholate; supplemented with Protease Inhibitor Complete). After one hour on ice, the lysates were pelletted at 15000xg, 15 minutes, at 4°C. The surnatant represents the nuclear lysate. An aliquot of the nuclear lysate was used for protein analysis with the Bradford kit (Bio-Rad) for protein quantification.

Proteins were separated in 6% SDS-polyacrylamide gel (classic Laemli conditions) and transferred to nitrocellulose membranes (Whatman GmbH), in Transfer Buffer (25 mM Tris, 192 mM Glicine, 20% Methanol), and trasfer was carried out overnight 15 Volts, at 4°C. Membranes were blocked using 5% non fat dried milk in PBST (0.1% (v/v) Tween 20 in 1x PBS) for 1 hour

at room temperature and incubated with a primary antibody diluted in PBST supplemented with 5% non fat dried milk. Membranes were immunoblotted with antibodies of interest for 1 to 2 hours. The antibodies used are describe above in the "Antibodies" section. After washing 3 times with PBST, membranes were incubated with peroxidase-conjugated secondary or anti-rabbit IgG (Pierce, 1:10000 dilution), in PBST with 5% non fat dried milk for 45 minutes at room temperature. After washing as above, the chemio-luminescent signals developed by ECL reagents (Millipore) were detected using films Amersham Hyperfilm ECL (GE Healthcare). Quantifications of the signals were performed using ImageJ software, normalizing each band intensity with the average grey value and then with the corresponding Cpsf73 band intensity. The relative intensity values were obtained comparing the normalized intensities of the Brm FL and BrmS bands present in one single sample.

Luciferase reporter assays

1,5 x 10^5 human neuroblastoma SH-SY5Y cells stably overexpressing the SOD1 or the SOD1 (G93A) variant, as well as 10^5 human HEK293 cells, were seeded in 24 mutiwells as described above. 24 hour after the seeding, cells were co-trasfected with a total of 0,75 µg of plasmid DNA, composed by the *Firefly* luciferase vector, a constant amount of the *Renilla* expression vector pRL-TK (50 ng/well, Promega), and a variable amount of the empty pGL2 Basic vector. In this

manner, all the vectors containing different stretches of the putative *SMARCA2* regulatory sequences were transfected as equimolar amounts. The pGL2-Promoter vector (Promega, a vector in which luciferase expression is driven by SV40 promoter) was transfected in parallel wells as a positive control, while the empty vector (pGL2-Basic) was transfected in parallel wells as a negative control. The experiments were carried out in duplicate. Luciferase expression was maintained for 24 hours, and then cells were lysed in 100 µg/well of the Passive Lysis Buffer 1x according to manufacturer's instructions. To obtain a complete lysis, the cells were subjected to a freeze-thaw cycle. 10 µl of the lysates were subjected to the luciferase assays, which was carried out using the Dual-Luciferase Reporter Assay System (Promega) and a Berthold luminometer (Berthold Inc.). The relative luminescence units (RLUs) were obtained normalizing the *Firefly* luciferase readings to the corresponding *Renilla* luciferase readings. The RLUs were further normalized to the bakground signal (constituted by the RLUs of the empty vectors).

Immunofluorescence

Cells were seeded on a glass coverslip at day 1. At day 2, cells were either fixed or transfected (and fixed on day 3). The cells were washed twice in PBS1x, and then fixed with 4% Paraformaldehyde (Sigma, diluted in PBS1x) for 10 minutes at room temperature. After washing, permeabilization was performed for 5 minutes on ice using cold CKS solution (20mM

HEPES pH 7.4; 3mM $MgCl_2$; 50mM NaCl; 300mM Saccharose; 0,2% Triton-X100). The coverslips were then washed twice with PBS1x-BSA 0,2% (w/v, Sigma). Blocking was performed with a PBS1x-Tween 0,05% solution supplemented with 10%FBS. The samples were then incubated with the following antobodies: mouse monoclonal anti-FLAG M2 (Sigma, diluted 1:200) and rabbit polyclonal anti-Brm (ab15597, Abcam, 1:100 dilution). The incubation with the primary antibodies was performed for 1 hour at 37°C, in a humified chamber. The coverslips were then washed three times with the PBS1x-BSA 0,2% solution. The incubation with the secondary antibody was performed in the same conditions of the primary antibodies. The secondary antibodies were: anti-mouse Alexa 488 (Molecular Probes, 1:2000 dilution) or anti-rabbit Alexa 488 (Molecular Probes, 1:2000 dilution). After the secondary antibody incubation, DAPI staining (4,6-diamidino-2-phenylindole, Sigma) was performed for 10 minutes at room temperature. After three washings with the PBS1x-BSA 0,2% solution, the coverslips were fixed using FluorSave Reagent (Calbiochem). The images were collected with a confocal Leica microscope.

3.6 Figures

Figure 1: **Human *SMARCA2* transcripts and Brm protein isoforms**

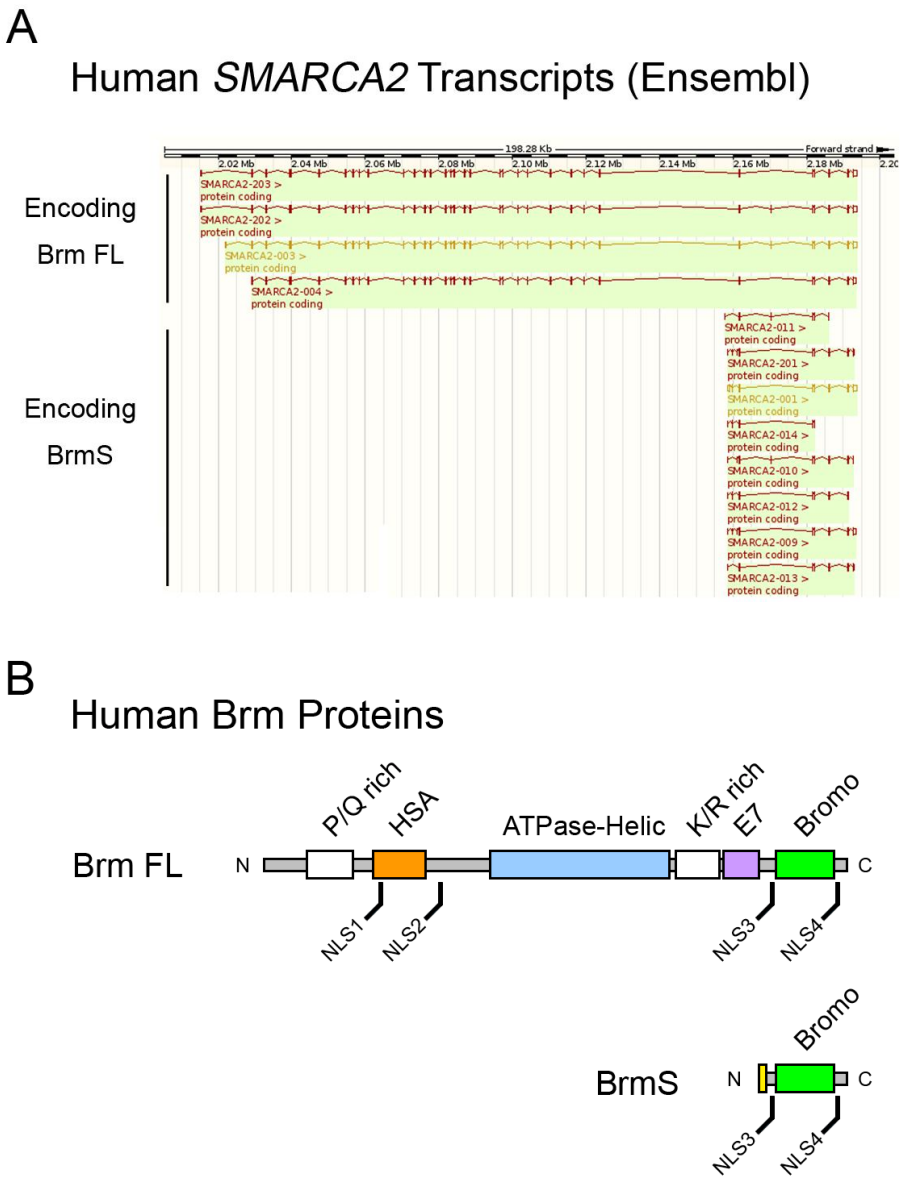


Fig. 1A: Schematic representation of the human SMARCA2 transcripts, as reported by Ensembl Genome Browser. Two classes of transcripts could be identified: the first group of transcripts encode the Brm FL isoforms, while the second group encodes BrmS isoforms.

Fig. 1B: Schematic representation of the domain composition of human Brm FL and BrmS protein isoforms, as predicted by various bioinformatic tools (see Matherials and Methods). P/Q rich= Proline/Glutamine rich domain, HSA = Helicase Sant domain, ATPase-Helic= ATPase and helicase domain, K/R rich= Lysine/Arginine rich domain, E7= domain of interaction with Rb protein, Bromo= bromodomain, NLS = Nuclear Localization Signals.

Figure 2: **The expression of human *SMARCA2* transcripts/Brm isoforms is impaired by mitochondrial stress**

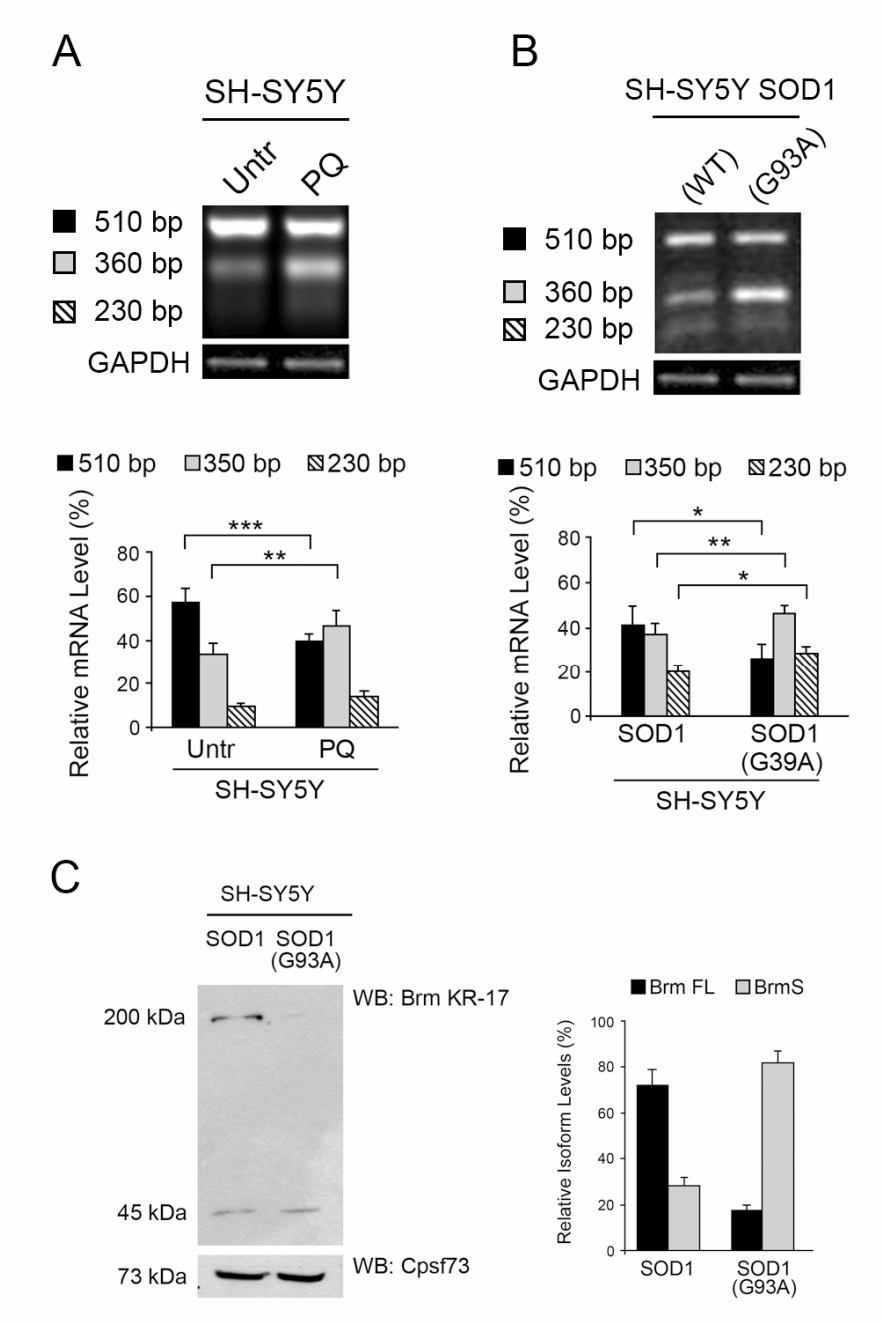


Fig. 2A: Results obtained from the RT-PCR analysis carried out on SH-SY5Y cells untreated or treated with Paraquat (PQ). The 510 bp band represents the transcripts encoding the Brm FL isoforms, while the 350 bp and the 230 bp bands represent the transcripts encoding the BrmS isoforms. *Upper panel*: a representative gel. *GAPDH* was used as a loading control. *Lower panel*: quantification of six independent experiments. The experimental variability is expressed as standard deviations T-test: **= $p < 0,01$, ***= $p < 0,001$.

Fig. 2B: Results obtained from the RT-PCR analysis carried out on SH-SY5Y cells overexpressing the wild-type (WT) or mutated (G93A) SOD1. The 510 bp band represents the transcripts encoding the Brm FL isoforms, while the 350 bp and the 230 bp bands represent the transcripts encoding the BrmS isoforms. *Upper panel*: a representative gel. *GAPDH* was used as a loading control. *Lower panel*: quantification of six independent experiments. The experimental variability is expressed as standard deviations T-test: *= $p < 0,05$, **= $p < 0,01$.

Fig. 2C: Results of the western blot carried out on SH-SY5Y SOD1 nuclear extracts. The ≈200 kDa band represents the Brm FL isoforms, while the ≈45 kDa represents the BrmS isoforms. Cpsf73 was used as a loading control. *Left*: a representative gel. *Right*: densitometry of three independent experiments.

Figure 3: **Bionformatic study of the human *SMARCA2* gene putative alternative promoter**

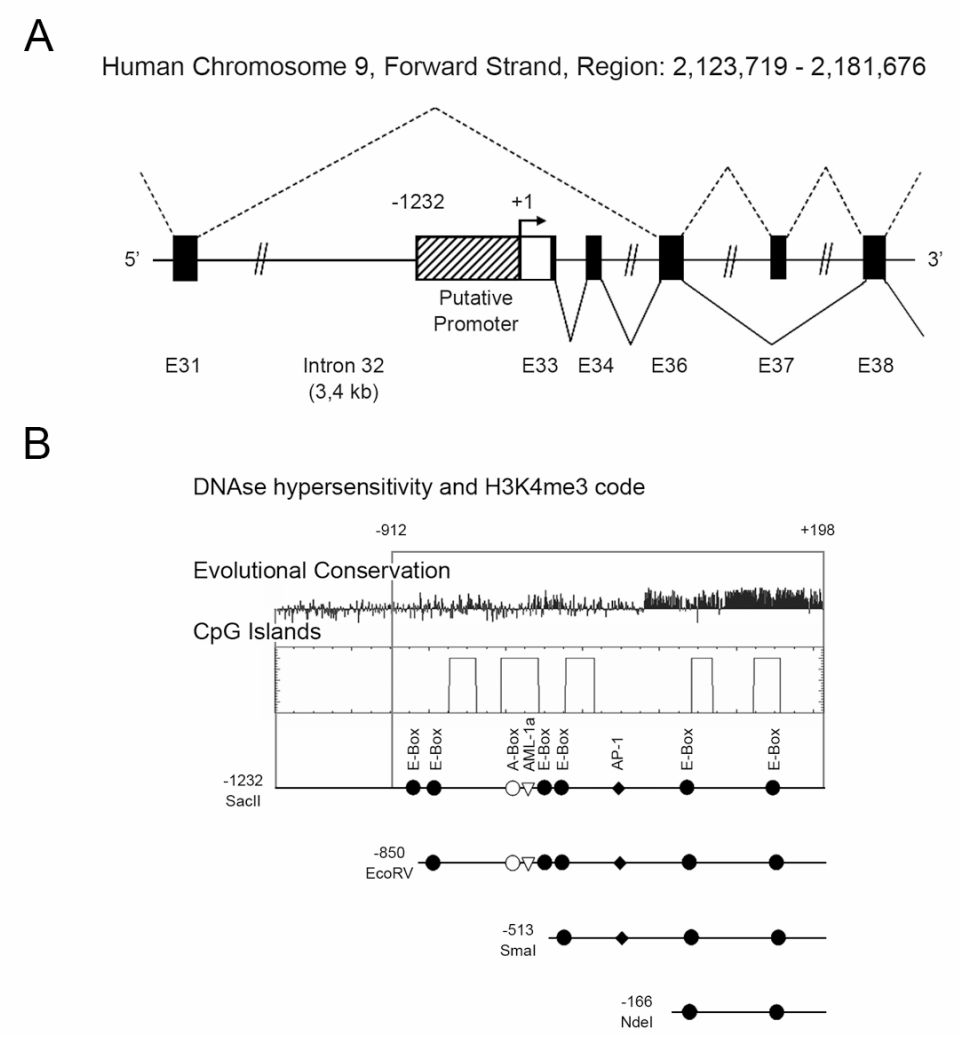


Fig. 3A: Schematic representation of the genomic region of interest, containing the putative human *SMARCA2* gene alternative promoter.

Fig. 3B: Bioinformatic analysis of the region of interest. +1 denotes the transcription start site, as reported by Ensemble Genome Browser. The evolutional conservation, as well as the presence of the H3K4me3 histone mark and the DNAse hypersensible region, were evaluated using UCSC Genome Browser. The CpG islands analysis was performed with Emboss CpGplot. Various softwares (TFSearch, MatInspector, TESS) were used for the analysis of the putative transcription factors binding sites.

Figure 4: **Characterization of the human *SMARCA2* alternative promoter**

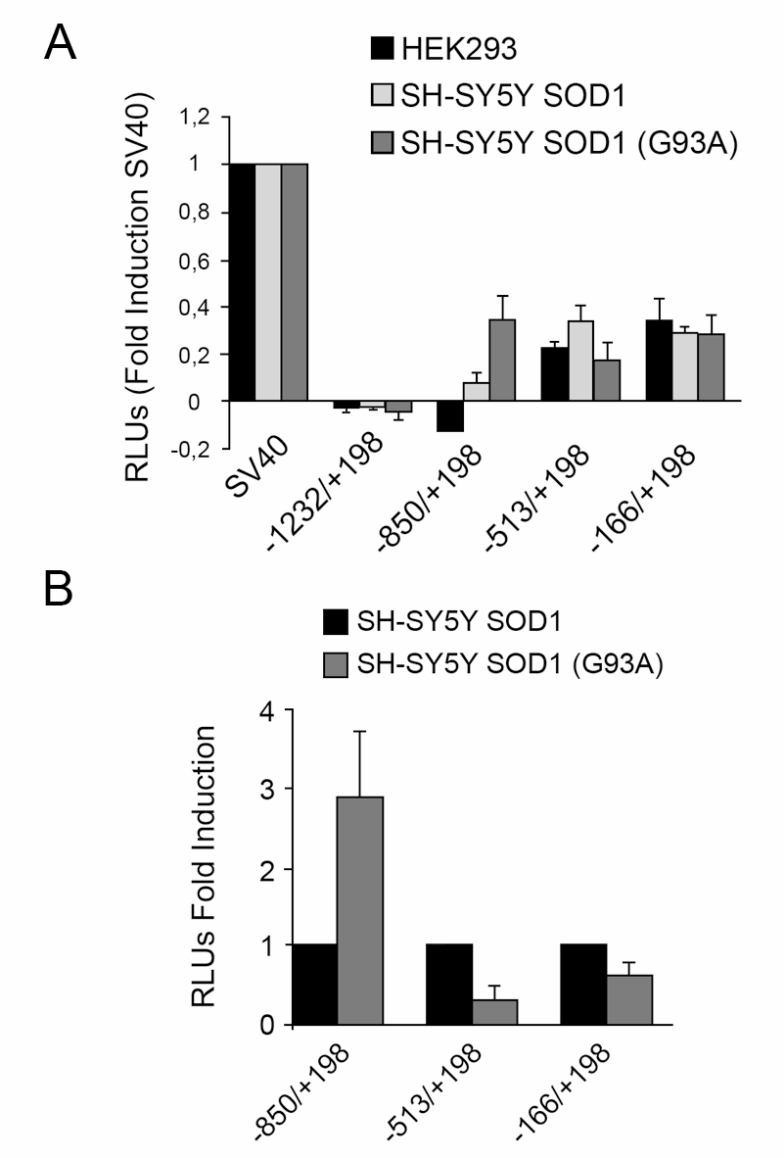
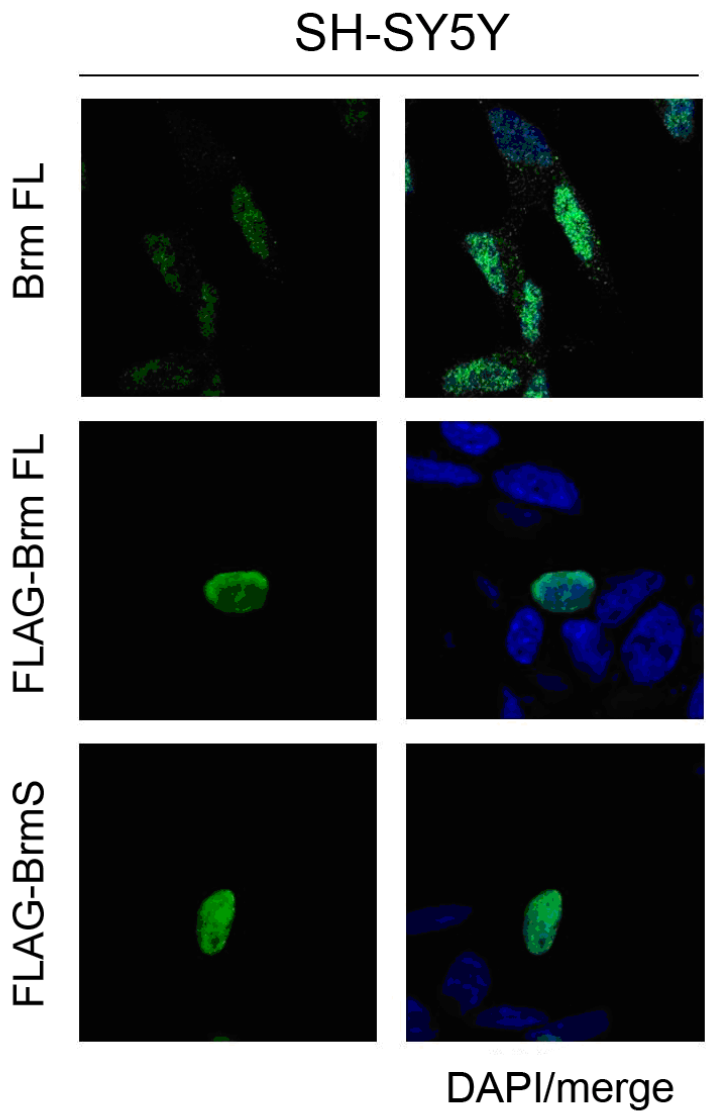


Fig. 4A: Results of the luciferase (luc) reporter assays carried on SH-SY5Y and HEK293 cells. The relative luciferase units (RLUs) were normalized to the background signal obtained from the corresponding empty vector (pGL2 Basic) signal. For each cell line, a second normalization was then operated setting the normalized over-background RLUs of the pGL2-SV40 construct to 1. The other constructs' RLUs were expressed as fractions of this value. The graph shows the results of three independent experiments carried out in duplicate. The experimental variability is expressed as standard deviations

Fig. 4B: Results of the luciferase (luc) reporter assays carried on SH-SY5Y SOD1 and SOD1 (G93A). The relative luciferase units (RLUs) were normalized to the background signal obtained from the corresponding empty vector (pGL2 Basic) signal. For each construct, a second normalization was then operated setting the normalized over-background RLUs of the SOD1 values to 1. The other constructs' RLUs were expressed as fractions of this value. The graph shows the results of three independent experiments carried out in duplicate. The experimental variability is expressed as standard deviations

Figure 5: **Brm FL and BrmS localize in the nucleus of human SH-SY5Y cells**



SH-SY5Y cells were seeded on glass coverslips, and the day after they were transfected with expression vectors containing the FLAG-tagged cDNAs of a Brm FL or BrmS isoforms. Following fixation, cells were subjected to immunofuorescence.

Left column: alexa488 signals of the indicated proteins
Right column: DAPI/merge

First row: endogenous Brm FL, as revealed by an antibody which specifically recognizes an epitope located in the N-terminal region of the full length isoforms.

Second row: transfected Brm FL isoform, as revealed by the FLAG epitope.

Third row: transfected BrmS isoform, as revealed by the FLAG epitope.

3.7 References

Bourachot B., Yaniv M., Muchardt C. *The activity of mammalian brm/SNF2alpha is dependent on a high-mobility-group protein I/Y-like DNA binding domain*. Mol Cell Biol.,1999;**19**:3931-9.

Bourdon J., Fernandes K., Murray-Zmijewski F., Liu G., Diot A., Xirodimas D.P., Saville M.K., Lane D.P. *p53 isoforms can regulate p53 transcriptional activity*. Genes & Dev., 2005;**9**:2122-2137.

Clapier C.R., Cairns B.R. *The biology of chromatin remodeling complexes*. Ann. Rev. Biochem., 2009; **78**:273-304.

Deaton M., Bird A. *CpG islands and the regulation of transcription*. Genes Dev., **25**:2011:1010-1022.

Fontana G.A., Rigamonti A., Jurdziak N., Lenzken S.C., Barabino S.M.L. *The chromatin-remodeling factor Brahma regulates the inclusion of alternative 3' terminal exons*. Manuscript in preparation.

Hargreaves D.C., Crabtree R.C. *ATP-dependent chromatin remodeling: genetics, genomics and mechanisms*. Cell. Res., 2011; **21**:396-420.

Kadam S., Emerson B.M. *Transcriptional specificity of human SWI/SNF BRM chromatin remodeling complexes*. Mol Cell, 2003; **11**:377-389.

Kimura K., Wakamatsu A., Suzuki Y., Ota T. *Identification and characterization of putative alternative promoters Diversification of transcriptional modulation : large-scale identification and characterization of putative alternative promoters of human genes*. Gen. Res., 2006; **16**:55-65

Kolasinska-Zwierz P., Down T., Latorre I., Liu T., Liu X.S., Ahringer J. *Differential chromatin marking of introns and expressed exons by H3K36me3*. Nat. Gen., 2009; **41**:376-381

Landry J.R., Mager D.L., Wilhelm B.T. *Complex controls : the role of alternative promoters in mammalian genomes*. Trends in Gen., 2003;**19**:640-648.

Lavigne M., Eskeland R., Azebi S., Saint-André V., Min Jang S., Batsché E., Ying Fan H., Kingston R.E., Imhof A., Muchardt C. *Interaction of HP1 and Brg1/Brm with the globular domain of histone H3 is required for HP1-mediated repression*. PLoS Gen., 2009; **5**:1-11.

Loe-Mie Y., Lepagnol-Bestel A.M., Maussion G., Doron-Faigenboim A., Imbeaud S., Delacroix H., Aggerbeck L., Pupko T., Gorwood P., Simonneau M., Moalic J.M. *SMARCA2 and*

other genome-wide supported schizophrenia-associated genes: regulation by REST/NRSF, network organization and primate-specific evolution. Hum Mol Genet., 2010; **19**:2841-57

Maracchioni A., Totaro A., Angelini D.F., Di Penta A., Bernardi G., Carri M.T., Achsel T. *Mitochondrial damage modulates alternative splicing in neuronal cells: implications for neurodegeneration*. J Neurochem., 207;**100**:142–153

Muchardt C., Yaniv M. *ATP-dependent chromatin remodelling: SWI / SNF and Co. are on the job*. J. Mol. Biol., 1999; **293**:187-198

Pal S., Gupta R., Kim H., Wickramasinghe P., Baubet V., Showe L.C., Dahmane N., Davuluri V. *Alternative transcription exceeds alternative splicing in generating the transcriptome diversity of cerebellar development*. Gen. Res., 2011; **21**:1260-1272.

Salbaum J.M., Weidemann A., Masters C.L., Beyreuther K. *The promoter of Alzheimer's disease amyloid A4 precursor gene*. Prog. Clin. Biol. Res., 1989; **317**:277-83.

Shi P., Gal J., Kwinter D.M., Liu X., Zhu H. 2010. *Mitochondrial dysfunction in amyotrophic lateral sclerosis*. Biochem Biophys Acta, 2010;**1802**:45–51

Sironi C., Teesalu T., Muggia A., Fontana G., Marino F., Savaresi S., Talarico D. *EFA6A encodes two isoforms with distinct biological activities in neuronal cells*. J.Cell Sci., 2009; **122**:2108-2118.

Steiger J.L., Bandyopadhyay S., Farb D.H., Russek S.J. *cAMP response element-binding protein, activating transcription factor-4 , and upstream stimulatory factor differentially control hippocampal $GABA_BR1a$ and $GABA_BR1b$ subunit gene expression through alternative promoters*. J. Neurosci., 2004;**24**:6115- 6126.

Timmusk T., Palm K., Lendahl U., Metsis M. *Brain-derived Neurotrophic Factor expression in vivo is under the control of neuron-restrictive silencer element*. J. Biol. Chem.. 1999;**274**:1078 -1084.

Yang M., Sun Y., Ma L., Wang C., Wu J., Bi A., Liao J. *Complex alternative splicing of the Smarca2 gene suggests the importance of smarca2-B variants*. Journal Canc., 2011; **2**:386-400.

Chapter 4

Conclusions

4.1 Summary

In the present thesis, I report the results obtained from my research activity carried out during the three years of the PhD program. My work was focused on the human protein Brahma (Brm, SNF2α), encoded by the *SMARCA2* gene. Brm is one of the two mutually exclusive alternative ATPases that could be present in the SWI/SNF-BAF (Brahma-Associated Factors) complexes, which are huge multi-enzymatic complexes that remodel the chromatin. Inside the chromatin-remodeling complex, the main task of Brm is to hydrolize the ATP to obtain the required energy to expose particular sites present in the chromatin or to alter the nucleosome composition. These modifications of the chromatin state have crucial impacts on gene expression (Clapier et al., 2009)

It has been proposed that Brm can also regulate the alternative splicing of pre-mRNA (AS), modulating the inclusion of internal cassette exons. More specifically, Brm exerts this activity interacting with Sam68 (an exon inclusion enhancer) and components of the spliceosome. In turn, this macro-molecular complex forms a “roadblock” to the RNA Polymerase II (RNA

Pol II) and slows down its processivity. Slowing down the polymerase processivity favours the use of weak splice sites by delaying the synthesis of downstream splice sites, thus facilitating the recognition of suboptimal exons. In the frame of this molecular mechanism, Brm acts as exon inclusion enhancer, and its activity is exerted cotranscriptionally on internal cassette exons (Batsché et al., 2006).

My work started taking advantage of previous results obtained from a genome-wide analysis of the cellular response to mitochondrial stress (Lenzken et al., 2011). In particular, this microarray analysis was aimed to identify the changes in the gene expression and in the AS patterns registered in human neuroblastoma cells in response to acute (treatment with Paraquat, PQ) and chronic (expression of the mutant SOD1(G93A) protein, which is one of the genetic causes of ALS) mitochondrial stress. This analysis revealed that changes in the transcription of specific genes are accompained by profound changes in the AS of genes involved in axon growth and guidance. Among the most downregulated genes, we found *SMARCA2*, the gene encoding Brm.

Following the validation of the microarray results, I was able to demonstrate the substantial consistency of the *SMARCA2* downregulation reported by the microarray. Interestingly, I found that the protein Brm is strongly downregulated in the SOD1 (G93A) overexpressing cells, and that the downregulation is more pronounced at the protein level (roughly

80-90% less protein respect to the control) respect to the mRNA level (50-60% less mRNA respect to the control).

I found that Brm downregulation is triggered by a mitochondrial stress-induced impairment of *SMARCA2* gene transcription. In particular, following luciferase reporter assay, I was able to map a “mitochondrial stress responsive” region in the putative *SMARCA2* promoter.

Taking advantage of the strong downregulation of Brm in the SOD1 (G93A) cells, I decided to use this cell line as a model to study the aftermath of Brm depletion. In particular, I focused my attention on the roles putatively played by Brm in a particular AS event, namely the splicing of the 3’ terminal (or last) exon. This AS event (which is also called “last exon switching” or “coding region/alternative poly(A) site”) is one of the major events that alters the transcript composition. As a matter of fact, modulating the choice of different 3’ terminal exons, the same gene can encode multiple transcripts that differ in their 3’ sequences (both coding and non-coding). In turn, this would result in an alteration of the mRNA stability (considering that different last exons have distinct poly(A) sites, which are targets of miRNAs regulations) or in proteins with different C-terminal regions (Proudfoot, 2011).

In our exon-sensitive microarray results, more than 400 AS events (in 242 genes) were identified, 35 of which affecting the last exon. Having assumed that Brm is a regulator of AS and that AS of the last exon is an important mechanism to alter the protein composition, I asked wheter Brm could modulate the

inclusion of alternative last exon in some genes involved in axon growth and guidance. I choose this cathegory because it is one of the main class in which many deregulated genes identified by the microarray fall (Lenzken et al., 2011) and also because axon retraction is one of the first hallmarks of neurodegeneration (Schimdt et al., 2009).

I found that Brm modulates the choice of the alternative last exons of many genes, such as *RPRD1A*, *SLC6A15*, *IGSF1*, *SETD3*, and *STRADA*. Interestingly, of six genes analyzed, five of them respond in the same way to Brm overexression and to Brm depletion. In particular, in all the Brm-responsive genes, Brm overexpression induce the preferential inclusion of the distal last exon, while Brm depletion (similarly to what is observed in the SOD1 (G93A) cells) induces the preferential inclusion of the proximal 3' terminal exon. These observations suggest that Brm may regulate the choice of the last exons of these five genes through a common mechanism.

In order to investigate the molecular mechanism by which Brm modulates these AS events, I carried out ChIP experiments on one of these genes, namely the *RPRD1A* gene. This gene encodes two transcripts variants, one terminating at exon 8 (proximal) and one terminating at exon 9/10 (distal). I found that Brm phisiologically localizes at the level of the proximal last exon. When Brm levels are low, namely in the SOD1 (G93A) cells, Brm fails to localize in the same region. This loss of localization is accompained by an accumulation of the "slow processive" RNA Polymerase II (RNA Pol II pSer5). These

observation suggest an inhibitory role for Brm in the choice of the last exon: as a matter of fact, the loss of Brm localization triggers the “terminal exon pausing” of the polymerase, an event which has been linked to exon inclusion (de la Mata et al., 2003) and that has been previously observed only in yeasts (Oesterreich et al., 2010). The observation that Brm could have an inhibitory role in the choice of the proximal last exon is sustained also by my ChIP-ReChIP data, that demonstrated that Brm, by a direct interaction with the BRCA1-associated RING domain 1 (Bard1) protein, recruits the Bard1-Cstf complex in this genomic localization. The Bard1-Cstf complex is an inhibitory complex that has been shown to repress the 3’ end processing of the pre-mRNA (Kleiman et al., 2001).
In conclusion, I propose a model that may explain the Brm-dependent inhibition of the choice of the proximal last exon. In this scenario, Brm-containing SWI/SNF-BAF complexes localize at the level of the proximal 3’ terminal exon. Brm then recruits the inhibitory Bard1-Cstf complex, which in turn inhibits the 3’ end processing of the nascent pre-mRNA. This mechanism fits in the frame of the “first come, first served” theory of AS (Beyer et al., 1988), which basically states that the decisions on the splicing outcomes shouls be made on the first come exon (in this case, the proximal last exon).

During my PhD thesis period, I also started a second project focused on the identification and in the characterization of an alternative *SMARCA2* promoter that controls the expression of

short isoforms of Brm. I called these isoforms “BrmS” (Brm short). These isoforms correspond to the C-terminal region of the Brm full length isoforms (Brm FL), which contains a bromodomain. Nevertheless, BrmS isoforms have a 33 aminoacids N-terminal region, instead Brm FL isoforms display an ATPase domain and other domains involved in protein-protein interactions.

I found that the transcripts encoding BrmS isoforms are generated by an alternative promoter. Interestingly, also this second *SMARCA2* promoter displays a “mitochondrial stress-responsive” region. However, the two *SMARCA2* promoters respond in an opposite way to the SOD1 (G93A) overexpression. As a matter of fact, the activity of promoter that generates the “full length” transcripts is downregulated by mitochondrial stress, while the promoter activity of the regulatory region that controls the expression of the “short” transcripts is upregulated by the same stressful stimulus.

I have validated these observations monitoring the abundance of the mRNAs and the protein levels of the two classes of isoforms, and I found that mitochondrial stress induces a change in the relative ratio of *SMARCA2* transcripts and Brm proteins.

The observation that both Brm isoforms localize in the nucleus may suggest the intriguing hypothesis that BrmS isoforms, through an interaction with acetylated histones mediated by the bromodomain (Lavigne et al., 2009), may modulate Brm FL activity.

4.2 Future perspectives

From the present thesis, a novel role for Brm in the regulation of AS emerges. Previously, it was demonstrated that Brm acts as a contranscriptional exon inclusion enhancer of internal cassette exons, and that this activity is exerted through an interaction with the slow processive (pSer5) RNA Pol II (Batsché et al., 2006). On the contrary, my data suggest that Brm, at the level of the last proximal exon, inhibits its cotranscriptional inclusion, and that a peak of enrichment of the pSer5 RNA Pol II is present only when Brm fails to localize in the same genomic region. My data, rather to be in contrast with previous evidencies, add a new role for Brm in the regulation of one specific AS event, namely the AS of last exon, which is known to be regulated by specific splicing mechanisms (Proudfoot et al., 2011).

The results of my ChIP-ReChIP experiments suggest that Brm may regulate the preferential inclusion of the distal 3' terminal exon through an interaction with the Bard1-Cstf complex, suggesting a novel mechanism of cotranscriptional splicing coupled to the termination of transcription and to the 3' pre-mRNA end processing. In the future, it will be of major interest to better characterize this model.

So far, it has been reported that Brm is downregulated in many types of human cancers (15-20% of solid tumors) through a miRNAs-mediated mechanism (Sakurai et al., 2011), and that

the rescuing of the epigenetic silencing of Brm is a promising anticancer strategy (Gramling et al., 2011). I demonstrated that Brm downregulation in human neuroblastoma cells is triggered by a mitochondrial stress-induced impaiment in *SMARCA2* promoter activity. This observation adds a novel mechanism, other than epigenetic silencing, that could cause Brm downregulation. In our laboratory, my collaborators are collecting microarray data regarding the mitochondrial stress-induced impairments in the expression of miRNAs. It will be very interesting to verify if the same miRNAs which downregulate Brm in tumor cells also participate to the downregulation of Brm in our cellular models. Similarly, as a parallel project, it could be also interesting to test the activities of the *SMARCA2* promoter constructs that I have generated in tumor cell lines, to check if Brm downregulation in tumors is not only due to epigenetic silencing but also to a transcriptional impairment. Interestingly, it has been very recently demonstrated that a transcriptional impairment of Brm expression, caused by two sequences insertion in Brm promoter, is associated to lung cancer risk (Liu et al., 2011).

From my data also emerges that mitochondrial stress impairs the Brm-dependent AS of the terminal exons of five out of six genes involved in the regulation of axon growth and guidance or, more generally, in the regulation of neuronal functions. Recent reports suggested that axon retraction is one of the first hallmarks of neurodegeneration, and that these proteins could

represent a novel therapeutic targets for Amyotrophic Lateral Sclerosis (ALS) clinical trials (Schmidt et al., 2009). In my screening, I found that genes like *RPRD1A* (or *p15RS*, a gene involved in the inhibition of the WNT/β-catenin pathway and/or in the regulation of termination of the RNA Pol II-dependent transcription) (Yang et al., 2010, Wu et al., 2010; Blakely et al., 2011; Ni et al., 2011), *SLC6A15* (a gene encoding a brain-restricted aminoacid transporter) (Takanaga et al., 2005) *IGSF1* (encoding a protein that plays crucial roles in cell-to-cell interactions) (Mazzarella et al., 1998) and *STRADA* (a regulator of neuronal polarity and synaptic organization) (Kim et al., 2010) respond in a dose-dependent fashion to Brm expression. The Brm-dependent, mitochondrial stress-induced alterations of the choice of the 3' terminal exons of these genes may generate proteins which display profound changes in their domain compositions (*data not shown*). In order to establish a possible link between Brm expression, these genes, and neurodegeneration, it will be of major interest to validate our observation in an *in vivo* system. A Brm knock-out mouse model is available (Reyes et al., 1998), and it exhibits a mild phenotype. No signs of neurodegeneration were reported in these mice, probably because the animals were sacrified at early adult stages, and neurodegeneration usually onsets in late adult life stages.

So far, no reports have specifically connected Brm to the onset of neurodegeneration and/or to other diseases which specifically alters neuronal functionality. However, alterations in

the expression of *SMARCA2* gene has been linked to the emergence of schizophrenia. This role has been established by a genome-wide survey of genes involved in the pathophysiology of this psychiatric disease (Leo-Mie et al., 2010), as well as by direct evaluation of the social behavior of Brm knock out mice (Koga et al., 2010). A direct evidence concerning the roles that Brm may play in the onset of neurodegenerative diseases, and more specifically in the onset of ALS, is still missing. However, the results reported in the present thesis, together with my preliminary observations regarding the Brm-dependent transcriptional regulation of the expression of *CXCR4* and *HGF* (*data not shown*), two genes linked to ALS onset (Luo et al., 2007; Kodoyama et al., 2007), suggest that this link may exist.

If direct causal links would be established between Brm, the deregulation in the expression of genes involved in axon growth and guidance, and neurodegeneration, then these results would fit in family of “RNA-related” genes whose functions are impaired in ALS pathology. As a matter of fact, a long list of genes which confer major risk of ALS onset are linked to various aspect of RNA biology, ranging from transcription (*ELP3, ANG2, STX3*), to splicing (*FUS, TDP43*) and editing (*SMN, TLS*) (Blitterswijk et al., 2010).

Finally, it is notewhorty to observe that the putative role of Brm in the control of the AS of the last exons of genes involved in axon growth and guidance fits in the frame of the crucial roles

played by Brm in the regulation of the neuron-specific gene expression. As a matter of fact, it has been demonstrated that Brm is enriched in neurons and plays crucial roles in regulating the expression of genes involved in neurogenesis (Olave et al., 2002), such as *Neurogenin1* (Wu et al., 2009).

In the present thesis, I also report the identification of a novel alternative promoter (AP), localized in one intron of the human *SMARCA2* gene. This promoter controls the production of the transcripts encoding the BrmS isoforms. Another research team reported the presence of isoforms of murine *SMARCA2* gene, but this paper focused on the murine brm isoforms derived only from alternative splicing events (Yang et al., 2011). BrmS isoforms display a C-terminal bromodomain, a domain involved in the interaction with the acetylated histones (Lavigne et al., 2009). BrmS isoforms differ from the Brm FL isoforms in the N-terminal region, because they do not display the N-terminal catalytic domains and the regions which allow to Brm FL to enter in the SWI/SNF-BAF complex (Muchardt et al., 1995). These observations suggest that BrmS may interact with acetylate histones, but also that this activity may be exerted outside of the chromatin remodeling complex. Interestingly, I have demonstrated that both Brm isoforms are nuclear, suggesting that a cross-talk between them is possible. From my analysis also emerges that the promoter activity of the regulatory region that control the expression of BrmS increases in response to SOD1 (G93A) overexpression, while the activity

of the promoter from which the full lenght isoforms originate is donwregulated by the same stressful stimulus. These data are consistent with the results obtained from the analysis of the relative abundance of the mRNAs and the protein levels of the two isoforms. Taken together, these observations suggest the intriguing hypothesis that BrmS isoforms may modulate Brm FL activity during mitochondrial stress, “buffering” the sites of interaction with acetylated histones. With my collaborators, we are exploring this hypothesis and trying to set up a screening system to identify a possible role of “dominant negative” for the BrmS protein variants.

4.3 References

Beyer A.L., Osheim Y.N. *Splice site selection, rate of splicing and alternative splicing on nascent transcripts*. Gen. Dev., 1988; **2**:754-765.

Batsché E., Yaniv M., Muchardt C. *The human SWI / SNF subunit Brm is a regulator of alternative splicing*. Nat. Struct. & Mol. Biol., 2006;**13**:22-29.

Blakely B.D., Bye C.R., Fernando C.V., Horne M.K., Macheda M.L., Stacker S.A., Arenas E., Parish C.L. *Wnt5a regulates midbrain dopaminergic axon growth and guidance*. PloS One, 2011; **6**:e18373

Blitterswijk M.V., Landers J.E. *RNA processing pathways in amyotrophic lateral sclerosis*. Neurogenetics, 2010; **11**:275-290.

Clapier C.R., Cairns B.R. *The biology of chromatin remodeling complexes*. Ann. Rev. Biochem., 2009; **78**:273-304.

de la Mata M., Alonso C. R., Kadener S., Fededa J.P., Blaustein M., Pelisch F., Cramer P., Bentley D., Kornblihtt A.R. *A slow RNA polymeras II affects alternative splicing in vivo*. Mol. Cell, 2003; **12**:525-532.

Gramling S., Rogers C., Liu G. Reisman D. *Pharmacologic reversal of epigenetic silencing of the anticancer protein BRM: a novel targeted treatment strategy*. Oncogene, 2011; **30**:3289-3294.

Kim J.S., Hung W., Narbonne P., Roy R., Zhen M. *C . elegans STRADα and SAD cooperatively regulate neuronal polarity and synaptic organization*. Development, 2010;**102**:93-102.

Kleiman F.E., Manley J.L. *The BARD1-Cstf-50 interaction links mRNA 3' end formation to DNA damage and tumor suppression*. Cell, 2001; **104**:743-753.

Kadoyama K., Funakoshi H., Ohya W., Nakamura T. *Hepatocyte growth factor (HGF) attenuates gliosis and motoneuronal degeneration in the brainstem motor nuclei of a transgenic mouse model of ALS*. Neurosci. Res.,2007; **59**:446-456.

Koga M., Ishiguro H., Yazaki S., et al. *Involvement of SMARCA2 / BRM in the SWI / SNF chromatin-remodeling complex in schizophrenia*. Hum.Mol Gen., 2009; **18**:2483-2494.

Lavigne M., Eskeland R., Azebi S., Saint-André V., Min Jang S., Batsché E., Ying Fan H., Kingston R.E., Imhof A., Muchardt C. *Interaction of HP1 and Brg1/Brm with the globular domain of*

histone H3 is required for HP1-mediated repression. PLoS Gen., 2009; **5**:1-11.

Lenzken S.C., Romeo V., Zolezzi F., Cordero F., Lamorte G., Bonanno D., Biancolini D., Cozzolino M., Pesaresi M.G., Maracchioni A., Snges R., Achsel T., Carrì M.T., Calogero R.A., Barabino S.M.L. *Mutant SOD1 and mitochondrial damage alter expression and splicing of genes controlling neuritogenesis in models of neurodegeneration*. Hum. Mut., 2011; **32**:168-182.

Loe-Mie Y., Lepagnol-Bestel A.M., Maussion G., Doron-Faigenboim A., Imbeaud S., Delacroix H., Aggerbeck L., Pupko T., Gorwood P., Simonneau M., Moalic J.M. *SMARCA2 and other genome-wide supported schizophrenia-associated genes: regulation by REST/NRSF, network organization and primate-specific evolution*. Hum Mol Genet., 2010; **19**:2841-57

Liu G., Gramling S., Munoz D., Cheng D., Azad A.K., Mirshams M., Chen Z., Xu W., Roberts H., Sheperd F.A:, Tsao M.S., Reisman D. *Two novel BRM insertion promoter sequence variants are associated with loss of BRM expression and lung cancer risk.* Oncogene,2011.; **30**:3295-3304.

Luo Y, Xue H., Pardo A.C., Mattson M.P., Rao M.S., Maragakis N.J. *Impaired SDF1 / CXCR4 signaling in glial progenitors derived from SOD1 G93A mice*. Jour. Neurosci. Res., 2007; **85**:2422-2432.

Mazzarella R., Pengue G., Jones J., Jones C., Schlessinger D. *Cloning and expression of an Immunoglobulin Superfamily Gene (IGSF1) in Xq25*. Mol. Microbiol., 1998;**162**:157-162.

Muchardt C., Sardet C., Bourachot B., Onufryk C., Yaniv M. *A human protein with homology to Saccharomyces cerevisiae SNF5 interacts with the potential helicase hbrm*. Nuc. Ac. Res., 1995; **23**:1127-1132.

Ni Z., Olsen J.B., Guo X., Zhong G., Ruan E. D., Marcin E., Young P., Guo H., Li J., Moffat J., Emili A., Greenblat J.F. *Control of the RNA polymerase II phosphorylation state in promoter regions by CTD interaction domain-containing proteins RPRD1A and RPRD1B*. Transcription, 2011; **25**:237-242.

Oesterreich F.C., Preibisch S., Neugebauer K.M. *Global analysis of nascent RNA reveals transcriptional pausing in terminal exons.* Mol Cell, 2010; **40**:571-581.

Olave I., Wang W., Xue Y., Kuo A., Crabtree G.R. *Identification of a polymorphic, neuron-specific chromatin remodeling complex.* Gen. Dev., 2002;**16**:2509-2517.

Proudfoot N.J. *Ending the message: poly(A) signals then and now.* Genes & Dev., 2011; **25:**1770-1782.

Reyes J.C., Barra J., Muchardt C., Camus A., Babinet C., Yaniv M. *Altered control of cellular proliferation in the absence of mammalian brahma (SNF2α)*. EMBO J., 1998;**17**:6979-6991.

Sakurai K., Furukawa C., Haraguchi T., Inada K. et al. *MicroRNAs miR-199a-5p and -3p target the Brm subunit of SWI/SNF to generate a double-negative feedback loop in a variety of human cancers*. Canc. Res., 2011; **75**:1680-1689.

Schmidt E.R., Pasterkamp R.J., van den Berg L.H. *Axon guidance proteins: novel therapeutic targets for ALS ?* Progr. Neurobiol., 2009;**88**:286-301.

Takanaga H., Mackenzie B., Peng J., Hediger M.A. *Characterization of a branched-chain amino-acid transporter SBAT1 (SLC6A15) that is expressed in human brain*. Biochem. Biophys.Res. Comm., 2005;**337**:892-900.

Wu M., Zhang Y., Wu N., Shen Y. *Histone marks and chromatin remodelers on the regulation neurogenin1 gene in RA induced neuronal differentiation of p19 cells*. J. Cell. Biochem., 2009; **107**:264-271.

Wu Y., Zhang Y., Zhang H., Ren F., Liu H., Zhai Y., Jia R., Yu J., Chang Z.. *p15RS attenuates Wnt/Beta-Catenin signaling by disrupting Beta-Catenin/TCF4 interaction*. J. Biol. Chem., 2010; **285**:34621-34631.

Yang G.Y., Luo Z. *Implication of Wnt signaling in neuronal polarization.* Dev Neurobiol., 2010; **18**:495-507.

Yang M., Sun Y., Ma L., Wang C., Wu J., Bi A., Liao J. *Complex alternative splicing of the Smarca2 gene suggests the importance of smarca2-B variants*. Journal Canc., 2011; **2**:386-400.

Acknowledgments

"La semplicità è l'estrema perfezione"

Leonardo da Vinci

Se si potesse effettuare una studio (o, come va di moda, una statistica) riguardante quale delle pagine di una tesi sia più letta o consultata dalla maggior parte dei lettori, penso che quest'ultima sarebbe di certo la pagina dei ringraziamenti. Lunghi elenchi di persone, nomi e fatti, nel quale perdersi, identificando chi è effettivamente nominato oppure no, e del perché si sia verificata una delle due sopra elencate circostanze.

La pagina dei ringraziamenti è esattamente quella su cui vi trovate ora, ma, stranamente, non vedrete nominativi. Fatto alquanto strambo, enigmatico, e raro. Sono fermamente convinto che le persone che mi conoscono sappiano benissimo quanto io possa essere strambo, ed enigmatico. Sul fatto del raro, alcune di queste persone potranno anche aggiungere "per fortuna", e poi alzare gli occhi al cielo. Ma per queste persone, gli elenchi di nomi sono solo appelli per i quali dichiararsi presente o assente in classe non ha valore. Per queste persone, vedere il proprio nome in fondo ad un libro non conta, perché possono vedere nei miei occhi, ogni giorno, la

riconoscenza che provo nei loro confronti. Per queste persone, io sto scrivendo queste ultime parole della mia tesi.

Voglio sentitamente ringraziare la mia famiglia, tutte le persone che mi sono sempre state vicino, i miei collaboratori, le persone per cui reputo un onore il solo stare accanto, tutti gli amici vicini, lontani, o assenti.

Lungi dal volermi paragonare all'illustre Personaggio, ho solo tentato di mantenere semplici queste mie parole di gratitudine, esattamente come dovrebbe sempre essere un ringraziamento sincero. Se poi queste mie parole sono anche state "estremamente perfette", credo che questo me lo dobbiate dire voi.

Grazie mille, ancora una volta.

Gabriele