Characterization of AGO2 complexes

Candidate: Emanuela Cipolletta
Academic year 2011/2012
XXV cycle

Ph.D. Program: "HUMAN BIOLOGY AND GENETICS"
Department of Cellular Biotechnology and Hematology
Sapienza University of Rome

Director of Doctoral Program:
Prof. Paola Londei Sapienza University of Rome

Scientific Tutor:
Prof. Giuseppe Macino Sapienza University of Rome

Board of Examiners:
Prof. Maria D'Erme Sapienza University of Rome
Prof. Manuela Helmer-Citterich University of Rome Tor Vergata
Prof. Fabrizio Loreni University of Rome Tor Vergata

Acknowledgments

I thank my Tutor Prof. Giuseppe Macino who introduced me to Argonaute proteins. I am grateful to Dr. Claudia Carissimi and Dr. Valerio Fulci for supporting my scientific growth.
I address thanks to Silvia Gioiosa, Ilaria Laudadio, Gianluca Azzalin, Gina Scapucci, Teresa Colombo for their helpful collaboration and suggestions. A kind thanks to Katia Broccoletti.

Index

Abstract

The AGO2 protein is part of a highly conserved family of RNA-binding protein called Argonaute. This protein family plays a key role in gene expression regulation in association with different classes of small non-coding RNA (miRNA, siRNA, piRNA). In addition to the post-transcriptional regulation in the cytoplasm, AGO2 protein acts also in the nucleus at the transcriptional level with a mechanism, in humans, not clearly identified. To elucidate the molecular functions of AGO2 in human cells, we performed an extensive analysis of both proteins and small RNAs associated to endogenous AGO2 protein by high-throughput technologies. We identified various AGO2-interactors not previously described. Interactors that can shed light on molecular functions exerted by AGO2 in novel as well in known biological processes. In particular, we provided several evidence regarding the mechanism of AGO2-mediated transcriptional regulation. Interestingly, we uncovered the interaction of AGO2 with the SWI/SNF chromatin remodeling complexes. The SWI/SNF is a multisubunit complex that alter the structure or the positioning of nucleosomes, thus modulating the access of regulatory proteins and transcription factors to DNA. Futhermore, we identified a new class of AGO2-associated small RNAs in the nucleus that showed a remarkable overlap with the binding sites of SWI/SNF complexes, regulatory regions in proximity of transcription start site and with chromatin marks associated with transcriptional active state. Taken together, our data suggests that nuclear AGO2 regulates the transcriptional initiation interacting with SWI/SNF complexes to ensure proper nucleosome positioning in proximity of promoter regions.

Introduction

The Argonaute protein family

Argonaute proteins constitute a highly conserved protein family whose members have a key role in gene-expression regulation pathways in association of small RNAs. Small RNAs such as short interfering RNAs (siRNAs), microRNAs (miRNAs) or Piwi-interacting RNAs (piRNAs) are anchored into specific binding pockets and guide Argonaute proteins to target RNA molecules for translational inhibition or degradation. In particular, a mainly factor that dictates the specific mechanism of silencing mediated by Argonaute proteins is the degree of complementarity between the small RNA and the RNA target. In details, a perfect complementary determines the endonucleolytic cleavage of RNA target, a process known as slicing. On the contrary, translational inhibition and mRNA destabilization are induced when a partial complementarity occurs.

The number of Argonaute proteins is highly variable between species, ranging from one in *S. pombe* to 27 in *C.elegans*. In mammals have been identified eight Argonaute proteins (Höck J, 2008). Based on sequence similarities, the Argonaute protein family can be divided phylogenetically into three subfamilies (Farazi TA, 2008). The largest subfamily comprises the AGOs, named after its founding member AGO1 in *A. thaliana.* The second subfamily comprises the Piwis, named after the *D. melanogaster* protein PIWI (P-element induced wimpy testis). The third subfamily, Class 3, consists exclusively of *C. elegans* proteins. The AGO proteins are ubiquitously expressed and bind to siRNAs or miRNAs, whereas the PIWI proteins are mostly restricted to the germ line and associate with piRNAs.

Characteristic structural feautures

Argonaute proteins are multidomain proteins that contain an N-terminal domain, and a PAZ, middle (MID) and PIWI domain (Fig.1a). The recent determination of the crystal structure of full-length human AGO2, consistent with previous studies of prokaryotic homologues, have

revealed a bilobate architecture, with the MID and PIWI domains forming one lobe, and the N-terminal and PAZ domains constituting the other (Fig.1b) (Schirle NT, 2012).
The interaction between small RNAs and AGOs occurs through several contact points on the protein, mediated by distinct domains. In particular, the small RNA 5' and 3' termini are recognized by the PAZ and MID domains, respectively. Once loaded the small RNA, Argonaute proteins silence RNA targets by slicing, activity that resides in the PIWI domain.

The **PAZ** domain contains a specific binding pocket that anchors the characteristic two nucleotides 3' overhang. Although there are no sequence-specific contacts, the base of the terminal nucleotide stacks against the aromatic ring of a conserved phenylalanine residue (Yan KS, 2003).

The **MID** domain contains a highly basic pocket, which specifically binds the characteristic 5' phosphate of small RNAs and anchors the small RNA onto Argonaute proteins (Ma JB, 2005). The MID domain has been implicated in protein-protein interactions (Till S, 2007) and, based on sequence similarities to the 7-methylguanine (m7G) cap-binding motif of eukaryotic translational initiation factor 4E (eIF4E), also in the binding of the m7G cap of target mRNA (Kiriakidou M, 2007).

The **PIWI** domain shows extensive homology to RNase H, an endoribonuclease that cleaves RNA–DNA hybrids (Parker JS, 2004a). Indeed, biochemical *in vitro* studies of Argonaute proteins from several model organisms have shown that some are endonucleases, and these are often referred to as 'slicers'. *In vitro* assays using RNA substrates complementary to exogenous siRNA (Liu J, 2004) as well as endogenous miRNAs (Meister G, 2004) have identified AGO2 as the only member of the human AGO subfamily with endonuclease activity. Moreover, based on mutation studies, the catalytic triad of human AGO2 has been characterized as D(597), D(669) and H(807) (Song JJ, 2004). Interestingly, human AGO3 is catalytic inactive even though the catalytic triad DDH is conserved. This indicates that other factors such as post-translational modifications or interactions with specific proteins may modify the activity of AGO proteins.

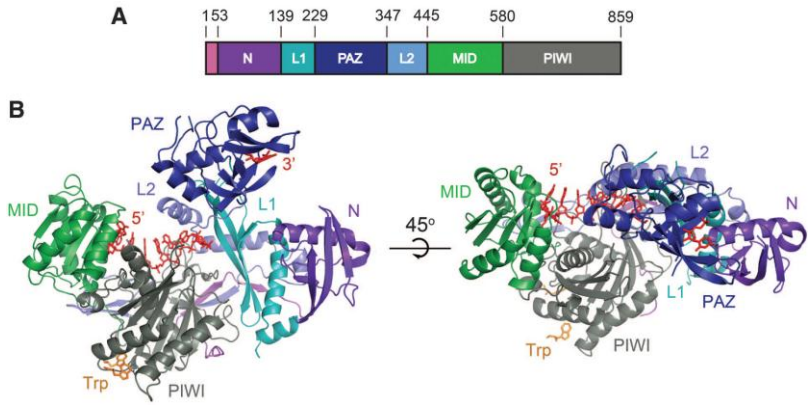

Figure 1 Structure of human AGO2. (A) Schematic of the AGO2 primary sequence. **(B)** Front and top views of Ago2 with the N (purple), PAZ (navy), MID (green), and PIWI (gray) domains and linkers L1 (teal) and L2 (blue). A generic guide RNA (red) can be traced for nucleotides 1 to 8 and 21 (Schirle NT, 2012).

Subcellular localization of AGO proteins

AGO proteins (AGOs) have been implicated in both post-transcriptional and transcriptional gene-expression regulation. This dual level of regulation reflects the localization of AGOs both in the cytoplasm and in the nucleus. The distinct subcellular localization of AGO proteins is mediated by accessory proteins and post-tradutional modifications, which co-ordinate AGOs activity with the plethora of cellular signals.

In the cytoplasm AGO proteins are enriched in distinct foci such as P-bodies and stress-granules. The former are cellular sites where occur mRNA turnover and storage (Eulalio A, 2007a), the latter are structures induced upon cellular stress and contain mRNAs stalled in the process of translation initiation (Leung AK, 2006). It has been reported that the AGO2 P-body localization is mediated by several accessory proteins such as GW182 (Eulalio A, 2007b), and influenced by post-tradutional modifications such as phosphorilation at serine-387 (Zeng Y, 2008).

The funding of AGO proteins in the nucleus of human cell is recent (Robb GB, 2005), and alternative nuclear trafficking mechanism are reported. Emerging evidence shows that the cytoplasmic and nuclear AGO functions are co-ordinated by a common set of proteins. Imp8, that is required for binding of AGO proteins to a variety of mRNA targets in the cytoplasm, modulates nuclear localization of AGO2 as well (Weinmann L, 2009). Similarly, TNRC6A, that interacts with AGOs and triggers translational repression and/or mRNA degradation in P-bodies, translocates AGO2 in and out the nucleus via its own recently discovered nuclear localization signals (NLS) and nuclear export signals (NES) (Nishi K, 2013).

Functions of AGO proteins

AGO-associated small noncoding RNAs. AGO proteins exert their regulatory activity in association with small noncoding RNA molecules that fall into two main classes: siRNAs and miRNAs.

The first class, siRNAs, derives from long double-stranded RNA (dsRNA) molecules that result from RNA virus replication, convergent transcription of cellular genes or mobile genetic elements, self-annealing transcripts or experimental transfection (Ghildiyal M, 2009). The endonuclease Dicer functions as a molecular ruler to cleave the dsRNA at ~21–25-nucleotide intervals. After Dicer-mediated cleavage, one strand of the siRNA duplex (the guide strand) is loaded onto an AGO protein that is the core component of the RNA-induced silencing complex (RISC). AGO loading takes place in the RISC-loading complex, a ternary complex that consists of an AGO protein, Dicer and a dsRNA-binding protein (known as TRBP in humans). During loading, the non-guide (passenger) strand is cleaved and ejected. AGO complexes uses the guide siRNA to associate with target RNAs containing perfectly complementary sequence and then catalyses the slicing of these targets. After slicing, the cleaved target RNA is released, and the RISC is recycled for another round of slicing (Zamore PD, 2000).

miRNAs, the second class, are encoded in the genome. These small RNAs are transcribed by polymerase II from endogenous miRNA genes resulting in a primary transcripts (pri-miRNAs) that contain ~65–70-nucleotide stem–loop structures. The hairpin structure is

excised in the nucleus by the Drosha–DGCR8 complex to yield a precursor miRNA (pre-miRNA). After its export to the cytoplasm, the pre-miRNA undergoes another endonucleolytic cleavage, which is catalyzed by Dicer, generating a miRNA–miRNA* duplex of ~21–25 nucleotides (where miRNA is the guide strand and miRNA* is the passenger strand). For most miRNAs, only one strand accumulates as mature miRNA. Such asymmetric loading is guided by the relative thermodynamic stability of the 5' ends of the small RNA duplex. The strand whose 5' end is less stably paired is preferentially incorporated into AGO complexes, whilst the passenger strand is subsequently degraded (Bartel DP, 2009).
In addition to this canonical miRNA biogenesis pathway, some alternative miRNA biogenesis pathways have recently been discovered. For example, the mirtron subclass is encoded in the introns of genes and is generated directly from byproducts of intron splicing and disbranching events (Berezikov E, 2007).

miRNAs generally interact with their mRNA targets through a limited base-pairing of only 2-7 nt (seed region) at the 5'end. With few exceptions, miRNAs-binding sites lie in the 3'UTR and are usually present in multiple copies. Typically, mRNA may have several putative targets for both the same and different miRNAs, and any particular miRNA may have hundreds of putative mRNA targets. Hence, given the substantial number of miRNAs, there exists a staggering combinatorial network of miRNA-mediated gene regulation (Selbach M, 2008).

Post-transcriptional regulation. The mechanism by which a small RNA regulates its mRNA target reflects both the specific AGO protein association and the extent of complementarity between the small RNA and the mRNA. In general, a fully complementary to their mRNA targets, typical of siRNAs, directs endonucleolytic cleavage of the mRNA, whereas a non perfect complementarity, characteristic of miRNAs, induces translation repression and mRNA degradation.

The mechanism by which AGO proteins mediate translational repression is still a matter of debate. Indeed, multiple mechanisms acting on translation initiation (Meister G, 2007), translation elongation (Petersen CP, 2006) and on the degradation of nascent polypeptides

(Nottrott S, 2006) are reported. Intriguingly, recent data suggests that the nuclear transcriptional history of an mRNA influences whether a miRNA represses its translation at the initiation or the elongation step (Kong YW, 2008).

Degradation of mRNAs associated to AGO-miRNA complexes involves deadenylation (removal of the poly(A) tail) (Eulalio A, 2007c), and/or decapping (removal of the m^7G cap structure) (Rehwinkel J, 2005). The exposure of mRNA termini led to exonucleolytic digestion from the 5' and 3'ends by the action of the exosome and of the exonuclease XRN1, respectively (Parker R, 2004b).

miRNAs might also silence their targets by sequestering mRNAs into P-bodies, which exclude the translation machinery, and releasing them upon specific signals allowing the re-enter into polyribosomes fraction for translation (Bhattacharyya SN, 2006).

Although miRNAs have long been thought to act by down-regulating protein synthesis, recent reports indicate that they can also stimulate translation under certain conditions (Vasudevan S, 2007). Thus, translational regulation by miRNAs would oscillates between repression and activation during the cell cycle: in proliferating cells they repress translation, whereas in G1/G0 arrest they mediate activation (Vasudevan S, 2007).

Transcriptional regulation. In addition to their post-transcriptional role in the cytoplasm, AGO proteins also function in the nucleus either repressing or activating transcription (Gagnon KT, 2012). In humans, the existence of nuclear RNAi pathways was first inferred from observations that nuclear miRNAs and siRNAs could cause cleavage of RNA targets, such as 7SK small nuclear RNA (Robb GB, 2005). In addition, endogenous human AGO2 was identified in the nucleus with highly specific antibodies (Rudel S, 2008).

The first report of human transcriptional gene silencing (TGS) utilized exogenous siRNAs to silence expression of an integrated GFP reporter driven by the eukaryotic translation elongation factor alpha (EF1α) promoter (Morris KV, 2004). The silencing at transcriptional level was confirmed by nuclear run-on analysis, which is the gold standard for differentiating silencing effects mediated by TGS from those of the post-transcriptional pathways.

Subsequently, other reports described the silencing of various genes, including E cadherin, RASSF1, TGFb receptor II, progesterone receptor, major vault protein, androgen receptor, cyclooxygenase-2, CDH1, and c-myc (Ting AH, 2005) (Castanotto D, 2005) (Janowski BA, 2007) (Kim JW, 2007) (Napoli S, 2009) (Green VA, 2011).

Mechanistic details about TGS have begun to emerge after the observation that siRNA-induced silencing at the EF1α promoter was sensitive to the Pol II inhibitor α-amanitin (Weinberg MS, 2006). This observation led to the proposal of two models to explain the mechanism of TGS: (a) the siRNA binds to DNA, facilitated by the opening of the DNA duplex by the transcription machinery; (b) the siRNA binds to nascent promoter-associated RNA. Work by Han *et al.* favored the latter by revealing a requirement for a promoter-associated transcript (Han J, 2007). The antisense strand of the siRNA was shown to target a EF1α promoter-associated RNA variant with an extended 5'UTR. In addition, the association was inhibited by RNase A but resistant to RNase H treatment, implicating an RNA-RNA interaction in TGS.

Other studies have also demonstrated a requirement for promoter-associated transcripts for small RNA-mediated gene silencing. Some of these studies have shown that, similar to EF1α, silencing requires sense strand transcripts (Hawkins PG, 2009), while others have found that antisense transcription through the targeted region is necessary (Schwartz JC, 2008).

The findings reported to date, support a model for TGS in which AGO-small RNAs complexes are guided to their targets by complementary base-pairing with low copy, promoter-associated Pol II transcript. This serves as a scaffold for the recruitment of chromatin-modifying complexes that favor heterochromatin formation and, therefore, decrease transcription of the targeted gene.

More recently, small RNA-targeting of sequences beyond the mRNA 3'terminus of the PGR gene has been demonstrated to induce TGS (Yue X, 2010). Targeting genomic regions outside promoter sequence involves a looping mechanism bringing the 3' terminus and the promoter into close proximity to allow modulation of promoter activity.

The report of transcriptional gene activation (TGA) showed increase expression of E-cadherin, p21, and VEGF on the transfection of promoter-targeting siRNAs into cultured cells

(Li LC, 2006). The mechanism by which transcriptional activation is achieved shares common features with the transcriptional silencing. The transcriptional opposite outcome depends on the specific region targeted by the small RNA and by the different set of modifying complexes recruited.

Emerging evidence has revealed that the TGA/TGS in human cells can be mediated not only by exogenous small RNA but also by endogenous ones. Consistently, it has been reported that a substantial fraction of human miRNAs are present in the nucleus, with even an abundance greater than their cytoplasmic levels, and, in some cases, shown to alter promoter activity. Cases such as miR-373 and miR-320 in which the former induces and the latter represses transcription (Place RF, 2008) (Kim DH, 2008). Moreover, it has been recently discovered endo-siRNAs also in mammals with a putative role in transcriptional regulation (Yang N, 2006). Many are the genomic sources of dsRNA triggers for endo-siRNA such as: structured loci that pair intramolecularly to produce long dsRNA, complementary overlapping transcripts and bidirectionally transcribed loci, protein-coding genes that associate with the cognate pseudogenes and from regions of pseudogenes that form inverted repeated structures (Ghildiyal M, 2009).

SWI/SNF chromatin remodeler complexes

In eukaryotic cells, genetic information encoded in over 1 meter of DNA is packaged into chromatin and compartmentalized in the nucleus. The basic unit of chromatin is the nucleosome, which consists of 146 base pairs of duplex DNA wrapped around a histone octamer comprised of two of each of the conventional histone proteins: H2A, H2B, H3 and H4. Highly related histone variants are also incorporated throughout the genome for regulatory purposes (Talbert PB, 2010). A fifth histone protein, H1, promotes higher order chromatin structures by encouraging condensation of neighboring nucleosomes from “beads on a string” to the 30 nm fiber. As this resulting fiber accounts for only ~25-fold of a 5000-fold DNA-to-nucleus compaction ratio, several other as yet unknown mechanisms must contribute to higher order compaction and nuclear organization.

Despite this intricate packaging, DNA must be accessible for critical cellular processes such as transcription, replication, recombination, and repair. DNA accessibility is facilitated by two classes of enzymes, ATP-dependent nucleosome remodelers and histone modifying enzymes, which work in concert to direct nucleosome dynamics. In this way, chromatin organization provides a packing solution and an opportunity for gene expression regulation.

Histone modifying enzymes modify the N-terminal tails of histone proteins (by acetylation, methylation, phosphorylation and ubiquitylation) to alter the structure of chromatin and provide binding sites for regulatory proteins. In fact, many chromatin-associated proteins contain domains that bind histone modifications such as the bromodomain that recognizes acetylated residues.

Chromatin remodeling complexes utilize the energy of ATP to disrupt nucleosome-DNA contacts, move nucleosomes along DNA, and remove or insert histone octamers. They thus make DNA/chromatin available to proteins that need to access DNA or histones directly during cellular processes. These enzymes play an essential role in several aspects of embryonic development including pluripotency, cardiac development, dendritic morphogenesis and self-renewal of neural stem cells. In the adult, deletion or mutation of these proteins often leads to tumorigenesis as a consequence of a dysregulated cell cycle control (Weissman B, 2009)

The best-studied family of chromatin remodeler is the evolutionarily conserved SWI/SNF family (Clapier CR, 2009). The human SWI/SNF complexes are large multi-subunit complexes approximately of 2MDa in size that use the energy of ATP hydrolysis to remodeler nucleosome. These complexes consist at least of 9-12 subunits including one of two mutually exclusive catalytic ATPase subunits (BRG1 and BRM); a set of highly conserved core subunits (SNF5, BAF155 and BAF170); and accessory subunits that are thought to contribute to the targeting, assembly and regulation of lineage-specific functions of the complexes. The combinatorial assembly of accessory subunits, some of which have multiple isoforms, enables multiple varieties of SWI/SNF complexes to exist both within a given cell and across cell types (Wu JI, 2009).

In mammals cells, the SWI/SNF family can be divided into BAF and PBAF subfamily. This subdivision is due to the exclusive presence of ARID1a-b only in BAF complexes and BAF180, BAF200, and BRD7 in PBAF complexes (Fig.2). These exclusive subunits provide functional specificity recruiting the complexes to specific loci and the association with specific proteins (Nie Z, 2000) (Thompson M, 2009).

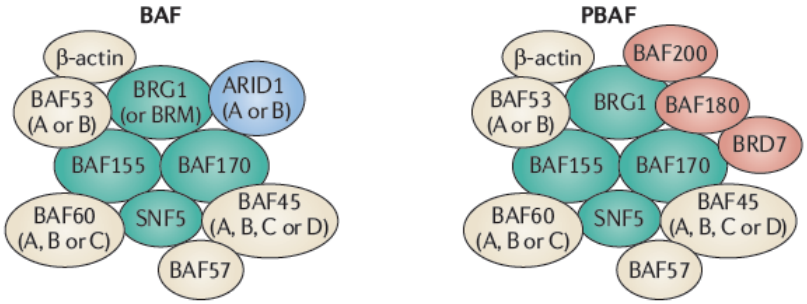

Figure 2 SWI/SNF complexes subfamilies. The human SWI/SNF complexes are divided into two subfamilies BAF and PBAF in base of exclusive subunits ARID1 (blue) and BAF 180 (red), respectively. Invariants proteins of the SWI/SNF complexes (core and ATPase subunits) are reported in green, whereas accessory proteins are shown in light brown (Wilson BG, 2011).

The multiple assembly of the subunits reflects the pleiotropic functions of the SWI/SNF complexes that range from DNA replication (Cohen SM, 2010), DNA repair (Osley MA, 2007) to gene expression regulation (Urnov FD, 2001).

To exert this multiplicity of roles, SWI/SNF complexes have to first contact specific genomic region either directly, contacting DNA and histone modifications, or indirectly by protein-protein interactions.

The direct recruitment is mediated by DNA binding domains (HMG-like and AT-rich interactive) and by domains that recognize specific histone modifications, such as acetylation

(bromodomains) and methylation (chromodomains), present in several complexes' subunits. Indeed, the SWI/SNF complexes are defined polymorphic readers of epigenetic modifications (Wu JI, 2009). Moreover, their large size (about 12-fold bigger than a nucleosome) should enable reading of multiple histone modifications on adjacent nucleosomes providing significant affinity for either targeting and retention.

Conversely, the indirect recruitment involves protein-protein interaction with histone modifying complexes (Zhang HS, 2000) and with transcriptional activators and repressors such as p53 (Lee D, 2002), cyclin E (Shanahan F, 1999), nuclear hormone receptors (Trotter KW, 2008).

One of primary functions of SWI/SNF complexes is to assist in gene regulation. A genome-wide analysis of SWI/SNF binding sites indicates that SWI/SNF complexes likely contribute to gene regulation through many different avenues. In fact SWI/SNF complexes are reported to bind promoters, enhancers, CTCF sites and many regions occupied by Pol II (Euskirchen GM, 2011). Furthermore, SWI/SNF complexes may facilitate looping interactions among these various elements. It has been shown *in vitro* that SWI/SNF can interact simultaneously with multiple DNA sites and generate loops between them (Bazett-Jones DP, 1999).

So, once targeted to specific promoters or enhancers, SWI/SNF complexes promote transition to and from the active and repressed chromatin state. These opposite effects are mediated by interactions with different types of coregulators. For example, SWI/SNF complexes interact with MYC, a transcription factor that regulate gene expression programmes during cell cycle progression, apoptosis and differentiation. In particular, the core subunit SNF5 directly interacts with MYC and is capable of cooperating in the activation of MYC target genes *in vitro* (Cheng SW, 1999). By contrast, the ATPase subunits BRG1 binds to RB and facilitates the repression of RB target genes, including E2Fs and CCND1 (Trouche D, 1997).

Transcriptional activation by SWI/SNF complexes is achieved by sliding or ejecting nucleosome allowing the transient exposure of binding sites for transcriptional activators and the stabilization of pre-initiation complex formation (Salma N, 2004). Additionally, SWI/SNF have been shown to be associated with regions downstream the promoter and to influence the

RNA Pol II promoter escape (Soutoglou E, 2002) and transcription elongation (Corey LL, 2003).

On the contrary, SWI/SNF repress transcription inducing a more compact chromatin conformation, and also allowing transcriptional repressors to bind to chromatin or facilitating other modifications of the chromatin such as deacetylation of histones.

Besides the role in transcriptional regulation, recent findings illustrate that SWI/SNF act also posttranscriptionally regulating not only the amount of mRNA synthesized from a given promoter but also the type of alternative transcript produced. SWI/SNF interact with components of the nascent pre-mRNP complex and with the transcription machinery (Cho H, 1998). These interactions cause a delay in the transcription elongation rate, which in turn affects splice site selection. Moreover, SWI/SNF associate with several components of the spliceosome and with Sam 68, an ERK-activated enhancer of variant exon inclusion (Batsché E, 2006). The posttranscriptional regulation of SWI/SNF seems to be independent from its chromatin remodeling activity. In fact, mutation of the ATPase domain of hBrm did not impair its effect on splicing regulation (Batsché E, 2006).

Aim of the Thesis

Studies over the past years have provided strong evidence regarding the multi-level regulation of gene expression operated by AGO proteins. In particular, human AGO2 protein, the only AGO endonucleolytically active, has been demonstrated to act either post-transcriptionally or at transcriptional level. It is likely that this multiplicity of roles is exerted by (a) the association with specific proteins that would modulate AGO2 activity in response to determined cellular cues, and target AGO2 to different cell compartments and to different protein complexes; (b) the small RNA loaded into AGO2 protein which confers specificity to AGO2 complexes for the RNA molecule to be targeted.

In order to elucidate the diversified roles of AGO2, we aimed at the characterization of both proteins and small RNAs associated to AGO2 protein by high-throughput technologies.

To provide more physiological results compare to previous studies effectuated in overexpression conditions, a monoclonal antibody against the endogenous human AGO2 protein was utilized.

From this extensive analysis, we will be able to provide further insights into known AGO2 functions and to identify new biological processes in which AGO2 is involved.

Experimental Procedures

Cell culture

HeLa S3 cells were grown in DMEM (Gibco) supplemented with 10% fetal bovine serum (Sigma), L-glutamine (Gibco) and penicillin-streptomycin (Gibco).
For the SILAC (stable isotope labeling by amino acids in cell culture) analyses, HeLa S3 cells were grown, in SILAC DMEM (PAA) supplemented with 10% dialyzed fetal bovine serum (PAA), nonessential amino acids L-glutamine and L-leucine (Sigma), penicillin/streptomycin, and either a) "light" amino acids, L-lysine and L-arginine (SIGMA) or b) "heavy" isotope labeled amino acids, L-lysine-$^{13}C_6$, $^{15}N_2$-HCl (Sigma-Aldrich/Isotec) and L-arginine-$^{13}C_6$, $^{15}N_4$-HCl (Sigma-Aldrich/Isotec). The cells were adapted for at least six generations to achieve complete incorporation of the isotope labeled amino acids before they were propagated to a scale needed for the experiment.

Extracts preparation

For total cell extract, cells were washed and scraped off the culture plates with PBS, then lysed in appropriates volumes ($20x10^6$ cells/ml) of IP-buffer (150 mM KCl; 25 mM NaCl; 2 mM EDTA; 0.5% NP40; 0.5 mM DTT; protease inhibitor (Sigma)) and incubated on ice for 20 min. After that, the lysate was cleared for 10 min at 16,000 g.
Cytosolic and nuclear extracts were prepared as following: cells were resuspended, firstly, in two volumes of ice-cold buffer I (0.3 M sucrose in 60 mM KCl; 15 mM NaCl; 5 mM MgCl2; 0.1 mM EGTA; 15 mM Tris-HCl pH 7.5; 0.5 mM DTT; protease inhibitors), then added other two volumes of ice-cold Buffer II (Buffer I with 0.4% NP-40) ($25x10^6$ cells/ml final), mixed gently, and placed on ice for 10 min. Afterwards, cell lysate was layered on 24 ml of a sucrose cushion (1.2 M sucrose in 60 mM KCl; 15 mM NaCl; 5 mM $MgCl_2$; 0.1 mM EGTA; 15 mM Tris-HCl pH 7.5; 0.5 mM DTT) and centrifuged in a pre-chilled swing-out rotor at 10,000 g

for 20 minutes at 4°C. The cytoplasmic fraction was removed, cell nuclei were lysed in IP-buffer ($20x10^6$ cells/ml) and cleared. The nuclear and cytosolic fractionation efficiency was evaluated by western blotting analysis. H1 and GAPDH proteins were used as controls of the nuclear and cytosolic fraction, respectively.

Western blotting analysis

Proteins were resolved on 6 or 10% SDS-PAGE gels and transferred to protran nitrocellulose membrane (Whatman) by semi-dry blotting. Blots were blocked with 5 or 10% nonfat dry milk in PBST, depending on the antibody that was subsequently used. Specific proteins were detected by using SuperSignal West Pico Chemiluminescent reagents (Thermo Scientific). The following antibodies were used: anti-AGO2 (clone 11A9 Ascenion); anti-BAFF155 (Abcam); anti-BRG1 (Abcam); anti-GAPDH (Cell Signaling); anti-H1 (Millipore); anti-tubulina (Sigma), goat-anti mouse and anti-rabbit IgG-HRP conjugated (Bio-Rad); anti-rat IgG-HRP conjugated (Jackson).

Chromatin binding assay

The chromatin binding assay was performed as described previously (Cernilogar FM, 2011). Briefly, $10x10^6$ HeLa S3 cells were washed with cold PBS. $1x10^6$ cells (control fraction, C) were resuspended in RIPA buffer (150 mM Tris-HCl, pH 8.0, 150 mM NaCl, 0.5% DOC, 0.1% SDS, 1% NP-40, Protease inhibitors) and left for 30 min. on ice. The remaining cells were lysed for 15 min on ice in cold CSKI buffer (10 mM PIPES, pH 6.8, 100 mM NaCl, 1 mM EDTA, 300 mM sucrose, 1 mM $MgCl_2$, 1 mM DTT, 0.5% Triton X-100, Protease inhibitors). The cell lysate was divided into two portions, which were centrifuged at 500 g at 4 °C for 3 min. The supernatants (S1 fraction) contain Triton-soluble proteins. One of the pellets was washed twice in CSKI buffer and then resuspended in RIPA buffer (the P1 fraction). The second pellet, after washing in CSKI buffer, was resuspended in CSKII buffer (10 mM PIPES, pH 6.8, 50 mM NaCl, 300 mM sucrose, 6 mM $MgCl_2$, 1 mM DTT, protease inhibitors), then treated with DNase I (Roche) for 30 min. at 25°C and extracted with 250 mM

NH_2SO_4 for 10 min at 25 °C. The sample treated with DNase and salt was then centrifuged at 1,200 g for 6 min. at 4 °C, and the supernatant (S2 fraction) and the pellet (P2 fraction) were collected. The P2 fraction was also resuspended in RIPA buffer. The equivalent of $1x10^6$ cells of all fractions were analyzed by western blotting.

Immunoprecipitations

SILAC experiments. Light and Heavy total-cell lysates were pre-cleared with Protein G-sepharose beads (Sigma) for 2 hr 4°C under constant rotation and then filtered 0.45 µm filter (Millipore). Concurrently, anti-AGO2 (clone 11A9 Ascenion) and anti-rat IgG (Sigma) antibodies were coupled to Protein-G-sepharose beads for 2 hr at 4°C. After that, protein concentration was determined using Bradford reagent (Bio-Rad) and equal amounts of proteins from Heavy and Light lysates were used for the subsequent immunoprecipitation (IP). An aliquot of total extracts was taken out (input). In a first experiment, from the Heavy lysate was conducted the specific IP for AGO2 protein whereas from the Light lysate was performed the control IP with nonspecific rat IgG. In a second experiment, the pairing of lysates and antibodies was swapped. After overnight (o.n.) IP, the beads were washed once with IP-buffer for 5 min., next three times with IP-wash buffer (50 mM Tris-HCl, pH 7.5; 300 mM NaCl; 5 mM $MgCl_2$; and 0.05% Nonidet P-40). An aliquot of each IP was taken out to control the immunoprecipitation efficiency by western blotting. At this moment, Heavy IP and Light IP were mixed together and processed as a single sample throughout the following elution step. The beads were equilibrated in elution buffer (300 mM NaCl, 50 mM Tris-HCl pH 8, 0.01% NP-40, 5 mM $MgCl_2$) and the proteins immunoprecipitated were eluted from the antibody matrix by competition with synthetic AGO2 peptide (GeneScript) in about 50 µl of elution buffer for 1 h at 22°C 750 rpm for two times. The eluate was collected, cleaned up from eventually remaining beads with Micro Bio-Spin Chromatography Columns (BIO-RAD) and subjected to mass spectrometry analysis.

Co-immunoprecipitations experiments. Whole-cell lysate (about $30x10^6$ cells) or nuclear fraction of HeLa S3 cells (about $80x10^6$ cells) were incubated o.n. at 4°C under constant

rotation with antibodies pre-coupled beads (anti-AGO2; anti-rat IgG; anti-BAF155 (Abcam); anti-mouse IgG (Sigma)). All IP samples were washed once with IP-buffer for 5 min. and three times with IP-wash buffer. SDS-PAGE sample buffer (Sigma) was added to the beads and the co-immunoprecipitated proteins boiled for 5 min. and analyzed by western blotting.
For the analysis of RNA-dependent interactions, two approaches were utilized. In the first, AGO2-IP and the control IgG-IP were equally distributed into two tubes and 50 µl of PBS containing 100 µg/ml RNAse A (Sigma) were added to one tube, whereas only PBS to the other tube. After incubation at 4°C for 1 hr, beads were denaturated by adding protein sample buffer. In the second approach, RNase was added directly at the co-immunoprecipitation step. An aliquot of the lysate (input) and IP with or whitout RNAse A were taken out and RNA extracted by TriReagent (Sigma) according to the manufacturer's protocol. The RNA degradation efficiency was evaluated by agarose gel electrophoresis and by RT-qPCR analysis of miR-21, miR-16 and GAPDH mRNA.

RNA Immunoprecipitation experiments (RIP). The purified nuclei of about $200x10^6$ HeLa S3 cells were lysated in IP-buffer supplemented with RNase inhibitors (Promega) and centrifugated at 16,000g at 4°C. Next, the nuclear lysate was pre-cleared by adding Protein G-sepharose beads for 2 hr 4°C under constant rotation and then filtered through 0.45 µm filter. Lysate aliquots were taken out for RNA and protein inputs.
Antibodies anti-AGO2 and anti-rat IgG were coupled to protein G sepharose beads for 2 h at 4°C in IP-buffer containing 1 mg/ml heparin (Sigma). Coupled sepharose beads were washed with IP-buffer and subsequently incubated with the pre-cleared nuclear lysate o.n. at 4°C. Immunoprecipitates were washed ones with IP-buffer 5 min. 4°C, and three times with IP-wash buffer 5 min. 4°C. After the final wash, the beads were reconstituted with IP wash buffer and an aliquot was taken out for western blot. The input sample and IPs were firstly treated with DNaseI-RNase-free 10 min. at 37°C, secondly proteinase K digested (Roche). Total RNA and co-precipitated RNA was extracted by phenol:chloroform:isopropyl alcohol and precipitated in ethanol at -80°C. The efficiency of nuclear-AGO2 immunopurified RNAs was controlled by qRT-PCR of miR-21 and miR-16.

RNA samples were send to a facility company for libraries preparation and sequencing. Sequencing was performed on Illumina platform.

qRT-PCR

RNA was reverse-transcribed with enhanced AMV reverse transcriptase (Sigma) using specific primers. Subsequently, cDNA was used as template for qPCR performed with iCycler (Bio-Rad) using SensiMix SYBR & Fluorescein (Bioline) according to the manufacturer's instructions. All real-time experiments were performed in triplicate. Data were normalized relative to measured levels of U4 small nuclear RNA.

Bioinformatics analysis on RNA-seq data

The adaptor sequences, attached during short library construction, were removed from the raw data. Afterwards, we evaluated the reads quality using FastQC. The reads were filtered out for microRNAs, tRNAs, rRNAs and other known non-coding sRNAs by mapping against miRBase (version 18) database and hand-cured non-coding database using Bowtie (Langmead B, 2009). The resulting filtered sequences were aligned to the hg19 assembly version of human genome. Alignment parameters were to restrict to perfectly matching reads in at least first 18 bases. Only the reads that map uniquely on the human genome were further analyzed. The reads were clustered by merging into a single cluster reads that map on the same strand of the genome with at least 1 nt of overlap After that, the genomic coordinates corresponding to the clusters were intersected with the genomic coordinates of the selected features generated from the ENCODE project (Affymetrix ENCODE Transcriptome project, 2009). The selected features were reported in Tab.1. In order to examine the significance of our data, for each of the selected genomic features a shuffle file was produced. The shuffle file was generated by repositioning the genomic coordinates of the feature randomly on the same chromosome. Next, the overlap ratio between our data with real (m value) and shuffled (n value) genomic features was computed. A ratio m/n <1 indicates a casual overlapping between our sequences

with the selected genomic features. A ratio m/n >1 suggests a likely correlation. Clustering and genomic intersections were performed using dedicate BedTools (version 2.14.2).

Selected Genomic Features		
Structural	**Euchromatin-associated**	**Heterochromatin-associated**
5000 upstream to genes	H2A.Z	H3K27me3
5000 downstream to genes	H3K04me1	H3K9me3
100 upstream to genes	H3K04me2	
5' UTR	H3K04me3	
3'UTR	H3K27ac	
Genes	H3K9ac	
Exons	H3K36me3	
Introns	Strong promoters	
First 40 nt transcribed	CpG islands	
Strong enhancers	P300 protein	
	TBP protein	
	TFIIF protein	
	BAF155 protein	
	BRG1 protein	

Table 1 Selected genomic features. The genomic features are divided into structural and associated to chromatin, active (euchromatin) and inactive (heterochromatin).

Results

Proteomic characterization of AGO2-associated proteins

As part of our efforts to characterize AGO2 complexes, we investigated the AGO2-associated proteins by SILAC/MS approach. The SILAC is a metabolic labeling strategy that uses stable isotope labeled amino acids in growth medium to encode cellular proteomes for quantitative analysis. To this end, HeLa S3 cells were grown in media containing the "light" amino acid or its isotopically labeled ("heavy") analogue. From the total lysate of "heavy" cells we immunopurified the endogenous AGO2 protein using a monoclonal antibody. As control, a parallel immunoprecipitation with aspecific IgG was performed on the "light" cells. Moreover, a swapping between antibodies and SILAC labeled cells was conducted.

The two biological replicas were concordant and provided us the identification of several known and novel AGO2-associated proteins. In Tab2 is reported the top list of co-purified proteins of one experiment. The proteins were ranked on the enrichment value of the specific immunoprecipitation over the control one.

As expected, we identified some of the proteins previously shown to associate with AGO2 which result implicated in transcriptional regulation and mRNA destabilization. Among the most enriched we found TNRC6A-B-C, MOV10, HuR, the subunit 1 of CCR4-NOT deadelynation complex and PABP. We identified other members of AGO subfamily (AGO1-AGO3) and several heterogeneous nuclear ribonucleoproteins, ribosomal proteins, RNA helicases.

Moreover, from our proteomic analysis emerged several nuclear proteins that were not known to be associated to AGO2.

With the aim to uncover if AGO2 associates with proteins that are part of specific macromolecular complexes or are involved in same pathways, we performed a Gene Ontology analysis using GOrilla tool. The ranked list of AGO2-associated proteins were

given as input to GOrilla with default running parameters. The output for cellular components category is reported in Fig.3 as a directed acyclic graph (DAG).

The analysis revealed various macromolecular complexes many of which are involved in the well-known post-transcriptional role of AGO2. Besides the cytoplasmatic complexes, from the analysis emerged with high significance nuclear complexes implicated in mRNA processing (mRNA cleavage and polyadenylation) and in chromatin remodeling. Intriguingly, our results revealed the association of AGO2 to the main family of chromatin remodeling complexes: the SWI/SNF family. Among the most highly enriched proteins in our AGO2-immunoprecipitations we found all the core proteins (BAF155, BAF170, SNF5), both the ATPase subunits (BRG1 and BRM) and several accessory proteins including the subunits signature of BAF (ARID1a-b) and PBAF complexes (BAF180). As results, our quantitative MS analysis suggests that AGO2 is broadly associated to several SWI/SNF complexes.

Proteins	Rank	Proteins	Rank	Proteins	Rank
BAF45D	1	FIP1	34	MYH9	67
SNF5	2	DSP	35	TUBGCP2	68
IGLC1	3	FAM83G	36	DDX3	69
BAF57	4	CDHF1	37	ACTA1	70
VPS35	5	VH-3 family (VH26)D/J	38	CLLD7	71
AGO2	6	PDHX	39	LSM12	72
TNRC6C	7	IGHV	40	KIAA0221	73
CSTF3	8	H2AFX	41	PSF	74
LAK1	9	CASP14	42	TBC1D5	75
BAF155	10	BAF180	43	KIAA0731	76
CSTF1	11	CPSF100	44	MOV10	77
CSTF2T	12	WDC146	45	NSEP1	78
ANX2	13	TNRC6A	46	RPS26	79
SMARCD2	14	BAF170	47	MYL6	80
ARID1A	15	AGO3	48	HNRNPK	81
KIAA0268	16	AGO1	49	P4HA	82
TNRC6B	17	CYPA	50	HNRNPM	83
CSNK1A1	18	ARID2	51	CTNNG	84
BCL7A	19	IGF2BP3	52	RPS24	85
SS18	20	C20orf99	53	CFL	86
BAF53A	21	PABPC4	54	HNRNPF	87
BAF60A	22	TFG	55	HNRNPA3	88
CLPX	23	PABP1	56	RPL1	89
KIAA0292	24	CDHF4	57	NCL	90
H3.3A	25	HGRG8	58	DDX5	91
ARID1B	26	CCR4-associated factor 1	59	DDX15	92
VPS26B	27	LMNB1	60	RPS6	93
BRG1	28	CSTF2	61	KIAA0264	94
CTAGE5	29	TUBG	62	MRLC2	95
H2BFD	30	MATR3	63	HSD14	96
BRM	31	HUR	64	HSP90AB1	97
DLAT	32	CPSF1	65	AUF1	98
H4/A	33	EDH17B4	66	DAP3	99

Table 2 SILAC/MS results. Top-list of AGO2-associated proteins ranked on fold enrichment over the control immunoprecipitation.

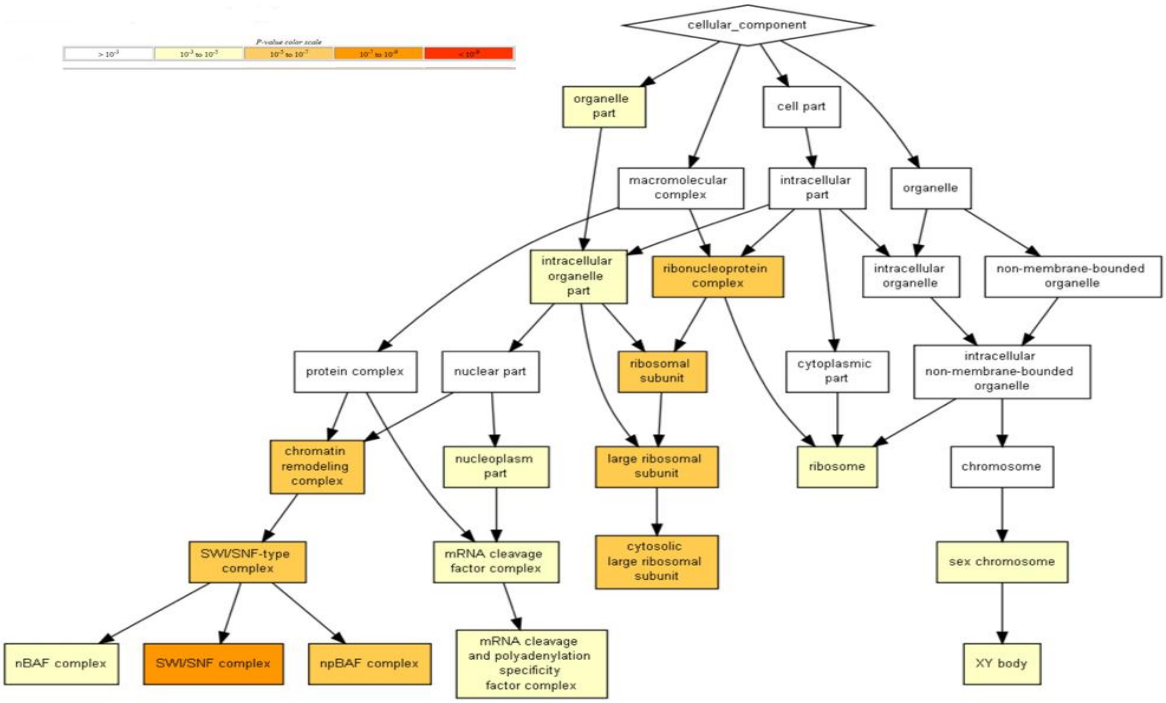

Figure 3 GO terms analysis (cellular component category) of AGO2-associated proteins. The resulting enriched GO terms are visualized using a DAG graphical representation with color coding reflecting their degree of enrichment.

Validation of AGO2-SWI/SNF complexes interaction

As the nuclear proteins emerged from the proteomic analysis were particularly interesting, consistent with evidence on the AGO2 role in transcriptional regulation, we focalized our study on AGO2-nuclear complexes. In particular, we proceeded to investigate the interaction between AGO2 and SWI/SNF chromatin remodeler complexes.

To confirm that AGO2 and SWI/SNF associate *in vivo*, immunoprecipitations and western blotting experiments on HeLa S3 cells were carried out. We immunoprecipitated AGO2 protein from whole cell extracts of unlabeled HeLa S3 and by western blotting verified the interaction with BAF155 (a core protein) and BRG1 (one of the two mutually exclusives ATPase) (Fig.4A). Moreover, we performed the reverse immunoprecipitation using antibody anti-BAF155 (Fig.4B). As shown in Fig.4A, both BAF155 and BRG1 co-immunoprecipitate specifically with AGO2 protein and, in a specular manner, AGO2 co-immunoprecipitates with the core subunits of SWI/SNF complexes.

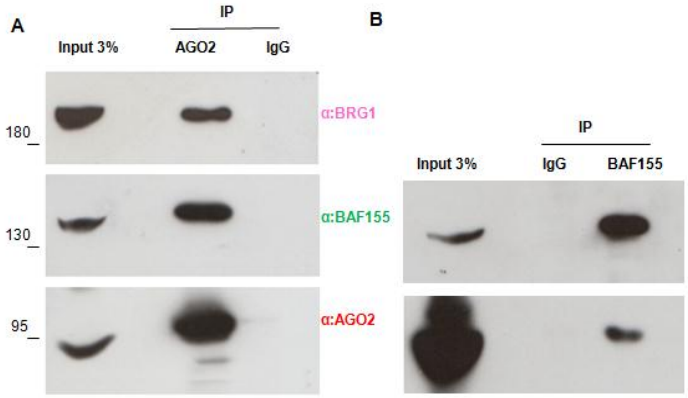

Figure 4 Human AGO2 interacts with SWI/SNF subunits BAF155 and BRG1. (A) HeLa S3 whole cell extracts were immunoprecipitated using anti-AGO2 or IgG (negative control) and analyzed by western blotting using antibodies anti-BAF155 (core subunits) and anti-BRG1 (one mutually exclusive ATPase). **(B)** Reverse co-immunoprecipitation using BAF155 antibody was blotted with anti-AGO2.

AGO2 is known to exert its multiple functions in union with both small and long RNA molecules. Therefore, to examine whether the AGO2-SWI/SNF complexes interaction was RNA dependent, we performed RNase A treatments on immunoprecipitates.

As shown in Fig.5A, the interaction of BAF155 and BRG1 with AGO2 was not impaired upon RNAse A treatment. Therefore, we demonstrated that the association is independent of RNA molecules and that involves a protein-protein interaction.

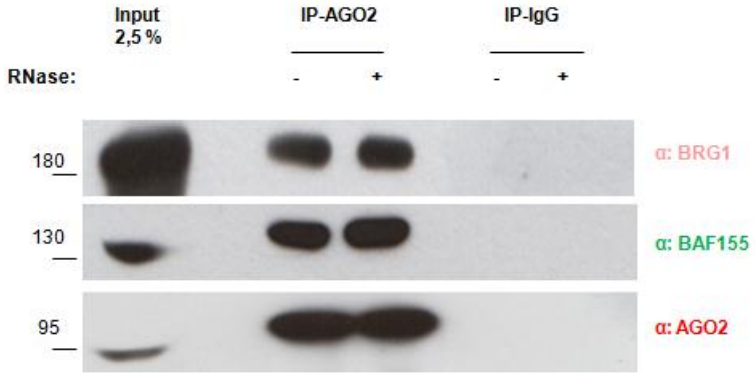

Figure 5 AGO2-SWI/SNF complexes interaction is RNA independent. In the experiment reported, AGO2-coimmunoprecipitates from HeLa S3 whole cell extracts were treated or not with RNase A and the association with BAFF155 and BRG1 were evaluated by western blotting.

SWI/SNF complexes interact specifically with AGO2 and not with AGO1

In several reports it has been demonstrated a role for AGO1 in transcriptional regulation (Kim DH, 2006). Moreover, in our proteomic analysis AGO1 was found associated with AGO2. Hence, to investigate if AGO1 interacts with SWI/SNF complexes too, we carried out the co-immunoprecipitation experiments using both AGO1 and BAF155 antibody (Fig.6). Neither immunoprecitating for AGO1 (Fig.6A) nor for BAFF155 (Fig.6B) it was possible to observe an interaction between these proteins.

Thus, our result suggests that only AGO2 specifically interacts with SWI/SNF complexes.

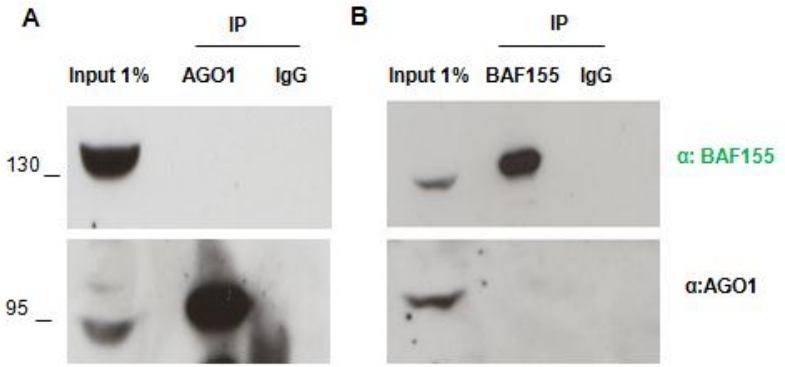

Figure 6 AGO1 does not interact with the SWI/SNF complexes. The interaction between AGO1 and BAF155 was tested by direct **(A)** and reverse **(B)** co-immunoprecipitation experiments in HeLa S3 cells followed by western blotting analysis.

AGO2 interacts with SWI/SNF complexes in the nucleus

SWI/SNF complexes have predominantly a nuclear localization in virtue of their activity in chromatin remodeling. In addition, AGO2 nuclear localization has been reported (Rüdel S, 2008). Hence, reasoning about a plausible concerted mechanism of action between AGO2 and SWI/SNF complexes it was necessary to confirm such interaction in the nucleus, besides the validations performed on whole cells extracts. To this aim, we performed a cell fractionation into cytosolic and nuclear fraction, followed by AGO2 immunoprecipiation from the nuclear fraction. The association with BAF155 was verified by western blot (Fig.7B). The purity of our cytosolic and nuclear fractions was determined by immunoblot analysis, as shows Fig.7A, the cytoplasmic marker GAPDH and the nuclear marker H1 were appropriately partitioned into cytosolic and nuclear fractions, respectively.

As shown in Fig.7B, our results confirm the interaction between AGO2 and BAF155 in the nucleus of HeLa S3 cells.

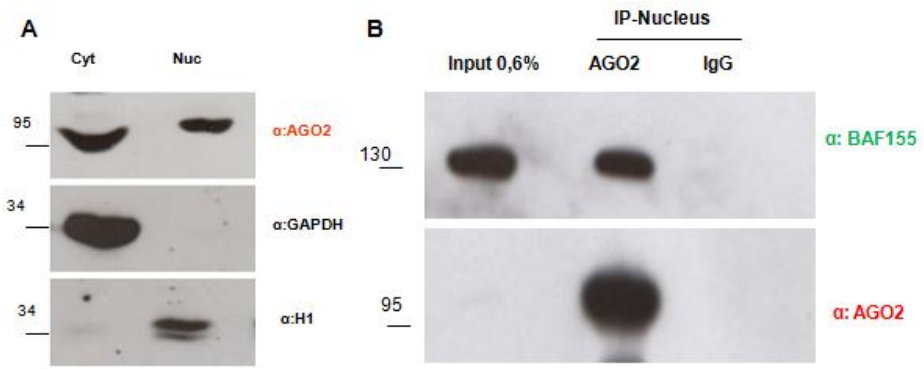

Figure 7 AGO2 and SWI/SNF complexes interact in the nucleus. (A) Control of cell fractionation by blotting for GAPDH and H1, cytosol and nuclear marker respectively. **(B)** AGO2-BAF155 co-immunoprecipitation from the nuclear fraction of HeLa S3 cells.

Characterization of AGO2-SWI/SNF complexes interaction in the nucleus

To evaluate the presence of AGO2 in different nuclear compartments, we carried out a chromatin fractionation experiments. This fractionation also known as chromatin-binding assay was adapted to HeLa S3 from Cernilogar FM *et al.* (Cernilogar FM, 2011) and the experimental scheme reported in Fig.8A. Briefly, this fractionation protocol rely on the separation of the chromatin-bound proteins (P1) from proteins that are not bound to chromatin (S1) in base on solubility to TritonX-100 of S1 fraction proteins. The chromatin-bound proteins are instead released only after high salt and DNase treatment (S2). The efficiency of the fractionation was assessed by the specific distribution of Tubulin, a chromatin unbound marker, and H1, a chromatin-bound marker (Fig.8B).

The AGO2 and BAF155 distribution profiles through the several resulting fractions were compared (Fig.8B). As expected, AGO2 was mainly detected in the S1 fraction, that includes the cytosolic and nucleoplasmic components, respect to BAF155. Interestingly, a substantial amount of AGO2 was revealed also in S2 fraction that specifically represents protein associated to chromatin as well as BAF155.

Given that we demonstrated the interaction of AGO2 with SWI/SNF also in nucleus and identified AGO2 in the chromatin fraction, we suggests that the AGO2 and SWI/SNF complexes may co-operate at chromatin level.

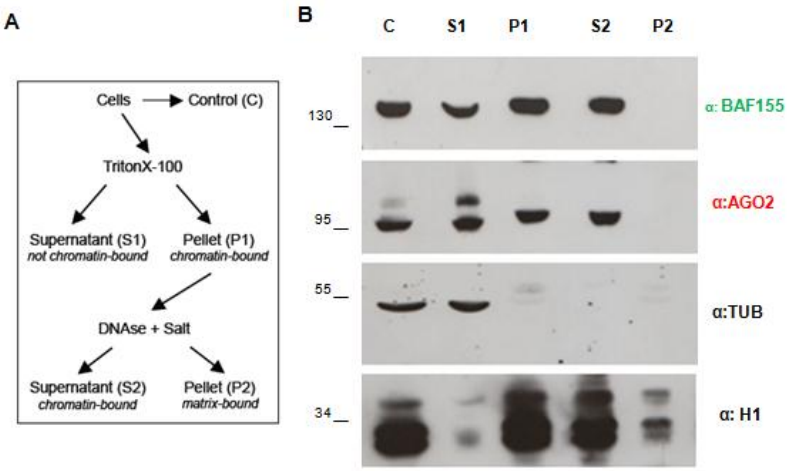

Figure 8 Chromatin binding assay. (A) Scheme of the procedure used to fractionate HeLa S3 cells (Cernilogar FM, 2011). Chromatin associated proteins should be found in fractions P1 and S2. **(B)** Equal amounts of each resulting fraction was analyzed by western blotting for the presence of the indicated proteins. Tubulin (TUB) serves as a chromatin unbound marker, whereas H1 is a chromatin bound marker that is fully released only after high salt and DNase treatment.

Analysis of nuclear AGO2-associated small RNA

The interaction of AGO2 with SWI/SNF complexes prompted us to hypothesize a concerted role of AGO2 and SWI/SNF in chromatin dynamics and possibly in gene expression regulation.

The current model for gene expression regulation exerted by AGO2 indicates that AGO2 utilizes small RNAs to recognize target molecules.

Therefore, in order to elucidate the mechanism through which AGO2 can bind to chromatin and recognize specific loci, we performed the analysis of small RNAs bound to nuclear AGO2 by RNA-seq.

To address this issue, we immunopurified the AGO2-bound small RNAs from the nuclear fraction of HeLa S3 cells. As control, a parallel immunoprecipitation using aspecific rat IgG was performed. RNAs extracted from the total nuclear RNA sample (Input) and from the IPs samples were subjected to deep-sequencing. Two independent biological replicas were effectuated.

The analysis and interpretation of the huge amount of data obtained from the deep-sequencing required the setup of a bioinformatic pipeline. The first step of our bioinformatic analysis was annotation of the reads. The reads were first aligned against miRBase and against a manually cured non-coding RNA database, in order to identify those reads which map to RNAs which are already known to be bound by AGO2 (miRNAs, few tRNA) as well as common contaminants (e.g. rRNAs). The sequences which could not be mapped to any of the above mentioned RNA classes were aligned to human genome. The percentages of different classes of non-coding sRNAs observed in the Input, in the AGO2-IP and IgG-IP samples are depicted in Fig.9.

..

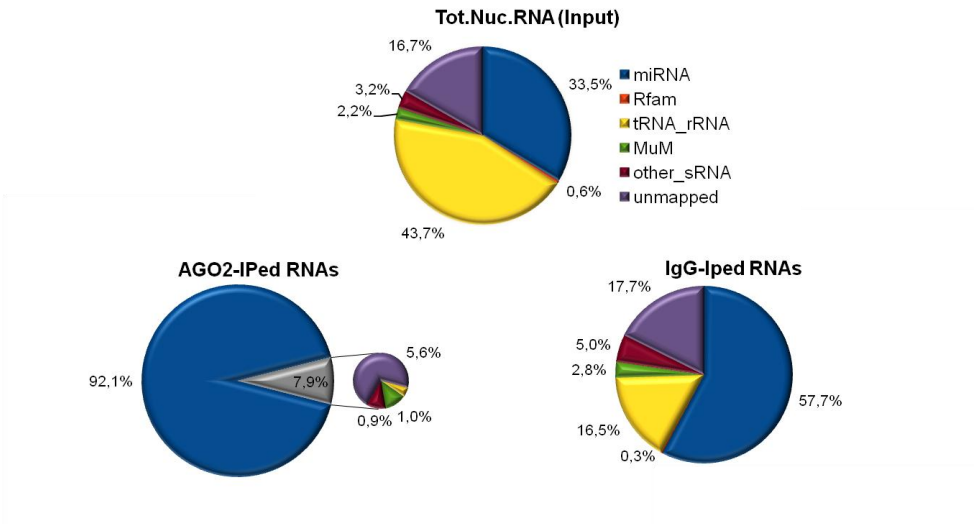

Figure 9 RNA-seq data annotation. The Pie charts illustrate the RNA classes present in the Input, AGO2-IP and IgG-IP samples. miRNA (microRNAs); tRNA (transfer RNAs); rRNA (ribosomal RNAs); Rfam (non-coding sRNAs present in Rfam database); MuM (multiple matching sequences); other sRNA (uncharacterized non-coding sRNAs); unmapped (reads no mapping on the human genome).

Only those reads mapping to a single locus were further analyzed, while Multiple Matching reads (MuM) were not investigated. However, no significant enrichment in MuM reads was observed in AGO2 sample compared to Input and IgG, ruling out the possibility that nuclear AGO2 preferentially binds repetitive sequences.

In the Input sample, the most abundant classes of RNAs were those of tRNA and rRNA followed by miRNA class. The less abundant were multiple matching and Rfam sequences. The IgG annotation profile was similar to the Input indicating that the immunoprecipitation procedure per se had no bias for a particular class of RNAs. On the contrary, the main class of RNA in the AGO2 sample was represented by that of miRNA. This specific miRNAs enrichment underlines the efficiency of our AGO2 immunoprecipitation, as miRNAs

constitute the principal class of non-coding RNA associated to AGO proteins. Moreover, from the reads annotation, in all samples emerged a class of low abundance, uniquely mapping sequences not included in any non-coding RNA databases. These sequences are reported in the pie chart as "other" RNAs.

Since endo-siRNAs have been discovered also in mammalian cells (Yang N, 2006) and mechanistic details about their function is lacking, we aimed at the discovery of new uncharactherized sRNA associated to AGO2. For this reason, we proceeded to investigate the "other" RNAs.

To evaluate whether these "other" RNAs may have a functional role or are purely background noise, a size analysis was performed (Fig.10). The size range considered was from 18 to 50 nt, where 18 is the minimum length required for alignment and 50 is the sequencing length.

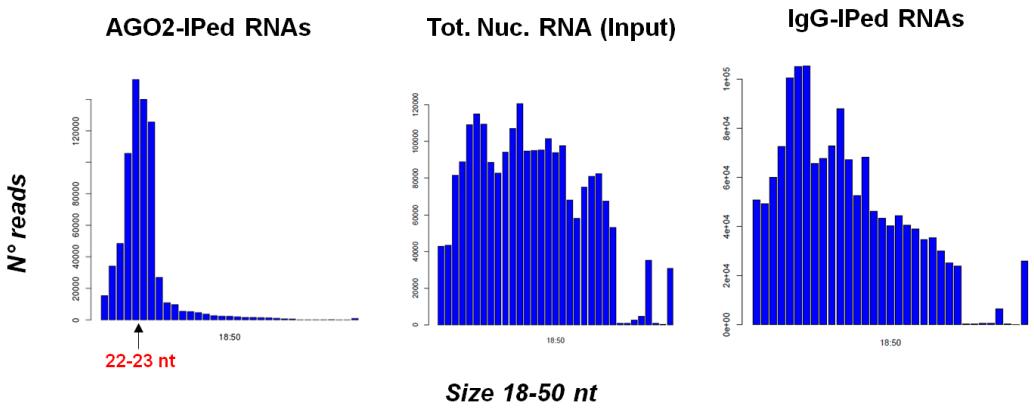

Figure 10 Size distribution of the uncharacterized RNAs (other RNA). Number of reads of "other" RNAs present in AGO2-IPed RNAs, Input and IgG-IPed RNAs were plotted in base to the size range.

The "other" RNAs of the Input and IgG samples showed a homogenous distribution in the considered range. On the contrary, the AGO2-IPed "other" RNAs displayed a size peak of 22-23 nt. Interestingly, this observed size represents the canonical length of small RNAs bound

to AGO proteins. This result led us to exclude the hypothesis of degradation products or aspecific materials, and corroborates the idea of a novel class of RNA associated to AGO2.

We wondered whether these "other" RNAs would be putative miRNAs. Therefore, we performed a mirRNA discovery analysis using miRanalyzer, but with negative results. This indicates that "other" RNAs represent an uncharacterized class of non-coding sRNAs associated to AGO2.

In light of this finding, we proceeded to analyze how the newly identified sRNA class was distributed along the genome. For each sample, overlapping reads on the same strand of the human genome were grouped together in a cluster. Based on an arbitrary value of 100 reads/cluster, clusters were divided into HIGH copy (>100 reads) and LOW copy (<100 reads).

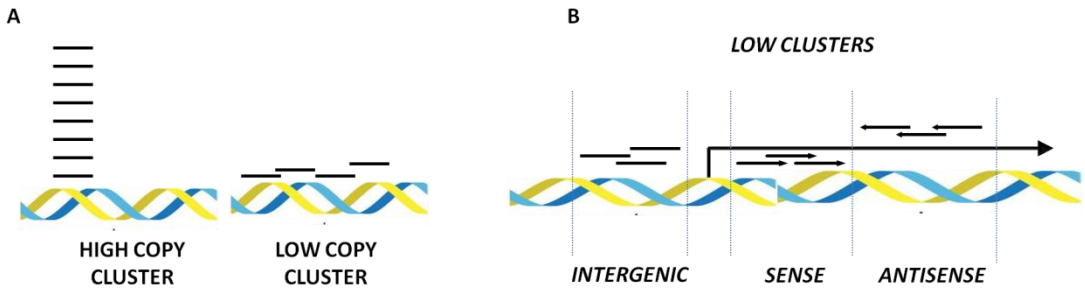

Figure 11 Reads clustering. (A) Clusters were divided into HIGH and LOW copy on the arbitrary value of 100 reads/cluster. **(B)** LOW clusters were subdivided into intergenic, sense and antisense. Sense and antisense distinction was in base on the reads orientation respect to an annotated transcript. Intergenic clusters mapped on region without any annotated transcript.

As evidenced from the schematic representation reported in Fig.11A, the HIGH and LOW clusters showed a different profile. The LOW clusters presented few reads covering a wider region of the genome. The HIGH clusters were characterized by numerous stacked reads with a relatively uniform 5' terminus, similar to a miRNAs cluster. We hypothesized that reads that

form HIGH clusters may be putative miRNA sequences not predicted by miRanalyzer algorithms. For this reason, we focused our analysis exclusively on the LOW clusters.

LOW clusters were divided into sense, antisense and intergenic (Fig.11B). Respectively sense and antisense clusters map on the same and on the opposite strand of an annotated transcript (ENSEMBL). Intergenic clusters map in regions without any annotated transcript.

Afterwards, we aimed at obtaining details about the genomic regions on which the new class of AGO2-associated RNAs map. With this intent, the genomic coordinates of clusters (sense, antisense and intergenic) of each sample (Input, AGO2-IP and IgG-IP) were intersected with the genomic coordinates of selected features (Tab.1), and the number of reads overlapping the features was calculated (m value). As we were interested in the transcriptional role of AGO2, besides structural features, were also chosen features associated with an active (euchromatin) and silenced (heterochromatin) transcriptional state of chromatin. Moreover, to test the significance of the observed overlap, the intersection of the genomic coordinates was effectuated also with shuffled features, thus obtaining the number of reads overlapping control shuffled features (n value). A overlap ratio (m/n) <1 indicates an overlap equal or lower than the one expected choosing random sequences, whilst a ratio m/n >1 is suggestive of a likely correlation. Even more importantly, the overlap ratios obtained for AGO2 sample were compared to those obtained for input and IgG samples. A comparable overlap ratio in the AGO2 and Input sample, even in the case of ratios > 1, would suggest that the observed overlap is the result of some experimental bias and not specific for AGO2-bound RNAs.

Collectively, the analysis showed that intergenic and to a lesser extent antisense RNAs associated to AGO2 mapped on euchromatin, regulatory regions in proximity of TSS (transcriptional start site). Some overlaps investigated between the intergenic clusters with euchromatin-associated features are reported in Fig.12. Input, AGO-IP and IgG-IP samples were plotted for each genomic feature in base to the overlap ratio.

Intriguingly, we observed a strong enrichment in AGO2 sample for sRNA corresponding to BAF155 and BRG1 binding sites. In addition, we found a considerable enrichment for binding regions of proteins implicated in transcription initiation like TBP protein (TATA

binding protein). We observed a remarkable enrichment for CpG islands, strong promoter, strong enhancers, for different histone modifications correlated with transcriptional activation near the TSS like H3K9ac.

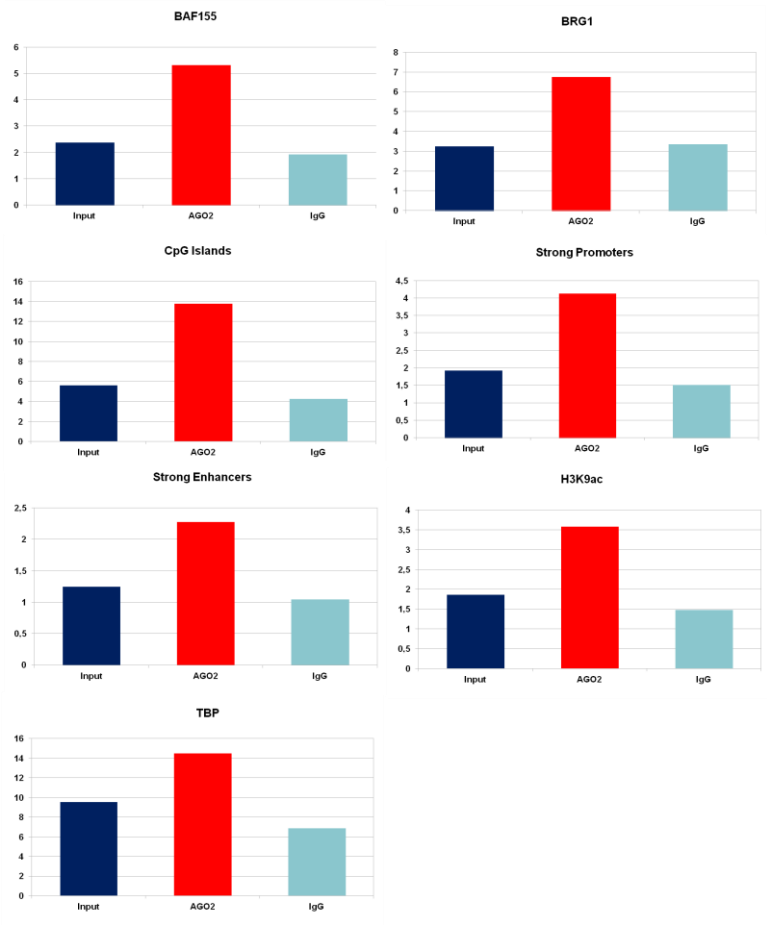

Figure 12 Overlap between Intergenic clusters with euchromatin-associated features. Input, AGO2-IP and IgG-IP samples were plotted for each reported features based on the overlap ratio.

By contrast, no significant correlation (overlap ratio minor than one) was observed for heterochromatin-associated features. In Fig.13 are reported the overlap ratios of intergenic clusters for the histone modification H3K27me3 which is associated with Polycomb-repressed regions (Simon JA, 2009) (Fig.13).

Regarding the analysis of the sense clusters, we observed a significant enrichment only for the structural features of 3'UTR (data not shown).

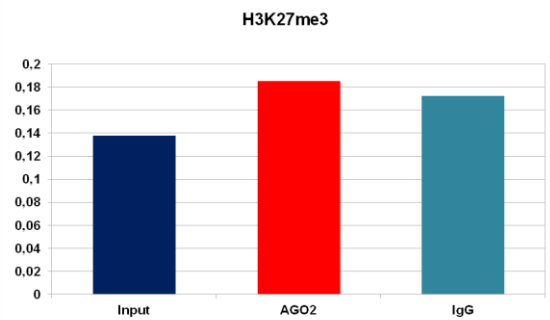

Figure 13 Overlap between intergenic clusters with heterochromatin-associated features. Input, AGO2-IP and IgG-IP samples were plotted for the heterochromatin-associated feature H3K27me3 based on the overlap ratio.

Since the new class of AGO2-associated small RNAs preferentially localizes on promoter regions, we were interested in examining their distribution at TSS level. Consistent with the AGO2-SWI/SNF interaction, we focused this analysis on promoters regulated by BAF155 using ChIP-seq data of Euskirchen GM. *et al*. (Euskirchen GM, 2011). For this analysis, all LOW clusters reads of AGO2-IP sample mapping on BAF155-bound promoters were considered. The reads were subdivided into sense and antisense relative to the direction of gene transcription. As can be seen in Fig.14, AGO2-associated RNAs surround the TSSs with a divergent orientations. The sense AGO2-bound RNAs (blue profile) mapped downstream of the TSSs, overlapping genic transcripts and peaking in abundance between +0 and +50 nucleotides downstream of the TSSs. The antisense AGO2-bound RNAs mapped upstream of the TSSs peaking between nucleotides -100 and -300 nucleotides.

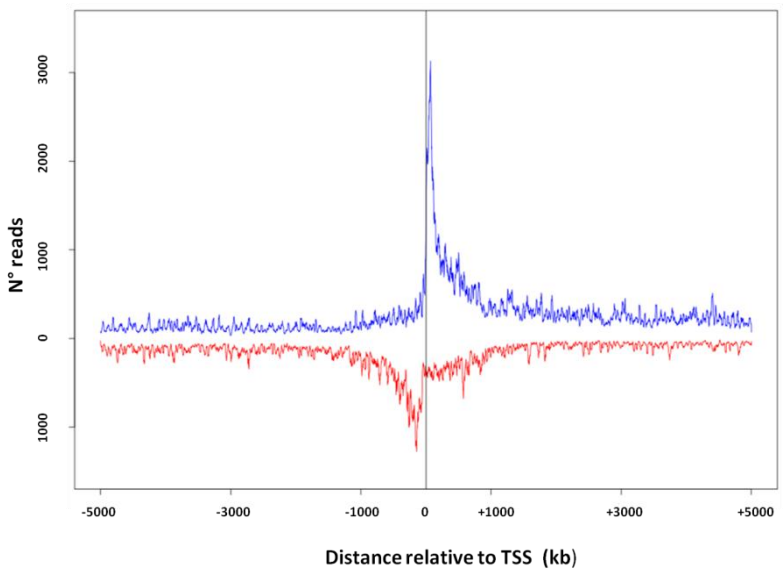

Figure 13 Alignment reads to BAF155 promoters. Low clusters reads of AGO2-IP sample were mapped on promoters presenting BAF155 binding sites and divided into sense (blue profile) and antisense (red profile) relative to direction of gene transcription.

Taken together, we identified a new class of non-coding RNAs associated to AGO2 in the nucleus which mapped to regulatory euchromatin regions implicated in the regulation of transcriptional initiation. Intriguingly, consistent with the identification of AGO2-SWI/SNF interaction, we observed a considerable enrichment for sRNAs corresponding to binding sites for two core components of SWI/SNF complexes. Moreover, the AGO2-associated sRNAs presented a bidirectional distribution surrounding the TSSs of promoters regulated by BAF155.

Discussion

A significant number of studies have revealed the key role of AGO2 in gene expression regulation. In fact, AGO2 has recently been demonstrated to regulate gene expression at the transcriptional level as well as post-transcriptionally in association with small RNAs which mediate the recognition of target molecules (Gagnon KT, 2012).

To shed light on the multiplicity of roles exerted by AGO2, we have performed an extensive analysis of both proteins and small RNAs associated to AGO2 protein by high-throughput technologies. In order to obtain a more physiological scenario on biological processes regulated by AGO2, compared to previous studies carried out in overexpression condition, we have immunopurified AGO2 using a monoclonal antibody against the endogenous AGO2 protein in a human cell line.

Collectively, our broad analysis provides an important contribution to understanding the molecular functions exerted by AGO2 both in new and known biological processes, for example, the modulation of RISC assembly and activity with endosomal trafficking (Lee YS, 2009). In this regard, we have identified the heterotrimer VPS35-VPS26-VPS29 (data reported in table 2) implicated in cargo recognition of the retromer complex. This complex plays a central role in the retrieval of different cargo proteins from the endosome to the trans-Golgi network. The exploration of this interaction will provide insights into the dynamics of RISC assembly and introduces the possibility of crosstalk or direct regulation between the RISC complex with pathways regulated by the retromer complex, including those controlling growth factors signalling, intracellular infection and intercellular transport.

Moreover, our results suggest that AGO2 is implicated also in different processes in the nucleus, in addition to those previously described, such as mRNA processing and chromatin remodelling.

In particular, we have provided several evidence regarding the mechanism through which AGO2 regulates the gene expression at the transcription level. To date, the mechanisms of AGO2-mediated transcriptional regulation has not been clearly elucidated.

Notably, we have demonstrated that nuclear AGO2 interacts via protein-protein association with SWI/SNF complexes which are the main family of chromatin remodeler. The SWI/SNF complexes play a fundamental role in transcription regulation mainly by altering the structure and the positioning of nucleosomes, thereby modulating the access of regulatory proteins to DNA (Saha A, 2006). Moreover, we have identified a novel class of AGO2-associated small RNAs in the nucleus that map to euchromatin regions implicated in the regulation of transcriptional initiation (e.g. CpG islands, promoters, enhancers) and that shows a remarkable overlap with the binding sites of SWI/SNF complexes. It is reasonable to suppose that these small RNAs would be operative in directing the recruitment of AGO2 in association with SWI/SNF complexes to chromatin.

Several models have been proposed regarding the modality of AGO2-binding on specific loci. One of the most intriguing postulates an RNA-RNA interaction between the AGO2-bound sRNA and a non-coding or coding RNAs spanning through the promoter region (Han J, 2007). The characterization of AGO2-bound long RNAs in the nucleus by RIP (RNA-immunoprecipitation) and the identification of the binding sites on RNA molecules by PAR-CLIP (Photoactivatable-Ribonucleoside-Enhanced Crosslinking and Immunoprecipitation) experiments, will enable to verify the model of AGO2 chromatin recruitment via RNA-RNA interaction.

Taken together, our data suggests that nuclear AGO2 regulates the transcriptional initiation interacting with SWI/SNF complexes to ensure proper nucleosome positioning in proximity of promoter regions. We propose to test this hypothesis in further experiments, for example by genome-wide nucleosome positioning in control cells and in cells in which AGO2 has been knocked down. Moreover, the analysis of the genomic loci bound by AGO2, through ChIP-AGO2 experiments, along with the identification of RNAs associated to SWI-SNF complexes might provide further details.

The SWI/SNF complexes show pleiotropic functions, which are conferred by the combinatorial assembly of several accessory subunits. Since we have observed that AGO2 is associated to several SWI/SNF complexes, we hypothesize that AGO2 would cooperate with SWI/SNF also in other biological processes. The recent report in *Drosophila* about the

involvement of AGO2 in DNA damage repair (Wei W, 2012) as well as the established role of SWI/SNF in the same pathway (Osley MA, 2007) are suggestive of a their concerted activity in this context, too.

Furthermore, our results raise many future issues. The first concerns the biogenesis mechanism of this new class of sRNAs associated to AGO2. Many would be the mechanisms underlying their generation, among which we hypothesize the divergent transcription. It has been reported that the most active genes, in organism as diverse as yeast and mammals, are bidirectionally transcribed and that this might have a regulatory function in transcription (Seila AC, 2009). Recent studies in mammals have shown that polymerase occupancy is nearly equivalent in both directions (Core LJ, 2008) and that the polymerase upstream and antisense to the directions of genes does not productively elongate after the site of pausing (Seila AC, 2008). Moreover, high-throughput analysis of the small RNA population from mouse ES cells (Seila AC, 2008) and from human fibroblasts (Core LJ, 2008) have identified a high density of RNA mapping in proximity of the TSS. These promoter-proximal small RNAs overlap the regions bound by the sense and antisense paused RNA polymerases suggesting that at least some might be by-products of so called "paused" RNAPII. The bidirectional distribution of our AGO2-associated small RNAs around the TSS of promoters bound by BAF155 would be consistent with the scenario above mentioned.

Another question, regarding the generation of the various promoter-proximal sRNA, is whether a DICER processing would be involved similar to endo-siRNAs identified in mouse oocytes (Watanabe T, 2008). The analysis of our new class of nuclear AGO2-associated RNAs in DICER mutant cells line displaying a hypomorphic phenotype (Cummins JM, 2006) will elucidate the DICER involvement in their processing.

References

Affymetrix ENCODE Transcriptome project. "Post-transcriptional processing generates a diversity of 5'-modified long and short RNAs." *Nature*, 2009: 457(7232):1028-32.

Bartel DP. "MicroRNAs: target recognition and regulatory functions." *Cell*, 2009: 136, 215–233.

Batsché E, Yaniv M, Muchardt C. "The human SWI/SNF subunit Brm is a regulator of alternative splicing." *Nat Struct Mol Biol*, 2006: 13(1):22-9.

Bazett-Jones DP, Côté J, Landel CC, Peterson CL, Workman JL. "The SWI/SNF complex creates loop domains in DNA and polynucleosome arrays and can disrupt DNA-histone contacts within these domains." *Mol Cell Bio*, 1999: 19(2):1470-8.

Berezikov E, Chung WJ, Willis J, Cuppen E, Lai EC. "Mammalian mirtron genes." *Mol Cel*, 2007: 28(2):328-36.

Bhattacharyya SN, Habermacher R, Martine U, Closs EI, Filipowicz W. "Stress-induced reversal of microRNA repression and mRNA P-body localization in human cells." *Cold Spring Harb Symp Quant Biol*, 2006: 71:513-21.

Castanotto D, Tommasi S, Li M, Li H, Yanow S, Pfeifer GP, Rossi JJ. "Short hairpin RNA-directed promoter in HeLa cells." *Mol Ther*, 2005: 12:179–183.

Cernilogar FM, Onorati MC, Kothe GO, Burroughs AM, Parsi KM, Breiling A, Lo Sardo F, Saxena A, Miyoshi K, Siomi H, Siomi MC, Carninci P, Gilmour DS, Corona DF, Orlando V. "Chromatin-associated RNA interference components contribute to transcriptional regulation in Drosophila." *Nature*, 2011: 480(7377):391-5.

Cheng SW, Davies KP, Yung E, Beltran RJ, Yu J, Kalpana GV. "c-MYC interacts with INI1/hSNF5 and requires the SWI/SNF complex for transactivation function." *Nat Genet*, 1999: 22(1):102-5.

Cho H, Orphanides G, Sun X, Yang XJ, Ogryzko V, Lees E, Nakatani Y, Reinberg D. "A human RNA polymerase II complex containing factors that modify chromatin structure." *Mol Cell Biol*, 1998: 18(9):5355-63.

Clapier CR, Cairns BR. "The biology of chromatin remodeling complexes." *Annu Rev Biochem*, 2009: 78:273-304.

Cohen SM, Chastain PD 2nd, Rosson GB, Groh BS, Weissman BE, Kaufman DG, Bultman SJ. "BRG1 co-localizes with DNA replication factors and is required for efficient replication fork progression." *Nucleic Acids Res*, 2010: 38(20):6906-19.

Core LJ, Waterfall JJ, Lis JT. "Nascent RNA sequencing reveals widespread pausing and divergent initiation at human promoters." *Science*, 2008: 322(5909):1845-8.

Corey LL, Weirich CS, Benjamin IJ, Kingston RE. "Localized recruitment of a chromatin-remodeling activity by an activator in vivo drives transcriptional elongation." *Genes Dev*, 2003: 17(11):1392-401.

Cummins JM, He Y, Leary RJ, Pagliarini R, Diaz LA Jr, Sjoblom T, Barad O, Bentwich Z, Szafranska AE, Labourier E, Raymond CK, Roberts BS, Juhl H, Kinzler KW, Vogelstein B, Velculescu VE. "The colorectal microRNAome." *Proc Natl Acad Sci U S A*, 2006: 103(10):3687-92.

Eulalio A, Behm-Ansmant I, Izaurralde E. "P bodies: at the crossroads of post-transcriptional pathways. ." *Nat. Rev Mol Cell Biol*, 2007a: 8: 9–22.

Eulalio A, Behm-Ansmant I, Schweizer D, Izaurralde E. "P-body formation is a consequence, not the cause, of RNA-mediated gene silencing." *Mol Cell Biol*, 2007b: 27(11):3970-81.

Eulalio A, Rehwinkel J, Stricker M, Huntzinger E, Yang SF, Doerks T, Dorner S, Bork P, Boutros M, Izaurralde E. "Target-specific requirements for enhancers of decapping in miRNA-mediated gene silencing." *Genes Dev*, 2007c: 21(20):2558-70.

Euskirchen GM, Auerbach RK, Davidov E, Gianoulis TA, Zhong G, Rozowsky J, Bhardwaj N, Gerstein MB, Snyder M. "Diverse roles and interactions of the SWI/SNF chromatin remodeling complex revealed using global approaches." *PLoS Genet*, 2011: 7(3):e1002008.

Farazi TA, Juranek SA, Tuschl T. "The growing catalog of small RNAs and their association with distinct Argonaute/Piwi family members." *Development*, 2008: 135(7):1201-14.

Gagnon KT, Corey DR. "Argonaute and the nuclear RNAs: new pathways for RNA-mediated control of gene expression." *Nucleic Acid Ther*, 2012: 22(1):3-16.

Ghildiyal M, Zamore PD. "Small silencing RNAs: an expanding universe." *Nat Rev Genet*, 2009: 10(2):94-108.

Green VA, Weinberg MS. "Small RNAinduced transcriptional gene regulation in mammals mechanisms,therapeutic applications, and scope within the genome." *Prog Mol Biol Transl Sci,* 2011: 102: 11–46.

Han J, Kim D, Morris KV. "Promoter-associated RNA is required for RNA-directed transcriptional gene silencing in human cells." *Proc Natl Acad Sci U S A*, 2007: 104(30):12422-7.

Hawkins PG, Santoso S, Adams C, Anest V, Morris KV. "Promoter targeted small RNAs induce long-term transcriptional gene silencing in human cells." *Nucleic Acids Res*, 2009: 37(9):2984-95.

Höck J, Meister G. "The Argonaute protein family." *Genome Biol*, 2008: 9(2): 210.

Janowski BA, Younger ST, Hardy DB, Ram R, Huffman KE, Corey DR. "Activating gene expression in mammalian cells with promoter-targeted duplex RNAs." *Nat Chem Biol*, 2007: 3:166–173.

Kim DH, Sætrom P, Snøve O, Rossi JJ. "MicroRNA-directed transcriptional gene silencing in mammalian cells." *Proc Natl Acad Sci U S A, 2008*: 105(42): 16230–16235.

Kim DH, Villeneuve LM, Morris KV, Rossi JJ. "Argonaute-1 directs siRNA-mediated transcriptional gene silencing in human cells." *Nat Struct Mol Biol*, 2006: 13(9):793-7.

Kim JW, Zhang YH, Zern MA, ROSSI JJ,Wu J. "Short hairpin RNA causes the methylation of transforming growth factor-beta receptor II promoter and silencing of the target gene in rat hepatic stellate cells." *Biochem Biophys Res Commun, 2007*: 359: 292–297.

Kiriakidou M, Tan GS, Lamprinaki S, De Planell-Saguer M, Nelson PT, Mourelatos Z. "An mRNA m7G cap binding-like motif within human Ago2 represses translation." *Cell*, 2007: 129(6):1141-51.

Kong YW, Cannell IG, de Moor CH, Hill K, Garside PG, Hamilton TL, Meijer HA, Dobbyn HC, Stoneley M, Spriggs KA, Willis AE, Bushell M. "The mechanism of micro-RNA-mediated translation repression is determined by the promoter of the target gene." *Proc Natl Acad Sci U S A*, 2008: 105(26):8866-71.

Langmead B, Trapnell C, Pop M, Salzberg SL. "Ultrafast and memory-efficient alignment of short DNA sequences to the human genome." *Genome Biol*, 2009: 10(3):R25.

Lee D, Kim JW, Seo T, Hwang SG, Choi EJ, Choe J. "SWI/SNF complex interacts with tumor suppressor p53 and is necessary for the activation of p53-mediated transcription." *J Biol Chem*, 2002: 277(25):22330-7.

Lee YS, Pressman S, Andress AP, Kim K, White JL, Cassidy JJ, Li X, Lubell K, Lim do H, Cho IS, Nakahara K, Preall JB, Bellare P, Sontheimer EJ, Carthew RW. "Silencing by small RNAs is linked to endosomal trafficking." *Nat Cell Biol*, 2009: 11(9):1150-6.

Li LC, Okino ST, Zhao H, Pookot D, Place RF, Urakami S, Enokida H, Dahiya R. " Small dsRNAs induce transcriptional activation in human cells." *Proc Natl Acad Sci USA, 2006*: 103: 17337-17342.

Liu J, Carmell MA, Rivas FV, Marsden CG, Thomson JM, Song JJ, Hammond SM, Joshua-Tor L, Hannon GJ. "Argonaute2 is the catalytic engine of mammalian RNAi." *Science*, 2004: 305(5689):1437-41.

Ma JB, Yuan YR, Meister G, Pei Y, Tuschl T, Patel DJ. "Structural basis for 5'-endspecific recognition of guide RNA by the A. fulgidus Piwi protein." *Nature*, 2005: 434:666-670.

Meister G. "miRNAs get an early start on translational silencing." *Cell*, 2007: 131:25-28.

Meister G, Landthaler M, Patkaniowska A, Dorsett Y, Teng G,Tuschl T. "Human Argonaute2 mediates RNA cleavage targeted by miRNAs and siRNAs." *Mol Cell*, 2004: 15:185-197.

Morris KV, Chan SW, Jacobsen SE, Looney DJ. "Small interfering RNA-induced transcriptional gene silencing in human cells." *Science*, 2004: 305(5688):1289-92.

Nie Z, Xue Y, Yang D, Zhou S, Deroo BJ, Archer TK, Wang W. "A Specificity and Targeting Subunit of a Human SWI/SNF Family-Related Chromatin-Remodeling Complex." *Mol Cell Biol*, 2000: 20(23):8879-88.

Nishi K, Nishi A, Nagasawa T, Ui-Tei K. "Human TNRC6A is an Argonaute-navigator protein for microRNA-mediated gene silencing in the nucleus." *RNA*, 2013: 19(1):17-35.

Nottrott S, Simard MJ, Richter JD. "Human let-7a miRNA blocks protein production on actively translating polyribosomes." *Nat Struct Mol Biol*, 2006: 13:1108-1114.

Osley MA, Tsukuda T, Nickoloff JA. "ATP-dependent chromatin remodeling factors and DNA damage repair." *Mutat Res*, 2007: 618(1-2):65-80.

Parker JS, Roe SM, Barford D. " Crystal structure of a PIWI protein suggests mechanisms for siRNA recognition and slicer activity." *EMBO J*, 2004a: 23:4727-4737.

Parker R, Song H. "The enzymes and control of eukaryotic mRNA turnover." *Nat Struct Mol Biol*, 2004b: 11(2):121-7.

Petersen CP, Bordeleau ME, Pelletier J, Sharp PA. "Short RNAs repress translation after initiation in mammalian cells ." *Mol Cell*, 2006: 21:533-542.

Place RF, Li LC, Pookot D, Noonan EJ, Dahiya R. "microRNA-373 induces expression of genes with complementary promoter sequences." *Proc Natl Acad Sci USA*, 2008: 105:1608-1613.

Rehwinkel J, Behm-Ansmant I, Gatfield D, Izaurralde E. "A crucial role for GW182 and the DCP1:DCP2 decapping complex in miRNA-mediated gene silencing." *RNA*, 2005: 11(11):1640-7.

Robb GB, Brown KM, Khurana J, Rana TM. "Specific and potent RNAi in the nucleus of human cells ." *Nat Struct Mol Biol*, 2005: 12(2):133-7.

Rüdel S, Flatley A, Weinmann L, Kremmer E, Meister G. "A multifunctional human Argonaute2-specific monoclonal antibody." *RNA*, 2008: 14(6):1244-53.

Saha A, Wittmeyer J, Cairns BR. "Chromatin remodelling: the industrial revolution of DNA around histones." *Nat Rev Mol Cell Biol*, 2006: 7(6):437-47.

Salma N, Xiao H, Mueller E, Imbalzano AN. "Temporal recruitment of transcription factors and SWI/SNF chromatin-remodeling enzymes during adipogenic induction of the peroxisome proliferator-activated receptor gamma nuclear hormone receptor." *Mol Cell Biol*, 2004: 24(11):4651-63.

Schirle NT, MacRae IJ. "The Crystal Structure of Human Argonaute2." *Science*, 2012: 336, 1037.

Schwartz JC, Younger ST, Nguyen NB, Hardy DB, Monia BP, Corey DR, Janowski BA. "Antisense transcripts are targets for activating small RNAs." *Nat Struct Mol Biol*, 2008: 15(8):842-8.

Seila AC, Calabrese JM, Levine SS, Yeo GW, Rahl PB, Flynn RA, Young RA, Sharp PA. "Divergent transcription from active promoters." *Science*, 2008: 322(5909):1849-51.

Seila AC, Core LJ, Lis JT, Sharp PA. "Divergent transcription: a new feature of active promoters." *Cell Cycle*, 2009: 8(16):2557-64.

Selbach M, Schwanhausser B, Thierfelder N, Fang Z, Khanin R,Rajewsky N. "Widespread changes in protein synthesis." *Nature*, 2008: 455: 58–63.

Shanahan F, Seghezzi W, Parry D, Mahony D, Lees E. "Cyclin E associates with BAF155 and BRG1, components of the mammalian SWI-SNF complex, and alters the ability of BRG1 to induce growth arrest." *Mol Cell Biol*, 1999: 19(2):1460-9.

Simon JA, Kingston RE. "Mechanisms of polycomb gene silencing: knowns and unknowns." *Nat Rev Mol Cell Biol*, 2009: 10(10):697-708.

Song JJ, Smith SK, Hannon GJ, Joshua Tor L. "Crystal structure of Argonaute and its implications for RISC slicer activity." *Science*, 2004: 305(5689):1434-7.

Soutoglou E, Talianidis I. "Coordination of PIC assembly and chromatin remodeling during differentiation-induced gene activation." *Science*, 2002: 295(5561):1901-4.

Talbert PB, Henikoff S. "Histone variants--ancient wrap artists of the epigenome." *Nat Rev Mol Cell Biol*, 2010: 11(4):264-75.

Thompson M. "Polybromo-1: The chromatin targeting subunit of the PBAF complex." *Biochimie*, 2009: 309e319.

Till S, Lejeune E, Thermann R, Bortfeld M, Hothorn M, Enderle D, Heinrich C, Hentze MW, Ladurner AG. "A conserved motif in Argonaute-interacting proteins mediates functional interactions through the Argonaute PIWI domain." *Nat Struct Mol Biol*, 2007: 14(10):897-903.

Ting AH, Schuebel KE, Herman JG, Baylin SB. "Short double-stranded RNA induces transcriptional gene silencing in human cancer cells in the absence of DNA methylation." *Nat. Genet, 2005*: 37: 906–910.

Trotter KW, Archer TK. "The BRG1 transcriptional coregulator." *Nucl Recept Signal*, 2008: 6:e004.

Trouche D, Le Chalony C, Muchardt C, Yaniv M, Kouzarides T. "RB and hbrm cooperate to repress the activation functions of E2F1." *Proc Natl Acad Sci U S A*, 1997: 94(21):11268-73.

Urnov FD, Wolffe AP. "Chromatin remodeling and transcriptional activation: the cast (in order of appearance)." *Oncogene*, 2001: 20(24):2991-3006.

Vasudevan S, Tong Y, Steitz JA. "Switching from repression to activation: microRNAs can up-regulate translation." *Science*, 2007: 318(5858):1931-4.

Watanabe T, Totoki Y, Toyoda A, Kaneda M, Kuramochi-Miyagawa S, Obata Y, Chiba H, Kohara Y, Kono T, Nakano T, Surani MA, Sakaki Y, Sasaki H. "Endogenous siRNAs from naturally formed dsRNAs regulate transcripts in mouse oocytes." *Nature*, 2008: 453(7194):539-43.

Wei W, Ba Z, Gao M, Wu Y, Ma Y, Amiard S, White CI, Rendtlew Danielsen JM, Yang YG, Qi Y. "A role for small RNAs in DNA double-strand break repair." *Cell*, 2012: 149(1):101-12.

Weinberg MS, Villeneuve LM, Ehsani A, Amarzguioui M, Aagaard L, Chen ZX, Riggs AD, Rossi JJ, Morris KV. "The antisense strand of small interfering RNAs directs histone methylation and transcriptional gene silencing in human cells." *RNA*, 2006: 12(2):256-62.

Weinmann L, Höck J, Ivacevic T, Ohrt T, Mütze J, Schwille P, Kremmer E, Benes V, Urlaub H, Meister G. "Importin 8 is a gene silencing factor that targets argonaute proteins to distinct mRNAs." *Cell*, 2009 : 136(3):496-507.

Weissman B, Knudsen KE. "Hijacking the Chromatin Remodeling Machinery: Impact of SWI/SNF Perturbations in Cancer." *Cancer Res*, 2009: 69(21):8223-30.

Wilson BG, Roberts CW. "SWI/SNF nucleosome remodellers and cancer." *Nat Rev Cancer*, 2011: 11(7):481-92.

Wu JI, Lessard J, Crabtree GR. " Understanding the words of chromatin regulation." *Cell*, 2009: 136: 200–206.

Yan KS, Yan S, Farooq A, Han A, Zeng L, Zhou MM. "Structure and conserved RNA binding of the PAZ domain." *Nature*, 2003: 426:468-474.

Yang N, Kazazian HH Jr. "L1 retrotransposition is suppressed by endogenously encoded small interfering RNAs in human cultured cells." *Nat Struct Mol Biol*, 2006: 13(9):763-7.

Yue X, Schwartz JC, Chu Y, Younger ST, Gagnon KT, Elbashir S, Janowski BA, Corey DR. "Transcriptional regulation by small RNAs at sequences downstream from 3' gene termini." *Nat Chem Biol*, 2010: 6(8):621-9.

Zamore PD, Tuschl T, Sharp PA, Bartel DP. "RNAi: double-stranded RNA directs the ATP-dependent cleavage of mRNA at 21 to 23 nucleotide intervals." *Cell*, 2000: 101(1):25-33.

Zeng Y, Sankala H, Zhang X, Graves PR. "Phosphorylation of Argonaute 2 at serine-387 facilitates its localization to processing bodies." *Biochem J*, 2008: 413(3):429-36.

Zhang HS, Gavin M, Dahiya A, Postigo AA, Ma D, Luo RX, Harbour JW, Dean DC. "Exit from G1 and S phase of the cell cycle is regulated by repressor complexes containing HDAC-Rb-hSWI/SNF and Rb-hSWI/SNF." *Cell*, 2000: 101(1):79-89.