

PhD degree in Systems Medicine (curriculum in Molecular Oncology)

European School of Molecular Medicine (SEMM),

University of Milan and University of Naples “Federico II”

Settore disciplinare: BIO/10

**MOLECULAR BASIS OF MYOSIN VI
ALTERNATIVE SPLICING IN CELLULAR
TRANSFORMATION**

Rossella Scotto di Perrotolo

R11748

IFOM, Milan

Tutor: Prof. Simona Polo

IFOM, Milan

University of Milan

PhD Coordinator: Prof. Giuseppe Viale

Anno accademico 2019-2020

To my grandparents Mario and Rosa

Table of Contents

List of abbreviations	4
Figures index	4
1. Abstract.....	11
2. Introduction.....	13
2.1. Myosin VI.....	13
2.1.1. Myosin VI: structure and alternative splicing	13
2.1.2. Myosin VI motor anchor activities	15
2.1.3. Myosin VI functions	16
2.1.4. Myosin VI in ovarian cancer	18
2.2. The pre-mRNA splicing process.....	22
2.2.1. The spliceosome catalyzes a two-step reaction.....	22
2.2.2. The splicing cycle	24
2.2.3. Alternative splicing: a way to expand the proteome.....	27
2.2.4. Alternative splicing in cancer	35
2.2.4.1. Splicing regulators as oncoproteins or oncosuppressors	35
2.2.4.2. Alternative splicing in cell proliferation and survival.....	37
2.2.4.3. Alternative splicing in cancer metabolism	38
2.2.4.4. Alternative splicing in angiogenesis	38
2.2.4.5. Alternative splicing in EMT.....	39
2.2.4.6. Alternative splicing and cell invasion	39
2.2.4.7. Alternative splicing in cancer drug resistance.....	40
2.2.4.8. Alternative splicing and signaling.....	40
2.2.4.9. Targeting splicing defects in cancer.....	41
3. Aim of the thesis.....	42
4. Materials and methods	43
4.1. Buffers.....	43
4.1.1. Phosphate-buffered saline (PBS).....	43
4.1.2. Tris-buffered saline - Tween (TBS-T).....	43
4.1.3. 10X Sodium Dodecyl Sulphate – PolyAcrylamide Gel Electrophoresis (SDS-PAGE) running buffer	43
4.1.4. 50X Tris-Acetate-EDTA (TAE).....	43
4.1.5. 1X Radioimmunoprecipitation assay (RIPA) buffer.....	43
4.1.6. Laemmli buffer	44

4.1.7. Immunofluorescence (IF) buffer	44
4.1.8. RNA immunoprecipitation buffers	44
4.1.8.1. Low salt lysis buffer.....	44
4.1.8.2. Nuclei lysis buffer	44
4.1.8.3. Low-salt wash buffer	45
4.1.8.4. High-salt wash buffer	45
4.1.8.5. LiCl wash buffer	45
4.1.8.6. Elution buffer	46
4.2. Reagents.....	46
4.2.1. siRNA.....	46
4.2.2. Primary antibodies.....	47
4.2.3. Secondary antibodies.....	48
4.3. Cloning techniques.....	49
4.3.1. Agarose gel electrophoresis	49
4.3.2. Transformation of competent cells	49
4.3.3. Minipreps.....	49
4.3.4. Large scale plasmid preparation	49
4.3.5. Constructs and plasmids	49
4.4. Cell lines.....	50
4.4.1. Commercial cell lines	50
4.4.2. Generated cell lines	51
4.5. Cell assays.....	51
4.5.1. Myosin VI isoform switch assay	51
4.5.2. Small interfering RNA (siRNA) transfection	51
4.5.3. Wound healing assay	52
4.5.4. DNA transduction.....	52
4.5.5. Caco-2 cysts formation	52
4.5.6. Immunofluorescence.....	53
4.6. Isoform detection by RT-PCR.....	54
4.7. qPCR	55
4.8. SDS-Polyacrylamide Gel-Electrophoresis (SDS-PAGE) and Immunoblot Analysis (IB)	55
4.9. UV-RIP: RNA immunoprecipitation after UV crosslinking.....	56
4.9.1. Lysate preparation	56
4.9.2. Immunoprecipitation.....	56
4.9.3. RNA isolation.....	57
4.9.4. RT-qPCR	57
4.9.4.1. Percent Input calculation.....	58
4.9.4.2. Fold enrichment calculation.....	58

4.10. <i>Affymetrix Human Transcriptome Array</i>	58
4.11. <i>Immunohistochemistry (IHC)</i>	59
5. Results	60
5.1. <i>Myosin VI isoform expression correlates with epithelial phenotype</i>	60
5.2. <i>Identification of signaling pathways involved in myosin VI alternative splicing</i>	64
5.2.1. PAR3 and ZO-1 do not affect myosin VI isoform switch.....	64
5.2.2. E-cadherin-mediated cell-cell contacts are determinant for myosin VI isoform switch	66
5.2.3. β -catenin influences myosin VI isoform expression.....	69
5.3. <i>Identification of splicing regulators involved in myosin VI alternative splicing</i>	72
5.3.1. Diverse splicing regulators are modulated when cells grow as monolayers	72
5.3.1.1. PTBP1 affects myosin VI alternative splicing in a cell type-specific manner	74
5.3.1.2. FUS is not involved in myosin VI alternative splicing	75
5.3.1.3. SRSF2 promotes myosin VI _{short} expression	76
5.3.1.4. RBFOX2 promotes myosin VI _{long} expression	77
5.3.1.5. KHSRP promotes myosin VI _{long} expression.....	81
5.3.2. Characterization of RNA protein interactions	84
5.3.2.1. Splicing regulator binding sites in myosin VI AS region.....	84
5.3.2.2. RBFOX2 interacts with myosin VI pre-mRNA	87
5.4. <i>Myosin VI alternative splicing regulation in ovarian cancer</i>	90
5.4.1. Myosin VI short is selectively expressed in ovarian cancer	90
5.4.2. SRSF2 and RBFOX2 regulates myosin VI AS in ovarian cancer	91
5.4.3. SRSF2 depletion restores myosin VI _{long} expression in ovarian cancer cells.....	92
5.4.4. SRSF2 depletion impairs ovarian cancer cell migration.....	95
6. Discussion	97
6.1. <i>E-cadherin and β-catenin are determinants for myosin VI isoform switch</i>	98
6.2. <i>RBFOX2 and KHSRP promote the expression of myosin VI_{long}</i>	99
6.3. <i>SRSF2 promotes the expression of myosin VI_{short}</i>	102
6.4. <i>Future perspectives</i>	104
Acknowledgement	106
7. References	107
Publications	128

List of abbreviations

AJ	Adherens junctions
AKT	RAC-alpha serine/threonine-protein kinase
aPKC	Atypical protein kinase C
ARID1A	AT-rich interactive domain-containing protein 1A
AS	Alternative splicing
ASE	Alternative splicing event
ATP	Adenosine triphosphate
BCL2L1	Bcl-2-like protein 1
BPS	Branch point sequence
BRAF	v-raf murine sarcoma viral oncogene homolog B1
CaM	Calmodulin
CBD	Cargo-binding domain
CCPs	Clathrin-coated pits
CD44	Cluster of Differentiation 44
CELF	CUG-BP and ETR-3-like factors
CFTR	Cystic fibrosis transmembrane conductance regulator
CLCa	Clathrin light chain A
CLKS	Cyclin-dependent like kinases
Dab2	Disabled homolog 2
EGFP	Enhanced Green Fluorescent Protein
ELK1	ETS Like-1 protein
EMT	Epithelial to Mesenchymal transition
ENAH	Protein enabled homolog
ESCs	Embryonic stem cells
ESE	Exonic splicing enhancer
ESFs	Engineered splicing factors
ESRP	Epithelial Splicing Regulatory Protein
ESS	Exonic splicing silencer
EWS	Ewing Sarcoma breakpoint region 1
FAS	Fas Cell Surface Death receptor
FBP2	Far-upstream Element-binding Protein 2
FGFR	Fibroblast growth factor receptors
FTE	Fallopian Tube Epithelium
FUS	Fused in sarcoma

GIPC	GAIP-interacting protein C-terminus
GSK-3β	Glycogen synthase kinase 3
HDAC6	Histone deacetylases 6
HER2	Human epidermal growth factor receptor 2
HGFR	Hepatocyte Growth Factor Receptor
Hip1R	Huntingtin Interacting Protein 1 Related
hnRNPs	Heterogeneous nuclear ribonucleoproteins
IF	Immunofluorescence
IHC	Immunohistochemistry
ISE	Intronic splicing enhancer
ISS	Intronic splicing silencer
KD	Knockdown
KH	K-homology
KHSRP	KH-type splicing regulatory protein
LI	Large insert
LIR	LC3-interacting region
LSm	Sm-like
MAP	Mitogen-activated protein
MBNL	Muscleblind-like
MDM4	Double minute 4 protein
MET	Mesenchymal to epithelial transition
miRNA	microRNA
MIU	Motif interacting with ubiquitin
MYC	Master Regulator of Cell Cycle Entry and Proliferative Metabolism
MyUb	Myosin VI ubiquitin-binding domain
NDP52	Nuclear dot protein 52 kDa
NOTCH1	Notch homolog 1
NOVA	Neuro-oncological ventral antigen
NUMB	Protein numb homolog
OSE	Ovarian surface epithelium
PANDAR	Promoter of CDKN1A Antisense DNA Damage Activated RNA
PAR	Partitioning defective
PBS	Phosphate-buffered saline
Pbx1	Transcription factor pre-B cell leukemia homeobox
PKM	Pyruvate kinase

pre-mRNA	precursor messenger RNA
PRMT5	Protein arginine methyltransferase 5
PSF	Polypyrimidine tract-binding protein-associated splicing factor
PTBP1	Polypyrimidine tract binding protein 1
PTEN	Phosphatase and tensin homolog
qPCR	Polymerase chain reaction
Rac1	Rac Family Small GTPase 1
RBFOX	RNA binding protein fox-1
RBM	RNA Binding Motif Protein
RGG	Arg-Gly-Gly repeats
RIPA	Radioimmunoprecipitation assay
RNP	Ribonucleoprotein
RON	Recepteur d'Origine Nantais
RRM	RNA recognition motif
RS domain	Serine-arginine rich domain
RT-PCR	Reverse transcriptase polymerase chain reaction
Sam68	Src-Associated substrate in Mitosis of 68 kDa
SDS-PAGE	Sodium Dodecyl Sulphate – PolyAcrylamide Gel Electrophoresis
SF1	Splicing factor 1
SF3B	Splicing factor 3B
SF3BP1	Splicing factor 3B subunit 1
SI	Small insert
siRNA	small interfering RNA
snRNAs	small nuclear RNAs
snRNPs	small nuclear RNPs
SNVs	Single nucleotide variants
SR	Serine and arginine-rich
SREs	Splicing regulatory elements
SRPKs	SR protein-specific kinases
SRSF	Serine/arginine-rich splicing factor
SS	Splice site
T6BP	TRAF6-binding protein
TAE	Tris-Acetate-EDTA
TBS-T	Tris-buffered saline - Tween
TCF/LEF	T-cell factor/lymphoid enhancer factor

TGF-β	Transforming growth factor β
TIP60	Tat-interactive protein 60
TJ	Tight junction
Tom1	Target of Myb protein 1
TP53	Tumor protein p53
Twist	Twist-related protein 1
U2AF	U2 auxiliary factor
UV-RIP	RNA immunoprecipitation after UV crosslinking
VEGFA	Vascular endothelial growth factor A
WNT	Wingless-related integration site
ZEB	Zinc finger E-box binding homeobox
ZO-1	Zonula occludens protein 1
ZRSR2	U2 small nuclear ribonucleoprotein auxiliary factor 35 kD subunit-related protein 2

Figures index

Figure 1. Myosin VI isoforms have different interactomes.....	15
Figure 2. Ovarian cancer cells are addicted to myosin VI _{short} for cell migration.....	20
Figure 3. Pre-mRNA consensus sequences and spliceosome composition in humans.....	22
Figure 4. Assembly and dissociation cycle of the spliceosome.....	25
Figure 5. Dynamics of the Spliceosomal RNA-RNA Rearrangements.....	26
Figure 6. Constitutive versus alternative splicing: the role of cis-acting elements and transacting factors	28
Figure 7. Protein structures and mechanism of action of SR proteins	30
Figure 8. The hnRNP family	31
Figure 9. Positional-dependent effect of RBFOX proteins in alternative splicing.....	34
Figure 10. Alternative splicing alterations in cancer.....	35
Figure 11. Myosin VI isoforms expression in different cell lines.....	60
Figure 12. Myosin VI isoform switch.....	61
Figure 13. EMT markers modulation in A549 sparse growing cells versus cell monolayers.....	62
Figure 14. Polarized Caco-2 cysts express myosin VI _{long} isoform.....	63
Figure 15. PAR3 and ZO-1 do not affect myosin VI isoform switch.....	65
Figure 16. E-cadherin depletion impairs myosin VI isoform switch	67

Figure 17. E-cadherin KD impairs isoform switch in a 3D system	68
Figure 18. PAR3 KD does not affect isoform switch also in a 3D system	69
Figure 19. Active β -catenin localization is affected by E-Cadherin depletion	70
Figure 20. β -catenin depletion impairs myosin VI _{short} expression in Caco-2 cells	71
Figure 21. Diverse splicing regulators are modulated when cells grow as monolayers	72
Figure 22. Validation of splicing regulators differentially expressed in the myosin VI isoform switch system.....	73
Figure 23. PTBP1 affects myosin VI alternative splicing in A549 but not in Caco-2 cells.....	74
Figure 24. FUS is not involved in myosin VI alternative splicing.....	75
Figure 25. SRSF2 promotes skipping of the large insert	77
Figure 26. RBFOX2 promotes inclusion of the large insert	78
Figure 27. Characterization of RBFOX2 overexpressing cell lines	79
Figure 28. RBFOX2 is not a limiting factor in myosin VI alternative splicing.....	80
Figure 29. RBFOX3 does not affect myosin VI AS	81
Figure 30. KHSRP promotes inclusion of the large insert.....	82
Figure 31. RBFOX2 and KHSRP do not show an additive effect	83
Figure 32. Myosin VI isoform expression in HepG2 cells.....	84

Figure 33. Splicing regulators binding sites in myosin VI pre-mRNA	86
Figure 34. RBFOX2 interacts with myosin VI pre-mRNA	88
Figure 35. RBFOX2 binding to UCGAUG motifs in myosin VI pre-mRNA.....	89
Figure 36. Myosin VI isoform expression in ovarian cancer	90
Figure 37. Myosin VI isoforms expression in ovarian cancer cell lines	91
Figure 38. Myosin VI AS regulators are conserved in the ovarian cell line OVCAR-5	92
Figure 39. SRSF2 depletion restores myosin VI _{long} expression	93
Figure 40. SRSF2 is the only member of the SR proteins family promoting myosin VI _{short} expression	94
Figure 41. SRSF3 and SRSF12 depletion do not induce the inclusion of myosin VI large insert	95
Figure 42. SRSF2 KD impairs ovarian cancer cell migration	96
Figure 43. Hypothetical model for myosin VI AS	105

1. Abstract

Scattered evidences suggest that Alternative Splicing (AS) in cancer may drive epigenetic reprogramming. AS can also sustain phenotypic plasticity of cancer cells, an important cue in adaptation to a changing microenvironment. However, the molecular mechanisms leading to the positive selection of pro-tumorigenic splice variants remain poorly characterized.

Aberrant AS of the actin motor protein myosin VI is observed in ovarian cancer, in which a positive selection of a short isoform supports tumor cell migration. Indeed, the inclusion or the skipping of an alternatively-spliced exon cassette, called the large insert, results in the generation of a long and a short isoform, respectively, causing different structural conformations. These conformations, in turn, determine specific interactomes of the long and short splice variants, and thus different biological functions, namely clathrin-mediated endocytosis and cell migration, respectively.

In this project, we aimed at uncovering the mechanism of myosin VI AS in epithelial cell lines to understand how its deregulation affects cancer progression. To this end, we set out to identify both the signaling pathways and the trans-acting regulators involved in the isoform choice. As cellular model systems, we selected A549 (lung adenocarcinoma) and Caco-2 (colorectal adenocarcinoma) epithelial cell lines that express both myosin VI isoforms and can switch on/off their expression according to the cell culture growth conditions. Our setup included biochemical and biological assays, and isoform expression analysis.

We found that, in sparse growth conditions, A549 and Caco-2 cells express predominantly the myosin VI_{short} isoform but switch to an almost exclusive expression of the myosin VI_{long} isoform upon reaching confluency with established cell-cell contacts. This process is fully reversible, as the dissociation of confluent cells and re-plating at low density switched back the isoform expression to myosin VI_{short}, mimicking the process that leads to the positive selection of myosin VI_{short} observed in ovarian cancer cells. Further analysis revealed that the presence of E-cadherin and β -catenin at cell-cell contacts in confluent cells is critical for the AS of myosin VI switch. We also identified three splicing regulators that directly contribute to myosin VI AS decision, namely RBFOX2, KHSRP and SRSF2: the first two promote the inclusion of the large insert and the latter induces its skipping.

Our study highlights an emerging role of AS in cancer cell migration and strongly supports the idea of an AS-reprogramming during cell transformation and acquisition of a malignant

phenotype. Specifically, our findings suggest the involvement of an E-cadherin/ β -catenin axis in the AS-reprogramming during the conversion from an epithelial- to a motile phenotype. Moreover, the identification of splicing factors that are involved in myosin VI AS prompt further investigations to possibly identify other mRNAs, sharing the same regulatory mechanism, that might, like myosin VI, be important in ovarian cancer metastatic program.

2. Introduction

2.1. Myosin VI

2.1.1. Myosin VI: structure and alternative splicing

Myosin VI belongs to a superfamily of motor proteins that hydrolyzes adenosine triphosphate (ATP) to move cargoes along actin filaments, and it is the only member that walks toward the minus end of the actin track (Wells et al. 1999). As a member of the myosin superfamily, myosin VI is composed by a conserved structure (**Fig. 1A**) that includes: a catalytic motor head, which is responsible for binding to actin and hydrolysis of ATP; a neck region containing the unique reverse gear (Ménétreay et al. 2005) and an IQ domain that binds the calcium sensor calmodulin (CaM); and a tail domain, which binds cargoes and anchors the protein to specific membrane compartments. The tail domain is the critical determinant of myosin VI functions and is composed of a N-terminal helical region and a C-terminal globular region called cargo-binding domain (CBD). Two ubiquitin binding domains lie in the tail domain: the motif interacting with ubiquitin (MIU) domain (Penengo et al. 2006) and the myosin VI ubiquitin-binding (MyUb) domain (Wollscheid et al. 2016; He et al. 2016). In particular, the MyUb interacts preferentially with K63-linked ubiquitin chains (He et al. 2016), a non-proteosomal signal that modulates several cellular pathways where myosin VI has been involved, including autophagy, signaling and membrane trafficking events (Tumbarello, Kendrick-Jones, and Buss 2013; Grumati and Dikic 2018; Niño, Sala, and Polo 2019).

A region called “large insert” (LI), located between the two ubiquitin binding domains (**Fig. 1A**, in orange), undergoes alternative splicing (Wollscheid et al. 2016; Buss et al. 2001). This region is encoded by an exon cassette composed of three exons: 29, 30 and 31. Partial (long isoform 1) or total (long isoform 3) inclusion of the exon cassette generates myosin VI_{long}, while skipping generates myosin VI_{short} (isoform 2) (**Fig. 1A**). Notably, the presence or absence of the large insert causes different structural conformations of myosin VI that in turn determine specific interactomes of the two isoforms (Wollscheid et al. 2016) (**Fig. 1B**). Indeed, exon 31, present in myosin VI_{long}, encodes for an alpha helix called α 2 linker that engages the two-helix bundle of the MyUb domain, masking the RRL interaction motif and forming a clathrin binding domain (Wollscheid et al. 2016). In the absence of the large insert (myosin VI_{short}), the RRL motif is exposed and myosin VI can interact with autophagy receptors such as optineurin (Sahlender et al. 2005). As a consequence, the two isoforms are involved in different biological functions: myosin VI_{long} plays a critical role in clathrin-mediated endocytosis [(Biancospino et al. 2019), attached at the end of this thesis] and is expressed in polarized epithelial cells (Wollscheid et al. 2016; Au et al. 2007; Buss et al.

2001), while myosin VI_{short} is essential in supporting cell migration (Wollscheid et al. 2016) and other functions (Tumbarello, Kendrick-Jones, and Buss 2013) (**Fig. 1B**). Noteworthy, while myosin VI_{short} presents no specific expression or localization, the long isoform is specifically found in polarized epithelial cells with well-developed apical microvilli (Buss et al. 2001). Intriguingly, when the RRL interaction surface is lost, myosin VI_{long} acquires an additional binding surface for clathrin, which was recently structurally solved (Biancospino et al. 2019). Our laboratory recently showed that, after myosin VI interaction, residues 46-61 of clathrin light chain a (CLCa) form a short α -helix that appears to complete the myosin VII_{long} three-helix bundle domain (Biancospino et al. 2019). The interaction between myosin VI_{long} and CLCa explains the specialized localization myosin VI_{long} at the apical surface of polarized Caco-2 cysts where this motor protein is required for fission of clathrin-coated pits (CCPs) (Biancospino et al. 2019).

Myosin VI also possesses a second alternatively spliced region, known as the small insert (SI), consisting of 9 amino acids (Buss et al. 2001). The expression and function of the SI have not been extensively addressed, but the myosin VI isoform containing the SI is expressed on PC12 cells, in which it plays a role in tethering secretory granules at the cortical actin network (Tomatis et al. 2013).

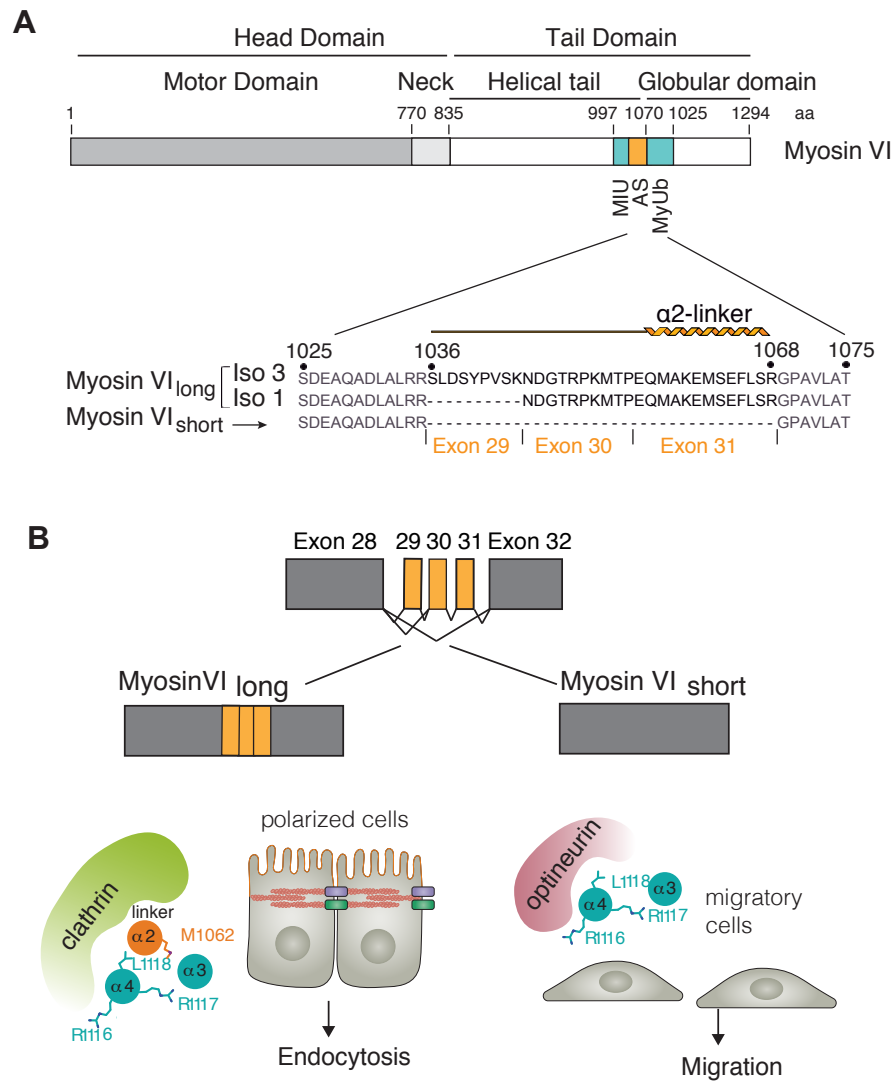


Figure 1. Myosin VI isoforms have different interactomes. **A.** The myosin VI domains include a N-terminal head domain, a C-terminal tail domain and two ubiquitin binding domains (MIU and MyUb). The large insert (LI) localization, sequence and coding exons are indicated (in orange). **B.** Myosin VI isoforms display different structural conformations and interaction partners. The presence of the α 2-linker in myosin VI_{long} (encoded by Exon 31) allows binding to clathrin (left), whereas its absence in myosin VI_{short} exposes a binding surface for other interactors (right).

2.1.2. Myosin VI motor anchor activities

Myosin VI, like the other myosins, can act as both motor and anchor, and different stimuli allow the conversion from one status to the other and vice versa. Under low loads, the protein acts as a processive transporter, but it is converted to an anchor when a sufficient load is applied (Oguchi et al. 2008).

Mechanistically, binding of apo-CaM (Ca²⁺-free) to the IQ motif and the CBD can maintain myosin VI in a back-folded inactive conformation. Subsequently, when the concentration of Ca²⁺ increases, its binding to the CaM induces a conformational change of myosin VI and allows the binding with its cargoes. In order to achieve myosin VI conversion into a motor, upon a decrease in Ca²⁺ levels, CaM stabilizes the lever arm and myosin VI gains thus the mechanical rigidity that is required for movement (Batters et al. 2016).

In this picture, the cargoes can have a double role: stabilizing the open conformation or acting as a trigger for the release of the back-folded state (Fili et al. 2017).

The opening of myosin VI conformation is required for the exposure of the region that determines myosin VI dimerization, which represents a fundamental step in the processive walking mechanism required for myosin VI function as an intracellular transporter (Phichith et al. 2009; Mukherjea et al. 2009; Park et al. 2006). Given the poor dimerization abilities of myosin VI at low local concentrations (Park et al. 2006), clustering of the monomers by cargo binding is fundamental, which can be achieved by different means that strictly depend on the structural features of the binding partners. Indeed, while the endocytic adaptor Disabled homolog 2 (Dab2) is able to bring together two monomers thanks to two myosin VI-binding sites (Yu et al. 2009), the same effect is obtained with cargo homodimerization, like in the cases of optineurin and nuclear dot protein 52 kDa (NDP52) (Fili et al. 2017; Phichith et al. 2009).

2.1.3. Myosin VI functions

Thanks to the capacity of acting both as an anchor and as a motor, as well as the ability to interact with different cargoes, myosin VI can play a role in different cellular processes. A plethora of cargoes, which reflect the multifunctionality of myosin VI, has been identified over the years with the help of biochemical and mass spectrometry approaches (O'Loughlin, Kendrick-Jones, and Buss 2020).

As previously mentioned, polarized epithelial cells express mainly myosin VI_{long}, which localizes at the apical membrane at the base of microvilli, together with the endocytic adaptors AP-2 and Dab2 (Buss et al. 2001; Collaco et al. 2010). A functional role for myosin VI in the endocytosis at the apical membrane was first demonstrated in small intestine epithelial cells, where the endocytosis of cystic fibrosis transmembrane conductance regulator (CFTR), an apically localized Cl⁻ channel, requires interaction with myosin VI, Dab-2 and clathrin (Swiatecka-Urban et al. 2004). Importantly, myosin VI directly binds CLCa (Biancospino et al. 2019) and Dab-2 (Yu et al. 2009) through two different interaction surfaces, providing a link between clathrin-mediated endocytosis and the actin cytoskeleton

(Wollscheid et al. 2016; Biancospino et al. 2019). Our laboratory showed that myosin VI_{long} selectively binds to the clathrin light chain, competing with Huntingtin Interacting Protein 1 Related (Hip1R) (Biancospino et al. 2019). CLCa is likely to recruit myosin VI to CCPs to provide the mechanical force that is needed to complete neck constriction and fission. Using a genetic reconstitution approach, we demonstrated a clear change in the distribution of CCP structures at the apical surface of Caco-2 cysts upon abrogation of myosin VI-CLCa binding, with an increased number of elongated CCPs in CLC-depleted cysts, as well as in cysts reconstituted with the I54D CLCa mutant which are unable to bind myosin VI (Biancospino et al. 2019).

Interestingly, myosin VI recruitment to endocytic sites and to apical microvilli through lipid interaction seems to be sensitive to the saddle-shaped curvature of the membranes (Rogez et al. 2019). Myosin VI_{long} isoform is also expressed in the nervous system: it is involved in synaptic plasticity of the cerebellum, where it mediates the removal of surface AMPA receptors from Purkinje cell spines through clathrin-mediated endocytosis (Wagner et al. 2019).

While myosin VI_{long} isoform plays a specific role in clathrin-mediated endocytosis, myosin VI_{short} is involved at later stages of the endocytic pathways through its interaction with GAIP-interacting protein C-terminus (GIPC) (Wollscheid et al. 2016), an endocytic adaptor that regulates receptor-mediated trafficking by transiently associating with a pool of endocytic vesicles before the maturation into early endosomes (Varsano et al. 2006).

Besides endocytosis, myosin VI has been involved in autophagy and Golgi complex network. Initial localization studies revealed that myosin VI is enriched in the Golgi apparatus, in particular at the trans Golgi where it co-localizes with optineurin (Sahlender et al. 2005; Buss et al. 1998). Optineurin and myosin VI seem to play a role in vesicle formation since myosin VI depletion results in a strong decrease of vesicle fusion events at the plasma membrane (Bond et al. 2011). Furthermore, in polarized MDCK cells, myosin VI and optineurin are present on a specialized recycling compartment, which works as a sorting station for proteins delivered to the basolateral domain (Au et al. 2007).

Myosin VI can interact with several autophagy receptors, such as optineurin and NDP52, through the RRL motif (Morriswood et al. 2007). These autophagy receptors possess a ubiquitin binding domain that allows recognition of substrates targeted for degradation, and a LC3-interacting region (LIR) deputed to the interaction with the autophagosomal membrane. Depletion of myosin VI leads to autophagosome accumulation due to defects in the maturation process and lysosome fusion (Tumbarello, Kendrick-Jones, and Buss 2013; Tumbarello et al. 2012). In this context, myosin VI appears to be involved in the delivery of

endosomal membranes that contain the interacting protein Target of Myb protein 1 (Tom1) to autophagosomes. Indeed, by docking to NDP52, TRAF6-binding protein (T6BP) and optineurin, myosin VI would promote autophagosome maturation and fusion with lysosomes (Tumbarello, Kendrick-Jones, and Buss 2013; Tumbarello et al. 2012). A recent structural study confirmed this scenario, indicating that myosin VI acts as a bridge between endosomes and autophagosomes, facilitating autophagosome generation (Hu et al. 2019). While it is clear that myosin VI interacts with endosomal Tom1 through the WWY motif located in the globular domain (Hu et al. 2019; Finan, Hartman, and Spudich 2011), the detailed interaction of the autophagy receptors with the MyUb remains elusive. In this context, ubiquitin is thought to play a key role: gel filtration experiments showed that K63- or M1-linked ubiquitin chains can induce the formation of a complex including myosin VI, the autophagy receptors and Tom1 (Hu et al. 2019).

Additionally, myosin VI was found to be recruited to damaged mitochondria upon ubiquitination of the outer membrane, where it promotes the assembly of actin cages that isolate dysfunctional organelles (Kruppa et al. 2018). However, this specific process seems to be independent from autophagy receptor recruitment.

Finally, myosin VI has been implicated in cell migration (Wollscheid et al. 2016; Chibalina et al. 2010; Yoshida et al. 2004; Geisbrecht and Montell 2002). The implication of myosin VI in the migratory potential of cancer cells and cell adhesion is not surprising, given the fact that regulation of the actin filament network is critical for cell motility. Initial data generated in the *Drosophila* ovary showed that myosin VI is required for border cell migration and the formation of membrane protrusions at the leading edge of moving cells (Geisbrecht and Montell 2002). Here, myosin VI was found in complex with armadillo and DE-cadherin. The same interaction was confirmed also in mammals, where a direct binding between myosin VI and E-cadherin was proved (Mangold et al. 2012). Myosin VI is recruited to cell-cell contacts by E-cadherin during the maturation of the adherens junctions (AJ), where it interacts with both E-cadherin and vinculin (Maddugoda et al. 2007). Moreover, a recent study showed that an increase in the tension on AJ leads to an increased anchoring of myosin VI to the actin filaments, where it mediates the activation of RhoA pathway in order to counteract the tensile stress and promote collective migration (Acharya et al. 2018).

2.1.4. Myosin VI in ovarian cancer

As previously mentioned, the first evidences for a role of myosin VI in cell migration were obtained in *Drosophila*, where, during oogenesis, a group of follicle cells (border cells)

migrate toward the oocyte in a process that requires a functional Jaguar (myosin VI orthologue), DE-cadherins and armadillo (Geisbrecht and Montell 2002). Later studies demonstrated, using both microarray expression analysis and immunohistochemistry, a strong overexpression of myosin VI in ovarian and prostate cancer tissues, which positively correlated with a clinically aggressive behavior (Dunn et al. 2006; Yoshida et al. 2004). Myosin VI knockdown (KD) was then found to inhibit ovarian cancer dissemination in nude mice (Yoshida et al. 2004), a process that shows similarities with border cells migration.

In order to evaluate a potential differential expression of myosin VI isoforms in cancer, our laboratory performed an *in silico* analysis of RNAseq data from The Cancer Genome Atlas to obtain the relative abundance of exon 31 as a discriminating marker for myosin VI_{long} (Wollscheid et al. 2016), (**Fig. 2A,B**). This analysis revealed that exon 31 skipping is a common event in cancer, where myosin VI_{short} is selectively expressed, especially in ovarian cancer (Wollscheid et al. 2016), (**Fig. 2C**). Analysis of isoform expression in high-grade primary ovarian cancers by RT-PCR showed that while normal ovarian samples (ovarian surface epithelium, OSE) express mainly myosin VI_{long}, all of the tumor samples and six out of seven ovarian-derived cell lines analyzed almost exclusively express myosin VI_{short}, suggesting a positive selection for this isoform in tumor progression (Wollscheid et al. 2016), (**Fig. 2D,E**).

The selective overexpression of myosin VI_{short} supports ovarian cancer cell migration (*in vitro*). Indeed, depletion of myosin VI in cells specifically overexpressing the short isoform, like SKOV-3 and HEY, severely impaired their migratory ability and wound closure (**Fig. 2F**), on the contrary, the same effect was not observed in OVCAR-5 cells which express all myosin VI isoforms (**Fig. 2G**), indicating that cells selectively expressing the myosin VI_{short} isoform become addicted to it for tumor cell migration (Wollscheid et al. 2016).

These findings are very exciting considering the emerging role of alternative splicing in cancer cell biology, where the expression of cancer-associated splice variants has been shown to promote cell transformation and the acquisition of a more aggressive phenotype (Biamonti et al. 2019; Sveen et al. 2016; Eswaran et al. 2013). The study of the expression of pro-tumorigenic isoforms in tumors is becoming relevant both to cancer classification and to the identification of therapeutic targets.

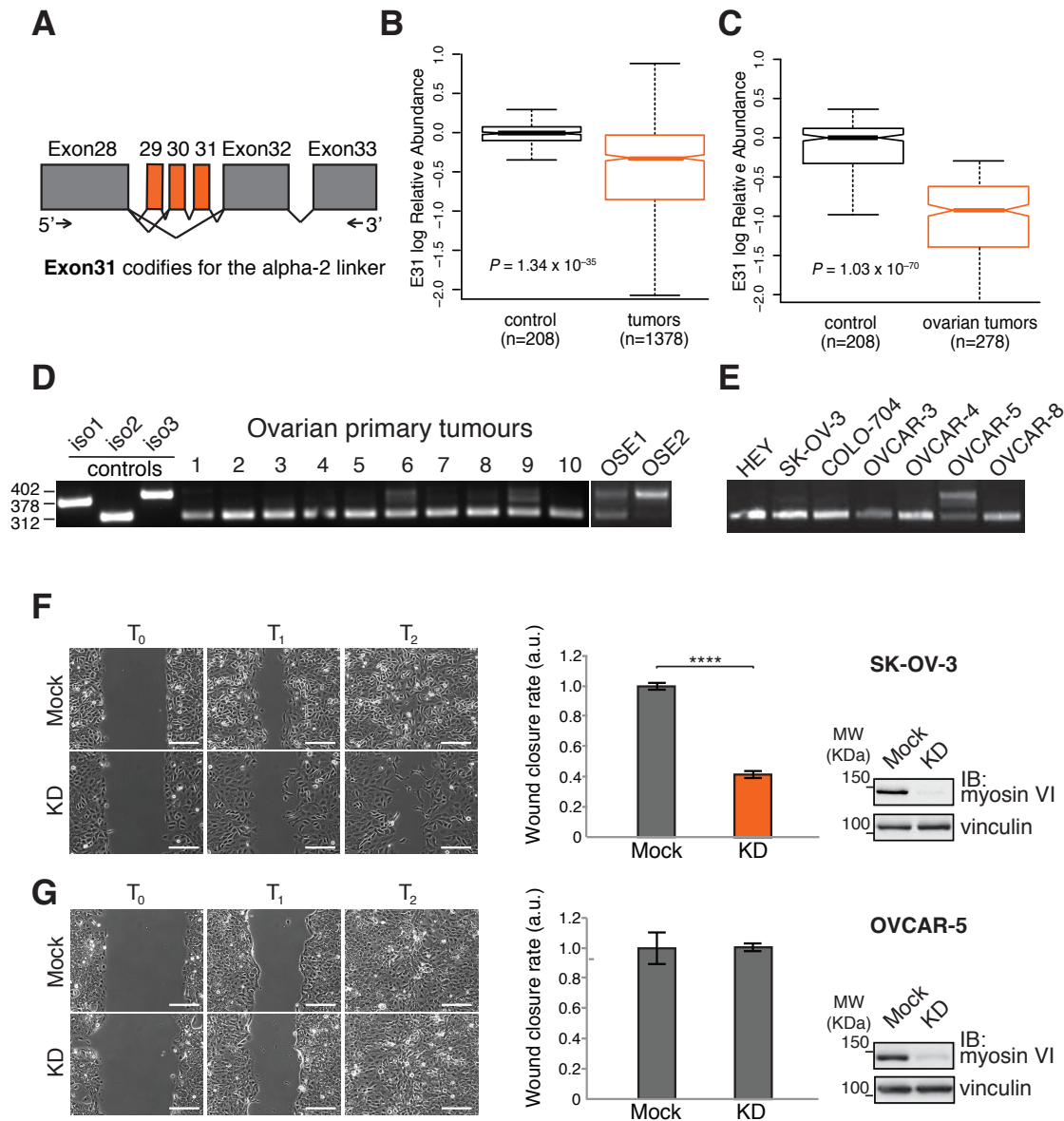


Figure 2. Ovarian cancer cells are addicted to myosin VI_{short} for cell migration. **A.** Schematic representation of the AS region amplified by RT-PCR. Grey boxes represent coding exons and alternative splicing events are depicted. Exon 31 that codes for the α 2-linker is represented in orange. Oligos used for the PCR mapped in Exon 28 and Exon 33 and are indicated by arrows. **B,C.** Box plots representing exon 31 log relative abundance (E31RA) in tumors and ovarian tumors, respectively. Box limits, 25th and 75th percentiles; center line, median; whiskers, 1.5 \times interquartile range (IQR); notches, \pm 1.58 IQR/sqrt(n). P value is from Wilcoxon rank-sum test, lower tail. Values are normalized to their medians in the tumor matched-control samples. E31RA of ovarian cancer and the average E31RA of normal samples from all available cancer types. **D.** RT-PCR analysis of cDNA prepared from the indicated primary cells. PCR controls are from plasmids carrying myosin VI of the different isoforms. Cells from normal tissue (OSE-Ovarian Surface Epithelium) were used as control to compare cancer versus normal myosin VI isoform expression. **E.** RT-PCR analysis of cDNA prepared from the indicated ovarian cancer cell lines. **F,G.** Wound healing assay. The indicated cell lines were transiently transfected with an siRNA targeting myosin VI to achieve KD of the protein. Left panel, sample images: T₀ first frame, T₁ and T₂

arbitrary time points identical for Mock and KD of the same cell lines. Scale bar: 200 μm . Central panel: quantification of the wound closure speed relative to control that represents the average of 60 different measures from three independent experiments +/- SD. Right panel: anti-myosin VI IB and anti-Vinculin IB performed at T_0 to allow quantification of the efficiency of the KD. Adapted from *Wollscheid et al., Nat. Struct. Mol. Biol. , 2016*.

2.2. The pre-mRNA splicing process

2.2.1. The spliceosome catalyzes a two-step reaction

Most genes in higher eukaryotes are transcribed as precursor messenger RNA (pre-mRNA) containing protein-coding segments called exons, which are interrupted by noncoding segments named introns (Sharp 2005, 1994). Splicing is a two-step reaction in which introns are removed and exons are joined together to form mature mRNAs (Chow et al. 1977; Berget, Moore, and Sharp 1977). RNA cleavage and ligation reactions are catalyzed by a ribonucleoprotein (RNP) machine, the spliceosome, that requires ATP hydrolysis (Papasaikas and Valcarcel 2016). The spliceosome is composed of five different RNP subunits, called small nuclear RNPs (snRNPs), and their associated protein co-factors (Will and Luhrmann 2011). These snRNP subunits are generated through the association of small nuclear RNAs (snRNAs), a group of non-coding and non-polyadenylated transcripts, with Sm or Sm-like (LSm) proteins (Matera and Wang 2014), as summarized in **Figure 3A**.

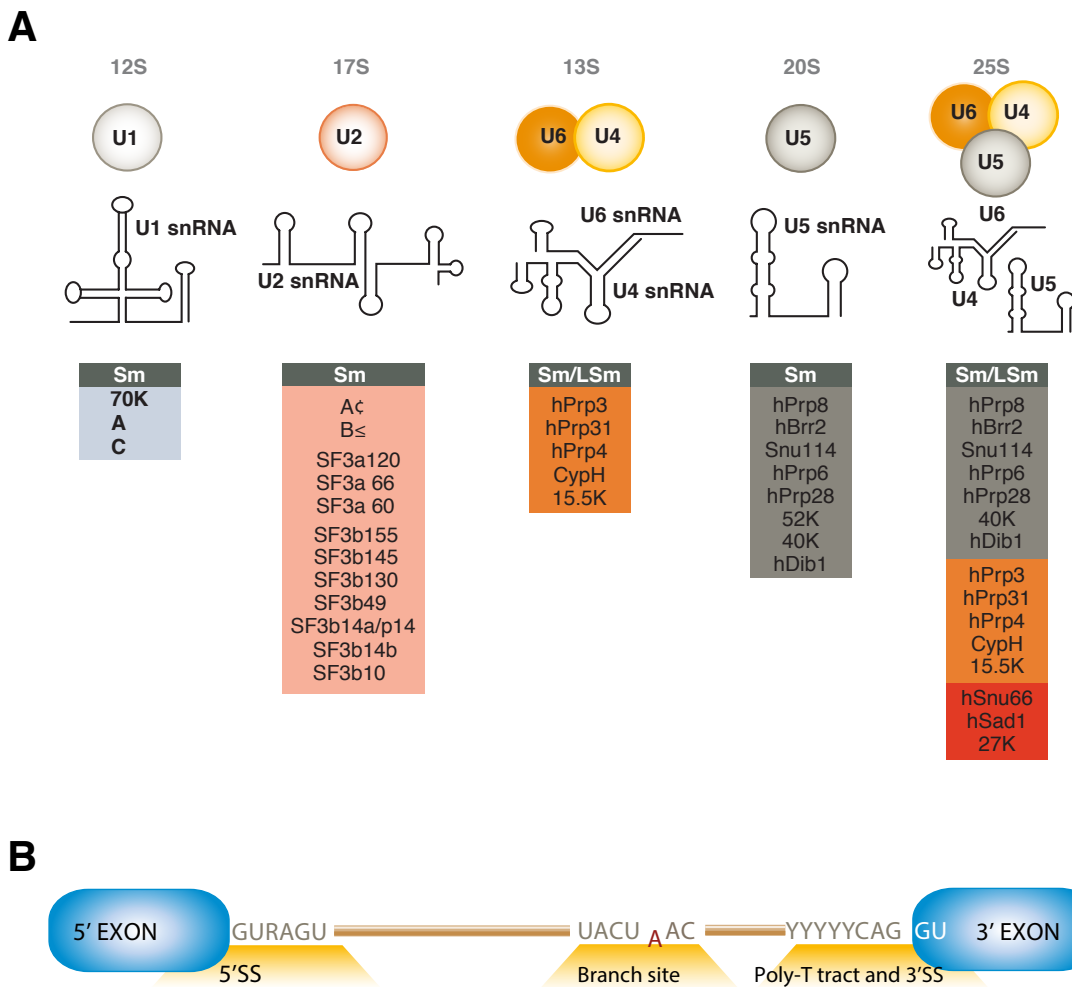


Figure 3. Pre-mRNA consensus sequences and spliceosome composition in humans. A. Protein composition and snRNA secondary structures of the major human spliceosomal snRNPs. The

spliceosomal snRNPs are represented as colored circles. All seven Sm proteins (B/B', D3, D2, D1, E, F, and G) or LSm proteins (Lsm2-8) are indicated by "Sm" or "LSm", respectively, on the top of the boxes showing the proteins associated with each snRNP. The U4/U6/U5 tri-snRNP contains two sets of Sm proteins and one set of LSm proteins. **B.** Schematic representation of consensus sequences recognized by the spliceosome: 5' splice site, branch site point, and 3' splice site. R is a purin, Y is a pyrimidine. Adapted from *Daguenet, Dujardin, and Valcarcel, EMBO Rep, 2015; Will and Luhrmann, Cold Spring Harb Perspect Biol, 2011.*

The spliceosome does not possess a preformed active site, but it is assembled *de novo* directly onto each intron through the interaction of U1, U2, U4/U6 and U5 snRNPs, as well as additional splicing factors (Will and Luhrmann 2011). Most metazoans possess two distinct pre-mRNA splicing machineries known as the major (U2-dependent) and minor (U12-dependent) spliceosomes, which recognize and excise either the major (U2-type) or the minor (U12-type) class of introns, respectively (Turunen et al. 2013). Within the minor spliceosome, U11, U12, U4atac, and U6atac replace the major spliceosome counterparts U1, U2, U4, and U6 snRNAs, respectively, while U5 snRNA is shared by both spliceosomes. Minor introns constitute only ~0.35% of all human introns and, therefore, our description will be focused on U2-dependent splicing (Verma et al. 2018).

The spliceosome can be considered a highly sophisticated ribonucleoprotein chaperone that forces consensus sequences of the pre-mRNA into arrangements that activate an RNA catalytic center and the following intron removal (Papasaikas and Valcarcel 2016).

The chemistry of splicing is quite straightforward: it consists of two transesterification reactions that in principle can occur without the assistance of protein co-factors (Valadkhan and Manley 2001). Group II introns, present in bacterial, mitochondrial and chloroplast genomes, are indeed able to self-splice, without spliceosome assistance (Martin and Koonin 2006), but the formidable differences in length and sequence of the eukaryotic introns make their removal dependent on the spliceosome (Papasaikas and Valcarcel 2016).

Intron removal relies on four consensus sequences that are recognized by spliceosome components: the exon/intron junctions at the 5' and 3' end of the intron, called 5' and 3' splice sites (SS), respectively, the branch point sequence (BPS) located between 18-40 nucleotides upstream of the 3'SS, and the polypyrimidine tract located between the BPS and the 3'SS (Daguenet, Dujardin, and Valcarcel 2015). In metazoans, these consensus sequences are very short and poorly conserved (**Fig. 3B**); thus, in order to compensate for the limited information contained in the splicing substrate itself, the correct folding of eukaryotic introns depends on a multitude on trans-acting factors that make up the

spliceosome. Spliceosomal proteins become essential for the correct recognition and pairing of the splice sites and their positioning to allow transesterification reactions to proceed (Will and Luhrmann 2011), and they represent an important layer of fine regulation.

2.2.2. The splicing cycle

Splicing is achieved by the spliceosome through a series of stepwise interactions between spliceosomal components and the pre-mRNA (**Fig. 4**) Initially, U1 snRNA pairs with the 5'SS, forming the so-called E-complex (E for early). U2 snRNA pairs with the branch point sequence, forming a duplex where the adenosine bulges out: this is the pre-spliceosome or complex A in which the interaction between U1 and U2 snRNPs brings the two splice sites close to each other (**Fig. 4A**). U2 snRNP engagement is assisted by the associated proteins splicing factor 1 (SF1) and U2 auxiliary factor (U2AF). Next, the preassembled U4-U6-U5 tri-snRNP is recruited, forming the pre-catalytic spliceosome or complex B, which undergoes conformational and compositional rearrangements with the displacement of U1 and U4 snRNPs that determine the transition to B_{act} (activated) complex (**Fig. 4B**). This complex is further remodeled into the catalytically active complex B*, which catalyzes the first transesterification reaction, generating complex C. After U2 snRNP conformational change and repositioning of the reaction intermediate, this complex catalyzes the second step of the splicing reaction. The post-spliceosomal complex then dissociates, releasing the mRNA product and the excised intron bound to U2-U5-U6 (**Fig. 4B**). These snRNPs will then be recycled in another splicing reaction (Daguenet, Dujardin, and Valcarcel 2015; Matera and Wang 2014; Papasaikas and Valcarcel 2016; Shi 2017; Will and Luhrmann 2011). During the splicing cycle, the spliceosome undergoes a series of conformational changes that are mainly represented by base pairings occurring between pre-mRNA and snRNA and between the different snRNAs. Assembly of spliceosomal complexes is supported by DExD/H-type RNA-dependent ATPases/helicases that assist a series of dynamic base-pairing rearrangements between snRNAs and pre-mRNA (Matera and Wang 2014).

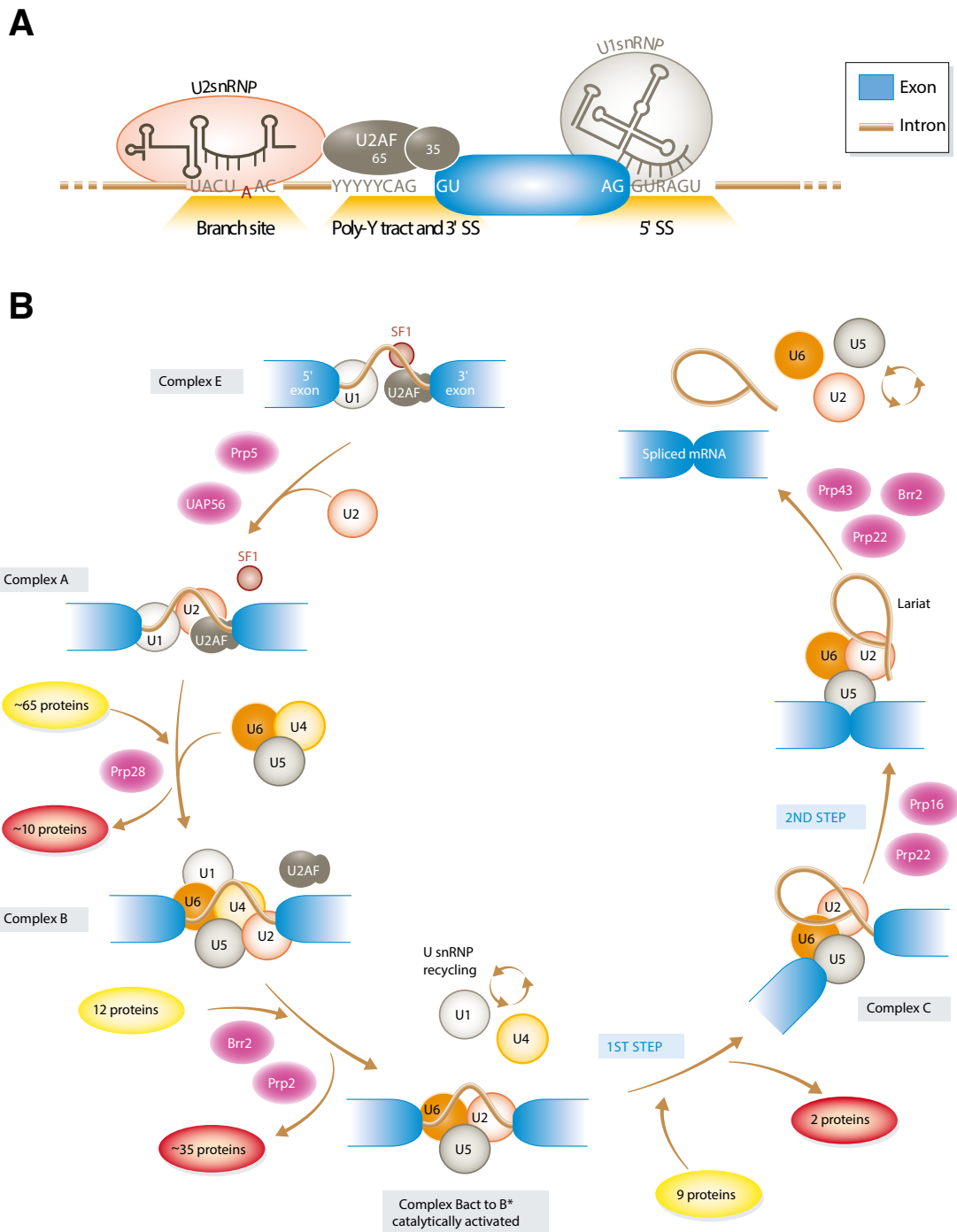


Figure 4. Assembly and dissociation cycle of the spliceosome. **A.** Splicing complex assembly is initiated by consensus sequence elements located at the exon (blue)/intron (brown) boundaries. U1 snRNP and U2AF recognize the 5' and 3' SS, respectively, inducing U2 snRNP recruitment at the branch point. **B.** The stepwise interaction of spliceosomal snRNPs during the removal of an intron from a pre-mRNA is depicted. Rearrangements and remodeling are assisted by ATPases/helicases to allow the progression through complexes E, A, B, and C. From *Daguenet, Dujardin, and Valcarcel, EMBO Rep, 2015.*

Before spliceosome activation, U1 and U2 snRNAs pair with the 5'SS and branch point, respectively. U6 conserved sequences, sequestered by base-pairing with U4, are an essential component of the spliceosome active site since they prevent a premature activation (Papasaikas and Valcarcel 2016; Lee and Rio 2015; Black and Steitz 1986). During the association of the tri-snRNP with the A complex, the helicase Prp28 disrupts the 5'SS- U1 snRNA interaction, allowing U1 replacement with U6 (Staley and Guthrie 1999). This event allows the interaction of the helicase Prp8 with the 5'SS, releasing Prp8 from its binding with U4, thus permitting U4/U6 unwinding (Pena et al. 2008; Reyes et al. 1999). Indeed, the U4/U6 interaction is disrupted by Brr2 helicase in a ATP-dependent manner, allowing U6 to refold (Mozaffari-Jovin et al. 2012), (**Fig. 5**). In this active conformation, U6 also forms base pairs with U2 and an intramolecular loop. The U5 snRNA is initially in contact with exon nucleotides close to the 5'SS, and then with nucleotides of the 3' exon, assisting in the correct positioning of the exons for the second trans-esterification reaction (Papasaikas and Valcarcel 2016; Lee and Rio 2015).

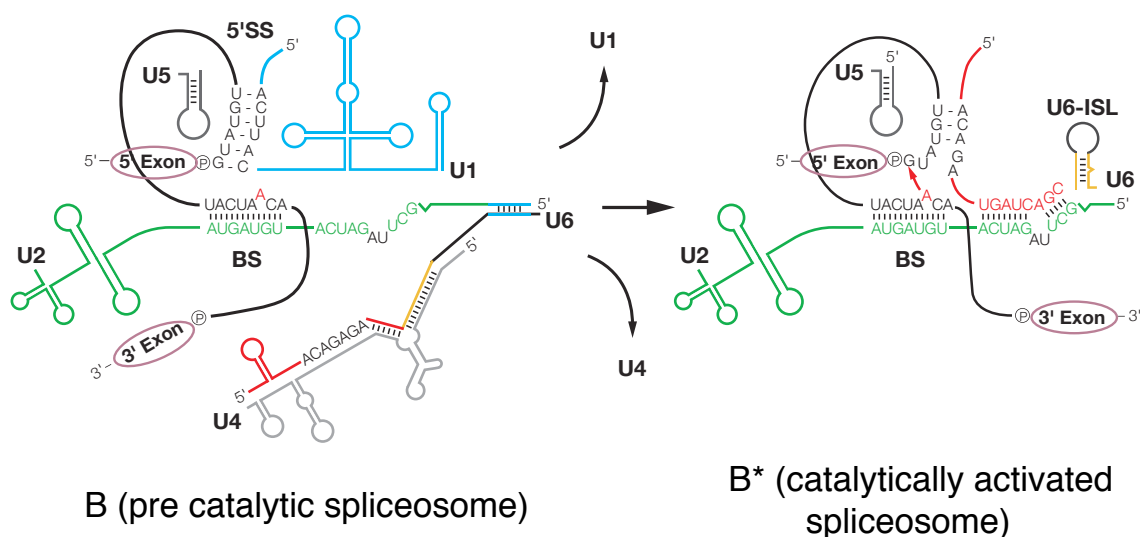


Figure 5. Dynamics of the Spliceosomal RNA-RNA Rearrangements. The network of RNA interactions, the pre-catalytic (left) and the catalytically activated (right) spliceosome. During spliceosome activation, U6 dissociates from U4 and forms base pairs with the 5'SS. Spliceosomal snRNAs are depicted with schematic secondary structures and are not drawn to scale; only loop I of U5 is shown. Adapted from *Alternative pre-mRNA Splicing, Theory and Protocols*, Edited by Stefan Stamm, Chris Smith, and Reinhard Lührmann, 2012.

2.2.3. Alternative splicing: a way to expand the proteome

In a typical multi-exon pre-mRNA, most exons are constitutive: they are flanked by splice sites that are always recognized by the spliceosome and are always included in the final mRNA product. Regulated exons, instead, possess splice sites that can be recognized or not by the spliceosome, and thus are included or excluded from the final mRNA, depending on different conditions. These latter exons, by definition, undergo alternative splicing. Alterations in splice site choice impacts on the mRNA and protein products of a gene, determining the inclusion or exclusion of a portion of coding sequence in the mRNA and giving origin to different “versions” of the same protein called protein isoforms, which differ in their peptide sequence and, hence, functions (Black 2003).

More than 95% of human genes undergo alternative splicing (Lee and Rio 2015), and approximately 60% of these transcripts encode functional protein isoforms (Leoni et al. 2011). Therefore, AS is considered an extraordinary mechanism to generate proteomic diversity and amplify the protein-coding capacity of the human genome by generating multiple protein isoforms from a single gene.

2.2.3.1. Cis-acting elements and trans-acting factors

Splicing of both constitutive and alternative exons is catalyzed by the same machinery, but in the case of alternatively spliced exons, their recognition depends on additional signals.

AS events can be classified based on how they affect the exonic structure of the mature mRNA, as shown in **Figure 6A**. Exons can be extended or shortened, altering thus the position of their splice sites; a regulated exon that can be included or not is called exon cassette. In certain cases, multiple cassette exons are mutually exclusive, with mechanisms enforcing the exclusive choice. Instead, missed intron removal is named intron retention.

Because of 5' and 3' SS consensus motif degeneration, the most critical step in the splicing reaction is the identification of functional splice sites from the numerous pseudo sites found in any pre-mRNA transcript. In general, splice sites close to the consensus sequences tend to be functionally stronger; and, instead, weaker splice sites may need a positive influence in order to be recognized (De Conti, Baralle, and Buratti 2013).

Indeed, most pre-mRNA molecules contain many auxiliary cis-acting regulatory elements that enhance or inhibit splice site recognition (Zhang, Li, et al. 2008). Elements promoting exon inclusion are defined exonic splicing enhancers (ESE) or intronic splicing enhancers (ISE) depending on their location. On the contrary, elements inducing exon skipping are called exonic or intronic splicing silencers (ESS or ISS, respectively). These sequences were initially considered important only for alternative splicing, but they turned out to play a role

also in constitutive splicing (Han et al. 2011). These cis-acting elements are recognized and bound by trans-acting factors, RNA binding proteins that can promote or prevent the recruitment of the spliceosome (Matera and Wang 2014; Black 2003). The best characterized categories of splicing regulators are serine and arginine-rich (SR) proteins and heterogeneous nuclear ribonucleoproteins (hnRNPs) (Busch and Hertel 2012). In general, SR proteins have been described as positive regulators of splicing, whereas hnRNP proteins are classified as negative regulators (**Fig. 6B**). However, more recently, the historical distinction between positive and negative regulators of AS has been overcome with increasing evidences showing how the behavior of these proteins may vary depending on position-specific binding effects. For example, SR proteins can also repress splicing, depending on where they bind within a pre-mRNA and an hnRNP such as the polypyrimidine tract binding protein 1 (PTBP1), can also activate splicing (Dvinge 2018; Dvinge et al. 2016).

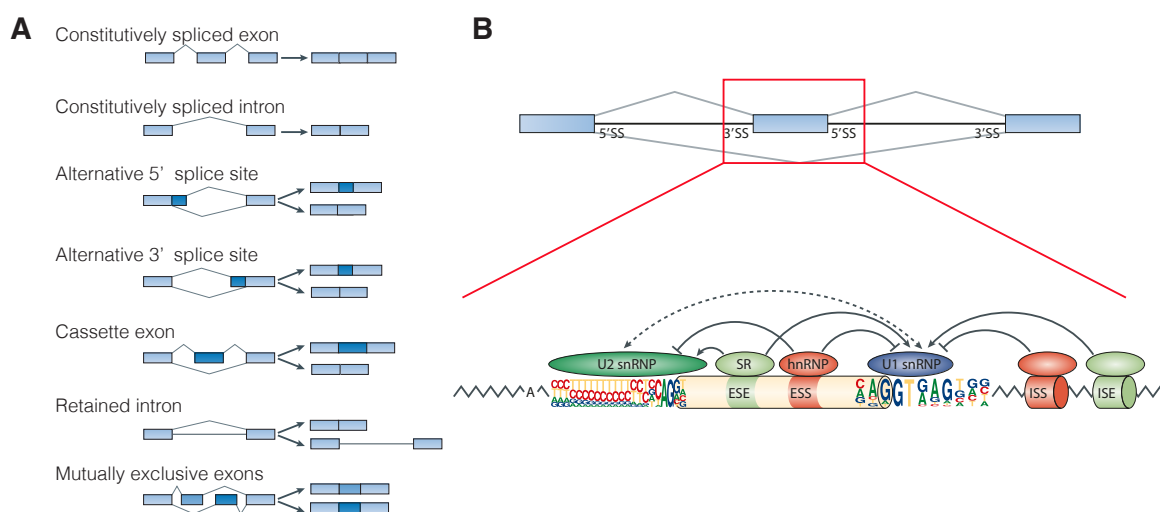


Figure 6. Constitutive versus alternative splicing: the role of cis-acting elements and trans-acting factors. **A.** Schematic representation of constitutive and alternative splicing events. Light blue: constitutive sequence that always forms part of the mature mRNA; mid-blue or dark blue: alternative sequence that can be either included or excluded in the mature mRNA. **B.** Cis-acting splicing regulatory elements (SREs) and trans-acting splicing factors regulate splice site choice. SREs are classified as exonic splicing enhancers (ESEs), intronic splicing enhancers (ISEs), exonic splicing silencers (ESSs), or intronic splicing silencers (ISSs) based on their relative locations and activities. These SREs specifically recruit splicing factors to promote or inhibit recognition of nearby splice sites, like SR proteins or hnRNPs. The consensus motifs of splice sites are shown in the colored pictograph, where the height of each letter represents nucleotide frequency in each position. Adapted from *Matera and Wang, Nat Rev Mol Cell Biol, 2014; Dvinge et al., Nat Rev Cancer, 2016.*

2.2.3.1.1. SR proteins

SR proteins constitute a subset of splicing regulators characterized by arginine-serine-rich domains involved in both constitutive and alternative splicing (Zhou and Fu 2013; Long and Caceres 2009; Fu 1995). In humans, 12 SR proteins have been identified, designated as serine/arginine-rich splicing factor (SRSF) 1–12, which can be divided into shuttling or non-shuttling SR proteins (**Fig. 7A**). Indeed, some of these proteins are exclusively located in the nucleus, while others can shuttle between the nucleus and the cytoplasm (Jeong 2017), as indicated in **Figure 7A**. At the structural level, SR proteins are characterized by one or two N-terminal RNA recognition motif (RRM) domains and a C-terminal serine-arginine rich domain (RS domain). The RS domain is constituted by at least 50 amino acids with >40% RS content, in the form of consecutive and extended arginine and serine dipeptides (Jeong 2017; Zhou and Fu 2013; Busch and Hertel 2012). These structural requirements and the ability to complement splicing reactions define the SR family members. Other RS domain-containing proteins that participate in splicing exist, but they do not possess all these features together (Jeong 2017; Busch and Hertel 2012; Fu 1995).

SR proteins are involved in early spliceosome assembly, facilitating recruitment and stabilization of U1 snRNP and U2AF at 5' and 3' SS, respectively (Bourgeois, Lejeune, and Stevenin 2004; Kohtz et al. 1994; Staknis and Reed 1994; Zuo and Maniatis 1996). Based on the exon definition model, the role of SR proteins in constitutive splicing would be due to simultaneous interaction with RNA sequences and with factors bound to splice sites, through a network of protein-protein contacts mediated by the RS domains of all these proteins (Bourgeois, Lejeune, and Stevenin 2004; Berget 1995).

Many biochemical studies in the past established the capacity of these factors to promote splice site selection when bound to ESEs, (Liu et al. 2000; Schaal and Maniatis 1999; Cavaloc et al. 1999; Liu, Zhang, and Krainer 1998; Tacke and Manley 1995). However, more recent studies on SRSF1, SRSF2 and SRSF7 revealed that the binding to an intronic sequence has an inhibitory effect, as schematized in **Figure 7 B,C** (Erkelenz et al. 2013; Dembowski et al. 2012). Moreover, since alternative splicing is a matter of competition, it has been shown that while SR protein (namely SRSF1 and SRSF2) binding to an alternatively spliced exon induces its inclusion, binding to a flanking constitutive exon induces skipping of the internal exon (**Fig. 7D**) (Luo et al. 2017; Han et al. 2011; Ghigna et al. 2005). Such competition occurs also between different SR proteins, meaning that binding of one SR protein may enhance or repress the binding of other family members; for example, SRSF2 depletion in MEF cells can increase or decrease SRSF1 binding to different target sites (Pandit et al. 2013; Zhou and Fu 2013). In addition, a position-dependent mechanism

has been proposed for SR protein AS regulation. An SR protein, like SRSF1, bound to an intronic region upstream of the functional 5' splice site interferes with exon definition, acting as a decoy or pseudoexon for the splicing machinery (Buratti et al. 2007).

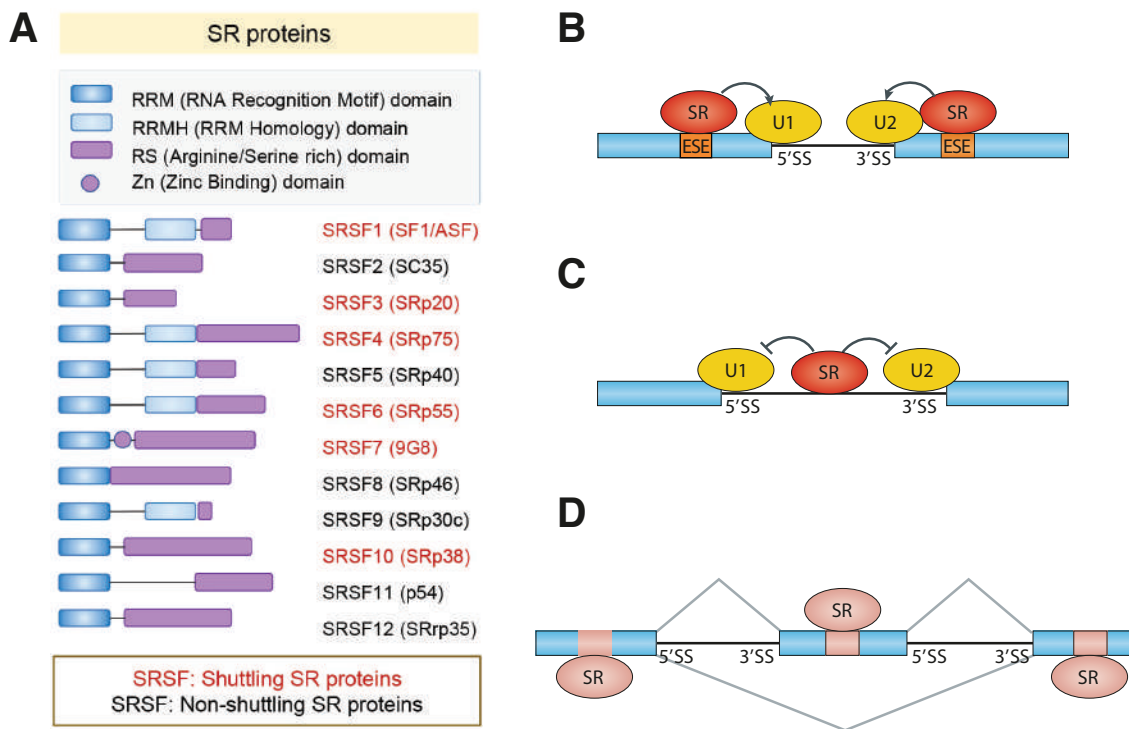


Figure 7. Protein structures and mechanism of action of SR proteins. **A.** The domain structures (RRM, RRMH, RS, and Zn) are denoted as shown in the box. Current names for SR proteins are SRSFs, but aliases are also indicated in the parenthesis. Among 12 core SR proteins, 6 (red letters) are reported to shuttle between the nucleus and the cytoplasm (shuttling SR proteins), whereas the others (black letters) have not been shown to have a shuttling activity (non-shuttling SR proteins). **B.** Position-dependent effects of SR proteins on splice site selection. SR protein binding to exonic splicing enhancer (ESE) stimulates the recognition of the nearby 5' and 3' splice sites by U1 and U2 snRNP, respectively. **C.** SR protein binding to intronic regions inhibits splicing, likely by interfering with the communication between functional 5' and 3' splice sites. **D.** SR protein-dependent exon inclusion or skipping. SR protein binding to the internal alternative exon promotes exon inclusion, whereas SR protein binding to the flanking competing exon(s) causes exon skipping. Adapted from Jeong, *Mol Cells*, 2017; Zhou and Fu, *Chromosoma*, 2013.

SR proteins undergo extensive phosphorylation in their SR domains on serine-arginine repeats by SR protein-specific kinases (SRPKs) and cyclin-dependent like kinases (CLKs), (Zhong et al. 2009; Yomoda et al. 2008; Hayes et al. 2006; Fukuhara et al. 2006). SRPKs constitute a unique family of kinases characterized by a long spacer sequence that separates the kinase domains in two lobes. SRPK1 is ubiquitously expressed, SRPK2 is expressed

mainly in the lung, while SPRK3 is expressed only in muscles. While SRPKs are detected both in the cytoplasm and in the nucleus, CLKs are constitutively localized in the nucleus, with enrichment in nuclear speckles (Zhou and Fu 2013). Based on this differential localization pattern, the two kinase families are thought to regulate SR protein activity in a coordinated fashion.

2.2.3.1.2. hnRNP proteins

The hnRNP proteins constitute a heterogeneous group of splicing regulators rather than a single family. The hnRNPs mainly bind splicing silencers, influencing both constitutive and alternative splicing (Lee and Rio 2015; Black 2003). Like SR proteins, hnRNPs possess a modular structure (Fig. 8), with RRM, K-homology (KH) domains or RNA-binding domains consisting of Arg-Gly-Gly repeats (RGG) that mediate specific interactions with the RNA and, in many cases, additional glycine-rich, acidic or proline-rich domains (Geuens, Bouhy, and Timmerman 2016; Busch and Hertel 2012).

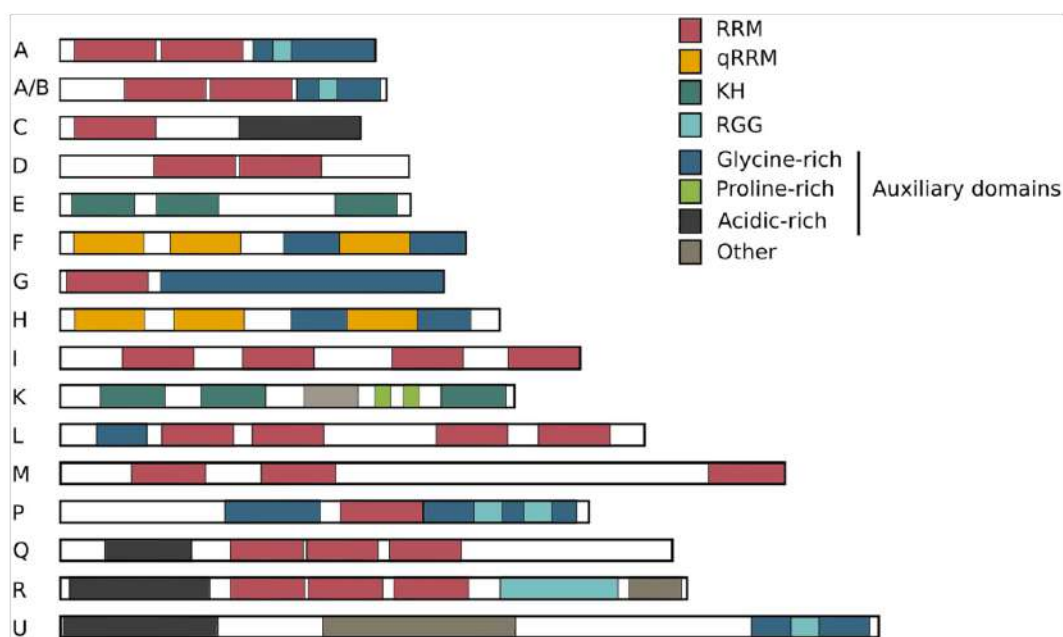


Figure 8. The hnRNP family. The hnRNPs have different molecular weights ranging from 34 to 120 kDa and are named alphabetically from hnRNP A1 to hnRNP U. As shown in the overview, several structural domains are shared between different family members. The members of the hnRNP family are built up of four unique RNA binding domains (RBDs): RRM RNA recognition motif, qRRM quasi-RNA recognition motif, KH K-homology domain, and RGG RNA-binding domain consisting of Arg-Gly-Gly repeats. The sizes of these 16 common hnRNPs are drawn relative to each other. Adapted from *Geuens, Bouhy, and Timmerman, Hum Genet, 2016.*

In general, hnRNP proteins inhibit splicing binding to exons but can have both positive and negative effects on splicing when bound to intronic cis-acting elements. For example, hnRNP L represses or enhances splicing by binding to CA-rich motifs in exons (Melton et al. 2007; Rothrock, House, and Lynch 2005) or introns (Hui et al. 2005; Hui et al. 2003), respectively. Also, hnRNP H promotes the selection of a nearby 5' splice site by binding to intronic G-rich sequences (Chou et al. 1999; Xiao et al. 2009), but represses splicing when it binds to exons (Chen, Kobayashi, and Helfman 1999; Crawford and Patton 2006; Mauger, Lin, and Garcia-Blanco 2008).

One of the most studied negative regulators of alternative splicing is the polypyrimidine tract-binding protein (PTBP1, also known as hnRNP I). PTBP1 preferentially binds to intron pyrimidine-rich elements and mediates tissue-specific regulation of exon splicing (Patton et al. 1991; Garcia-Blanco, Jamison, and Sharp 1989). The cis-acting elements bound by PTBP1 are represented by sequences containing 15–25 pyrimidines, with a preference for pyrimidine tracts containing UCUU, CUCUCU (Reid et al. 2009; Ray et al. 2009; Perez, McAfee, and Patton 1997). General rules for PTBP1 mechanism of action are difficult to derive because of the variability in the location of PTB binding sites relative to its targeted exons. Most studies are focused on PTBP1-mediated AS repression, where PTBP1 may bind either exonic or intronic regions. PTBP1 intronic binding sites are often found within the pyrimidine tract of the 3' splice site (Amir-Ahmady et al. 2005; Shen et al. 2004), but also downstream of the 5' splice site of the repressed exon, like in the cases of exon 3 of the α -tropomyosin transcript, c-Src exon N1 and FGFR2 Exon IIIb (Wagner and Garcia-Blanco 2002, 2001; Chou et al. 2000; Carstens, Wagner, and Garcia-Blanco 2000; Gooding, Roberts, and Smith 1998; Chan and Black 1997).

PTBP1 and its paralog PTBP2 have been well studied for their role in neurogenesis. In embryonic stem cells (ESCs), PTBP1 controls the splicing of the transcription factor pre-B cell leukemia homeobox (Pbx1), inducing skipping of exon 7 and thus preventing the expression of the neuronal isoform Pbx1a in favor of the shorter Pbx1b. Exon 7 is de-repressed as ESCs differentiate into neuronal progenitor cells (NPCs) and then into mature neurons (Linares et al. 2015). PTBP1 also represses the expression of its paralog PTBP2, stimulating skipping of its exon 10 and leading to nonsense-mediated-decay (NMD); (Boutz et al. 2007; Makeyev et al. 2007; Spellman, Llorian, and Smith 2007). Upon mitotic exit and transition to NPCs, the downregulation of PTBP1 allows the expression of PTBP2, which is required for neuronal development and survival (Vuong, Black, and Zheng 2016; Keppetipola et al. 2012). Later in the maturation process, also PTBP2 is down-modulated

and synaptogenesis takes place. Overall, the changes in the expression of the two PTBs define two transitions in splicing regulation, one early in neuronal differentiation when PTB is replaced with PTBP2, and one later in neuronal maturation when PTBP2 expression is reduced (Vuong, Black, and Zheng 2016).

2.2.3.1.3. Tissue-specific splicing factors

A third heterogeneous group of splicing regulators is composed of tissue-specific splicing factors, including neuro-oncological ventral antigen (NOVA), RNA binding protein fox-1 (RBFOX), Muscleblind-like (MBNL), and CUG-BP and ETR-3-like factors (CELF) family proteins, Hu proteins, TIA1/TIAR, and probably many more yet to be characterized (David and Manley 2008).

The proteins of the NOVA family, namely NOVA-1 and NOVA-2, are neuron-specific RNA-binding proteins. NOVA-1 is expressed primarily in the hindbrain and ventral spinal cord, whereas the expression of NOVA-2 is mainly restricted to the neocortex, even if tumors outside of the nervous system may trigger their ectopic expression (Zhou et al. 2014; Racca et al. 2010; Buckanovich, Posner, and Darnell 1993). More recently, NOVA-2 was demonstrated to be expressed in endothelial cells both in physiological and in pathological conditions (Belloni et al. 2019; Angiolini et al. 2019; Giampietro et al. 2015). NOVA-1 and NOVA-2 contain three KH domains, among them, KH3 is necessary and sufficient for specific RNA binding (Zhou et al. 2014). *In vitro* binding experiments have shown that NOVA proteins interact with clusters of YCAY elements (where Y is a pyrimidine) on pre-mRNA to regulate alternative splicing (Jensen et al. 2000; Lewis et al. 1999).

The RBFOX protein family is constituted by three closely related splicing regulators: RBFOX1, RBFOX2 and RBFOX3, that recognize and bind (U)GCAUG motifs mainly located in introns flanking tissue-specific alternative exons or in 3'- untranslated region (UTR) in mRNAs (Zhang, Zhang, et al. 2008; Jin et al. 2003). RBFOX1 is expressed in the brain, the heart and in skeletal muscles, RBFOX3 is expressed almost exclusively in the brain, while RBFOX2 is expressed ubiquitously (Kuroyanagi 2009).

UGCAUG motifs are enriched in proximity of tissue specific alternative exons in brain (especially microexons), muscle, epithelial and mesenchymal cells, in breast cancer, in mouse embryonic stem cells (poison exons), and in distal intronic regions (Li et al. 2015; Lovci et al. 2013; Shapiro et al. 2011; Lapuk et al. 2010; Bland et al. 2010; Yeo et al. 2009; Wang et al. 2008; Das et al. 2007).

Experimental and bioinformatics analyses demonstrated a position-dependent effect of RBFOX proteins in alternative splicing. As shown in **Figure 9**, binding of RBFOX to an

UGCAUG motif located upstream of the alternative exon generates exon skipping, while binding downstream induces exon inclusion (Weyn-Vanhentenryck et al. 2014; Lovci et al. 2013; Barash et al. 2010; Yeo et al. 2009; Wang et al. 2008; Jin et al. 2003). In the first case, RBFOX proteins may inhibit binding of the spliceosome to the BPS or 3'SS (Zhou, Baraniak, and Lou 2007), while in the second case the proteins can positively influence the recruitment of U1 snRNP (Huang et al. 2012).

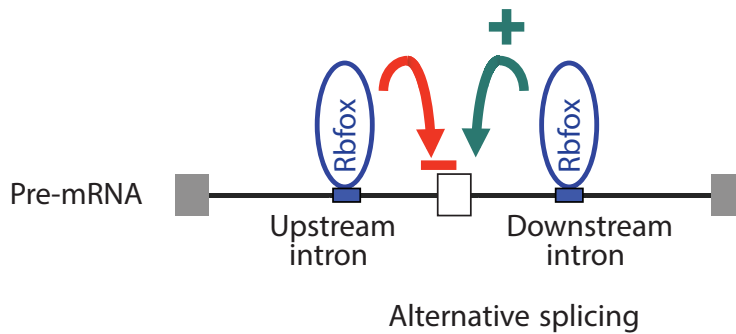


Figure 9. Position-dependent effect of RBFOX proteins in alternative splicing. Binding of RBFOX proteins upstream of the adjacent alternative exon inhibits its inclusion, while binding downstream enhances it. Adapted from Conboy, *Wiley Interdiscip Rev RNA*, 2016.

2.2.4. Alternative splicing in cancer

2.2.4.1. Splicing regulators as oncoproteins or oncosuppressors

Alternative splicing is a finely regulated process, precisely controlled during development and differentiation. In fact, deregulation of alternative splicing can lead to the production of aberrant protein isoforms that can contribute to diseases, especially cancer. Genome-wide studies have identified thousands of cancer-associated splice variants involved in all of the different hallmarks of cancer, including proliferation, differentiation, cell cycle, metabolism, apoptosis, motility, invasion, and angiogenesis (**Fig. 10**). Splicing alterations that promote tumorigenesis can arise from mutations in cis-acting elements or trans-acting factors, but more often, they can also arise from an altered expression or from the activity of the splicing regulators, which can thus act as oncogenes or oncosuppressors (Urbanski, Leclair, and Anczukow 2018; Zhang and Manley 2013; Shapiro et al. 2011; He et al. 2009; Venables et al. 2008).

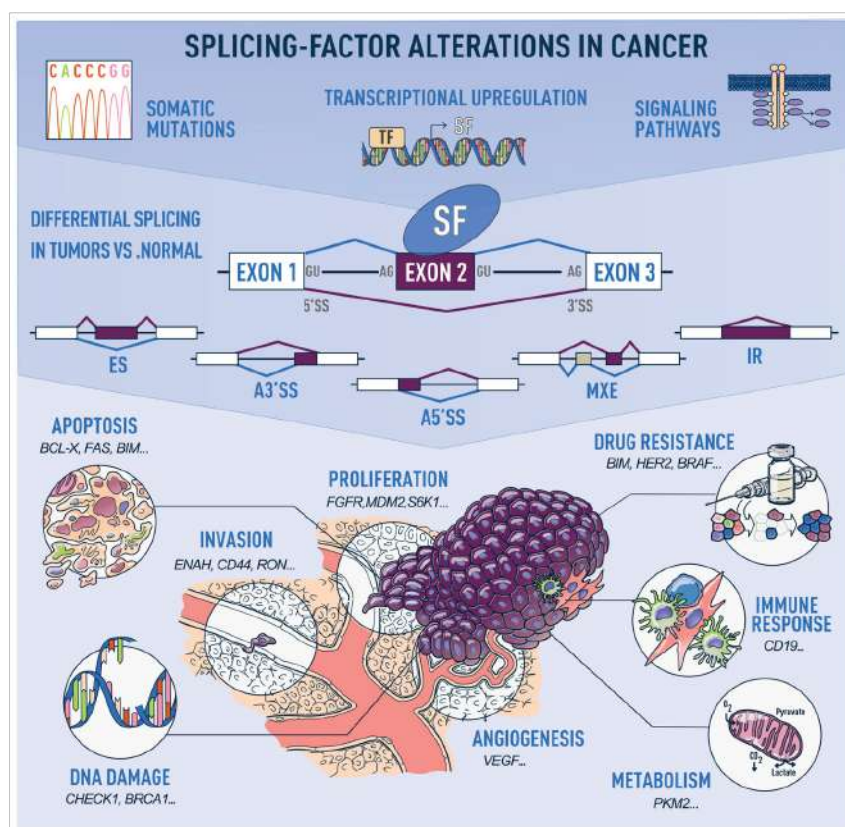


Figure 10. Alternative splicing alterations in cancer. Human tumors exhibit recurrent mutations or changes in the levels of splicing regulators due to copy number changes, or alterations in the transcriptional, post-transcriptional, or post-translational regulation of splicing factors in response to changes in signaling (top panel). These changes in the levels of the splicing factor lead to alterations in the splicing of their downstream targets. Misregulated splicing of isoforms involved in key cellular pathways contributes to tumor initiation and progression. Examples of cancer hallmarks and associated tumor isoforms are indicated (bottom panel). From Urbanski, Leclair, and Anczukow, *Wiley Interdiscip Rev RNA*, 2018.

RNAseq analyses of tumor and paired normal tissues from the same individuals have provided insights into the effects of somatic mutations on splicing (Jung et al. 2015; Supek et al. 2014). Tumor suppressor genes are enriched for single nucleotide variants (SNVs) that induce intron retention; this event, in most cases, causes the formation of premature stop codons that prevent the expression of the affected genes. The most commonly affected genes resulted to be tumor protein p53 (*TP53*), AT-rich interactive domain-containing protein 1A (*ARID1A*, encoding for a chromatin remodeler) and phosphatase and tensin homolog (*PTEN*) (Jung et al. 2015; Supek et al. 2014). On the other hand, among the mutations known to activate proto-oncogenes, there are alterations causing exon 14 skipping in the Hepatocyte Growth Factor Receptor (*HGFR*) gene and the activation of a cryptic splice site in the Notch homolog 1 (*NOTCH1*) gene (transmembrane receptor involved in development processes), resulting in an aberrantly acting form of the protein in chronic lymphocytic leukemia (Puente et al. 2015; Collisson EA 2014; Ma et al. 2003). However, a substantial proportion of somatic mutations affecting splicing of oncogenes is constituted by synonymous mutations affecting ESE (gain of function) or ESS (loss of function) sequences that are located within 30 nt of exon boundaries (Supek et al. 2014).

As previously mentioned, RNA splicing factors themselves are recurrently affected by somatic mutations that alter the final protein. In hematological malignancies, mutations in Splicing factor 3B subunit 1 (*SF3BP1*), U2AF1, SRSF2 and U2 small nuclear ribonucleoprotein auxiliary factor 35 kDa subunit-related protein 2 (*ZRSR2*) alter their capacity to bind RNA. These mutations occur in heterozygosis at restricted residues, conferring a different RNA binding specificity and resulting in a gain-of-function mechanism rather than a loss-of-function of the protein (Inoue and Abdel-Wahab 2016; Alsafadi et al. 2016; Kim et al. 2015; Ilagan et al. 2015; DeBoever et al. 2015; Darman et al. 2015). Instead, loss-of-function mutations in the splicing regulator RNA Binding Motif Protein 10 (*RBM10*) are found in lung adenocarcinoma, colorectal carcinoma, pancreatic adenocarcinoma and in intraductal papillary mucinous neoplasm (Giannakis et al. 2016; Witkiewicz et al. 2015; Collisson EA 2014; Imielinski et al. 2012). In lung cancer, missense or truncating mutations in *RBM10* are associated with a reduced level of the protein, changes in cell proliferation and alternative splicing of its targets. For instance, *RBM10*-mediated AS regulation of Protein numb homolog (*NUMB*) is disrupted in lung cancer with an increase in the expression of the pro-proliferative isoform (Zhao et al. 2017; Hernandez et al. 2016).

However, in most solid tumors, mutations in splicing factors are not a frequent event; more often it is their differential expression or activity that is responsible for splicing deregulation

(Urbanski, Leclair, and Anczukow 2018; Anczukow and Krainer 2016). For instance, SRSF1 is frequently upregulated in breast, lung, colon and bladder tumors; in breast cancer cells its overexpression promotes transformation by enhancing cell proliferation and decreasing apoptosis (Anczukow et al. 2012; Karni et al. 2007; Ghigna et al. 2005). Similarly, overexpression of other SR proteins has been found across different tumor types, especially SRSF3 and SRSF6 (Urbanski, Leclair, and Anczukow 2018). Analogously, hnRNP proteins can act as oncogenes as well. In glioblastoma, hnRNPA2/B1 overexpression causes skipping of Recepteur d'Origine Nantais (*RON*) exon 11, creating an oncogenic isoform involved in cell motility (Golan-Gerstl et al. 2011); hnRNP K can act both as a tumor suppressor and as an oncogene, while hnRNP M is implicated in regulating Epithelial to Mesenchymal transition (EMT) in breast epithelial cells, influencing the splicing of Cluster of Differentiation 44 (CD44) (Urbanski, Leclair, and Anczukow 2018; Xu et al. 2014).

2.2.4.2. Alternative splicing in cell proliferation and survival

Several alternative splicing events (ASEs) lead to the production of tumor-associated isoforms that function in the different hallmarks of cancer and are positively selected during tumorigenesis. For instance, fibroblast growth factor receptors (FGFR) 1/2/3 are involved in cell proliferation and migration both during development and wound healing. They are formed by a cytoplasmic kinase domain, a transmembrane domain and an extracellular ligand binding domain constituted by three immunoglobulin domains. This last region undergoes alternative splicing (under the control of hnRNP H1/F and ESRP1/2) of two mutually exclusive exons, 8 and 9, to generate the epithelial isoforms FGFR-IIIb or the mesenchymal FGFR-IIIc, with the latter promoting tumor growth and invasion (Gong 2014). Furthermore, a common mechanism through which alternative splicing influences tumorigenesis is the switch, in the expression of proteins involved in apoptosis, toward proapoptotic isoforms. For instance, the alternative splicing of Bcl-2-like protein 1 (BCL2L1), a member of the Bcl-2 family, generates isoforms BCL-XL, which prevents apoptosis, and BCL-XS, which induces cell death. The latter is produced through the usage of an alternative 5' splice site in exon 2, an event stimulated by Src-Associated substrate in Mitosis of 68 kDa (Sam68) and RBM25 and counteracted by SRSF1 (Bielli et al. 2014; Zhou et al. 2008; Paronetto et al. 2007; Boise et al. 1993). An increased expression of BCL-XL and a decreased expression of BCL-XS are detected in lymphoma, glioma, myeloma and neuroblastoma (Li et al. 2016; Tu et al. 1998; Xerri et al. 1996; Dole et al. 1995). Similarly, the Fas Cell Surface Death (FAS) receptor involved in the extrinsic pathway of apoptosis,

possesses an alternatively spliced isoform, sFAS, which is soluble and cannot promote cell death since it is produced upon skipping of exon 6, which encodes for the transmembrane domain (Cheng et al. 1994). Exon 6 inclusion is promoted by Ewing Sarcoma breakpoint region 1 (EWS), hnRNP A1 and TIA1 (Paronetto et al. 2014; Oh et al. 2013; Izquierdo et al. 2005), while RBM5 and PTBP1 induce exon skipping (Bonnal et al. 2008). The same happens with Caspase 2, initiator of apoptosis and tumor suppressor: skipping of exon 9, promoted by SRSF3, produces the proapoptotic isoform Casp-2L, while its inclusion, promoted by RBM5, produces a premature stop codon and the antiapoptotic Casp-2S isoform (Jang et al. 2014; Fushimi et al. 2008).

2.2.4.3. Alternative splicing in cancer metabolism

Alternative splicing influences not only cell proliferation and survival, but also processes like metabolism; the best-known example is the pyruvate kinase (PKM) alternative splicing. PKM is a key glycolytic enzyme, possessing two isoforms generated by mutually exclusive exons. Inclusion of exon 9 produces the constitutively active isoform PKM1, whereas inclusion of exon 10 generates PKM2 (Dayton et al. 2016). PKM2 expression depends on exon 9 repression by PTBP1, hnRNP A1/2 and/or exon 10 inclusion by SRSF3 (Dayton et al. 2016; David et al. 2010; Clower et al. 2010). PKM2 levels are increased in many solid tumors and its depletion inhibits tumor progression in ovarian, gastric, colon, liver, and esophageal cancer models (Wang et al. 2017; Shiroki et al. 2017; Chao et al. 2017; Liu et al. 2015).

2.2.4.4. Alternative splicing in angiogenesis

The Vascular endothelial growth factor A (*VEGFA*) gene encodes a growth factor that stimulates blood vessel formation through proliferation and migration of endothelial cells. The related transcript undergoes different alternative splicing events, resulting in isoforms of variable length. Inclusion of the variable exons 6a, 6b, 7a, and 7b produces the VEGFA_{xxx} isoforms (“xxx” indicates the final number of the amino acids); instead, inclusion of exon 8b in place of 8a produces the antiangiogenic VEGFA_{xxx}b isoforms (Bates et al. 2002; Houck et al. 1991). A decreased expression of VEGFA₁₆₅b is reported in melanoma and prostatic tumors, as well as a shift from VEGFA₁₆₅b to VEGFA₁₆₅ in colon cancer and squamous cell carcinoma (Biselli-Chicote et al. 2017; Varey et al. 2008; Pritchard-Jones et al. 2007; Woolard et al. 2004).

2.2.4.5. Alternative splicing in EMT

Gene expression programs implicated in tumor dissemination are the same as those participating in embryonic development and wound healing; one of these is EMT, a process in which cells dedifferentiate, lose their epithelial phenotype and acquire mesenchymal traits. In this context, splicing regulation can drive critical aspects of EMT-associated phenotypical changes (Biamonti et al. 2019; Bonomi, Gallo, et al. 2013).

CD44 is a transmembrane glycoprotein involved in cell division, survival and adhesion; inclusion or exclusion of the variable exons 6-10 generates CD44v or CD44s isoforms, respectively (Screaton et al. 1993). Epithelial Splicing Regulatory Protein 1 (ESRP1), hnRNPA1 and SRSF2 promote CD44v splicing, while hnRNP L inhibits it (Chen et al. 2017; Loh et al. 2015; Loh et al. 2014; Brown et al. 2011). Epithelial cells express mainly CD44v, but after undergoing EMT they switch to CD44s, thus enhancing cell migration in breast and ovarian cancer models (Brown et al. 2011). On the other hand, CD44v isoforms are frequently expressed at high levels across different cancer types, making the understanding of its role in cancer still debated (Urbanski, Leclair, and Anczukow 2018). However, CD44v8-10 has been demonstrated to interact with cystine transporter xCT, increasing the levels of reduced glutathione, enhancing the ability of cancer cells to survive after exposure to reactive oxygen species (ROS) damage, thus suggesting a reason for the overexpression of epithelial isoforms in some cancers (Ishimoto et al. 2011).

2.2.4.6. Alternative splicing and cell invasion

The tyrosine kinase receptor RON controls invasion of the extracellular matrix. Skipping of exon 11 results in the production of Δ RON, a constitutively active isoform upregulated in cancers (Ghigna et al. 2005; Zhou et al. 2003). SRSF1, hnRNP H and hnRNP A2 promote skipping, while hnRNP A1 promotes inclusion (Bonomi, di Matteo, et al. 2013; Lefave et al. 2011; Golan-Gerstl et al. 2011; Ghigna et al. 2005). Another example is the Rac Family Small GTPase 1 (Rac1), a Rho GTPase involved in cell motility and proliferation. Inclusion of exon 3b, an event promoted by SRSF1, produces the constitutively active isoform Rac1b with impaired GTP hydrolysis (Goncalves et al. 2014; Beausoleil et al. 2009; Singh et al. 2004; Fiegen et al. 2004; Matos, Collard, and Jordan 2003). Rac1b is expressed in breast, thyroid, colorectal and lung tumors (Faria et al. 2016; Goncalves et al. 2014; Zhou et al. 2013; Schnelzer et al. 2000).

2.2.4.7. *Alternative splicing in cancer drug resistance*

Alternative splicing is also one of the mechanisms enabling drug resistance to cancer cells. For instance, the human epidermal growth factor receptor 2 (HER2), frequently amplified or overexpressed in breast tumors, possesses the constitutively active isoform $\Delta 16\text{HER2}$, generated through skipping of exon 20 (Wada, Yagihashi, and Naito 2016; Castiglioni et al. 2006; Kwong and Hung 1998). This event is stimulated by SRSF3 and is prevented by hnRNP H1 (Gautrey et al. 2015). Expression of $\Delta 16\text{HER2}$ enhances cell proliferation and invasion, and decreases sensitivity to the HER-2-targeting antibody Trastuzumab, upregulating Bcl-2 and activating the Sarcoma-family kinase (SRC) (Tilio et al. 2016; Huynh and Jones 2014; Alajati et al. 2013; Cittelly et al. 2010; Mitra et al. 2009; Castiglioni et al. 2006).

2.2.4.8. *Alternative splicing and signaling*

Despite the increasing knowledge in the AS mechanism, the upstream signaling pathways that determine the expression of the different isoforms are still poorly characterized.

A known example of this regulation is the well characterized Wntless-related integration site (WNT)/ β -catenin pathway, whose activation leads to the transcriptional upregulation of PTBP1; moreover, the glycogen synthase kinase 3 (GSK-3 β), which is part of the same cascade, was found to phosphorylate the splicing factors SRSF2 and polypyrimidine tract-binding protein-associated splicing factor (PSF) (Aguilera et al. 2016; Heyd and Lynch 2010; Hernandez et al. 2004). Additionally, during Transforming growth factor β (TGF- β)-induced EMT, ZEB1/2 transcription factors repress the expression of ESRPs, affecting CD44, FGFR and p120 catenin splicing (Horiguchi et al. 2012).

Alterations in signaling pathway regulation that occur during tumorigenesis can have a strong impact on alternative splicing decisions, modulating the proteome of the cell. An example is given by the over-activation of the mitogen-activated protein (MAP) kinase pathway RAS/RAF/MEK/ERK, a frequent event in epithelial cancers that leads to activation of the transcription factors ETS Like-1 (ELK1) and Master Regulator of Cell Cycle Entry and Proliferative Metabolism (MYC), which in turn induces PTBP1 expression. Increased PTBP1 levels influence the alternative splicing of RAC1, NUMB and PKM (Hollander et al. 2016). Moreover, ERK activation leads to Sam68 phosphorylation in colorectal cancer cells, which increases its binding to the SRSF1 transcript, causing an intron retention event that results in an enhanced expression of the protein (Valacca et al. 2010). Another example is given by the RAC-alpha serine/threonine-protein kinase (AKT), which was found to

phosphorylate SRSF1 in lung cancer cells, promoting thus skipping of the exon cassette 3-6 in the caspase-9 transcript and generating the antiapoptotic isoform (Shultz et al. 2010).

2.2.4.9. Targeting splicing defects in cancer

In recent years, the possibility of splicing modulation as a new type of cancer therapy has been intensively investigated. Natural compounds deriving from bacteria, such as pladienolides, herboxidienes and spliceostatins, directly bind the Splicing factor 3B (SF3b) complex of the U2 snRNP, interfering with spliceosome assembly. These compounds lack chemical stability, thus several derivatives were generated, like E7107, spliceostatin A and sudemycins (Bates et al. 2017; Effenberger, Urabe, and Jurica 2017). E7107 was tested in solid tumors in phase I clinical trials, but unfortunately some episodes of optic neuritis as collateral effects led to its suspension; further studies and refinements are thus required to understand the reasons behind the toxicity (Eskens et al. 2013). Clinical trials are currently ongoing for an orally available compound, H3B-8800, which modulates SF3b in hematopoietic malignancies with mutations in splicing factors (Seiler et al. 2018).

Interestingly, cancers with certain molecular characteristics are more sensitive to drugs targeting the spliceosome; this is the case of MYC-driven cancers since they are characterized by an increased transcription and thus enhanced splicing activity, or cancer cells bearing core spliceosomal mutation, which are preferentially killed by these drugs through a synthetic lethality mechanism (Shirai et al. 2017; Obeng et al. 2016; Lee et al. 2016; Lee and Abdel-Wahab 2016; Hsu et al. 2015).

In addition, inhibition of phosphorylation of SR proteins by SRPKs or CLKs kinases has been considered a potential therapeutic target. Preclinical studies using inhibitors of these kinases showed a limited anti-cancer activity, with the exception of T-025, a CLK inhibitor that suppresses cancer growth in tumors that overexpress CLK2 or MYC (Wang et al. 2020; Iwai et al. 2018). Another emerging target is Protein arginine methyltransferase 5 (PRMT5), a methyltransferase that methylates Sm proteins, against which pharmacological inhibitors have been developed. Glioblastoma and chronic myeloid leukemia cells, that rely on PRMT5 activity to repress tumor suppressors, have been shown to be sensitive to PRMT5 KD or inhibition (Smith et al. 2018). Inhibition of PRMT5 methyltransferase activity with the small molecule PJ-68 has been shown to affect chronic myeloid leukemia stem cells, reducing their survival and serial clonal capacity (Jin et al. 2016).

The most commonly used approach to alter splicing relies on antisense oligonucleotides (ASO), which modulate splicing by pairing with cis-acting elements or splice sites. ASO have been approved by the Food and Drug Administration (FDA) in the U.S. for treatment

of monogenic diseases like spinal muscular atrophy and preclinical results have been achieved by targeting oncogenic splice variants (Wang et al. 2020). For example, skipping of Double minute 4 protein (MDM4) exon 6 by ASO reduces protein expression and prevents the growth of melanoma and diffuse large B cell lymphoma, enhancing the response to MAPK inhibitors (Dewaele et al. 2016). Another fine approach used to manipulate splicing is the use of engineered splicing factors (ESFs) formed by sequence-specific RNA binding modules and effector domains (like RS domains). ESFs have been used to shift the splicing of BCL-X toward the proapoptotic BCL-XS isoform (Wang et al. 2009).

3. Aim of the thesis

Previous studies from our laboratory identified myosin VI AS deregulation as a critical step in cancer cell migration, especially in ovarian tumors, however, leaving undefined how the alternative splicing of myosin VI is regulated.

The aim of this study was to analyze the molecular mechanism underlying myosin VI AS in cancer cells in order to identify the signaling pathway and the splicing regulators involved in myosin VI isoform choice. To this end, we set up a suitable alternative splicing cellular model, in which the expression of endogenous myosin VI isoforms can be modulated through specific cell culture growth conditions.

4. Materials and methods

4.1. Buffers

4.1.1. Phosphate-buffered saline (PBS)

NaCl	137 mM
KCl	2.7 mM
Na ₂ HPO ₄	10 mM
KH ₂ PO ₄	2 mM

4.1.2. Tris-buffered saline - Tween (TBS-T)

NaCl	137 mM
KCl	2.7 mM
Tris-HCl pH 7.4	25 mM
Tween-20	0.1%

4.1.3. 10X Sodium Dodecyl Sulphate – PolyAcrylamide Gel Electrophoresis (SDS-PAGE) running buffer

Glycine	192 mM
Tris-HCl pH 8.3	250 mM
SDS	1%

4.1.4. 50X Tris-Acetate-EDTA (TAE)

Tris base	2 M
Acetic acid	1 M
EDTA pH 8	10 mM

4.1.5. 1X Radioimmunoprecipitation assay (RIPA) buffer

Tris-HCl pH 8	50 mM
NaCl	150 mM
NP-40	1%
SDS	0.1%
Deoxycholic acid	1%

500X protease inhibitor cocktail (Calbiochem), 20 mM sodium pyrophosphate pH 7.5, 50 mM sodium fluoride, 2 mM PMSF and 10 mM sodium orthovanadate were added to

the buffer immediately before usage.

4.1.6. Laemmli buffer

SDS	8%
Tris-HCl pH 6.5	200 mM
Glycerol	40%
DTT	400 mM
Bromophenol blue	0.4%

The Laemmli buffer was prepared as a 4X stock solution and stored at -20°C.

4.1.7. Immunofluorescence (IF) buffer

BSA	8%
Triton X-100	200 mM
Tween-20	40%
NaN ₃	400 mM
PBS	1x

The IF buffer was freshly prepared for each immunofluorescence experiment.

4.1.8. RNA immunoprecipitation buffers

4.1.8.1. Low salt lysis buffer

Hepes pH 7.5	50 mM
NaCl	10 mM
Glycerol	10%
NP40	0.2%
Triton X-100	1%
Sodium deoxycolate	0.1%

500X protease inhibitor cocktail (Calbiochem) and 40 U/ml RNase OUT (ThermoFisher) were added to the buffer immediately before usage.

4.1.8.2. Nuclei lysis buffer

Tris-HCl pH 8	10 mM
NaCl	150 mM

Triton X-100	1%
Sodium deoxycholate	0.1%
Tris-HCl pH 8	10 mM
NaCl	150 mM

500X protease inhibitor cocktail (Calbiochem) and 40 U/ml RNase OUT (ThermoFisher) were added to the buffer immediately before usage.

4.1.8.3. Low-salt wash buffer

SDS	0.1%
Triton X-100	1%
EDTA	2 mM
Tris-HCl pH 8	20 mM
NaCl	150 mM

500X protease inhibitor cocktail (Calbiochem) and 40 U/ml RNase OUT (ThermoFisher) were added to the buffer immediately before usage.

4.1.8.4. High-salt wash buffer

SDS	0.1%
Triton X-100	1%
EDTA	2 mM
Tris-HCl pH 8	20 mM
NaCl	500 mM

500X protease inhibitor cocktail (Calbiochem) and 40 U/ml RNase OUT from (Thermofisher) were added to the buffer immediately before usage.

4.1.8.5. LiCl wash buffer

LiCl	0,25 M
NP40	1%
Sodium deoxycholate	1%
Tris-HCl pH 8	10 mM
EDTA	1 mM

500X protease inhibitor cocktail (Calbiochem) and 40 U/ml RNase OUT (ThermoFisher) were added to the buffer immediately before usage.

4.1.8.6. Elution buffer

Tris-HCl pH 8	10 mM
EDTA	1 mM
SDS	1%

40 U/ml RNase OUT (ThermoFisher) was added to the buffer immediately before usage.

4.2. Reagents

4.2.1. siRNA

Transient KD was performed with the following siRNAs (Thermo Fisher Scientific):

E-cadherin: sense: ACACUGCCAACUGGCUGGAGAUUAA
 antisense: UUAUUCUCCAGCCAGUUGGCAGUGU

ZO-1: sense: GGCAAGAGAAGAACCAGAUUUUUAU
 antisense: AUAAAUAUCUGGUUCUUCUCUUGCC

PTBP1: sense: UGAACGGGCACAAGCUGCACGGGAA
 antisense: UUCCCGUGCAGCUUGUGCCCGUUCA

RBFOX2: sense: GGAGCUGUAUAUGGUCCGGAGUUUAU
 antisense: AUAACUCCGGACCAUAUACAGCUCC

RBFOX3: sense: UGGGACGAUCGUAGAGGGACGGAAA
 antisense: UUCCGUCCCUCUACGAUCGUCCCA

SRSF2: sense: UCGAAGUCUCGGUCCCGCACUCGUU
 antisense: AACGAGUGCGGGACCGAGACUUCGA

KHSRP: sense: CCGGAAUGAGUACGGAUCUCGGAUU
 antisense: AAUCCGAGAUCCGUACUCAUUCGG

SRSF1: sense: GGACAUUGAGGACGUGUUCUACAAA
 antisense: UUUGUAGAACACGUCCUCAUGUCC

SRSF2 #2: sense: UCGCCUUCGUUCGCUUUCACGACAA
 antisense: UUGUCGUGAAAGCGAACGAAGGCGA

SRSF3: sense: UCGUCGCCCUCGAGAUGAUUAUCGU
 antisense: ACGAUAAUCAUCUCGAGGGCGACGA

SRSF4: sense: GGCCGAGAUAAAUAUGGCCCUCCUA
 antisense: UAGGAGGGCCAUAUUUAUCUCGGCC

SRSF5: sense: GAUAUGGACGGAUAAAGAGAUAUUGA
antisense: UCAAUAUCUCUUAUCCGUCCAUAUC

SRSF6: sense: GGCACAGAAAUAAAUGGCAGAAAUA
antisense: UAUUUCUGCCAUUUAUUUCUGUGCC

SRSF7: sense: CCCGACGUCCCUUGAUCCAAAUGA
antisense: UCAUUUGGAUCAAAAGGGACGUCGGG

SRSF8: sense: GGAUCUCACUACAGCUCAUCUGGUU
antisense: AACCAGAUGAGCUGUAGUGAGAUC

SRSF9: sense: GAGAAGGACUUGGAGGACCUGUUCU
antisense: AGAACAGGUCCUCCAAGUCCUUCUC

SRSF11: sense: UGGUGAUUCACUAAGAGAAUCCAAA
antisense: UUUGGAUUCUCUACUGAAUCACCA

SRSF12: sense: CCAGAUCUCGAGUUAUCGUCAUAA
antisense: UUAUGACGAUAACUUCGAGAUCUGG

For FUS depletion, the following oligo (Dharmacon) was used:

GAUCAAUCCUCCAUGAGUA

For SRSF10 depletion, the following oligos (Ambion) were used:

SRSF10: sense: GUCUUUUGAUUACAACUAUTT
antisense: AUAGUUCUAAUCAAAAAGACCG

4.2.2. Primary antibodies

Protein target	Species	Supplier	Code	IB	IF	IHC
GFP	Rabbit	Sigma	G1544	1:5000		
Myosin VI	Rabbit	Wollscheid et al. 2016	1296	1:2000		1:500
Myosin VI _{long}	Rabbit	Wollscheid et al. 2016	L30	1:500		1:400
Myosin VI _{short}	Rabbit	Wollscheid et al. 2016	S007			1:200
E-cadherin	Mouse	BD Biosciences	610181	1:1000	1:400- 1:100*	
ZO-1	Rabbit	Thermo Fisher	40-2200	1: 1000	1:250- 1:100*	
β-catenin	Mouse	SCBT	E5		1:400	

active β -catenin	Mouse	Merck-Millipore	8E7		1:200	
RBFOX2	Rabbit	Bethyl Laboratories	A300-864A	1:2000		
SRSF2	Mouse	Gift from J. Stevenin	1SC-4F11	1:3000		
H3	Rabbit	Abcam	ab1791	1:1000		
GAPDH	Mouse	Santa Cruz Biotechnology	sc-32233	1:1000		
PAR3	Rabbit	Millipore	07-330	1:1000		
FUS	Rabbit	Gift from A. Bachi		1:2000		
RBFOX3	Mouse	Millipore	MAB377	1:500		
KHSRP	Rabbit	Cell Signaling	E2E2U	1:1000		
vimentin	Mouse	BD Biosciences	550513	1:1000		
tubulin	Mouse	SIGMA	T5168	1:4000		

*the indicated dilution was used for immunofluorescence of cysts.

4.2.3. Secondary antibodies

Protein target	Species	Supplier	Code	IB	IF
anti-mouse IgG HRP	Goat	Bio-Rad	1721011	1:10000	
anti-rabbit IgG HRP	Goat	Bio-Rad	1706515	1:10000	
anti-rabbit Alexa488	Donkey	Thermo Fisher	A-21206		1-100*
anti-mouse Alexa488	Donkey	Thermo Fisher	A21202		1:400-1:100*
anti-rabbit Cy3	Donkey	Jackson Lab	711-165-152		1:800-1:100*
anti-mouse Cy3	Donkey	Jackson Lab	715-165-150		1:100*

4.3. Cloning techniques

4.3.1. Agarose gel electrophoresis

DNA samples were loaded on 0.8%-2% agarose gels along with DNA markers (1 kb or 100 bp DNA Ladder, NEB). Gels were made in TAE buffer containing Syber Safe, 1:10000 (Invitrogen), according to the manufacturer's instructions, and run at 100-120 V until desired separation was achieved.

4.3.2. Transformation of competent cells

Competent TOP10 (Invitrogen) cells were incubated with DNA on ice for 30 minutes and then subjected to a heat shock for 30 seconds at 42°C. The cells were re-incubated on ice for additional 5 minutes and, then, 300 µl of SOC medium were added and the cells were incubated at 37°C for 1 hour before plating them onto agar plates with appropriate antibiotics. The plates were incubated overnight at 37°C.

4.3.3. Minipreps

Individual colonies were used to inoculate 10 ml LB (containing the appropriate antibiotics) and grown overnight at 37°C. The suspensions were centrifuged for 10 minutes at 4000 rpm at 4°C. Minipreps were performed with the Wizard Plus SV Minipreps Kit (Promega) following the manufacturer's instructions. The plasmids were eluted in 50 µl nuclease-free water.

4.3.4. Large scale plasmid preparation

Cells containing transfected DNA were expanded into 200 ml cultures overnight. Plasmid DNA was isolated from these cells using the Macherey-Nagel Maxi-prep kit according to the manufacturer's instructions.

4.3.5. Constructs and plasmids

The pEGFP-RBFOX2 vector was generated by subcloning the RBFOX2 coding sequence, obtained by HindIII/BamHI double digestion of the pFLAG-CMV4-RBFOX2 vector (kindly provided by Franco Pagani), into the pEGFP vector.

The pEGFP-RBFOX3 vector was generated by subcloning the RBFOX3 coding sequence, obtained by HindIII/KpnI double digestion of the pFLAG-CMV4-RBFOX3 vector (kindly provided by Franco Pagani), into the pEGFP vector.

The pLVX-EGFP, pLVX-EGFP-RBFOX2 and pLVX-EGFP-RBFOX3 vectors were generated starting from pEGFP, pEGFP-RBFOX2 and pEGFP-RBFOX3 vectors, respectively.

The EGFP, EGFP-RBFOX2 or EGFP-RBFOX3 DNA sequences were PCR amplified using the following oligos:

Forward: CTCAAGCTTCGAATTCATGGTGAGCAAGGGCGAG

Reverse: TAGAGTCGCGGGATCCATCAGTTATCTAGATCCGGTGG

and then cloned into EcoRI/BamHI-linized pLVX vectors using the InFusion system (Takara Clontech).

The pLVX-EGFP-PTBP1 vector was generated starting from the pCMV-Myc-PTBP1 vector (Addgene Plasmid #23024). The PTBP1 DNA sequence was PCR amplified using the following oligos:

Forward: CGGTACCGCGGGCCCGGATGGACGGCATTGTCCCAGA

Reverse: GATCCGGTGGATCCCCTAGATGGTGGACTTGGAGAAGGA

and then cloned into the SmaI-linized pLVX-EGFP vector using the InFusion system.

The pLVX-FUS vector was generated starting from the pcDNA1-FUS vector (kindly provided by Pier Paolo di Fiore). The FUS DNA sequence was PCR amplified using the following oligos:

Forward: CGGTACCGCGGGCCCGGATGGCCTCAAACGATTATACCC

Reverse: GATCCGGTGGATCCCTTAATACGGCCTCTCCCTGCG

and then cloned into the SmaI-linized pLVX-EGFP vector using the InFusion system.

The pLVX-EGFP-SRSF2 vector was generated starting from the pcDNA3.1-SC35-cMyc vector (Addgene #44721). The SRSF2 DNA sequence was PCR amplified using the following oligos:

Forward: CGGTACCGCGGGCCCGGATGAGCTACGGCCGCCCCCCTC

Reverse: GATCCGGTGGATCCCTTAAGAGGACACCGCTCCTTCCTTTCAGG

and then cloned into the SmaI-linized pLVX-EGFP vector using the InFusion system.

4.4. Cell lines

4.4.1. Commercial cell lines

A549 cells were cultured in Dulbecco's Modified Eagle Medium (DMEM) containing 10% FBS and 2 mM L-Glutamine.

Caco-2 cells were maintained in Dulbecco's modified MEM, supplemented with 20% FBS S.A., 0.1 mM non-essential amino acids, and 2 mM L-glutamine.

OVCAR-5 cells were cultured in Roswell Park Memorial Institute (RPMI) 1640 medium containing 10% FBS N.A. and 2 mM L-Glutamine.

SKOV-3 cells were maintained in Mc COY's 5A medium supplemented with 10% FBS S.A.

4.4.2. Generated cell lines

A549 and Caco-2 cells stably expressing an shRNA against PAR3 were generated through lentiviral infections with the pLKO-shPAR3 vector (kindly provided by Pier Paolo di Fiore).

A549 and Caco-2 cells stably expressing RBFOX2 were generated through lentiviral infection with the pLVX-EGFP-RBFOX2 vector.

A549 cells stably expressing FUS, PTBP1, RBFOX3 or SRSF2 were generated through lentiviral infections with the corresponding pLVX-EGFP- vectors.

After the infection, cells were selected with puromycin for three days and the resulting bulk population was used for further assessment.

The A549 mouse E-cadherin cell line was generated through lentiviral infection with the mE-cadherin-GFP/sh-hE-Cadherin vector kindly provided by Giorgio Scita. GFP- positive cells were sorted by Fluorescence-activated cell sorting (FACS).

4.5. Cell assays

4.5.1. Myosin VI isoform switch assay

Cultured cells were maintained at low confluency (~50%) and collected as sparsely growing cells or plated in 6-well plates with 500 000 cells/well. A549 and Caco-2 cells reached confluency on the third day while OVCAR-5 cells reached confluency of the fifth day; they were collected or kept confluent for additional days. In the disaggregation assay, cells were kept confluent for five days, then detached, plated and maintained at low density (500.000 cells in a 10-cm plate).

4.5.2. Small interfering RNA (siRNA) transfection

For transient KD, cells were transfected with RNAiMax (Thermo Fisher Scientific) according to the manufacturer's instructions. siRNAs were used at 5-20 nM concentration, depending on the target expression and oligo efficiency. Cells were transfected twice, first in suspension and the day after in adhesion. The third day, sparsely growing cells were diluted to maintain low confluency and harvested the day after. Cells plated for confluency were seeded in 6-well plates during the first transfection and two days later, when they reached confluency, they were harvested at different time points.

4.5.3. Wound healing assay

600 000 cells/well were transfected with RNAiMax (Thermo Fisher Scientific) in suspension according to the manufacturer's instructions and seeded in a 12-wells plate. The next day cells were transfected in adhesion and they reached confluency the third day. The fourth day, the wound was obtained by scratching the cell monolayers with the sterile tip of a pipette. The multiwell was then transferred into the incubator chamber of the microscope for the acquisition of time-lapse digital frames every 5 minutes for 24 hours.

Live-imaging was performed using a Nikon automatic microscope equipped with closed heating and CO₂ perfusion devices. The area covered by the cells over time was calculated with MatLab software and is expressed in $\mu\text{m}^2/\text{s}$.

At the end of the time lapse, cells were harvested for RNA extraction in order to assess the efficiency of KD by quantitative polymerase chain reaction (qPCR) and the expression of myosin VI isoforms by reverse transcriptase polymerase chain reaction (RT-PCR), as described in the next sections.

4.5.4. DNA transduction

For the generation of cell lines that stably express exogenous DNA, DNA transduction using lentiviruses was performed.

For the generation of lentiviral particles, HEK-293T cells were transfected with 10 μg of the transfer DNA together with plasmids encoding GAG, POL, ENV (VSVG), and REV retroviral proteins using the calcium phosphate transfection method. The DNA plasmids and 240 mM CaCl₂ were mixed with HBS solution in order to generate DNA/CaPO₄ complexes. One day after the transfection, the medium was replaced with fresh medium in order to concentrate the virus. On the third day, the viral supernatant was collected, passed through PVDF 0.45 μm Millipore filters, and used to infect target cells after adding 8 $\mu\text{g}/\text{ml}$ polybrene. In order to select cells that had stably integrated the exogenous DNA, two days after the infection, puromycin was added to the medium for 72 hours (1.5 $\mu\text{g}/\text{ml}$ for A549 cells, 6 $\mu\text{g}/\text{ml}$ for Caco-2 cells).

4.5.5. Caco-2 cysts formation

Caco-2 single-cell suspensions were embedded in 50% Matrigel (Corning) and 1 mg/mL collagen mixture, and plated on a 8-well ibidi chamber (Cat.No: 80826) at a density of 9×10^3 cells/well. The suspension was allowed to solidify for 30 minutes at 37°C, and overlaid with 200 μl of medium. The medium was changed every two days. On the sixth day, cholera toxin

(Sigma Cat.No:C8052) was added at a concentration of 0,1 µg/ml and cysts were collected or fixed the day after.

For the generation of E-cadherin KD cysts, Caco-2 cells were transiently transfected with E-cadherin oligos. Cells were transfected twice, first in suspension and the day after in adhesion. Two days after the siRNA transfection, single cell suspensions were embedded as previously described.

4.5.6. Immunofluorescence

For IF analysis, cells plated on coverslips were rinsed twice with 1x PBS and fixed with 4% paraformaldehyde for 10 minutes at room temperature, or in 100% methanol at -20°C for 10 minutes (in the case of active β -catenin staining). After fixation, the samples were rinsed twice with 1x PBS and permeabilized at room temperature (RT) with 0.5% Triton-X100 in 1x PBS for 15 minutes. In order to reduce aspecific antibody interaction, blocking was performed with 2% BSA in PBS for 1 hour. The incubation of the primary antibodies was performed overnight at 4°C in 1x PBS containing 0.1% Triton X-100 and 1% BSA in order to increase antibody penetration in cell monolayers. After two washes in 1x PBS, the samples were incubated with secondary antibodies conjugated with fluorophores for 1 hour at RT. The samples were then washed twice with 1x PBS and incubated for 10 minutes at RT with DAPI (Sigma-Aldrich, cat. D9542), diluted 1:5000 in 1X PBS, in order to stain the cell nuclei. The coverslips were mounted in a glycerol solution (20% glycerol, 50 mM Tris pH 8.4) to avoid mechanical deformation of the sample. Images were captured using a Leica inverted SP2 microscope with a laser scanning confocal system.

For IF analysis of Caco-2 cysts, cysts grown in matrigel in ibidi chambers were fixed with 4% PFA for 25 minutes at RT, washed three times in PBS-glycine (0.75%), once in 1x PBS and permeabilized with 0,5% Triton X-100 for 20 minutes at RT. Blocking was performed with 10% goat serum in IF buffer for 1 hour at RT. The samples were incubated overnight in appropriate primary antibody dilutions in IF buffer at RT, then washed three times in IF buffer and incubated at RT for 1 hour with secondary antibodies diluted 1:100 in IF buffer. The samples were then washed again three times with IF buffer and incubated with DAPI for 15 minutes. Finally, cysts were washed in 1x PBS and post-fixed in 4% PFA for 10 minutes. Images were captured using a Leica inverted SP2 microscope with a laser scanning confocal system.

4.6. Isoform detection by RT-PCR

The expression of myosin VI isoforms in cell lines was assessed by RT-PCR. RNA from cultured cells or cysts was isolated with Maxwell RSC instrument (Promega) according to the manufacturer's protocols. Concentration and purity were determined by measuring optical density at 260 nm and 280 nm using a Nanodrop spectrophotometer. cDNA was generated from 200-500 ng of RNA in 20 µl reactions using the Applied Biosystems™ High-Capacity cDNA Reverse Transcription Kit (Thermo Fisher Scientific). The obtained cDNA was used as a template for PCR reactions with primers flanking the alternatively spliced region:

Forward (Exon 28): CCGAGCTCATCAGTGATGAGGC

Reverse (Exon 33): CCAAGCATGATACTTTTAGTCTCC

The equivalent of 50 ng of RNA was used for each reaction.

PCR reactions were performed using GoTaq® DNA Polymerase (Promega) as follows:

	ul
5x buffer	10
dNTPs 10 mM	1
Forward (10 uM)	2.5
Reverse (10 uM)	2.5
template	2-5
GoTaq	0,5
H ₂ O	up to 50

with the thermocycler parameters:

95 °C	5'	Denaturation
95 °C	45s	30 cycles
55 °C	45s	
72 °C	30s	
72 °C	10'	Extension
4 °C		

The resulting amplicons were loaded on 2.5% agarose gels along with 100 bp DNA Ladder. Gels were made in TAE buffer containing Syber Safe, 1:10000 (Invitrogen), according to the manufacturer's instructions, and run at 100-120 V until desired separation was achieved. For the expression of NUMB isoforms, the same protocol was used with the following PCR primers:

Forward (Exon 8): TGCTCCGATGACCAAACCAG

Reverse (Exon 10): CACCTCTTCTAACCATCGGTC

4.7. qPCR

For qPCR analysis, samples were analyzed by the Real Time PCR facility using TaqMan probes. cDNA was prepared as previously described, and reactions were performed in triplicates, using the equivalent of 5 ng RNA for each one as a template. The amplicon expression in each sample was normalized against the RPLP0 mRNA content.

The following Taqman assays (ThermoFisher) were used:

NOVA1:Hs01105684_m1; PTBP1:Hs00243060_m1; FUS:hs01100224_m1;
RBFOX2:Hs00204814_m1; RBFOX3:Hs01370654_m1; KHSRP:hs01100863_g1;
ELAVL1:Hs00171309_m1; MBNL2:Hs00255065_m1; KHDRBS3:Hs00938827_m1;
RBM10:Hs01016417_g1; RBM17:hs00998713_m1; RBM38:Hs00955734_g1;
RBM41:Hs01547530_m1; SRSF2:Hs00427515_g1; SRSF3:Hs00751507_s1;
SRSF7:Hs01032696_g1; 18S:Hs99999901_s1; RPLP0:hs99999902_m1;
SRSF1:hs00199471_m1; SRSF4:Hs00900675_m1; SRSF5:Hs00199477_m1;
SRSF6:Hs05331162_g1; SRSF8:Hs00259455_s1; SRSF9:Hs01596548_g1;
SRSF10:Hs00986049_g1; SRSF11:Hs00951159_m1; SRSF12:Hs00365534_m1.

4.8. SDS-Polyacrylamide Gel-Electrophoresis (SDS-PAGE) and Immunoblot Analysis (IB)

Cells were washed with 1x PBS, pelleted and either frozen at -80 C or directly processed. Cell pellets were lysed in RIPA buffer, sonicated with Bioruptor Plus (low intensity, 20s ON, 20s OFF, 5 cycles), and incubated for 10 minutes on ice. Lysates were cleared by centrifugation at 16 000 g for 20 minutes at 4°C. Protein concentration was measured by the Bradford assay (Biorad) following the manufacturer's instructions.

The samples were denatured by adding 4X Laemmli buffer and by heating at 95°C for 10 minutes. Proteins were then separated on precast gradient gels (4–20% TGX precast gel, Bio-Rad) by SDS-PAGE and transferred onto nitrocellulose membranes by Transblot Turbo (BIO-RAD). Ponceau staining was used to determine the efficiency of the protein transfer onto the membrane. Membranes were blocked for 1 hour (or overnight) in 5% milk in TBS supplemented with 0.1% Tween (TBS-T). After blocking, the membranes were incubated with

the primary antibody, diluted in TBS-T 5% milk, for 1 hour at RT, or overnight at 4°C. After several washes, the membranes were incubated with the appropriate horseradish peroxidase (HRP)-conjugated anti-mouse or anti-rabbit secondary antibodies (GE Healthcare) diluted

in 1% TBS-T milk for 1 hour at RT. After a number of washes, the bound secondary antibody was detected with ECL (GE Healthcare). Western blots were visualized using films (GE Healthcare) or Chemidoc (Bio-Rad).

4.9. UV-RIP: RNA immunoprecipitation after UV crosslinking

Adapted from Enhancer RNAs: Methods and Protocols, Methods in Molecular Biology, vol. 1468 (Schaukowitch et al. 2014).

4.9.1. Lysate preparation

About 50-60 million cells for each condition were washed with PBS and crosslinked on plates, in ice-cold PBS at 150 mJ/cm² using a 254 nm Ultraviolet Crosslinker. The cells were then collected and spun down in a benchtop centrifuge at 2000 rpm for 5 minutes at 4°C. The PBS was removed and cell pellets were resuspended in ice-cold low-salt lysis buffer (5 ml for ~50 M cells) and incubated on a rotating platform at 4°C for 10 minutes to allow cell lysis while preserving the nuclei. The suspensions were centrifuged at 2000 rpm for 10 minutes at 4°C, the supernatants were removed and the nuclei pellets were lysed in 1 ml of nuclei lysis buffer. In order to eliminate genomic DNA, 500 U/ml of DNase I (Roche, cat. no. 04 536 282 001) were added to the lysates, together with MgCl₂ and CaCl₂ cofactors to a final concentration of 10 mM and 1 mM, respectively. The lysates were incubated on a rotating wheel for 15 minutes at RT and another 45 minutes at 4°C, and were then cleared by centrifugation at 15 000 rpm for 10 minutes at 4°C.

Cleared lysates were transferred into new tubes and the DNase was inactivated by the addition of 4 mM EDTA pH 8 and 2 mM EGTA pH 8. The lysates were diluted 1:2 in nuclei lysis buffer and the protein concentration was determined using Bradford assay (Biorad) according to the manufacturer's instructions.

4.9.2. Immunoprecipitation

RBFOX2 was immunoprecipitated by incubating 1 mg of nuclear lysate with anti-RBFOX2 (Bethyl) or IgG rabbit (Sigma) as a negative control on the rotating wheel overnight at 4°C. 50 µg of lysate (1/20) were saved as Input for RNA extraction, and 20-50 µg of lysate were saved as protein Input. The next day, protein G beads were washed twice with 1x PBS, equilibrated in nuclei lysis buffer and added to the lysates, followed by incubation for 2 hours at 4°C to allow the formation of immune-complexes.

Precipitated complexes were washed as follows: twice with low-salt wash buffer, twice with high-salt wash buffer, twice with LiCl wash buffer, and once with TE. Before discarding the

last wash buffer, 1/10 of the beads were saved for protein analysis: the proteins were eluted by incubation in Laemmli buffer for 10 minutes at 95°C and loaded on a precast gradient gel (4–20% TGX precast gel, Bio-Rad) together with Input and supernatants, and analyzed by IB. For RBFOX2 protein detection, which has the same molecular weight as heavy-chain IgG-fragments (approx. 50 kDa), Clean-Blot™ IP Detection Reagent (HRP) was used (ThermoFisher).

4.9.3. RNA isolation

For protein-RNA complexes elution, 9/10 of the beads were incubated in elution buffer at 65 °C for 10 minutes at 1000 rpm to keep the beads in motion. The beads were spun down and the supernatants was transferred into new tubes; the elution was then repeated for a second time. Input samples were diluted in elution buffer as well. In order to release the immunoprecipitated RNA, after the addition of 7 µl of Proteinase K (20 mg/mL, Roche), the samples were incubated for 2 hours at 50°C in order to allow protein degradation.

RNA was then isolated by phenol-chloroform extraction with ethanol precipitation at -80°C overnight. The next day, the RNA pellets were washed and resuspended in RNase-free water. Concentration and purity were determined by measuring the optical density at 260 nm and 280 nm using a Nanodrop spectrophotometer.

4.9.4. RT-qPCR

RNA samples (8 µl) were treated with DNase (RQ1 Promega) according to the manufacturer's instructions and retro-transcribed (in 20 µl reactions) using the Applied Biosystems™ High-Capacity cDNA Reverse Transcription Kit (Thermo Fisher Scientific). An equivalent of about 100 ng of RNA (based on Input RNA quantification) was used in PCR reactions performed as previously described (see *Isoforms detection by RT-PCR*), using primers specific for myosin VI pre-mRNA that amplify a fragment spanning intron 31-exon 32 junction:

Forward: CCATTCTTAATCTCATCTGTCAAAGAAAC

Reverse: AGCAGGACCTCTGTAATTATTACAC

For qPCR analysis, samples were analyzed by the Real Time PCR facility using TaqMan probes. The reactions were performed in duplicates, using an equivalent of about 50 ng of RNA as cDNA template.

The following TaqMan probe, spanning exon32-intron32 junction, was used to assess RBFOX2/myosin VI pre-mRNA binding by percent input method:

Exon 32- Intron 32

Forward: GGAAATATGCAGAACTACGTGATACCA

Reverse: GGTATTTATCACCTATCACTATGAACATTTTCTATTTTGA

Probe: CCAAACACTTACCACAAGAAG

The following TaqMan probes, located in introns 29 and 31, were used to assess RBFOX2 binding to its UGCAUG motifs by calculating fold enrichment. As a negative control, a TaqMan probe targeting an upstream unrelated region (spanning exon 25-intron 25 junction) was also used.

Intron 29 (FOX2 site1)

Forward: ACAAACCTCACAATATATGAATTTTCTAATCTAGATGTAAT

Reverse: TTTGTAATCAACATGAGATCACTCAGAATTTTAATATATTTTAA

Probe: TCACTGCATGCCTTATTT

Intron 31 (FOX2 site2)

Forward: TGTTTTTCTAGTAGCTGTTTCCGGTTT

Reverse: CCAGAAACATGAAGAAAACATTGACTAAAGA

Probe: AAGAAAGCATGCATAAGTTT

Exon 25- Intron 25

Forward: ACCCGAGATGAATAAACAGATCAAGAATC

Reverse: CATAAAAAGTCAAAGTCATTCGGGTTGA

Probe: TTGATGGCCAAAATTAAGGTATG

4.9.4.1. Percent Input calculation

For each immunoprecipitated RNA fraction, the percent input was calculated as follows:

RNA Input: 1/20

Step 1: Adjusted input to 100%= Ct Input – log₂ 20

Step 2: Percent input= 100*2^{(Adjusted input - Ct (IP))}

4.9.4.2. Fold enrichment calculation

For each immunoprecipitated RNA fraction, the fold enrichment was calculated as follows:

Step 1: Nonspecific adjustment= (Ct IP) - (Ct mock)

Step 2: Fold enrichment = (2^{-DDCt})

In this way, the negative control (mock) is set to 1.

4.10. Affymetrix Human Transcriptome Array

RNA from A549 cells grown in different conditions was extracted as previously described, Gene expression was analyzed through Affymetrix GeneChip Human Transcriptome Array

2.0. The array contains more than 7.3 million probes: 70% of which cover exons for coding transcripts and the remaining 30% cover exon-exon splice junctions and non-coding transcripts. The array enables the analysis of both gene expression and alternative splicing variants. For gene expression studies, the fold change of confluent versus sparse cells, and disaggregated versus confluent cells was evaluated.

4.11. Immunohistochemistry (IHC)

The human samples described above were fixed in 4% paraformaldehyde and embedded in paraffin by Leica automatic processor. Then, 4 μm -thick serial sections were dissected using microtome. For IHC analysis on paraffin sections, paraffin was removed with xylene and the sections were rehydrated in graded alcohol. Antigen retrieval was carried out with preheated target retrieval solution (sodium citrate buffer: 10mM sodium citrate, 0.05% Tween 20, pH 6.0) for 45 minutes. Tissue sections were stained with H_2O_2 (DAB Substrate Kit peroxidase, Vector Lab) for quenching of endogenous peroxidases, blocked with FBS in PBS for 60 minutes and incubated overnight with primary antibodies. The antibody binding was detected using a polymer detection kit (GAR-HRP, Microtech) followed by a diaminobenzidine chromogen reaction (Peroxidase substrate kit, DAB, SK-4100; Vector Lab). All sections were counterstained with Mayer's hematoxylin and visualized using a bright-field microscope.

5. Results

5.1. Myosin VI isoform expression correlates with epithelial phenotype

To identify possible cellular models suitable for *in vitro* alternative splicing studies, we screened different cancer cell lines. A myosin VI isoform switch toward the expression of the long isoform occurs during *in vitro* epithelial cell polarization (Biancospino et al. 2019; Wollscheid et al. 2016; Dance et al. 2004; Buss et al. 2001). Thus, we harvested cells grown in sparse conditions or as cell monolayers and analyzed their expression of the myosin VI isoforms through RT-PCR (**Fig. 11A**). Confirming previous observations, we found that MDA-MB-231 and SKOV-3 exclusively express myosin VI_{short} (Wollscheid et al. 2016), while few epithelial cell lines switch on/off their isoform expression based on the culturing conditions (**Fig. 11B**): MCF10A from breast, A549 from lung adenocarcinoma, Caco-2, from colorectal adenocarcinoma, and OVCAR-5 from ovarian adenocarcinoma.

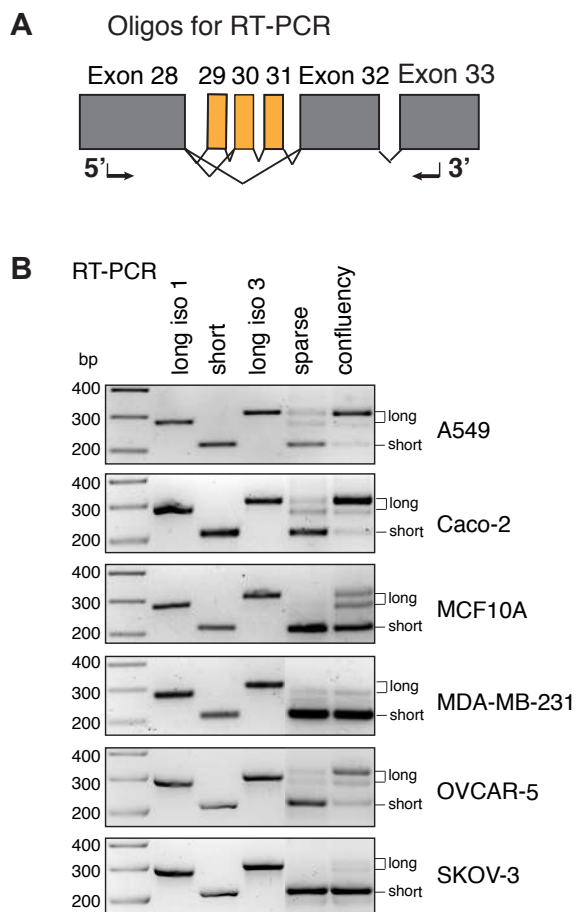


Figure 11. Myosin VI isoform expression in different cell lines.

A. Representation of the region amplified by RT-PCR. Coding exons are represented by boxes and alternatively spliced exons are depicted in orange. Arrows indicate oligonucleotide position. **B.** Cells grown in sparse growth conditions were allowed to reach confluency. RT-PCR results of cDNA prepared from the indicated conditions and cell lines are presented below the scheme. Controls (long iso1, short and long iso3) are from plasmids bearing the region of the different isoforms.

We repeated the experiment with A549 and Caco-2 cells to evaluate the timing of the switch, harvesting the cells one, three and five days after that confluency was reached. Both cell lines start to express myosin VI_{long} isoform already on the first day but they switch to an almost exclusive expression of the myosin VI_{long} isoform on the fifth day (**Fig. 12A**).

Immunofluorescence analysis performed on the third day shows that cell monolayers present mature cell-to-cell contacts as demonstrated by the apical localization of the tight junction (TJ) protein Zonula occludens protein 1 (ZO-1), and the basal localization of the AJ protein E-Cadherin (**Fig. 12B**). Importantly, the isoform conversion is fully reversible, as confluent cells dissociated and re-plated at low density switched back to myosin VI_{short} expression in three or four days (**Fig. 12C**). Importantly, the isoform switch from myosin VI_{long} to VI_{short} mimics the process that likely leads to the positive selection of myosin VI_{short} observed in primary ovarian cancer cells (Wollscheid et al. 2016).

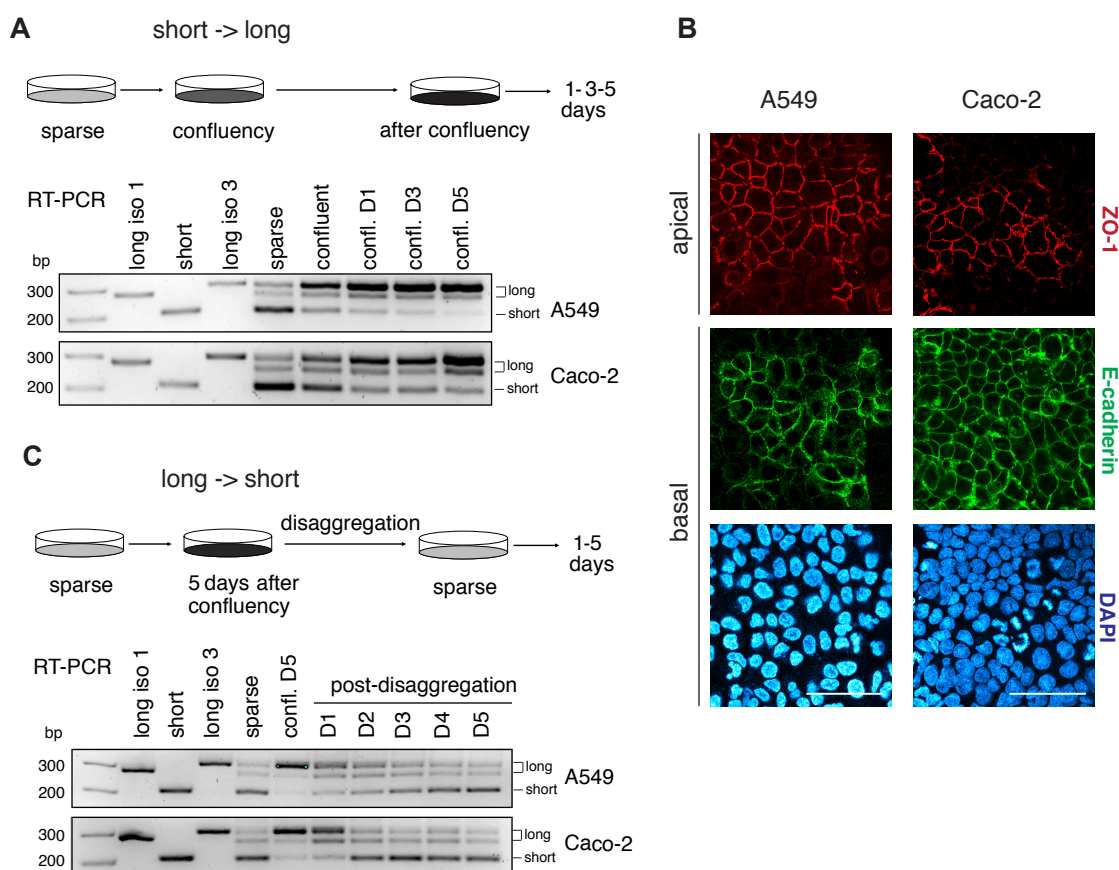


Figure 12. Myosin VI isoform switch. **A.** Cells grown in sparse growth conditions were allowed to reach confluency and kept in culture for one, three or five days. mRNA was extracted and RT-PCR was performed from the indicated conditions and cell lines. Controls (long iso1, short and long iso3) are from plasmids bearing the region of the different isoforms. **B.** IF of A549 and Caco-2 cells maintained in confluency for three days showing the apical localization of ZO-1, the nuclei (DAPI) and E-cadherin basal position. Scale bar: 50 μ m. **C.** Cells grown in sparse growth conditions were allowed to reach full confluency. After five days of growth in confluent conditions, cells were harvested or dissociated and re-plated at low density. mRNA was extracted and RT-PCR was performed from the indicated conditions and cell lines. Controls (long iso1, short and long iso3) are from plasmids bearing the region of the different isoforms.

We further analyzed the same samples by immunoblot, confirming the upregulation of myosin VI_{long} in confluent cells also at protein level, using an isoform-specific antibody generated in our laboratory (**Fig 13**). While Caco-2 cells are largely used as a model system to study cell polarity (Jaffe et al. 2008; Sun et al. 2008), A549 cells are less characterized in this context. A549 cells express low levels of E-cadherin, but they upregulate the protein upon confluency (**Fig. 13**); vimentin, instead, is downregulated in confluent growth conditions (**Fig. 13A**) and its expression rises again after disaggregation (**Fig. 13B**). Thus, our A549 AS switch *in vitro* system could be seen as a proxy of Epithelial to Mesenchymal Transition (EMT), Mesenchymal to Epithelial Transition (MET) phenomenon, where a polarized cell monolayer resembles an epithelial state and a sparse cell culture system resembles a mesenchymal state.

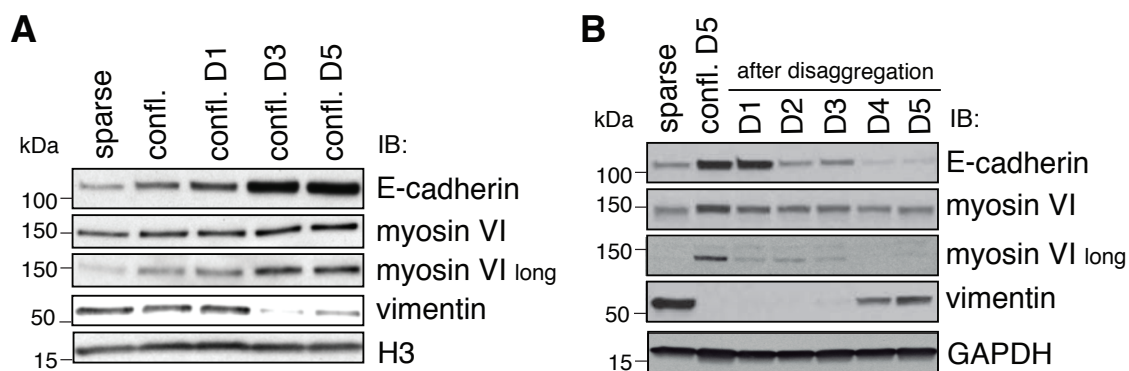


Figure 13. EMT marker modulation in A549 sparse cells versus cell monolayers. **A.** Cells grown in sparse growth conditions were allowed to reach confluency and were then kept in culture for additional days (one, three or five days). Cells were collected and analyzed by IB with the indicated antibodies. **B.** Cells grown in sparse growth conditions were allowed to reach full confluency. After five days of confluency, cells were harvested or dissociated and re-plated at low density. Cells were collected and analyzed by IB with the indicated antibodies.

In order to evaluate myosin VI isoform expression in a well-established epithelial polarized system, we moved from cell monolayer to three-dimensional (3D) cultures. For this purpose, we used Caco-2 intestinal epithelial cells, which have the capacity to form polarized cysts when plated as a single cell suspension embedded in a 3D matrix (Jaffe et al. 2008). Caco-2 single cells undergo subsequent cell divisions in which the apical domain of each cell is maintained at the centre of the growing structure, resulting in a cyst-like structure. After six days, cholera toxin administration speeds up the formation of a fluid-filled central lumen by activating cAMP signalling (Jaffe et al. 2008), (**Fig. 14A**). As can be observed in **Figure**

14B, the Caco-2 cysts present a correct and fully polarized architecture, as shown by the E-cadherin staining at cell-cell contacts and the apical staining of ZO-1. RT-PCR analysis of RNA extracted from fully-formed Caco-2 cysts showed that there is a switch toward myosin VI_{long} during cyst formation (**Fig. 14C**). Thus, using both model systems, cell monolayers and cysts, we are able to modulate an endogenous AS event without genetic manipulation or drug treatment, working close to physiological conditions.

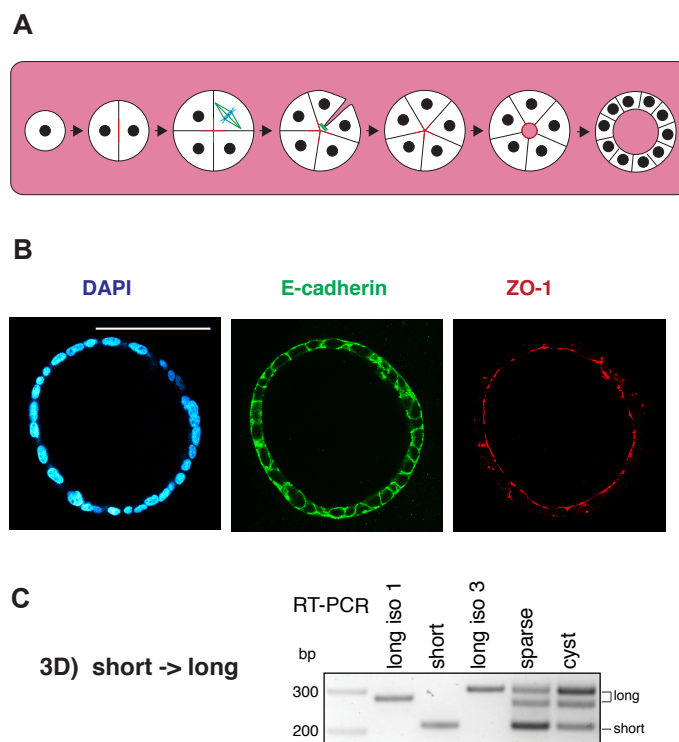


Figure 14. Polarized Caco-2 cysts express myosin VI_{long} isoform. **A.** Schematic representation: Caco-2 cells were cultured for one week in Matrigel to allow cyst formation. **B.** IF of a polarized Caco-2 cyst stained with the indicated markers. Medial sections of confocal images are taken. Scale bar: 100 μ m. **C.** RT-PCR results of cDNA obtained from Caco-2 cells grown in sparse growth conditions or Caco-2 cells cultured for one week in Matrigel.

5.2. Identification of signaling pathways involved in myosin VI alternative splicing

5.2.1. PAR3 and ZO-1 do not affect myosin VI isoform switch

The observation that cells expressing the myosin VI_{short} isoform switch to myosin VI_{long} when grown as monolayers with established cell-cell contacts, or form fully polarized cysts, led us to hypothesize that signaling pathways triggered by cell-cell junctions and cell polarity might be the drivers of myosin VI AS decision.

The establishment of the epithelial architecture and polarity is a cascade of sequential molecular events (Zihni et al. 2016). This process initiates by the establishment of cell-cell contacts mediated by E-cadherin and is followed by the recruitment of adherens- and TJ proteins (such as catenins and ZO-1, respectively) and the Partitioning defective (PAR) complex, formed by PAR3–PAR6–aPKC (**Fig. 15A**). Finally, PAR complex activation leads to atypical protein kinase C (aPKC)-mediated apical differentiation (Zihni et al. 2016). Therefore, we decided to interfere with this process by downregulating the three key players, namely E-cadherin (AJ), ZO-1/2 (TJ), and PAR3 (apical-basal differentiation).

The specific expression of myosin VI_{long} in polarized cells containing microvilli (Biancospino et al. 2019; Wollscheid et al. 2016; Dance et al. 2004; Buss et al. 2001) prompted us to first evaluate the contribution of cell polarity achieved by the PAR Complex. To this end, we generated A549 and Caco-2 cell lines stably expressing an shRNA against PAR3, and we evaluated myosin VI isoform expression in this system. The shRNA construct was very efficient in depleting the protein in both cell lines (**Fig. 15B**); nonetheless, PAR3 depletion did not show any effect on myosin VI AS (**Fig. 15C**), suggesting that the activity of the PAR complex is not the signal triggering myosin VI AS switch.

In order to investigate if TJ formation, which precedes polarity, could be the signal that triggers myosin VI_{long} expression, we performed a simultaneous transient KD of both of the tight junction proteins ZO-1 and ZO-2. Both proteins were efficiently silenced as visible in the immunoblot analysis, IB (**Fig. 15D**) and myosin VI isoform expression was assessed by RT-PCR (**Fig. 15E**). As with PAR3 depletion, we did not score any impairment of myosin VI isoform switch in the double ZO-1/2 KD, since myosin VI_{long} is normally observed in confluent cells (**Fig. 15E**). We concluded that neither the PAR complex nor ZO-1/2 play a major role in myosin VI AS decision.

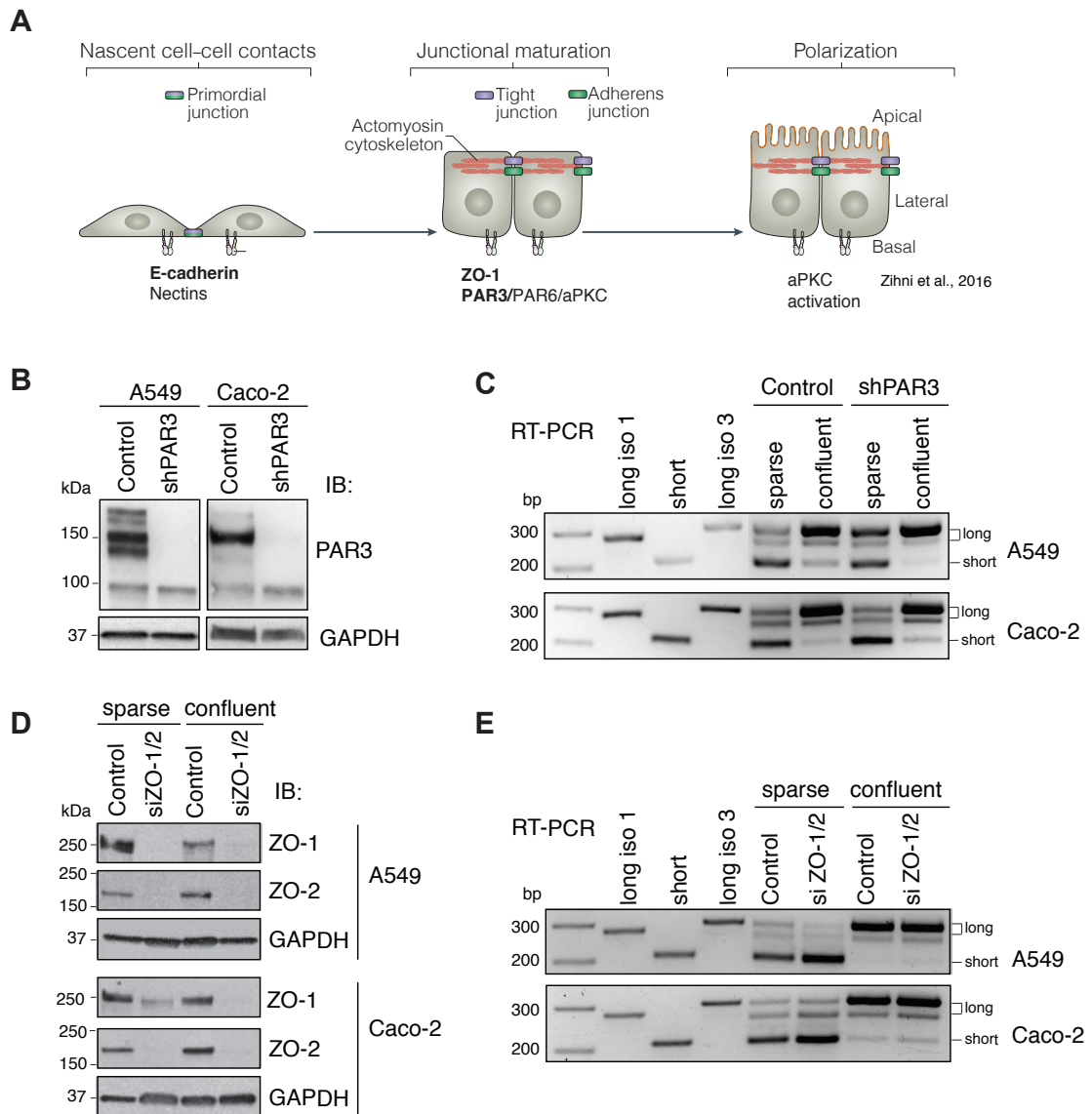


Figure 15. PAR3 and ZO-1 do not affect myosin VI isoform switch. **A.** Epithelial cells form cell–cell junctions by assembling a primordial junction initiated by E-cadherin, leading to the recruitment of TJ proteins like ZO-1. Subsequent increases in the recruitment of signaling proteins induces junctional maturation, with differentiation of TJ and AJ. Finally, polarization is induced by activation of CDC42, which binds to the partitioning defective 3 (PAR3)–PAR6–atypical protein kinase C (aPKC) complex, leading to the activation of the kinase and development of a polarized cell surface. Adapted from *Zihni et al., Nat Rev Mol Cell Biol, 2016*. **B.** IB showing PAR3 depletion in the indicated cell lines. **C.** RT-PCR results of cDNA prepared from A549 or Caco-2 PAR3 KD cells grown in sparse conditions or five days after that confluency was reached. **D.** IB showing ZO-1/2 depletion in the indicated cell lysates upon siRNA transfection. **E.** RT-PCR results of cDNA prepared from A549 or Caco-2 ZO-1/2 KD cells grown in sparse conditions or after that confluency was reached.

5.2.2. E-cadherin-mediated cell-cell contacts are determinant for myosin VI isoform switch

The major player in the initiation of cell-cell contacts is E-cadherin that, together with nectins, mediates the assembly of primordial junctions, allowing the subsequent recruitment of TJ proteins and the PAR complex (Zihni et al. 2016). We then asked whether the signal driving myosin VI AS may be triggered by primordial players. We thus depleted E-cadherin by transient siRNA transfection and evaluated the myosin VI isoform expression in this context in both A549 and Caco-2 cell lines (**Fig. 16 A,B**). We confirmed that A549 cells express low levels of E-cadherin when grown in sparse conditions and up-regulate the protein once they generate a monolayer; while Caco-2 cells express equal levels of E-cadherin in both growth conditions (**Fig. 16A**).

Transient E-cadherin depletion resulted in impairment of myosin VI AS in both A549 and Caco-2 cells, showing a stronger effect one day after that confluency was reached (**Fig. 16B**). Three days later, myosin VI_{long} expression was still impaired in Caco-2 cells while the effect was weakened in A549 cells. We also noticed that prolonged E-cadherin depletion in A549 cells has a strong negative impact on cell viability, complicating possible interpretation of the results obtained at late time points. Overall, these observations suggest that the establishment of primordial E-cadherin-mediated cell-cell contacts is the initial signal required for the switch from myosin VI_{short} to myosin VI_{long} isoform.

E-Cadherin depletion in Caco-2 cells not only disrupted myosin VI AS switch but also generated alterations in the architecture of the cell monolayer (**Fig. 16C**). E-cadherin depletion in Caco-2 cells three days after confluency was reached led to a mislocalization of the TJ marker ZO-1, which was observed in a medial stack where it is normally absent (**Fig. 16D**), in agreement with the role of E-cadherin in driving the correct formation of TJ together with mature AJ (Zihni et al. 2016). Immunofluorescence of A549 E-cadherin depleted cells at later time points was precluded as the cells appeared sick and only loosely attached to the coverslips.

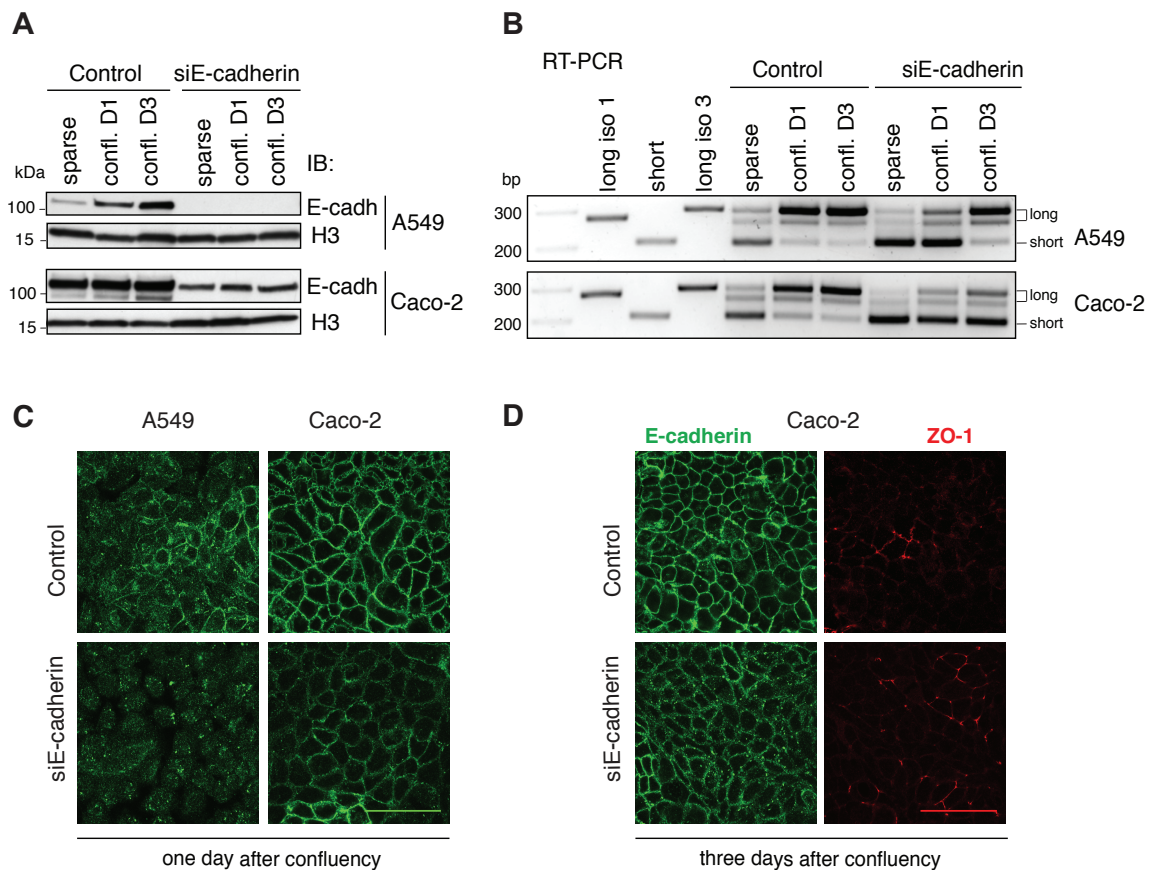


Figure 16. E-cadherin depletion impairs myosin VI isoform switch. **A.** IB showing E-cadherin depletion in A549 or Caco-2 cell lysates upon siRNA transfection. **B.** RT-PCR results of cDNA prepared from A549 or Caco-2 E-cadherin KD cells grown in sparse conditions and one (D1) or three (D3) days after that confluency was reached. **C.** Immunofluorescence analysis showing E-cadherin localization in E-cadherin-depleted A549 or Caco-2 cells cultured one day after that confluency was reached. Scale bar: 50 μ m **D.** Immunofluorescence analysis showing ZO-1 localization in Caco-2 E-cadherin KD cells three days after that confluency was reached (medial stack). Scale bar: 50 μ m.

As myosin VI AS switch was observed also in Caco-2 cysts, we evaluated the effect of E-cadherin depletion in this 3D model. Also in Caco-2 cysts, E-cadherin depletion, confirmed by qPCR (**Fig. 17A**), caused a severe impairment of myosin VI isoform switch (**Fig. 17B**). Notably, similar to what we scored in cell monolayers, ZO-1 was mislocalized in E-cadherin depleted cells showing signal both at the apical and basal surfaces (**Fig. 17C**). Upon E-cadherin depletion, myosin VI apical localization was partially lost with an increased signal in the cytoplasmic compartment (**Fig. 17D**). This observation is likely due to the loss of myosin VI_{long}, which is usually localized together with clathrin at the apical surface (Biancospino et al. 2019; Wollscheid et al. 2016).

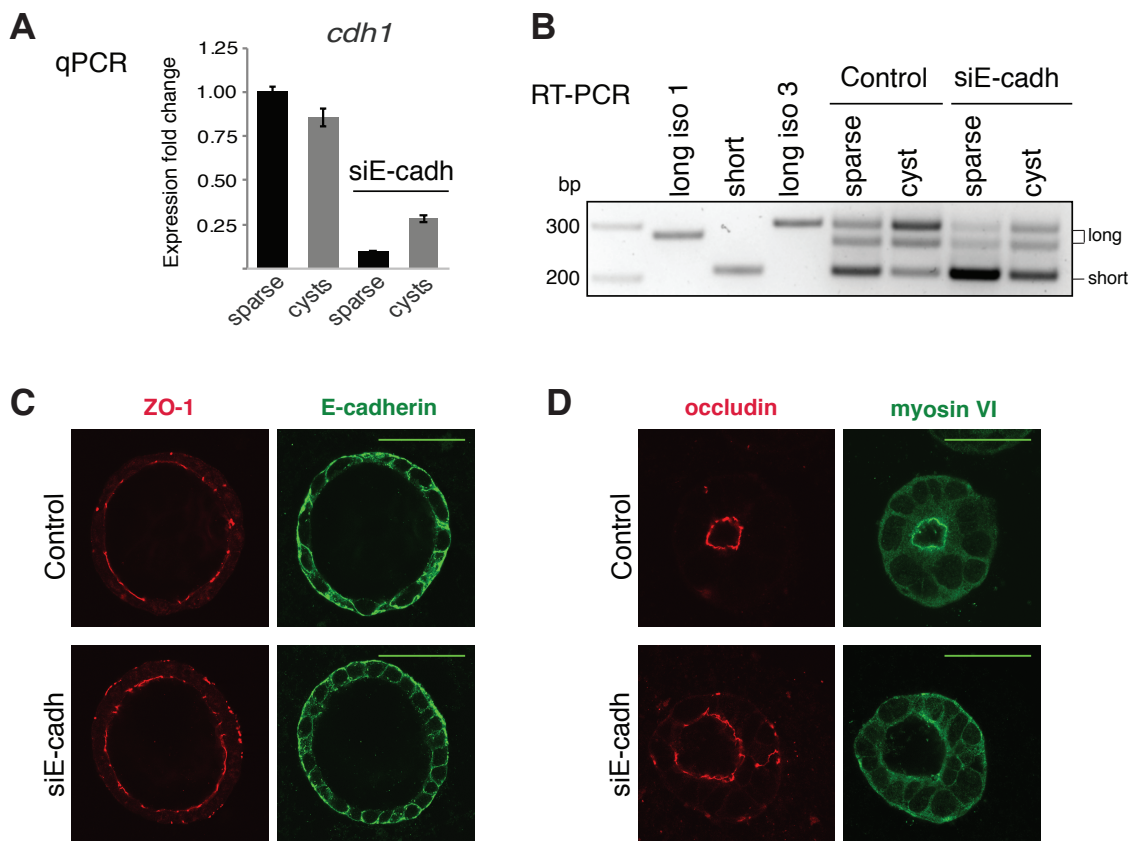


Figure 17. E-cadherin KD impairs isoform switch in a 3D system. **A.** qPCR showing the downregulation of E-cadherin transcript before (Time 0, sparse) and after Caco-2 cyst formation (one week in matrigel). **B.** RT-PCR results of cDNA prepared from control and E-cadherin-depleted Caco-2 cysts. **C.** Immunofluorescence analysis showing ZO-1 mislocalization in E-cadherin-depleted Caco-2 cyst. Scale bar: 100 μ m **D.** Immunofluorescence analysis showing myosin VI localization in E-cadherin-depleted Caco-2 cysts. Scale bar: 50 μ m.

These observations are specific for E-cadherin depletion, as cysts generated with the stable shPAR3 Caco-2 cell line (**Fig. 18A**) behave as control cysts and switch to the expression of myosin VI_{long} (**Fig. 18B**). Unlike E-cadherin KD, PAR3 depletion did not cause ZO-1 mislocalization (**Fig. 18C**), and myosin VI is normally localized at the apical surface (**Fig. 18D**).

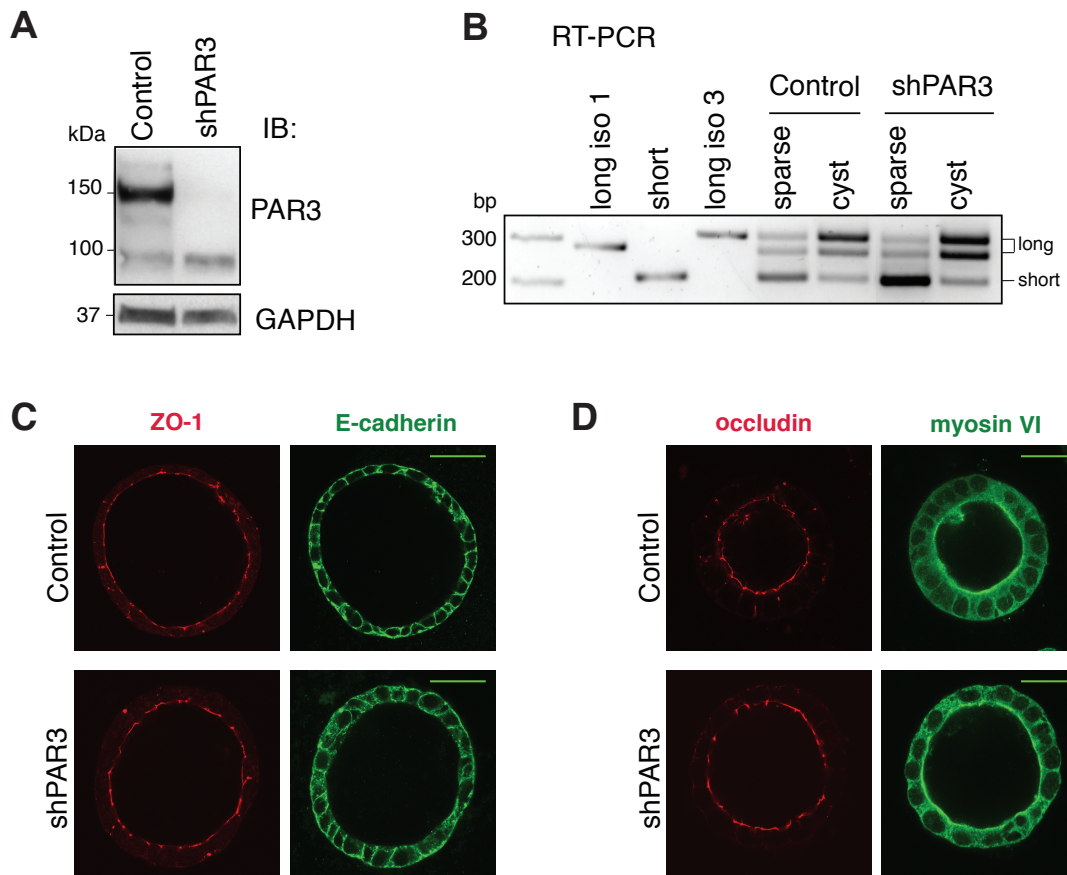


Figure 18. PAR3 KD does not affect the isoform switch in a 3D system. **A.** IB showing stable PAR3 depletion in Caco-2 cells. **B.** RT-PCR results of cDNA prepared from control and PAR3-depleted Caco-2 cysts. **C.** Immunofluorescence analysis showing ZO-1 localization in PAR3-depleted Caco-2 cyst. Scale bar: 50 μ m. **D.** Immunofluorescence analysis showing myosin VI localization in PAR3-depleted Caco-2 cysts. Scale bar: 50 μ m.

5.2.3. β -catenin influences myosin VI isoform expression

Having observed that E-cadherin depletion impairs myosin VI_{long} expression both in cell monolayers and in cells growing as cysts, we then focused our attention on β -catenin, a well-characterized partner of E-cadherin at cell-cell contacts. In epithelial tissues, β -catenin is sequestered at the membrane, preventing its translocation to the nucleus. When β -catenin reaches the nucleus, for example after WNT stimulation, it exerts a co-transcriptional activity by binding the T-cell factor/lymphoid enhancer factor (TCF/LEF) family of transcription factors (Jeanes, Gottardi, and Yap 2008). Thus, we decided to analyze the possible contribution of β -catenin in E-cadherin-depleted cells, which showed defects in myosin VI AS switch.

First, we analyzed β -catenin localization by immunofluorescence in our cellular models. As expected, in control cell monolayers, β -catenin localized at cell-cell junctions together with E-Cadherin (**Fig. 19A**). In contrast, in cell monolayers of E-cadherin-depleted cells we observed β -catenin nuclear staining in both A549 and Caco-2 cells (**Fig. 19A**). This nuclear staining suggests the presence of a pool of active β -catenin, thus we asked whether the effect observed after E-cadherin KD in myosin VI AS was due to the localization of the specific form of active non-phosphorylated (on Ser37 or Thr41) β -catenin using a specific antibody (Liu et al. 2002). In control cells, active β -catenin showed nuclear localization in sparse growth conditions, which was strongly reduced after that confluency was reached (**Fig. 19B**). In contrast, in confluent E-cadherin KD cells the nuclear active β -catenin signal clearly persisted (**Fig. 19C**).

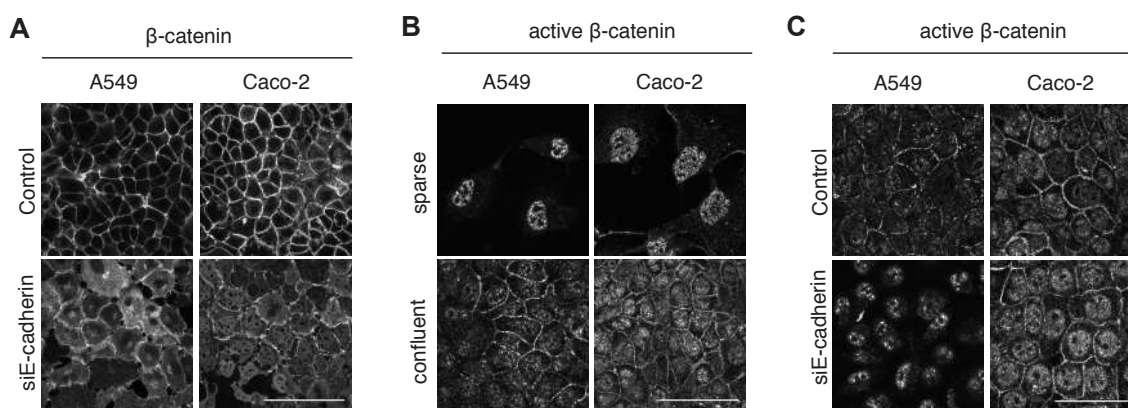


Figure 19. Active β -catenin localization is affected by E-Cadherin depletion. **A.** Immunofluorescence analysis showing β -catenin localization in E-cadherin-depleted A549 or Caco-2 cells cultured one day after that confluency was reached. Scale bar: 50 μ m. **B.** Immunofluorescence analysis showing active β -catenin localization in A549 and Caco-2 cells grown in sparse growth conditions or cultured one day after that confluency was reached. In cells grown in sparse growth conditions, active β -catenin is almost exclusively nuclear. Scale bar: 50 μ m. **C.** Immunofluorescence analysis showing active β -catenin localization in E-cadherin-depleted A549 or Caco-2 cells cultured one day after that confluency was reached. Scale bar: 50 μ m.

The presence of nuclear β -catenin in confluent E-cadherin-depleted cells may suggest that the impairment in myosin VI AS switch observed could be due to the persistence of nuclear active β -catenin rather than a direct effect of the E-cadherin absence. We hypothesized that in not perturbed sparse growth conditions, the active nuclear β -catenin might be required for the expression of myosin VI_{short}, promoting skipping of the large insert. In contrast, in

confluent cells, where β -catenin mostly localizes at the plasma membrane together with E-cadherin, this nuclear activity on AS would be reduced, allowing the inclusion of the exon cassette. We thus depleted β -catenin by siRNA transfection (**Fig. 20A**) and analyzed myosin VI isoform expression by RT-PCR (**Fig. 20B**). As predicted by our hypothesis, β -catenin downregulation strongly affected myosin VI AS in Caco-2 cells, with an increased expression of myosin VI_{long} in cells growing in sparse conditions (**Fig. 20B**). The same effect was not observed in A549 cells (**Fig. 20B**). A possible explanation for this discrepancy could be that A549 cells respond to β -catenin downregulation by activating additional signaling pathways. This phenotype and its explanation are currently under analysis.

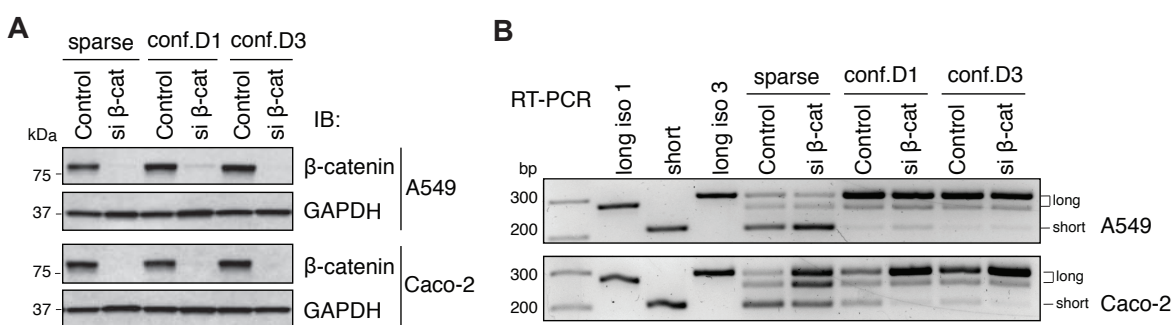


Figure 20. β -catenin depletion impairs myosin VI_{short} expression in Caco-2 cells. **A.** IB showing β -catenin depletion in the indicated cell lysates upon siRNA transfection. **B.** RT-PCR analysis of cDNA prepared from A549 or Caco-2 β -catenin-depleted cells grown in sparse conditions and one or three days after that confluency was reached.

To further test our hypothesis, we are currently producing cells expressing an inducible form of a constitutively active β -catenin (Barth, Stewart, and Nelson 1999). This cellular tool will allow us to manipulate the nuclear pool of β -catenin in both sparse and confluent cells and evaluate its effect on myosin VI AS. In addition, and in order to identify downstream β -catenin targets that could influence AS decision, another investigator in our laboratory recently started an RNAseq analysis of Caco-2 cells grown in sparse and confluent conditions in the presence or absence of β -catenin.

5.3. Identification of splicing regulators involved in myosin VI alternative splicing

5.3.1. Diverse splicing regulators are modulated when cells grow as monolayers

In order to identify the splicing regulators that might be involved in myosin VI AS switch, we undertook an unbiased approach looking for RNA-binding protein genes that are regulated in our cellular model systems. By taking advantage of the Affymetrix microarray technology, we performed a gene expression analysis of RNA extracted from three A549 samples: cells grown in sparse conditions, cells grown for five days after that confluency was reached, and confluent cells disaggregated and maintained at low-density culture for five days (**Fig. 21A**). Two biological replicates were analyzed showing consistent and reproducible results. We obtained a list of differentially expressed genes and focused our attention on the RNA-binding proteins involved in splicing. We obtained a list of 16 genes, the majority of which was down-modulated upon confluency (**Fig. 21B**).

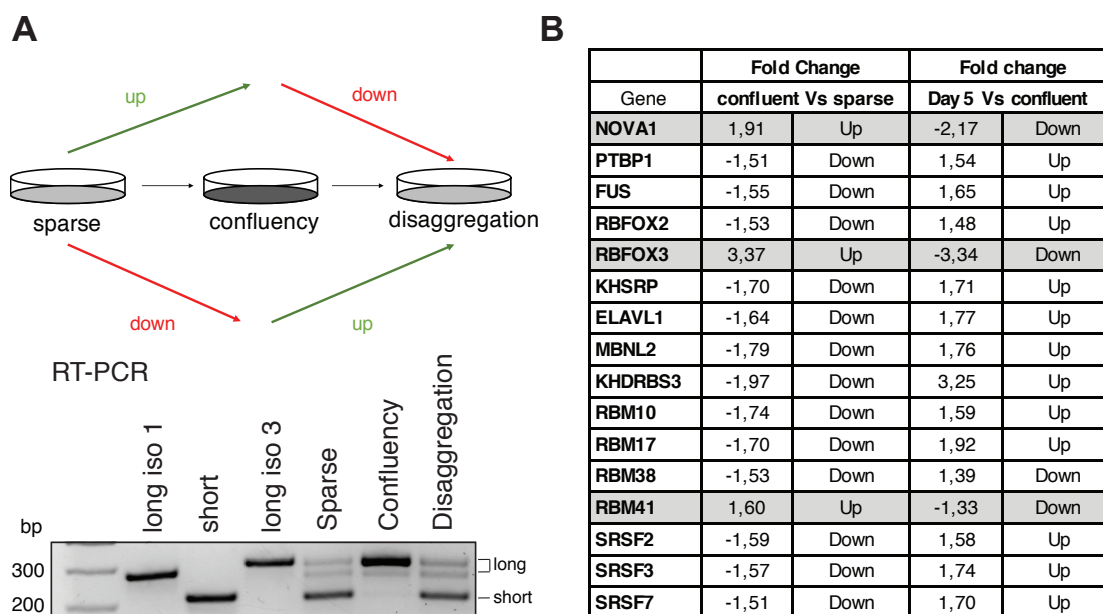


Figure 21. Diverse splicing regulators are modulated when cells grow as monolayers. **A.** RNA samples from A549 cells in the indicated conditions were analyzed by Affimetrix. Lower panel: RT-PCR results of cDNA prepared from cells grown in the indicated conditions. **B.** List of AS modulators that appeared to be regulated in our Affimetrix analysis with their relative fold changes (log2).

We then performed a qPCR validation of these candidates. In order to select the strongest candidates, we evaluated their expression also in Caco-2 cells, which, like A549 cells, undergo an AS switch of myosin VI (**Fig. 11B**). These two cell lines share the same behavior in myosin VI isoform expression, thus we expected that the potential RNA binding proteins

involved in the AS event would be regulated in the same manner upon confluency. Indeed, 13 out of 16 RNA-binding protein genes are equally regulated in both cell lines (**Fig. 22A** and **22B**).

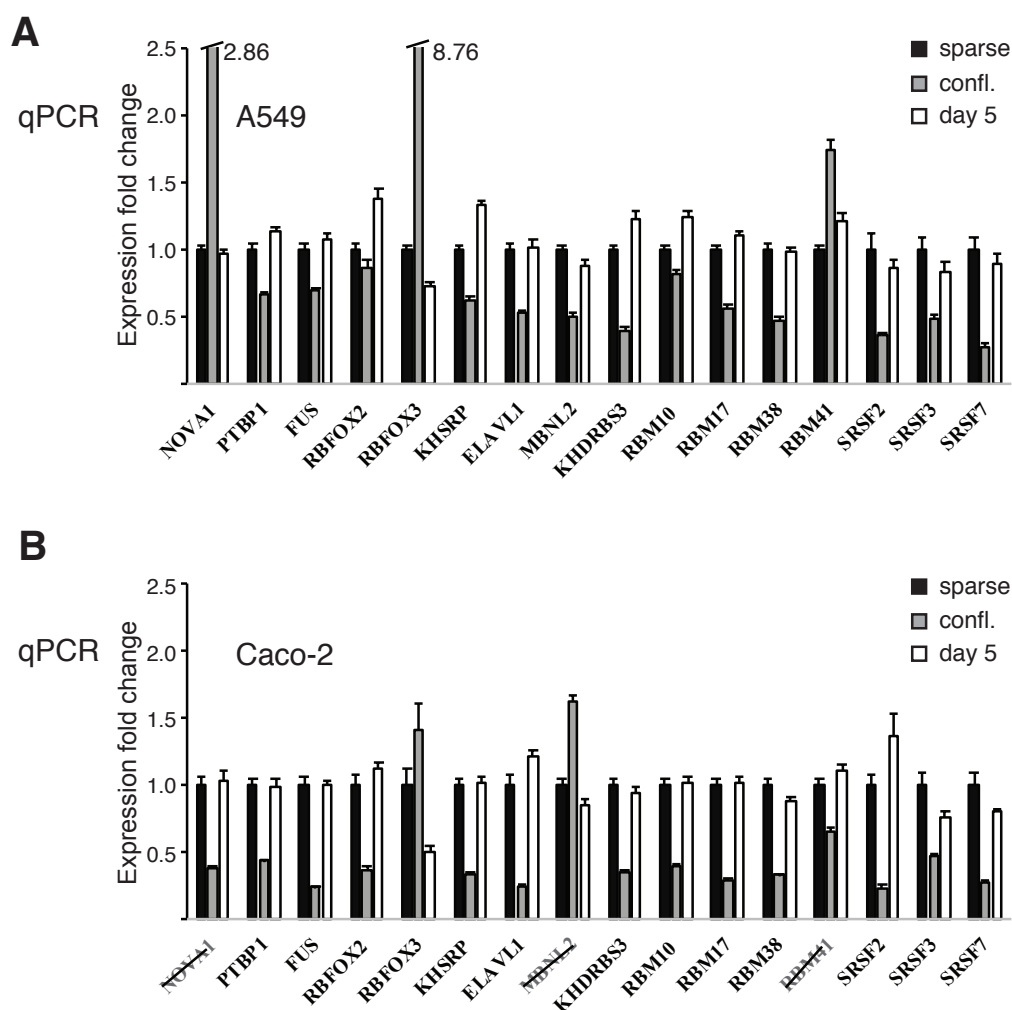


Figure 22. Validation of splicing regulators differentially expressed in the myosin VI isoform switch system. **A.** qPCR results representing the validation of splicing regulators found to be regulated in A549 cells. **B.** qPCR results representing the validation of the same splicing regulators in Caco-2 cells. In strikethrough, splicing factors that show a different behavior than that of A549 cells.

A double approach, a transient KD by siRNA transfection and over-expression of the protein in a stable and constitutive manner, was then employed to evaluate the importance of these splicing regulators (PTBP1, FUS, SRSF2, RBFOX2, and KHSRP) in myosin VI AS decision, as described in the next sections.

5.3.1.1. PTBP1 affects myosin VI alternative splicing in a cell type-specific manner

PTBP1 was found to be downregulated in confluent growing conditions when the myosin VI_{long} isoform is expressed. PTBP1 is a member of the hnRNP protein family of alternative splicing regulators that preferentially bind splicing silencers and promote exon skipping (Busch and Hertel 2012). Although largely known as a negative regulator of alternative splicing, PTBP1 can also stimulate exon inclusion by binding polypyrimidine stretches downstream the alternatively spliced exon cassette and close to the flanking constitutive exon (Busch and Hertel 2012; Llorian et al. 2010; Boutz et al. 2007).

In order to address the role of PTBP1 in myosin VI AS, we performed PTBP1 KD by siRNA transfection, which resulted in an efficient depletion in both A549 and Caco-2 cells, as evaluated by qPCR (Fig. 23A). In A549 cells, we scored a strong reduction of myosin VI_{long} expression in cells grown as monolayers (Fig. 23B), suggesting that PTBP1 stimulates the inclusion of the large insert in this cell line. Instead, PTBP1 depletion in Caco-2 cells did not affect myosin VI AS. The differences observed between the cell lines could suggest that PTBP1 action might be cell type-specific (Fig. 23B).

Next, we wondered if an overexpression of the protein in A549 cells could lead to the opposite effect, inducing myosin VI_{long} in cells growing in sparse conditions, thus we generated A549 cells stably expressing Enhanced Green Fluorescent Protein (EGFP)-PTBP1 by lentiviral infection (Fig. 23C). Despite reaching a good expression level, stable transfection of PTBP1 did not have any impact on myosin VI isoform switch (Fig. 23D).

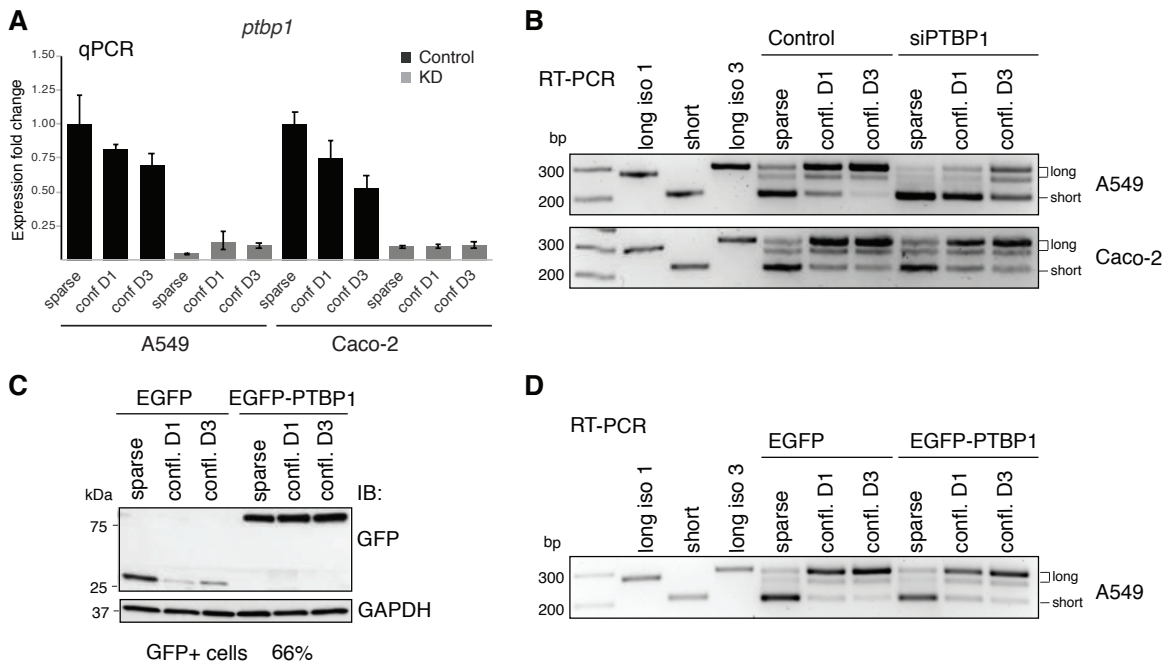


Figure 23. PTBP1 affects myosin VI alternative splicing in A549 but not in Caco-2 cells. A. qPCR showing the downregulation of PTBP1 transcript upon siRNA transfection in A549 or Caco-2 cells. **B.**

RT-PCR results of cDNA prepared from A549 or Caco-2 PTBP1 KD cells grown in sparse conditions and one or three days after that confluency was reached. **C.** IB showing EGFP-PTBP1 overexpression in A549 or Caco-2 cell lysates. Percentage of GFP-positive cells, calculated by immunofluorescence analysis of more than 200 cells, is indicated above the panel. **D.** RT-PCR results of cDNA prepared from A549 cells overexpressing PTBP1 grown in sparse conditions and one or three days after that confluency was reached.

5.3.1.2. *FUS is not involved in myosin VI alternative splicing*

Our Affimetrix analysis revealed a downregulation of Fused in sarcoma (FUS) from sparse growth conditions to confluent monolayers. FUS is a member of the FET protein family and it is a ubiquitously expressed RNA binding protein (Tan and Manley 2009) that preferentially binds intronic GU-rich motifs and regulates alternative splicing (Masuda et al. 2015; Lagier-Tourenne et al. 2012; Rogelj et al. 2012).

In order to understand if FUS is involved in myosin VI AS, we depleted the protein by siRNA transfection in A549 cells and assessed myosin VI isoforms expression by RT-PCR. Despite reaching an efficient depletion, measured by immunoblot analysis, IB (**Fig. 24A**), FUS KD did not affect myosin VI AS (**Fig. 24B**). Similar to protein depletion, overexpression of EGFP-FUS (**Fig. 24C**) had no impact on myosin VI isoform switch (**Fig. 24D**).

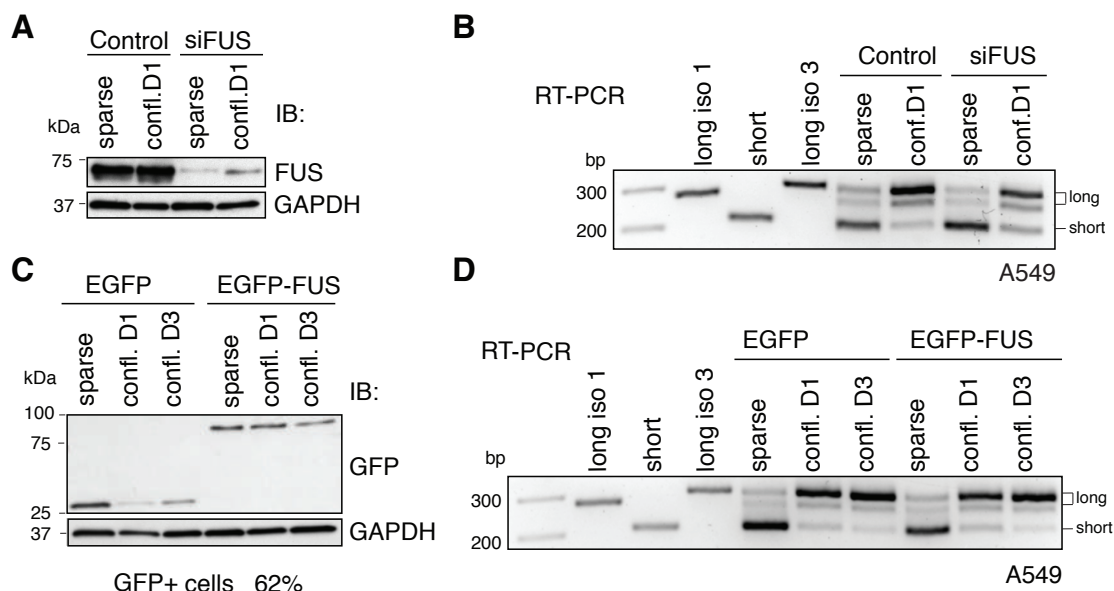


Figure 24. FUS is not involved in myosin VI alternative splicing. **A.** IB showing FUS depletion in the indicated cell lysates upon siRNA transfection. **B.** RT-PCR results of cDNA prepared from A549 FUS KD cells grown in sparse conditions or one day after that confluency was reached. **C.** IB showing EGFP-

FUS overexpression in A549 cell lysates. Percentage of GFP-positive cells is indicated. **D.** RT-PCR results of cDNA prepared from A549 cells overexpressing FUS grown in sparse conditions and one or three days after that confluency was reached.

5.3.1.3. SRSF2 promotes myosin VI_{short} expression

In our Affimetrix analysis, SRSF2 was shown to be downregulated when cells were grown as confluent monolayers. SRSF2 is a member of the SR protein family which is involved in both constitutive and alternative splicing (Jeong 2017; Fu and Maniatis 1992). Although generally described as positive regulators of splicing, SR proteins may also act as negative regulators (Pandit et al. 2013).

In order to address the role of SRSF2 in myosin VI AS, we performed SRSF2 KD by siRNA transfections. We efficiently depleted SRSF2 in both A549 and Caco-2 cell lines, as confirmed both by qPCR (**Fig. 25A**) and IB (**Fig. 25B**). Interestingly, we observed that SRSF2 depletion induces the expression of myosin VI_{long} already in sparsely grown cells, both in A549 and Caco-2 cells (**Fig. 25C**). This observation suggests that SRSF2 inhibits the inclusion of the large insert in sparse growth conditions.

Next, we wondered if overexpression of the protein could lead to the opposite effect, namely if SRSF2 overexpression could induce the expression of myosin VI_{short} in cell monolayers. By lentiviral infection, we tried to generate A549 and Caco-2 cell lines stably expressing EGFP-SRSF2. Unfortunately, SRSF2 overexpression had a dramatic impact on cell viability: Caco-2 cells died a few days after the infection while selected A549 cells expressed the fusion protein at very low levels only in 40% of the population (**Fig. 25D**, see expression of EGFP-SRSF2 compared to the endogenous SRSF2). This effect might be due to the induction of apoptosis caused by SRSF2-mediated expression of the pro-apoptotic splice variants of caspases-8 and -9 and Bcl-x (Merdzhanova et al. 2008). These problems precluded further analysis on myosin VI AS.

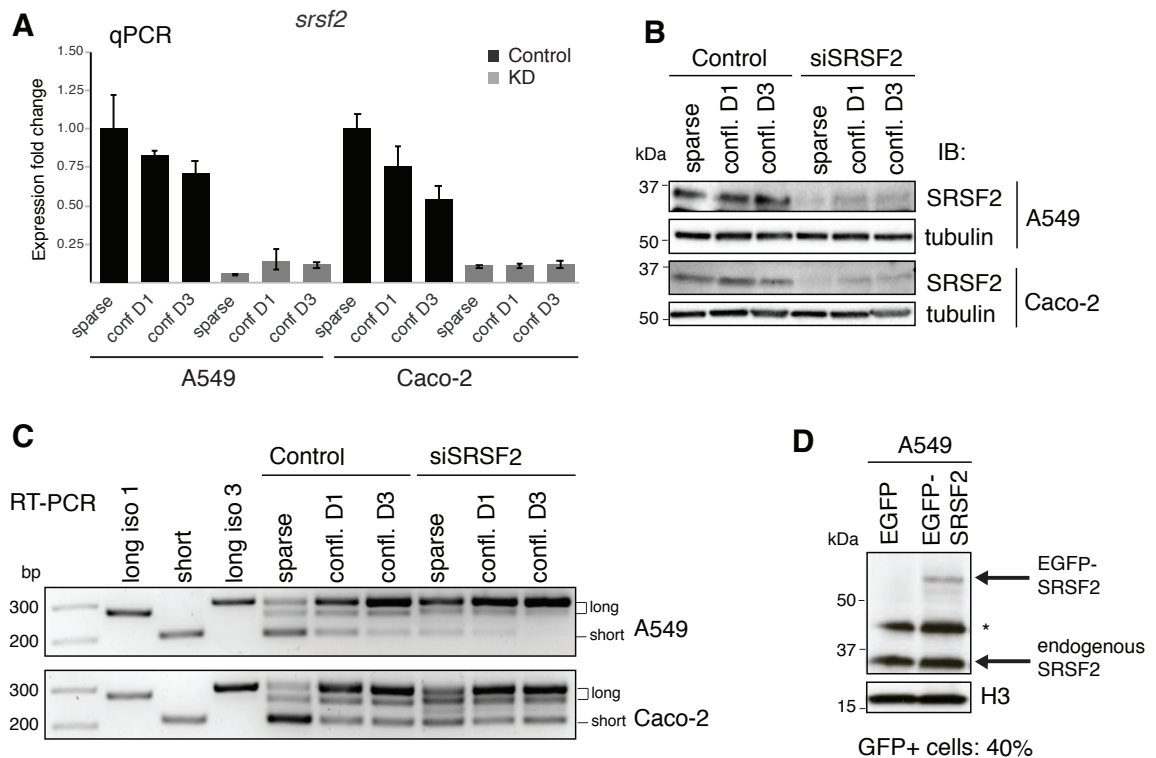


Figure 25. SRSF2 promotes skipping of the large insert. **A.** qPCR showing the downregulation of SRSF2 transcript upon siRNA transfection in A549 or Caco-2 cells. **B.** IB showing SRSF2 depletion in the indicated cell lysates upon siRNA transfection. **C.** RT-PCR results of cDNA prepared from A549 or Caco-2 SRSF2 KD cells grown in sparse conditions and one or three days after that confluency was reached. **D.** IB showing EGFP-SRSF2 expression in A549 cell lysates. Endogenous SRSF2 and fusion EGFP-SRSF2 are indicated. (*) un-specific signal. Percentage of GFP-positive cells is indicated.

5.3.1.4. RBFOX2 promotes myosin VI_{long} expression

Our Affymetrix analysis revealed a slight downregulation of RBFOX2 in confluent monolayers. RBFOX2 is a member of the Fox-1 family of tissue-specific alternative splicing regulators (Conboy 2017; Kuroyanagi 2009; Zhang, Zhang, et al. 2008). To address the role of RBFOX2 in myosin VI AS, we achieved its depletion by siRNA transfection. Efficient RBFOX2 KD (**Fig. 26A**) caused a severe impairment in myosin VI AS switch in both A549 and Caco-2 cells (**Fig. 26B**). In sparse growth conditions, RBFOX2 depletion led to a selective expression of myosin VI_{short}, while the switch toward the expression of myosin VI_{long} was strongly reduced (**Fig. 26B**). However, three days after that confluency was reached, we observed that cells start to express myosin VI_{long} even in the absence of RBFOX2, suggesting that RBFOX2 depletion is not sufficient to prevent the expression of

myosin VI_{long}. Thus, additional factors may contribute together with RBFOX2 in promoting myosin VI_{long} expression.

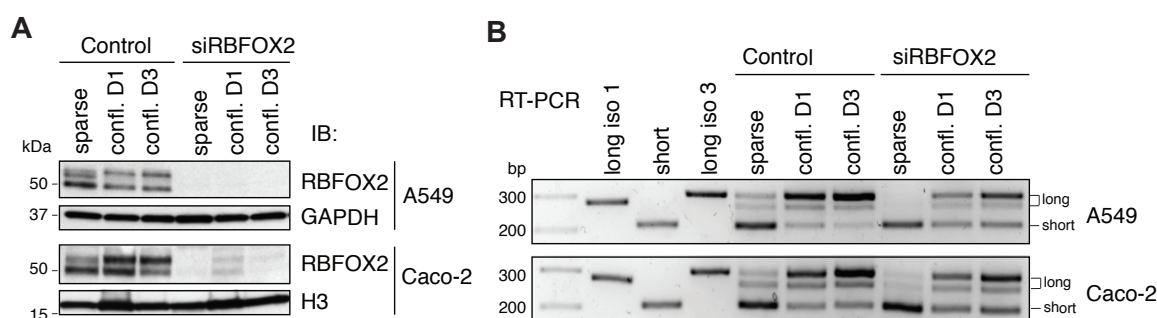


Figure 26. RBFOX2 promotes inclusion of the large insert. **A.** IB showing RBFOX2 depletion in the indicated cell lysates upon siRNA transfection. **B.** RT-PCR results of cDNA prepared from A549 or Caco-2 RBFOX2 KD cells grown in sparse conditions and one or three days after that confluency was reached.

Next, we wondered if overexpression of the protein could lead to the opposite effect, namely if RBFOX2 overexpression could induce the expression of myosin VI_{long} in cells growing in sparse conditions. By lentiviral infection, we generated A549 and Caco-2 cell lines stably expressing EGFP-RBFOX2, in which the fusion protein shows the typical nuclear localization of RBFOX2 (**Fig. 27A**). To be sure about the functionality of the tagged protein, we assessed its effect by looking at NUMB AS, a well-known RBFOX2 target (Lu et al. 2015). RBFOX2 is known to induce the skipping of exon 9 of NUMB, promoting the expression of a NUMB short isoform (Lu et al. 2015). Accordingly, in A549 cells that stably express EGFP-RBFOX2 we observed an increase in NUMB short isoform, confirming that our construct is functional (**Fig. 27B**). Moreover, RBFOX2 depletion in A549 cells presents the opposite effect, stimulating exon inclusion and increasing expression of the NUMB long isoform (**Fig. 27B**), thus validating previous results also in this cell line.

During the initial characterization of A549 and Caco-2 cells that stably express EGFP-RBFOX2, we noticed that the ectopic overexpression of RBFOX2 in A549 cells led to a strong downregulation of E-cadherin, which is a key determinant of myosin VI AS (**Fig. 27C**). The fact that E-cadherin depletion per se affects myosin VI AS complicated the interpretation of the data obtained with this cell line.

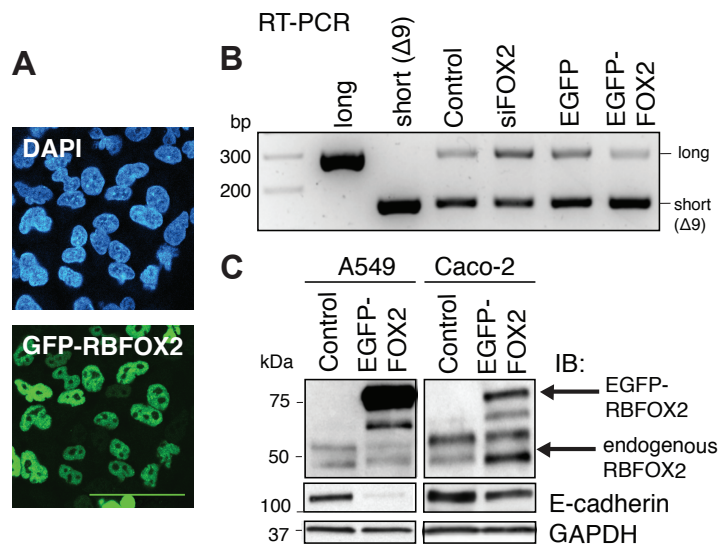


Figure 27. Characterization of RBFOX2 overexpressing cell lines. **A.** Immunofluorescence analysis showing ectopic EGFP-tagged RBFOX2 localization in A549 cells. Scale bar: 50 μ m. **B.** RT-PCR results of cDNA prepared from A549 indicated samples in order to assess NUMB isoforms expression. **C.** IB showing RBFOX2 overexpression in A549 or Caco-2 cells with relative E-cadherin levels.

In order to overcome this problem, we took advantage of a modified A549 cell line that expresses mouse E-cadherin under the control of the LTR promoter, and in which the endogenous human E-cadherin is depleted simultaneously by the expression of a specific human shRNA. Our results revealed that the ectopic mouse E-cadherin shows a physiological localization at cell-cell contacts (**Fig. 28A**), is expressed close to the endogenous level (**Fig. 28B**), and more importantly, is not modulated by RBFOX2 overexpression (**Fig. 28C**). In this controlled system, we found that RBFOX2 overexpression in A549 and Caco-2 cells (**Fig. 28C**) has no effect on myosin VI AS (**Fig. 28D**). Thus, RBFOX2 overexpression alone is not sufficient to induce myosin VI_{long} expression.

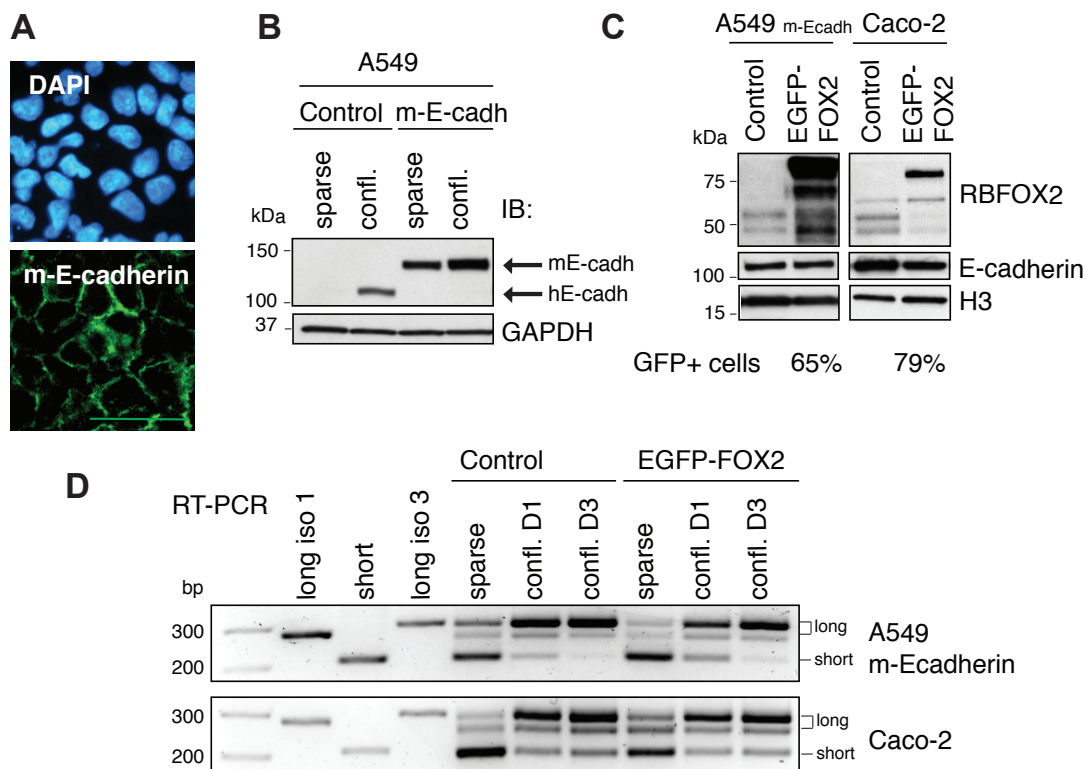


Figure 28. RBFOX2 is not a limiting factor in myosin VI alternative splicing. **A.** Immunofluorescence analysis showing ectopic mouse E-cadherin localization in A549 cells. Scale bar: 50 μ m. **B.** IB showing endogenous (Control) and mouse-E-cadherin expression in A549 cell lysates. **C.** IB showing EGFP-RBFOX2 overexpression in A549 or Caco-2 cell lysates. Percentage of GFP positive cells is indicated. **D.** RT-PCR results of cDNA prepared from A549 m-E-cadherin or Caco-2 cells overexpressing RBFOX2 grown in sparse conditions and one or three days after that confluency was reached.

We next wondered if the effect of RBFOX2 in myosin VI AS is specific for this member of the Fox-1 family. RBFOX1 and RBFOX3 are mainly expressed in the nervous system (Duan et al. 2016), but RBFOX3 expression has been described also in lung tissue (Kim et al. 2016; Kim et al. 2017). In our Affimetrix analysis, we scored an upregulation of RBFOX3 when cells were grown in confluent conditions, thus we decided to analyze the contribution of RBFOX3 on myosin VI AS. We depleted RBFOX3 in A549 (which originates from lung adenocarcinoma) cell monolayers and generated a stable cell line overexpressing the protein. Caco-2 cells were not evaluated since RBFOX3 is barely expressed in this cell line, as measured by qPCR.

IB against RBFOX3 revealed that the protein is only slightly expressed in A549 cells (Fig. 29A, band at ~50 kD in the control lane). RBFOX3 was efficiently depleted one day after

that confluency was reached, while the KD was weaker three days after growth in confluent conditions (**Fig. 29 A,B**). In these conditions, neither RBFOX3 KD nor EGFP-RBFOX3 overexpression (**Fig. 29A**) had any impact on myosin VI AS (**Fig. 29 C,D**), thus confirming that RBFOX2 is a specific member of the RBFOX family and that it promotes exon inclusion.

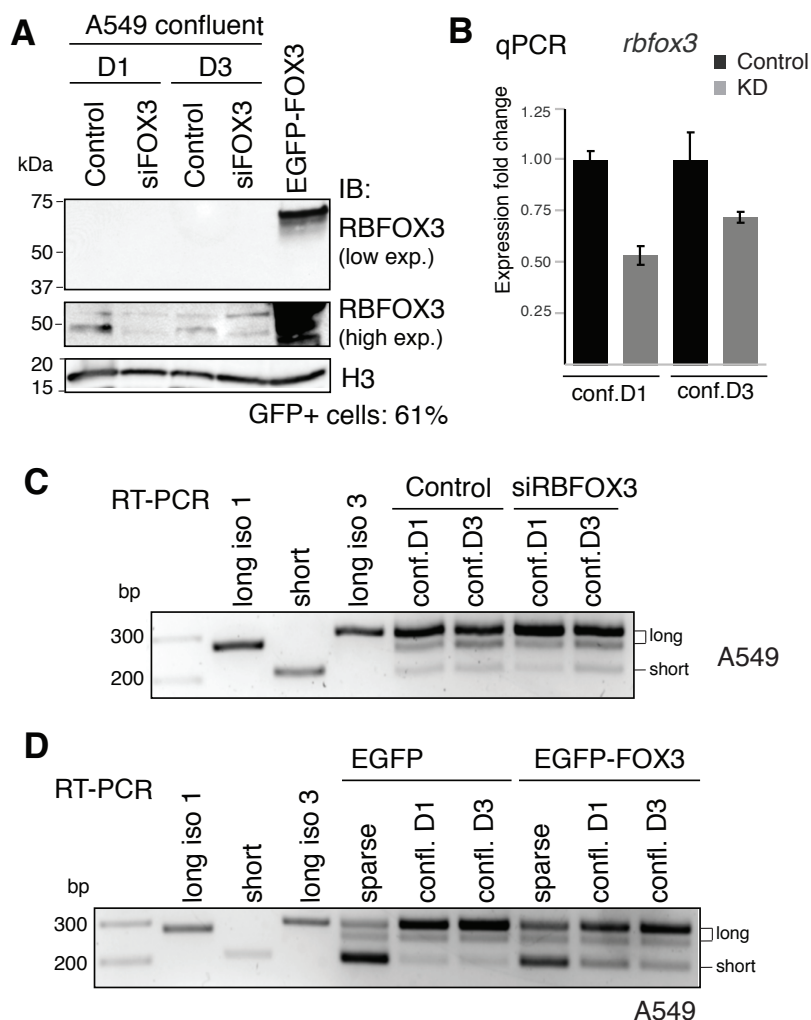


Figure 29. RBFOX3 does not affect myosin VI AS. **A.** IB showing RBFOX3 depletion or overexpression in the indicated A549 cell lysates. Percentage of GFP positive cells is indicated. **B.** qPCR showing the downregulation of RBFOX3 transcript upon siRNA transfection in A549 confluent cells. **C.** RT-PCR results of cDNA prepared from A549 RBFOX3 KD cells grown one or three days in confluent conditions. RBFOX3 depletion does not affect myosin VI AS. **D.** RT-PCR results of cDNA prepared from A549 cells overexpressing RBFOX3 grown in sparse conditions and one or three days after that confluency was reached.

5.3.1.5. KHSRP promotes myosin VI_{long} expression

In our Affimetrix analysis, KH-type splicing regulatory protein (KHSRP) was found to be downregulated in confluent monolayers. KHSRP is an RNA binding protein that controls

gene expression at different levels: transcription, pre-mRNA splicing, mRNA localization, and miRNA maturation (Gherzi et al. 2014). KHSRP was discovered as a splicing regulator able to bind intronic enhancers to promote exon inclusion (Min et al. 1997), and more recently it has been described to control AS of pre-mRNAs involved in EMT, both in a positive and in a negative manner (Puppo et al. 2016).

To investigate if KHSRP plays a role in myosin VI AS, we performed transient KD by siRNA transfection. As visible in **Figure 30**, KHSRP KD caused a severe impairment in myosin VI AS switch in A549 cells. Upon efficient depletion of the protein (**Fig. 30A**), the expression of myosin VI_{long} was completely abolished in sparsely growing cells, and the switch toward the expression of myosin VI_{long} was strongly reduced both one and three days after that confluency was reached (**Fig. 30B**). These observations suggest that KHSRP is a positive regulator of myosin VI AS. We could not analyze the effect of KHSRP depletion in Caco-2 cells because the KD was not efficient in this cell line.

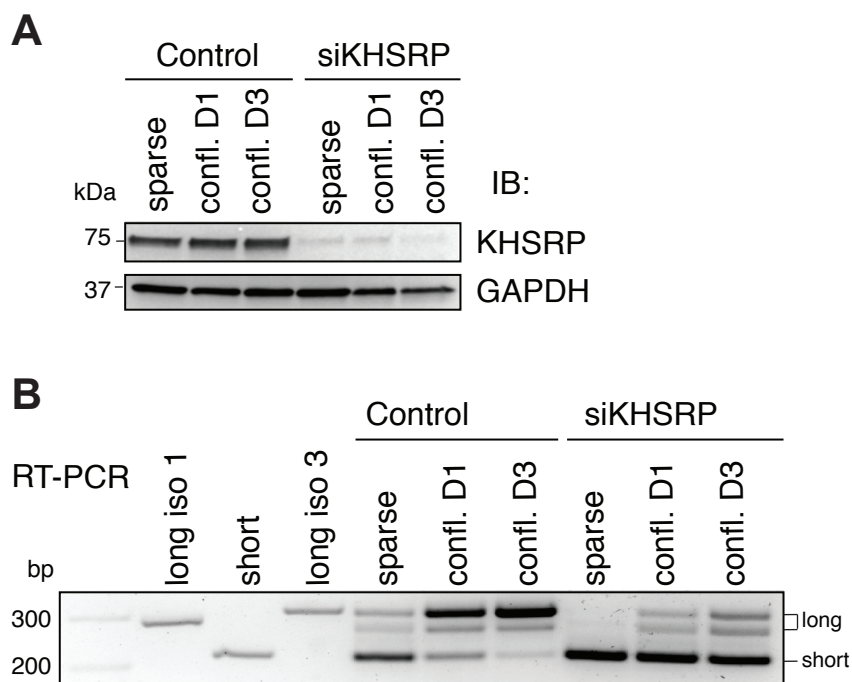


Figure 30. KHSRP promotes inclusion of the large insert. **A.** IB showing KHSRP depletion in the indicated A549 cell lysates upon siRNA transfection. **B.** RT-PCR results of cDNA prepared from A549 KHSRP KD cells grown in sparse conditions and one or three days after that confluency was reached.

In order to evaluate if KHSRP and RBFOX2 have an additive effect in promoting the expression of myosin VI_{long}, we depleted both proteins simultaneously. As shown in **Figure 31**, the simultaneous depletion of RBFOX2 and KHSRP mostly phenocopies the effect of

the single KHSRP KD. The absence of an additive effect might suggest that the two proteins act together to promote the inclusion of the large insert.

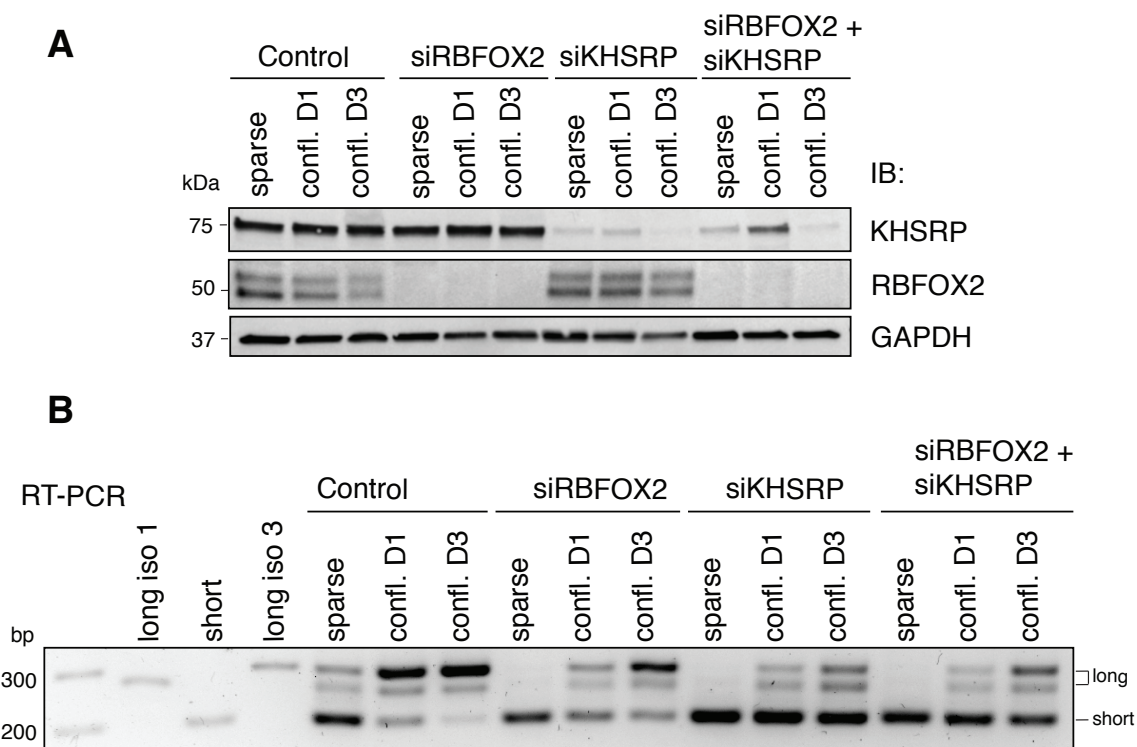


Figure 31. RBFOX2 and KHSRP KD does not show an additive effect. **A.** IB showing RBFOX2 and KHSRP depletion in the indicated A549 cell lysates upon siRNA transfection. **B.** RT-PCR results of cDNA prepared from A549 RBFOX2 and KHSRP-depleted cells grown in sparse conditions and one or three days after that confluency was reached.

5.3.2. Characterization of RNA protein interactions

5.3.2.1. Splicing regulator binding sites in myosin VI AS region

Our screening of splicing factors led to the identification of RBFOX2, KHSRP and SRSF2 as key regulators of myosin VI AS. In order to deeper analyze this phenotype, we asked whether they exert their effect directly by binding myosin VI pre-mRNA. First, we looked at already available CLIP (Crosslinking Immunoprecipitation) data deposited in the ENCODE (Encyclopedia of DNA Elements) Portal database (<https://www.encodeproject.org>) (Davis et al. 2018). In this portal, there are currently 102 eCLIP experiments present for 73 diverse RBP (two biological replicates) in HepG2 (hepatocellular carcinoma) and K562 (chronic myeloid leukemia) cells (Van Nostrand et al. 2016). Since we observed that HepG2 cells express all three myosin VI isoforms (**Fig. 32**), we looked at the eCLIP data obtained with the splicing factors involved in myosin VI AS region to scan their binding profiles between exon 28 and exon 32 of myosin VI.

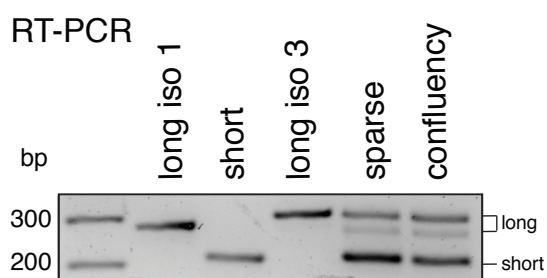


Figure 32. Myosin VI isoform expression in HepG2 cells. HepG2 cells grown in sparse conditions were allowed to reach confluency. RT-PCR results of cDNA prepared from the indicated conditions and cell lines are showed.

In the case of RBFOX2, two predominant peaks were observed; one downstream of the exon 29 (found in both replicates); and a second one located at the very end of the large insert before exon 32 (found in a single replicate) (**Fig. 33A**). Both regions correspond to two UGCAUG intronic motifs, which are the well-known binding sites for Fox-1 proteins (Jin et al. 2003). Similarly, in the experiments performed by immunoprecipitating KHSRP, two predominant peaks were observed in the myosin AS region, one before exon 29 and a second one before exon 32 (**Fig. 33B**). The first peak corresponds to a UGGGU motif, G-rich target sequences that are usually bound by KHSRP (Garcia-Mayoral et al. 2008). The second peak

appears to be the same as that of RBFOX2 and might correspond to the same UGCAUG motif, which has been shown to be bound also by KHSRP (Markovtsov et al. 2000). No experiments were available against SRSF2, probably due to the lack of a validated antibody. Thus, we searched for potential binding sites based on the SRSF2 mechanism of action reported in the literature. SRSF2 promotes myosin VI exon skipping, and it has been proposed that SRSF2 promotes skipping of alternative exons by binding to a flanking constitutive exon (Han et al. 2011). Later, Lou and colleagues (Luo et al. 2017) showed that SRSF2 alternative skipped exons are flanked by constitutive exons characterized by an enrichment in a TCCTG motif, thus, we searched for similar motifs in the myosin AS region. Strikingly, we identified the very same motif in exon 32, which is the downstream flanking constitutive exon of the alternatively splice exon cassette. Thus, we hypothesize that SRSF2 binding to exon 32 might induce skipping of the large insert (**Fig. 23**). A schematic representation of these *in silico* analyses is reported in **Figure 33C**.

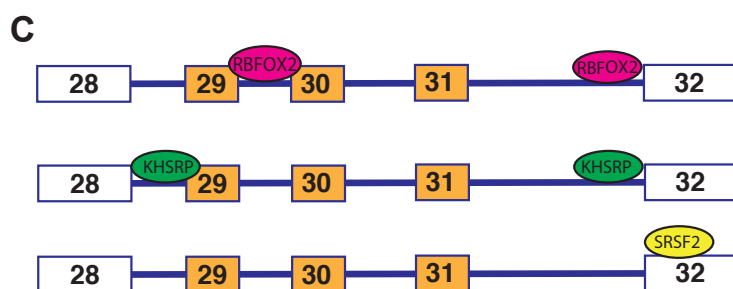
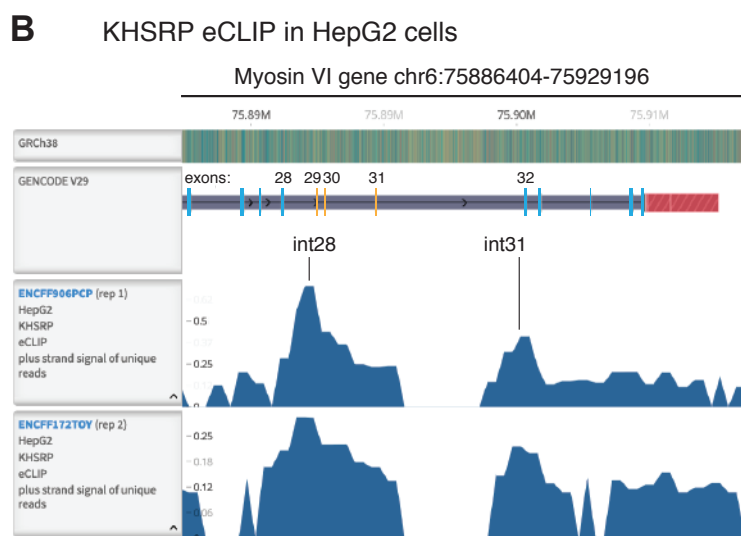
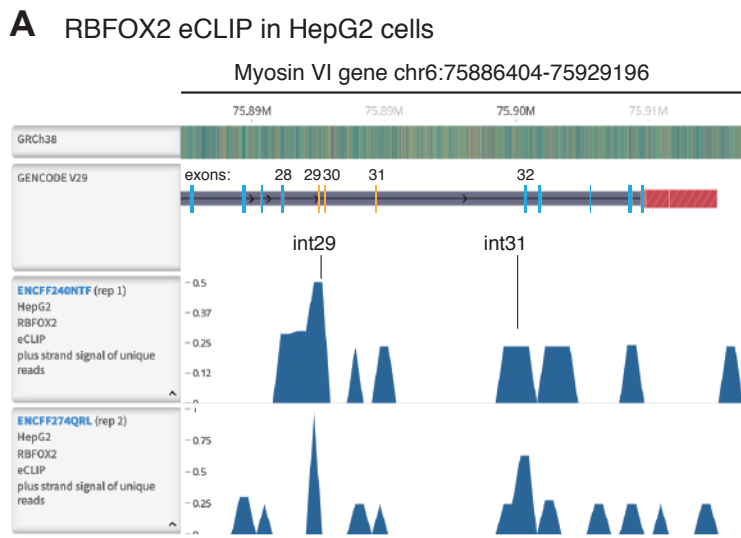


Figure 33. Splicing regulators binding sites in myosin VI pre-mRNA. **A.** Representation of eCLIP peaks in the myosin VI transcript upon RBFOX2 immunoprecipitation in Hep2G cells (Van Nostrand et al., 2016). Data are extracted from the ENCODE project database. **B.** Representation of eCLIP peaks in the myosin VI transcript upon KHSRP immunoprecipitation in HepG2 cells (Van Nostrand et al, 2016). Data are extracted from the ENCODE project database. **C.** Schematic representation of RBFOX2, KHSRP and SRSF2 binding sites in the myosin VI AS region. Binding sites are located within 300 nt from from splice sites.

5.3.2.2. RBFOX2 interacts with myosin VI pre-mRNA

In order to determine the physical interaction between RBFOX2 and myosin VI pre-mRNA, we performed RNA immunoprecipitation after UV crosslinking (UV-RIP). Briefly, cell plates were irradiated with UV light to crosslink protein/RNA complexes and the specific RBFOX2/RNA complexes were immunoprecipitated from nuclear lysates by using an anti-RBFOX2 antibody (**Fig. 34A**). Thereafter, the protein fraction was analyzed by IB whereas the RNA fraction by RT-PCR.

Immunoprecipitation of RBFOX2 efficiently depleted the protein from the lysate, as shown in the IB of **Figure 34B**. The RT-PCR analysis detected an amplification of a myosin VI pre-mRNA fragment (spanning intron31-exon32 junction) from the RNA fraction associated with RBFOX2 but not from the IgG negative control (**Fig. 34C**). We then quantified the binding of myosin VI pre-mRNA to RBFOX2 by RT-qPCR with the percent input method, using a TaqMan probe recognizing exon32-intron32 junction. About 1% of the myosin VI pre-mRNA resulted to be bound to RBFOX2, while we scored no significant enrichment in the IgG negative control (**Fig. 34D**). Aspecific 18S RNA was not immunoprecipitated by RBFOX2 (**Fig. 34D**), validating the entire procedure. Taken together, these results confirm that RBFOX2 interacts directly with myosin VI pre-mRNA, and exerts a direct effect on the AS region.

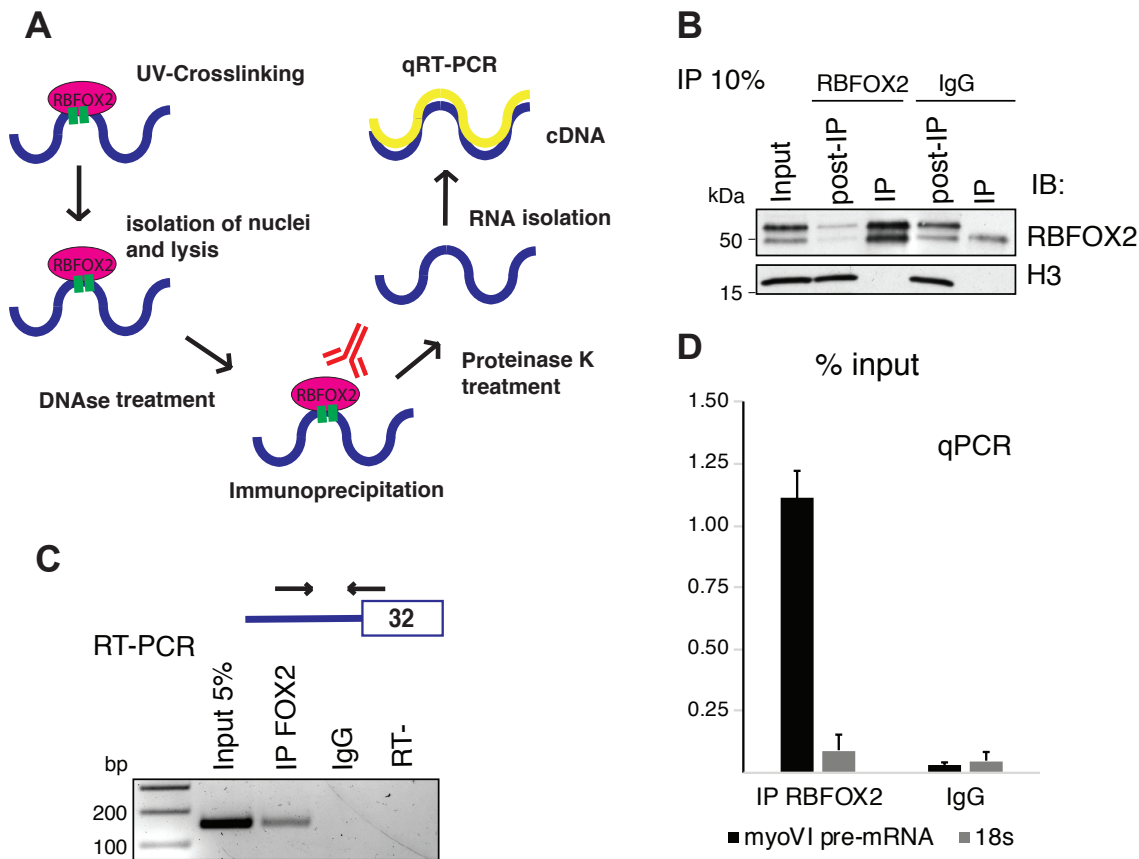


Figure 34. RBFOX2 interacts with myosin VI pre-mRNA **A.** UV-RIP steps for quantification of RBFOX2 binding to myosin VI pre-mRNA *in vivo*. UV cross-linking, isolation of nuclei and lysis, DNase treatment for removal of gDNA, RBFOX2 immunoprecipitation with a specific antibody, protein degradation, RNA isolation, and first strand RT followed by qPCR. **B.** IB showing RBFOX2 immunoprecipitation from A549 nuclear lysate, IgG used as control. **C.** RT-PCR results of cDNA prepared from RNA fraction associated with RBFOX2 or IgG negative control. Primers specific for myosin VI intron31-exon32 junction were used. **D.** qPCR quantification from cDNA prepared as in B. Percent input method is used to quantify RBFOX2-myosin VI pre-mRNA binding.

To further refine our analysis, we designed specific TaqMan probes centered on the two potential RBFOX2 binding sites, identified in the ENCODE database (**Fig. 35A**). As a negative control, a TaqMan probe located upstream of the large insert at exon 25-intron 25 boundary was designed. We then quantified the binding of RBFOX2 to these different regions of myosin VI pre-mRNA by RT-qPCR, represented as fold enrichment respect to IgG negative control. A ~190 fold enrichment was observed for RBFOX2 on the binding site located in intron 29 (the UGCAUG motif), and a ~65 fold enrichment on the binding site located in the intron 31 (**Fig. 35B**). To note, no binding was scored upstream of the AS

region, thus showing that RBFOX2 specifically binds the internal boundaries of the myosin VI AS region, namely the UGCAUG motifs in the intron downstream of exon 29 and in the intron before exon 32 (Fig. 33).

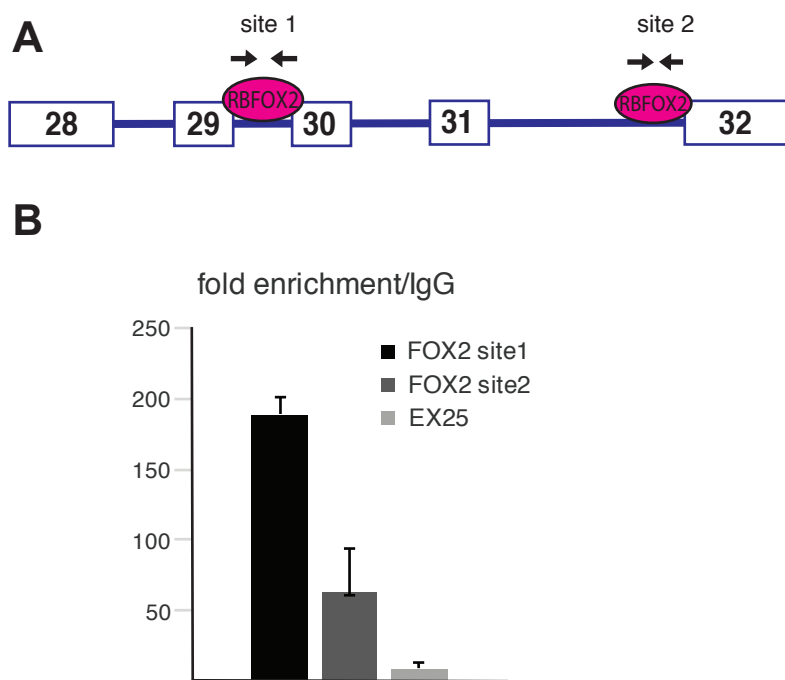


Figure 35. RBFOX2 binding to UGCAUG motifs in myosin VI pre-mRNA. **A.** Schematic representation of the TaqMan assays designed to target RBFOX2 binding sites in myosin VI AS region. **B.** Fold enrichment method is used to quantify RBFOX2 binding to different regions of myosin VI pre-mRNA by RT-qPCR after RBFOX2 immunoprecipitation. IgG negative control is set to 1. Results are shown as the average of two independent experiments.

5.4. Myosin VI alternative splicing regulation in ovarian cancer

5.4.1. Myosin VI short is selectively expressed in ovarian cancer

We recently described myosin VI_{short} overexpression in ovarian primary tumors (Wollscheid et al. 2016). Here, we analyzed ovarian tissues by immunohistochemistry, using isoform-specific antibodies generated in our laboratory. By using a myosin VI antibody that recognizes all myosin isoforms (myosin total), we observed that myosin VI is expressed in both normal Fallopian tube epithelium (FTE) and in ovarian primary tumors (**Fig. 36**, upper panel). However, when looking at specific myosin isoforms, we scored a selective isoform expression. Normal tissue preferentially expresses myosin VI_{long} isoform (**Fig. 36**, lower panel), in contrast, the tumor samples show a strong and selective expression of the short isoform. Thus, we confirmed our previous observation using RNA-based approaches (Wollscheid et al. 2016), and we detected selective expression of the myosin VI_{short} protein isoform in ovarian tumor tissues by immunohistochemistry.

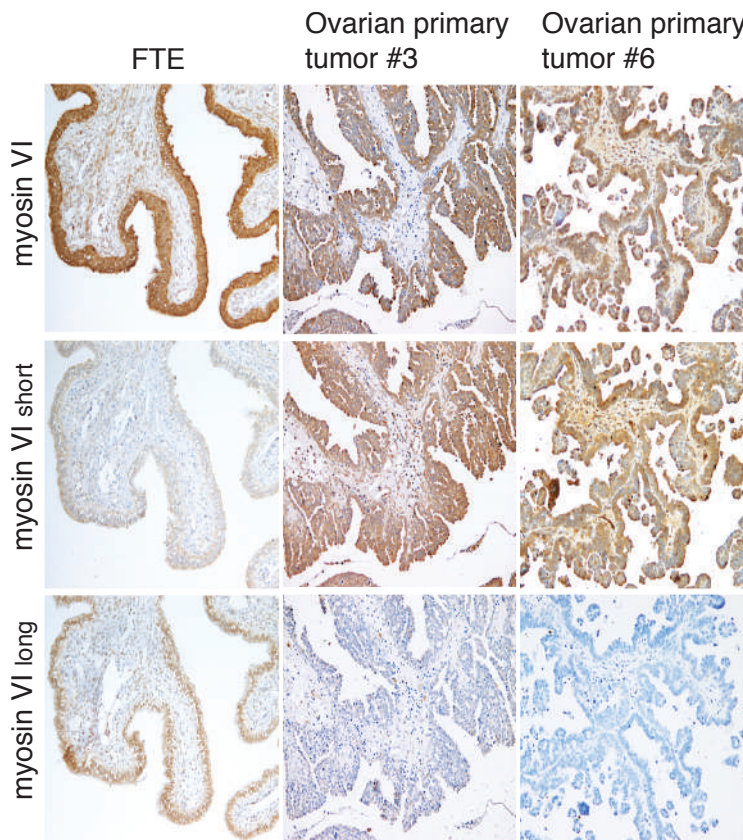


Figure 36. Myosin VI isoform expression in ovarian cancer. Ovarian primary tumors and normal tissue (fallopian tube epithelium, FTE, as control) were embedded in paraffin and sections were prepared for IHC analysis. The sections were stained with antibodies specific for myosin VI total (1296), myosin VI_{short}, or myosin VI_{long} in order to compare their expression in tumor and normal tissue [in collaboration with Ugo Cavallaro and Federica Pisati].

5.4.2. SRSF2 and RBFOX2 regulates myosin VI AS in ovarian cancer

We decided to evaluate the contribution of SRSF2, RBFOX2 and KHSRP in myosin VI AS regulation in ovarian cancer cells. To this end, we first analyzed the behavior of myosin VI AS in a panel of ovarian cancer cell lines (**Fig. 37**). Confirming previous data (**Fig. 11** and Wollscheid et al. 2016), the adenocarcinoma cell line OVCAR-5 was the only one expressing all myosin VI isoforms, which undergoes myosin VI isoform switch when cultured as monolayers. As none of the other ovarian cell lines tested expressed the long isoform, even when cultured as monolayers, we selected the OVCAR-5 cell line for further characterization.

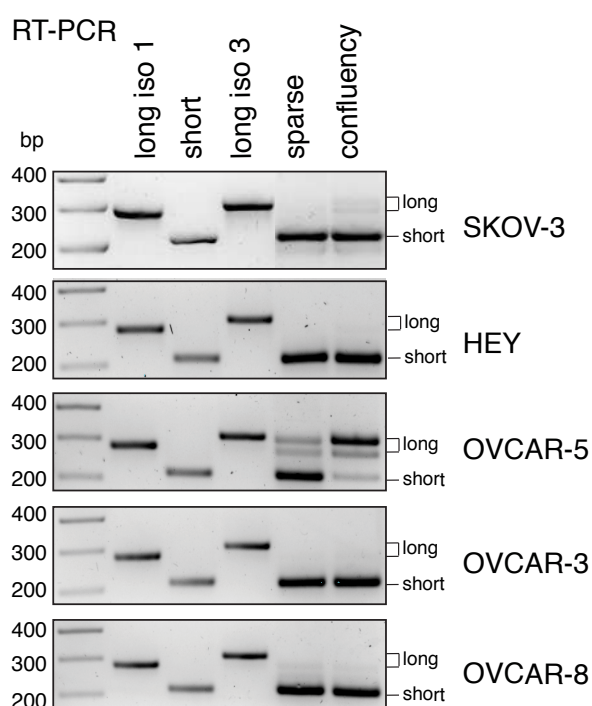


Figure 37. Myosin VI isoforms expression in ovarian cancer cell lines. Ovarian cancer cells grown in sparse conditions were allowed to reach confluency. RT-PCR results of cDNA prepared from the indicated conditions and cell lines are showed.

The positive effect of RBFOX2, as well as the negative effect of SRSF2, on myosin VI AS were confirmed also in OVCAR-5 cells (**Fig. 38**). Upon RBFOX2 KD (**Fig. 38A**), we observed a strong impairment in the expression of myosin VI_{long} (**Fig. 38B**). Instead, KHSRP depletion showed a weaker effect when compared with RBFOX2 (**Fig. 38B**). This effect was also weaker than the one observed in A549 cells (**Fig. 30**), thus indicating that KHSRP action on myosin VI AS could be tissue-specific. Instead, SRSF2 KD (**Fig. 38A**) induced the

expression of myosin VI_{long} in cells grown in sparse conditions (**Fig. 38B**), as previously observed in A549 and Caco-2 cells (**Fig. 25**).

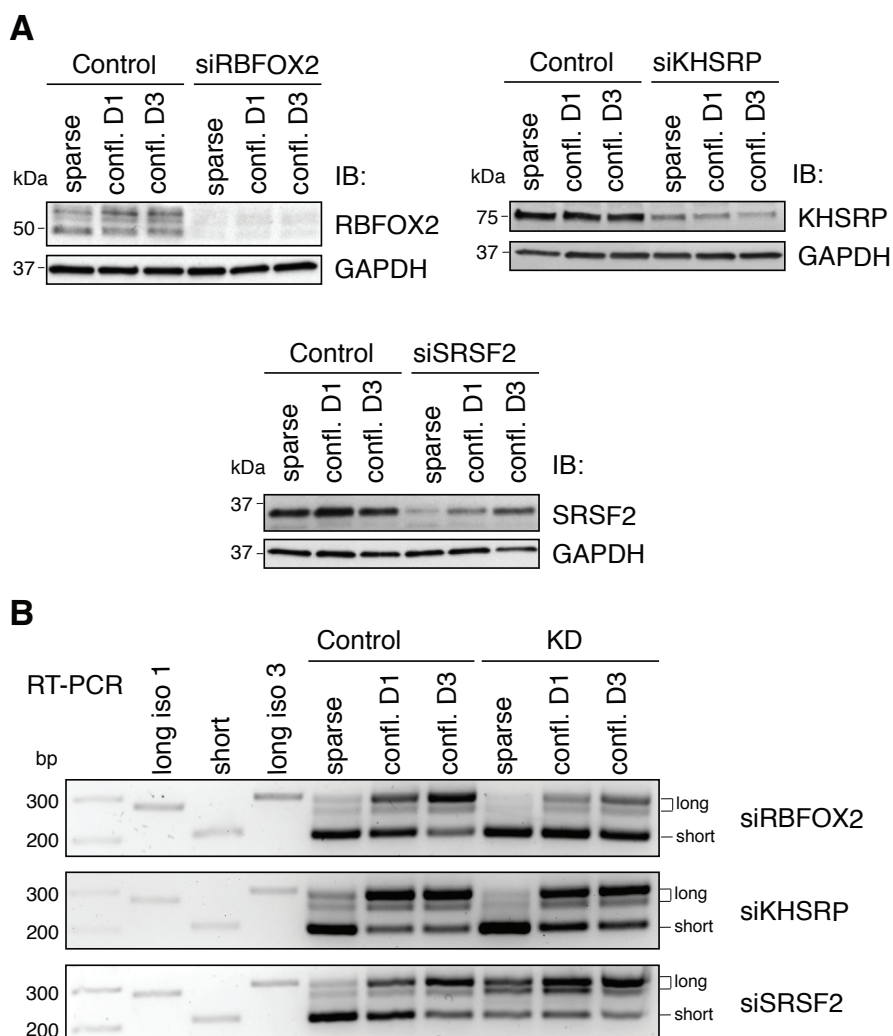


Figure 38. Myosin VI AS regulators are conserved in the ovarian cell line OVCAR-5. A. IB of the indicated proteins from OVCAR-5 cell lysates, control or upon siRNAs transfection. **B.** RT-PCR results of cDNA prepared from A549 RBFox2, KHSRP, or SRSF2-depleted cells grown in sparse conditions and one or three days after that confluency was reached.

5.4.3. SRSF2 depletion restores myosin VI_{long} expression in ovarian cancer cells

As observed in **Figure 37**, the majority of the ovarian cancer cell lines tested express only the myosin VI_{short} isoform; and these cells do not switch to the expression of the long isoform even if they are kept in confluent growth conditions. Thus, we wondered if depletion of SRSF2 was sufficient to restore the expression of the myosin VI_{long} isoform in those cell lines. Strikingly, depletion of SRSF2 in SKOV-3 cells led to the expression of the myosin VI_{long} isoform (**Fig. 39A**). Moreover, using our specific anti-myosin VI_{long}, we were able to confirm the expression of the corresponding long isoform protein also by IB (**Fig. 39B**).

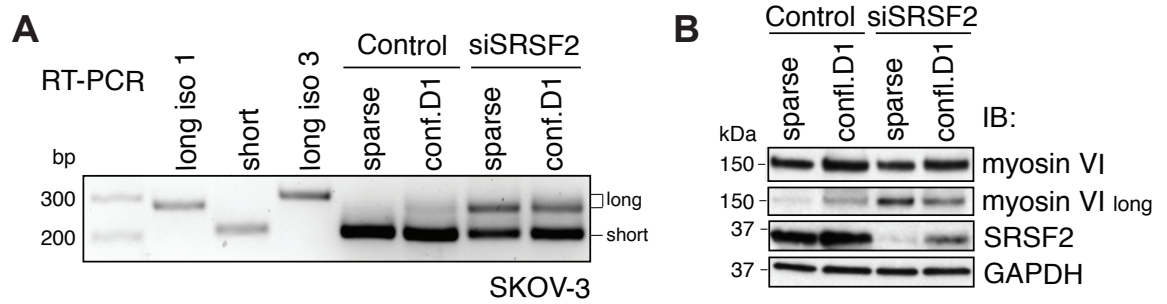


Figure 39. SRSF2 depletion restores myosin VI_{long} expression. **A.** RT-PCR results of cDNA prepared from SKOV-3 SRSF2 KD cells grown in sparse conditions or one day after that confluency was reached. **B.** IB showing SRSF2 depletion and myosin VI long expression in the indicated SKOV-3 cell lysates.

Prompted by the strong effect of SRSF2 on myosin VI alternative splicing, we decided to test all members of the SR protein family in order to address if SRSF2 affects our AS event in a specific manner. In order to identify both positive and negative effects on splicing, we decided to take advantage of OVCAR-5 cells, which express all myosin VI isoforms (**Fig. 37, 38**). Among the different SR proteins that were efficiently depleted as shown by qPCR (**Fig. 40A**), SRSF2 KD was unique as it simultaneously induced the expression of myosin VI_{long} isoform and decreased the expression of myosin VI_{short} (**Fig. 40B**), thus, it might be responsible for the positive selection of the short isoform that occurs in ovarian cancer development.

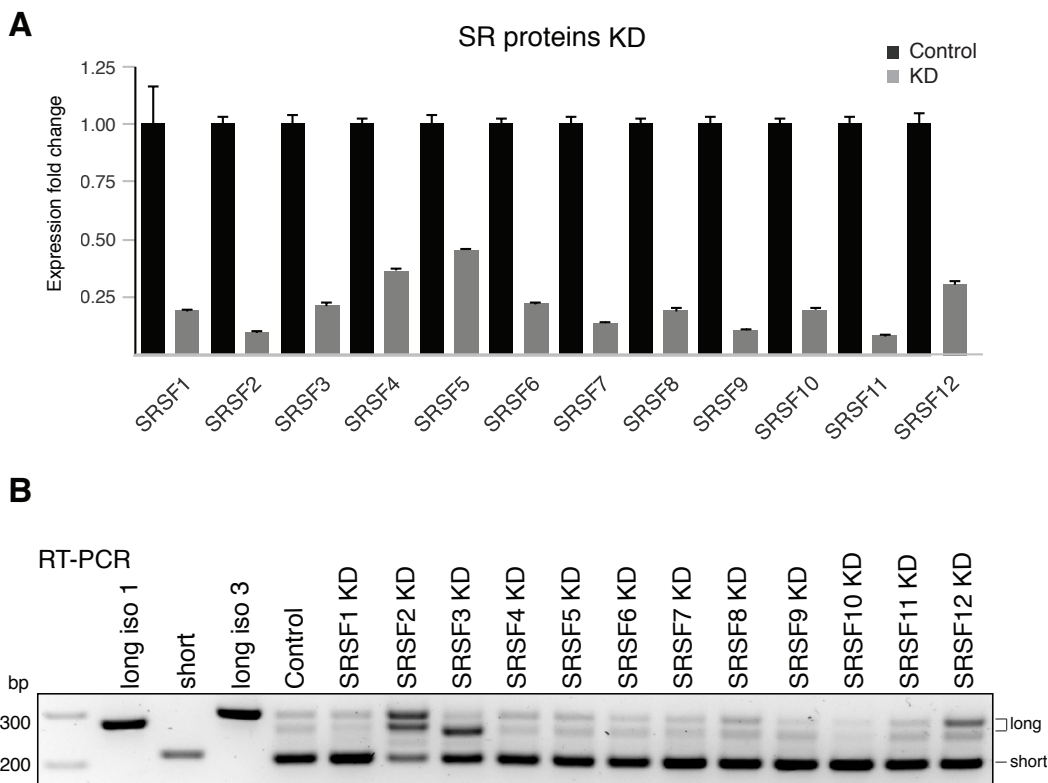


Figure 40. SRSF2 is the only member of the SR protein family that promotes myosin VI_{short} expression. **A.** qPCR showing the downregulation of SR protein transcripts upon siRNA transfection in OVCAR-5 cells. **B.** RT-PCR results of cDNA prepared from OVCAR-5 cells grown sparse.

Since SRSF3 and SRSF12 depletion resulted in a mild induction of the isoforms 1 and 3, respectively, in OVCAR-5 (**Fig. 40**), we decided to deplete the proteins also in SKOV-3 ovarian cells (**Fig. 41A**), which express only myosin VI_{short} (**Fig. 37**). While we confirmed the role of SRSF2 in myosin VI AS by using a second siRNA oligo, neither SRSF3 nor SRSF12 depletion induced the expression of myosin VI_{long}, as shown by RT-PCR (**Fig. 41B**). Thus, SRSF2 appears to have a unique role in myosin VI AS that is not shared by any other member of the SR proteins family.

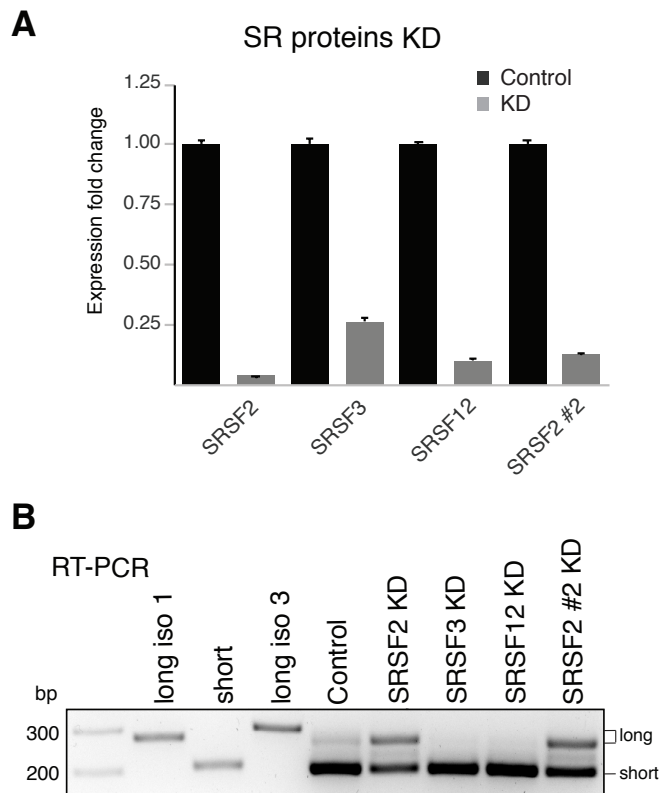


Figure 41. SRSF3 and SRSF12 depletion does not induce the inclusion of myosin VI large insert. A. qPCR showing the downregulation of SR protein transcripts upon siRNA transfection in SKOV-3 cells. **B.** RT-PCR results of cDNA prepared from SKOV-3 cells grown in sparse conditions.

5.4.4. SRSF2 depletion impairs ovarian cancer cell migration

We previously demonstrated that myosin VI_{short} is overexpressed in ovarian primary tumors and in most ovarian cancer cell lines, where it dictates addiction for tumor cell migration (Wollscheid et al. 2016), (**Fig. 2**). Now that we have identified SRSF2 as a key player involved in myosin VI AS, we wondered if its depletion would have an impact on cell migration as well. We hypothesized that SRSF2 KD could affect cell migration by reducing the expression of the myosin VI_{short} isoform, thus mimicking the migratory defect scored upon myosin VI KD and observed in cells expressing selectively myosin VI_{short} (**Fig. 2**). In order to address this scenario, we performed SRSF2 KD in SKOV-3 cells (**Fig. 42 A,B**) and analyzed their motility through wound healing assay (**Fig. 42 C,D**). This experiment was performed by my colleague Carlos Niño. SKOV-3 cells selectively express myosin VI_{short} (**Fig. 42B**) and are deficient in cell migration when myosin VI is depleted (**Fig. 42 C,D**), as previously shown (Wollscheid et al. 2016). As previously observed, SRSF2 KD induced the

expression of myosin VI_{long} (**Fig. 41B**), but more importantly, and as predicted from our hypothesis, knocking down SRSF2 severely inhibited migration and impaired wound closure (**Fig. 42 C,D**). SRSF2 KD showed a stronger phenotype than myosin VI KD, most likely due to SRSF2 acting on multiple splicing events that are possibly related with cell migration. These results might suggest that SRSF2 drive the expression of the myosin VI pro-oncogenic isoform (short), and additional AS events, in ovarian tumors, promoting ovarian cancer cell migration.

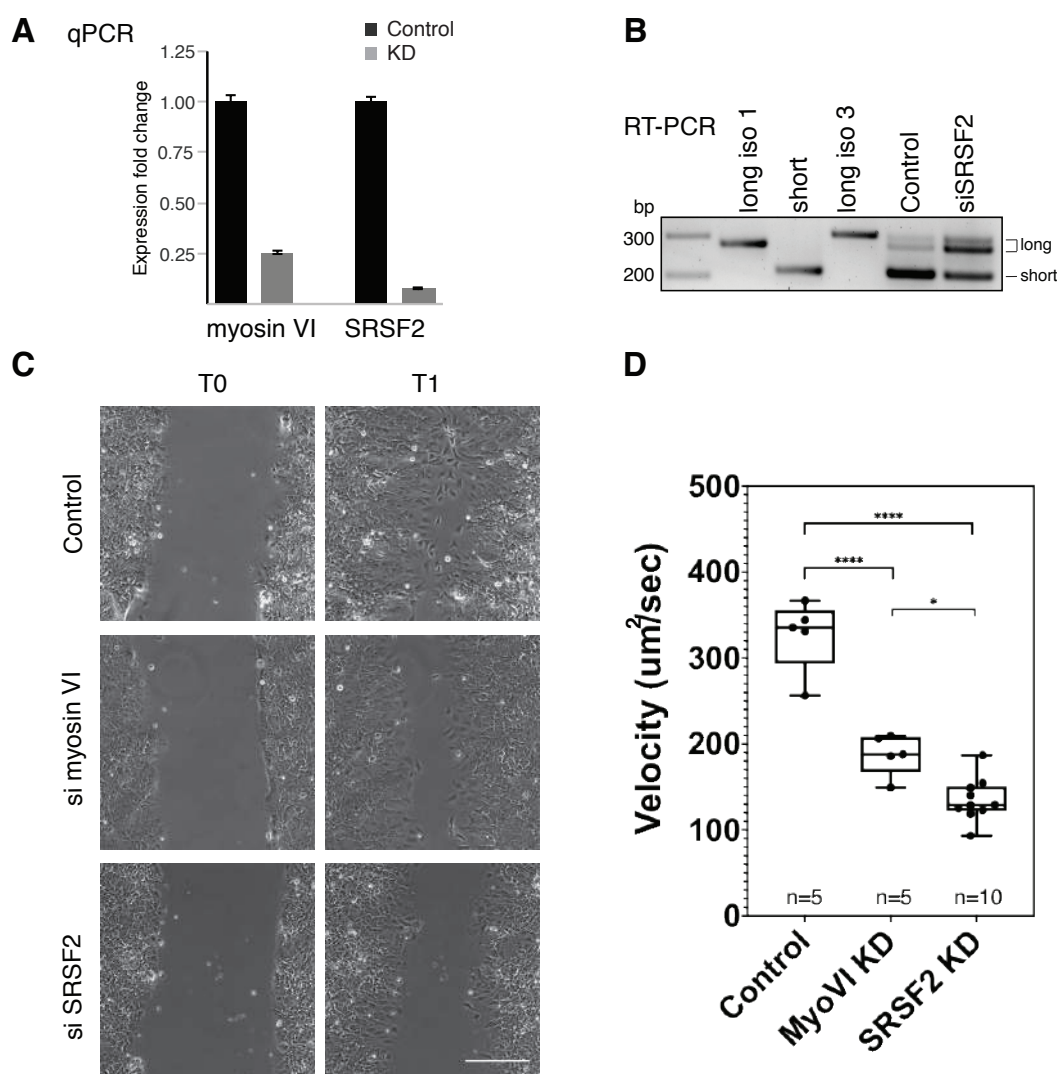


Figure 42. SRSF2 KD impairs ovarian cancer cell migration. **A.** qPCR showing the downregulation of myosin VI and SRSF2 transcripts upon siRNA transfection in SKOV-3 cells. **B.** RT-PCR results of cDNA prepared from SRSF2-depleted SKOV-3 cells 24 hours after wounding. **C.** Wound healing assay. SKOV-3 cells were transiently transfected with siRNA targeting myosin VI or SRSF2 to achieve KD of the proteins. Time lapse images: T₀ first frame, T₁ arbitrary time-point (8 hours after wound) identical for Mock and KDs. Scale bar: 200 μ m. **D.** Quantification of the speed of the cell populations expressed in μ m²/s. n=measurements of different fields from the same cell population.

6. Discussion

Regulation of gene transcription has been established as the major determinant for cellular fate and fitness, but post-transcriptional layers of gene regulation are equally important. It is well established that AS increases the capacity of the human genome by generating multiple protein isoforms from a single gene, and plays an important role in the regulation of gene expression and cellular fate. Indeed, the expression of oncogenic protein isoforms is one of the mechanisms that confers survival advantage to cancer cells (Sveen et al. 2016; Eswaran et al. 2013; Venables et al. 2009). Additionally, AS has been shown to be involved in the development of resistance to targeted therapies, as in the case of v-raf murine sarcoma viral oncogene homolog B1 (BRAF) inhibitors for the treatment of melanoma (Poulikakos et al. 2011).

Despite the increasing interest in these processes, little is known about how extracellular signals are transduced into the nucleus to regulate AS decision. One study showed that activation of MAPK signaling regulates v5 exon inclusion in the CD44 transcript, a process mediated by the RNA binding protein Sam68 (Matter, Herrlich, and König 2002). A broader effect has been attributed to the activation of the EGF signaling pathway, which regulates SR protein activity and induces, as a consequence, a massive AS reprogramming (Zhou et al. 2012). It is predicted that a better knowledge of the signaling pathways involved in the choice of cancer specific isoforms would lead to the identification of novel potential therapeutic targets.

Our laboratory identified the short isoform of myosin VI (myosin VI_{short}) as a cancer specific isoform (Wollscheid et al. 2016). In this thesis, we have established a simple way to study the AS switch between the two myosin VI isoforms, which has allowed us to identify several components of its regulation.

In our cellular systems, A549 and Caco-2 cells, myosin VI_{short} is the predominant isoform when cells are grown in low-density conditions and present mesenchymal characteristics (**Fig. 12 and 13**). Epithelial tissue displays a cuboidal morphology and presents a polar organization maintained by tight cell-cell interconnections. Mesenchymal cells lack these features and display higher motility and invasiveness. Interestingly, the expression of myosin VI_{short} is strongly reduced when cells start to organize into cellular monolayers, mimicking the architecture of a tissue. Once the cells are fully organized in a monolayer, they express almost exclusively the long isoform of myosin VI. The process is totally reversible as the dissociation of confluent cells and re-plating at low density switches back the myosin VI_{long} expression to myosin VI_{short}. These cellular models function as a proxy of

EMT and MET and mimic the events occurring during cancer development, allowing us to use myosin VI as an AS paradigmatic event for the identification of the regulatory signals involved in AS reprogramming during cancer progression.

6.1. E-cadherin and β -catenin are determinants for myosin VI isoform switch

We first analyzed the involvement of adhesion and polarity factors in our AS model and we uncovered a key role for E-cadherin in the regulation of myosin VI isoform choice (**Fig. 16B** and **17B**). To our knowledge, this is the first evidence of a role for E-cadherin mediated cell-to-cell contacts in the regulation of splicing. Interestingly, myosin VI and E-cadherin co-localize during the morphogenesis of epithelial cell-to-cell contacts with evidences of a functional link between these two proteins (Maddugoda et al. 2007). E-cadherin is part of the AJ, where it binds β -catenin and p120 (Niño, Sala, and Polo 2019; Meng and Takeichi 2009). β -catenin, instead, plays a role not only at the AJ, but also in the nucleus, where, through the association with different transcription factors (e.g., TCF family) it regulates gene transcription (Cadigan and Waterman 2012).

It has been shown that E-cadherin can affect the downstream β -catenin signaling to the nucleus in a cell-density dependent manner (Maher et al. 2009). In confluent cells, E-cadherin sequesters β -catenin at cell junctions and promotes a fast turnover of the protein; while in sub-confluent status this effect is reduced and active β -catenin migrates to the nucleus where it acts as a transcriptional activator (Maher et al. 2009).

We were able to reproduce this observation in our cellular systems where the level of nuclear active β -catenin strongly decreases when cells are grown as monolayers. Strikingly, this decrease was not observed in monolayers of E-cadherin-depleted cells where we observed the persistence of myosin VI_{short} expression (**Fig. 19**). We thus wondered whether the impairment of myosin VI isoform switch observed after E-cadherin KD was due to the persistence of β -catenin in the nucleus. As predicted by our hypothesis, β -catenin depletion induces the expression of myosin VI_{long} in sparse condition in Caco-2 cells (**Fig. 20**). Unfortunately, we did not observe the same effect in A549 cells, possibly due to the fact that they might respond to β -catenin downregulation by activating additional signaling pathways in order to maintain their mesenchymal state. We are currently testing our concept by overexpressing an N-terminal truncated constitutively active mutant of β -catenin (Barth, Stewart, and Nelson 1999; Barth et al. 1997). If our hypothesis is correct, the expression of this nuclear and constitutively active β -catenin form would promote the expression of myosin VI_{short} even in cell monolayers.

Our data raise the idea of a functional role of β -catenin in AS regulation and scattered evidence sustains it. β -catenin/TCF4 signaling in colorectal cancer cells induces the upregulation of the SR-protein SRSF3, which promotes the expression of the tumor-associated splice variant CD44E (Gonçalves, Matos, and Jordan 2008). Stable S37A β -catenin overexpression in HCT116 colorectal cancer cells promotes the expression of the $\Delta 5-6$ isoform of the receptor ER- β that lacks the ligand binding domain and is implicated in tumorigenesis (Lee et al. 2006). β -catenin was also found to directly bind different splicing regulators, among which FUS in DLD1 colorectal cancer cells (Sato et al. 2005). To identify AS events regulated by β -catenin, we performed a high-throughput RNA-sequencing, using RNA extracted from Caco-2 cells grown in sparse and confluent conditions, and in conditions of β -catenin depletion. Analysis is ongoing, but visual inspection of myosin VI AS cassette confirmed the data shown here. In the long run, we expect to obtain a complete picture of the AS governed by β -catenin and to identify its downstream effectors, which may be the ones identified as myosin VI AS regulators SRSF2, RBFOX2 and KHSRP.

6.2. RBFOX2 and KHSRP promote the expression of myosin VI_{long}

To identify RNA-binding proteins involved in the myosin VI AS, we took advantage of an Affymetrix gene expression analysis coupled with siRNA depletion that allowed the identification of RBFOX2, KHSRP and SRSF2 as key determinants for myosin VI AS decision. RBFOX2 and KHSRP exert a positive effect on myosin VI AS, stimulating the inclusion of the large insert and thus the expression of myosin VI_{long} (**Fig. 26** and **Fig. 30**). Interestingly, the activity of the splicing factors appears to be cell type-specific. The effect of RBFOX2 is stronger in the ovarian cell line OVCAR-5 (**Fig. 38**), while that of KHSRP seems to be predominant in the lung adenocarcinoma A549 cells (**Fig. 30**), reflecting the tissue specificity of splicing regulation (Baralle and Giudice 2017).

Notably, we identified two UGCAUG intronic consensus motifs for RBFOX proteins (Jin et al. 2003), both localizing in proximity to intron-exon boundaries in the myosin VI AS region (**Fig. 33**). To evaluate RBFOX2 binding to these sites, we performed UV-RIP analysis on A549 cells and validated both sites (**Fig. 34** and **35**). Thus, our data are in line with the position-dependent effect of RBFOX, whereby this splicing regulator binds downstream of the alternative-spliced exons (in our case exons 29 and 31) and stimulates their inclusion (Yeo et al. 2009).

The role of RBFOX2 in cancer is controversial. By knocking down 78 potential splicing factors in five different cell lines, RBFOX2 was identified as a driver of the mesenchymal splicing signature (Venables et al. 2013). In the same study, a signature of 43 exons, differentially spliced in epithelial versus mesenchymal samples in human colon and ovary samples, was identified among 283 alternatively spliced cassette exons (Venables et al. 2013). Unfortunately, myosin VI AS was not included in the analysis. In a different study, RBFOX2 was shown to regulate EMT-associated skipped exons in Twist-related protein 1 (Twist)-HMLE transformed cells. RBFOX2 depletion in Twist-HMLE transformed cells shifted the cell morphology from spindle-shaped to cobblestone-like, a feature of epithelial cells (Shapiro et al. 2011). Additionally, TGF- β -induced EMT upregulates RBFOX2 in NMuMG and PY2T murine cells, leading to increased expression of the RBFOX2-dependent isoforms of cortactin, Pard3 and dynamin2. Depletion of RBFOX2 does not revert the induction of EMT, but reduces the invasive potential of the cells (Braeutigam et al. 2014). Induction of EMT by TGF- β administration causes, in different cell lines, the RBFOX2-dependent exclusion of the TGF- β -activated kinase 1 (TAK1) variable exon 12, generating a constitutively active short isoform of this kinase (Tripathi et al. 2019).

On the other hand, RBFOX2 has also been associated with the maintenance of an epithelial phenotype. Like for myosin VI, RBFOX2 regulates AS of FGFR2 and promotes skipping of the mesenchymal exon IIIc and the inclusion of the epithelial exon IIIb. Moreover, RBFOX2 is involved in promoting a switch from FGFR2 exon IIIc to IIIb during MET when 293T cells are maintained in confluent growth conditions (Baraniak, Chen, and Garcia-Blanco 2006).

By comparing alternative splicing events from ovarian or breast tumors with their corresponding normal tissues, it was possible to identify about 100 ASE classified as cancer-associated events (Venables et al. 2009). Interestingly, several skipped exons in these tumors were enriched for RBFOX binding motif in the downstream intron and are affected by RBFOX2 depletion, like in the case of myosin VI. RBFOX2 was downregulated in ovarian cancer samples, while a selective expression of a shorter RBFOX2 inactive isoform was identified in breast tumors, indicating that loss of RBFOX2 splicing function is associated with tumorigenesis (Venables et al. 2009). Based on this last study and our results, we hypothesize that the selective expression of myosin VI_{short} observed in ovarian tumors could be linked to RBFOX2 downregulation. To test this idea, we are currently setting-up immunohistochemistry analysis in collaboration with the pathologist Claudio Tripodo at IFOM to evaluate the expression of RBFOX2 in ovarian primary tumors.

In our cellular systems, RBFOX2 promotes the expression of myosin VI long isoform, which can be considered an epithelial splice variant. However, RBFOX2 overexpression has no effect on myosin VI AS, indicating that this protein is not a limiting factor and that more regulators are implicated in the isoform choice. Indeed, we found that also KHSRP promotes the inclusion of the large insert (**Fig. 30**). KHSRP is an RNA binding protein containing four KH domains originally identified as both a transcription factor, namely Far-upstream Element-binding Protein 2 (FBP2) (Davis-Smyth et al. 1996), a splicing regulator (Min et al. 1997) and microRNA (miRNA) maturation factor (Briata et al. 2016; Gherzi et al. 2014). The splicing function of KHSRP has not been extensively investigated. KHSRP binds c-SRC pre-mRNA, stimulating the inclusion of the neural specific N1 exon (Hall, Huang, and Black 2004; Min et al. 1997). Moreover, KHSRP, together with hnRNP H, was found to be involved in the alternative splicing of the RPL3 gene (Russo et al. 2011). Instead, a higher number of studies investigated the splicing function of the KHSRP orthologue in *Drosophila*, P-element somatic inhibitor (PSI); in flies this protein represses splicing of the P-element third intron IVS3 in somatic cells (Sun et al. 2008; Labourier et al. 2002; Adams, Tarnig, and Rio 1997; Siebel, Admon, and Rio 1995; Siebel, Kanaar, and Rio 1994).

Interestingly, KHSRP has been described as a negative regulator of EMT. Upon EMT induction by TGF- β administration in NMuMg cells, KHSRP expression is silenced through the induction of miR-27b-3p. Importantly, KHSRP depletion phenocopies the effects of TGF- β administration in NMuMg cells (Puppo et al. 2016). KHSRP counteracts TGF- β -induced EMT, acting both on miRNA maturation (miR-192-5p prevents the expression of EMT transcription factors like Zinc finger E-box binding homeobox (ZEB) 1/2) and on alternative splicing. RNAseq analyses performed upon KHSRP overexpression in NMuMg cells identified 763 KHSRP-dependent alternative splicing events, mainly affecting proteins involved in cytoskeleton organization, cell junction assembly and GTPase-mediated signal transduction. Myosin VI was not identified among the targets that include the epithelial isoforms of CD44, FGFR2 and Protein enabled homolog (ENAH) (Puppo et al. 2016).

KHSRP binds its targets through four KH domains that constitute a central nucleic acid binding region, and this region is flanked on both sides by flexible regions containing motifs that are possibly involved in the interaction with other proteins (Gherzi et al. 2014). The KH domains are defined by a ~70 amino acids sequence that folds into an $\alpha\beta\alpha\beta\alpha$ structure (Musco et al. 1996). Each of these domains recognizes a short RNA sequence with low affinity, thus RNA recognition of diverse targets is likely to be mediated by a combinatorial

action of the different domains (Garcia-Mayoral et al. 2007). Among the different KH domains, KH2 preferentially binds AU-rich sequences, KH1 and KH3 prefer G and G/U rich sequences, while the KH4 domain has a mild preference for GA-rich motifs (Garcia-Mayoral et al. 2008; Garcia-Mayoral et al. 2007). In silico analysis of KHSRP eCLIP data revealed that the protein binds myosin VI large insert in two different regions (**Fig. 33B**). The first region bound by KHSRP, upstream of myosin VI exon 29, corresponds to a UGGGU motif, G-rich target sequences that are usually bound by KHSRP (Garcia-Mayoral et al. 2008). The second region appear to be the same as that of RBFOX2, probably corresponding to the same UGCAUG motif that has been shown to be bound also by KHSRP (Markovtsov et al. 2000). While further experiments are needed to consolidate these data, RBFOX2 and KHSRP likely cooperate to promote the inclusion of the large insert of myosin VI.

6.3. SRSF2 promotes the expression of myosin VI_{short}

SRSF2, also known as SC35, is a 35 kD member of the SR proteins and was first identified using a monoclonal antibody directed against partially purified spliceosomes (Fu and Maniatis 1990). SRSF2 possesses a single N-terminal RNA recognition domain and a prototypical C-terminal RS domain. To regulate both constitutive and alternative splicing, the RRM specifically binds cis-regulatory elements on the pre-mRNA both in exons and in introns, even if binding to exonic splicing enhancers has been the most studied mechanism (Hallay et al. 2006; Zahler et al. 2004; Zhu, Mayeda, and Krainer 2001).

In our myosin VI AS system, SRSF2 acts as a negative regulator. It has been reported that SR proteins promote skipping of an alternative exon by binding to a flanking constitutive exon and reinforcing its recognition (Han et al. 2011). SRSF2-excluded exons were shown to be associated with the enrichment of the TCCTG motif in the flanking constitutive exon (Luo et al. 2017). Interestingly, the same motif is present in the constitutive exon 32 of myosin VI, which is right downstream of the AS region, leading to the hypothesis that SRSF2 binding to this constitutive flanking exon might stimulate the skipping of the large insert. Unfortunately, the lack of a good antibody for immunoprecipitation precluded UV-RIP analysis and, possibly for the same reason, no eCLIP data is available for SRSF2. We are currently evaluating the possibility of using a GFP-tagged version to circumvent this issue and validate our model.

Since SRSF2 works as negative regulator, it might be inactivated once confluency is reached, allowing exon recognition and the inclusion of the large insert. Although SRSF2 mRNA levels were reduced when cells reached confluency, the protein levels remained

stable (**Fig. 25B**), suggesting that other mechanisms, like post-translational modifications might be involved in the reduction of protein activity when confluency is reached. As previously mentioned, SR proteins undergo extensive phosphorylation in their SR domains on serine-arginine repeats by SRPKs and CLKs (Zhong et al. 2009; Yomoda et al. 2008; Hayes et al. 2006; Fukuhara et al. 2006). Additionally, SRSF2 was also found to be phosphorylated by the serine/threonine kinase GSK3- β in the RS domain through an *in vitro* kinase assay (Hernandez et al. 2004). Moreover, SRSF2 is regulated also by an acetylation on lysine 52, which is mediated by the acetyltransferase Tat-interactive protein 60 (TIP60), promoting the degradation of the protein. This activity is counteracted by Histone deacetylases 6 (HDAC6) that, instead, stabilizes the protein (Edmond et al. 2011). This kind of post-translational modifications might change SRSF2 affinity for myosin VI pre-mRNA or alter its capacity to interact with other partners in sparse versus confluent conditions.

A potential role for SRSF2 in cancer development has been investigated in hematological malignancies. SRSF2 frequency of mutations reaches 28-47% in myelodysplastic syndromes and most of the alterations occur at proline 95, which is often mutated into histidine (P95H) (Wassie et al. 2014; Patnaik et al. 2013; Meggendorfer et al. 2012; Makishima et al. 2012). As demonstrated in SRSF2-P95H mutant mice, P95H mutation alters SRSF2 normal RNA binding specificity, giving rise to an aberrant gain of function property for the mutant protein (Kim et al. 2015).

SRSF2 was also found to be overexpressed and hyperphosphorylated in 65% of lung adenocarcinomas and in 91% of squamous cell lung carcinomas, correlating with larger tumor size and advanced stage (Gout et al. 2012). In the same cell types, SRSF2 can behave also as a tumor suppressor. Indeed, SRSF2 supports apoptosis induced by cisplatin (Edmond et al. 2011) and DNA-damaging agents by shifting the splicing pattern of caspase-8, caspase-9, Flip and Bcl-x genes in favor of pro-apoptotic splice variants (Merdzhanova et al. 2008). This mechanism probably explains why we failed in generating cell lines in which SRSF2 was overexpressed. It is tempting to speculate that during lung adenocarcinoma development, cancer cells have to overcome this SRSF2-dependent apoptosis. SRSF2 is also upregulated in about 89% of hepatocellular carcinoma samples analyzed by immunohistochemistry, an upregulation that is associated with the expression of cancer-related splice variants and with poor prognosis (Luo et al. 2017).

Only a few studies have been performed on ovarian cancers and which have showed that SRSF2 can act as both an oncogene and a tumor suppressor. On the one hand, SRSF2 has been described as pro-apoptotic in epithelial ovarian cancer cells; in this context, after platinum-based chemotherapy, SRSF2 binding to caspase 9 RNA is reduced, determining resistance to apoptosis (Pellarin et al. 2020). On the other hand, differently from its behavior in lung cancer (Edmond et al. 2011), in p53 positive ovarian cancer cells, SRSF2 has been implicated in cisplatin resistance by cooperating with the non-coding RNA Promoter Of CDKN1A Antisense DNA Damage Activated RNA (PANDAR) and preventing p53 Ser15 phosphorylation (Wang et al. 2018). Again, this dual role as an oncogene or a tumor suppressor underlines the complexity of AS and its actors.

In the case of myosin VI AS, SRSF2 acts as an oncogene, promoting the expression of the short isoform, which is positively selected by ovarian cancer cells to support cell migration (Wollscheid et al. 2016). As we previously described, the AS of myosin VI is deregulated in ovarian cancer, where cells, by losing the expression of the epithelial isoform myosin VI_{long}, become addicted to the expression of myosin VI_{short} for tumor cell migration (Wollscheid et al. 2016). While our initial screening was performed in Caco-2 and A549, we confirmed the effect of RBFOX2 and SRSF2 depletion in an ovarian cancer cell line, OVCAR-5 (**Fig. 38**), indicating that the mechanism of regulation we discovered is highly conserved. Importantly, by silencing SRSF2, we were able to promote the expression of myosin VI_{long} in SKOV-3, an ovarian cancer cell line which exclusively expresses myosin VI_{short} (**Fig. 39**), further confirming that SRSF2 is a determinant player in the selective expression of myosin VI_{short} observed in tumors. Interestingly, SRSF2 depletion in this cell line is also able to slow down cell migration, thus supporting the idea that AS de-regulation contributes to cancer metastasis (**Fig. 42**).

6.4. Future perspectives

As a long-term goal, we intend to determine whether SRSF2, RBFOX2 and KHSRP influence splicing outputs are regulated by E-cadherin/ β -catenin signaling. A potential scenario for such regulation is depicted in **Figure 43** and relies on the availability of active β -catenin. By a mechanism yet to be defined, the absence of active nuclear β -catenin in cell monolayers would inactivate SRSF2 binding to myosin VI pre-mRNA, allowing the recognition of the exon cassette by the spliceosome (modulated by KHSRP and RBFOX2) and thus the inclusion of the large insert (**Fig. 43**). Since, in our system, there are no major changes in the protein levels of RBFOX2 and SRSF2, it is likely that their ability to bind

myosin VI pre-mRNA in sparse versus confluent conditions may depend on post-translational modifications that may change localization and/or interaction with additional proteins. This speculative model needs to be addressed with the help of biochemical approaches and other tools that we are generating.

Another long-term goal of this project will be to shed light on the AS reprogramming during tumorigenesis by looking at genes that share a regulation similar to the one of myosin VI. In this direction, we intend to determine additional targets of SRSF2 by evaluating global isoform gene expression profiles after SRSF2 downregulation by RNAseq, like in the case of β -catenin.

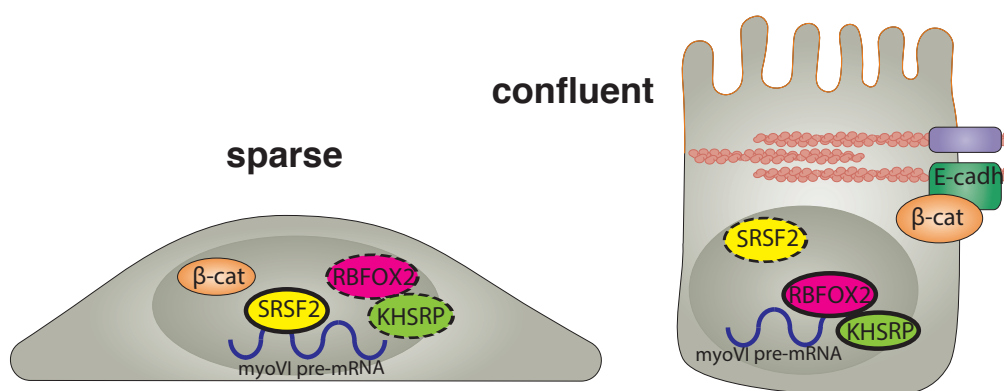


Figure 43. Hypothetical model for myosin VI AS. In cells growing in sparse conditions, β -catenin is transcriptionally active in the nucleus; myosin VI pre-mRNA is bound by SRSF2, which induces skipping of the large insert. In cell monolayers, β -catenin is sequestered at AJ while, in the nucleus, SRSF2 binding to myosin VI pre-mRNA is disrupted, and RBFOX2 and KHSRP facilitate exon recognition, allowing thus the inclusion of the large insert.

Acknowledgement

I would like to thank all the people who made this work possible.

Above all, I would like to thank my supervisor, Prof. Simona Polo, for giving me the opportunity to join her lab and for her constant mentoring during these four years. Her enthusiasm has always been a great inspiration for all the lab.

I also would like to thank Carlos Niño, a brilliant postdoc I have the fortune to work with, for sharing with me the project and his knowledge, and for training me.

I am grateful to all the members of Polo group for the scientific and human support, for being always helpful and kind. A special thanks to our staff scientist Elena Maspero for always having all the answers to our questions.

Moreover, I have to thank AIRC for the trust they put in supporting our project.

I would like to thank all my family, in particular my loving parents, for always supporting my choices unconditionally. I am also grateful to all my friends, especially for their ability to see in me what I often fail to see.

Last, but not least, I want to thank my partner Francesco. The perfect phrase to describe how he supported me in these years of PhD is from my favourite author, JRR Tolkien:

“I can't carry it for you, but I can carry you.”

7. References

1. Acharya, B. R., A. Nestor-Bergmann, X. Liang, S. Gupta, K. Duszyc, E. Gauquelin, G. A. Gomez, S. Budnar, P. Marcq, O. E. Jensen, Z. Bryant, and A. S. Yap. 2018. 'A Mechanosensitive RhoA Pathway that Protects Epithelia against Acute Tensile Stress', *Dev Cell*, 47: 439-52.e6.
2. Adams, M. D., R. S. Tarnag, and D. C. Rio. 1997. 'The alternative splicing factor PSI regulates P-element third intron splicing in vivo', *Genes Dev*, 11: 129-38.
3. Aguilera, O., M. Munoz-Sagastibelza, B. Torrejon, A. Borrero-Palacios, L. Del Puerto-Nevado, J. Martinez-Useros, M. Rodriguez-Remirez, S. Zazo, E. Garcia, M. Fraga, F. Rojo, and J. Garcia-Foncillas. 2016. 'Vitamin C uncouples the Warburg metabolic switch in KRAS mutant colon cancer', *Oncotarget*, 7: 47954-65.
4. Alajati, A., N. Sausgruber, N. Aceto, S. Duss, S. Sarret, H. Voshol, D. Bonenfant, and M. Bentires-Alj. 2013. 'Mammary tumor formation and metastasis evoked by a HER2 splice variant', *Cancer Res*, 73: 5320-7.
5. Alsafadi, S., A. Houy, A. Battistella, T. Popova, M. Wassef, E. Henry, F. Tirode, A. Constantinou, S. Piperno-Neumann, S. Roman-Roman, M. Dutertre, and M. H. Stern. 2016. 'Cancer-associated SF3B1 mutations affect alternative splicing by promoting alternative branchpoint usage', *Nat Commun*, 7: 10615.
6. Amir-Ahmady, B., P. L. Boutz, V. Markovtsov, M. L. Phillips, and D. L. Black. 2005. 'Exon repression by polypyrimidine tract binding protein', *Rna*, 11: 699-716.
7. Anczukow, O., and A. R. Krainer. 2016. 'Splicing-factor alterations in cancers', *Rna*, 22: 1285-301.
8. Anczukow, O., A. Z. Rosenberg, M. Akerman, S. Das, L. Zhan, R. Karni, S. K. Muthuswamy, and A. R. Krainer. 2012. 'The splicing factor SRSF1 regulates apoptosis and proliferation to promote mammary epithelial cell transformation', *Nat Struct Mol Biol*, 19: 220-8.
9. Angiolini, F., E. Belloni, M. Giordano, M. Campioni, F. Forneris, M. P. Paronetto, M. Lupia, C. Brandas, D. Pradella, A. Di Matteo, C. Giampietro, G. Jodice, C. Luise, G. Bertalot, S. Freddi, M. Malinverno, M. Irimia, J. D. Moulton, J. Summerton, A. Chiapparino, C. Ghilardi, R. Giavazzi, D. Nyqvist, D. Gabellini, E. Dejana, U. Cavallaro, and C. Ghigna. 2019. 'A novel L1CAM isoform with angiogenic activity generated by NOVA2-mediated alternative splicing', *Elife*, 8.
10. Au, J. S., C. Puri, G. Ihrke, J. Kendrick-Jones, and F. Buss. 2007. 'Myosin VI is required for sorting of AP-1B-dependent cargo to the basolateral domain in polarized MDCK cells', *J Cell Biol*, 177: 103-14.
11. Baralle, F. E., and J. Giudice. 2017. 'Alternative splicing as a regulator of development and tissue identity', *Nat Rev Mol Cell Biol*, 18: 437-51.
12. Baraniak, A. P., J. R. Chen, and M. A. Garcia-Blanco. 2006. 'Fox-2 mediates epithelial cell-specific fibroblast growth factor receptor 2 exon choice', *Mol Cell Biol*, 26: 1209-22.
13. Barash, Y., J. A. Calarco, W. Gao, Q. Pan, X. Wang, O. Shai, B. J. Blencowe, and B. J. Frey. 2010. 'Deciphering the splicing code', *Nature*, 465: 53-9.
14. Barth, A. I., A. L. Pollack, Y. Altschuler, K. E. Mostov, and W. J. Nelson. 1997. 'NH2-terminal deletion of beta-catenin results in stable colocalization of mutant beta-catenin with adenomatous polyposis coli protein and altered MDCK cell adhesion', *J Cell Biol*, 136: 693-706.
15. Barth, A. I., D. B. Stewart, and W. J. Nelson. 1999. 'T cell factor-activated transcription is not sufficient to induce anchorage-independent growth of epithelial cells expressing mutant beta-catenin', *Proc Natl Acad Sci U S A*, 96: 4947-52.
16. Bates, D. O., T. G. Cui, J. M. Doughty, M. Winkler, M. Sugiono, J. D. Shields, D. Peat, D. Gillatt, and S. J. Harper. 2002. 'VEGF165b, an inhibitory splice variant of vascular endothelial growth factor, is down-regulated in renal cell carcinoma', *Cancer Res*, 62: 4123-31.

17. Bates, D. O., J. C. Morris, S. Oltean, and L. F. Donaldson. 2017. 'Pharmacology of Modulators of Alternative Splicing', *Pharmacol Rev*, 69: 63-79.
18. Batters, C., D. Brack, H. Ellrich, B. Averbeck, and C. Veigel. 2016. 'Calcium can mobilize and activate myosin-VI', *Proc Natl Acad Sci U S A*, 113: E1162-9.
19. Beausoleil, E., C. Chauvignac, T. Taverne, S. Lacombe, L. Pognante, B. Leblond, D. Pallares, C. D. Oliveira, F. Bachelot, R. Carton, H. Peillon, S. Coutadeur, V. Picard, N. Lambeng, L. Desire, and F. Schweighoffer. 2009. 'Structure-activity relationship of isoform selective inhibitors of Rac1/1b GTPase nucleotide binding', *Bioorg Med Chem Lett*, 19: 5594-8.
20. Belloni, E., A. Di Matteo, D. Pradella, M. Vacca, C. D. R. Wyatt, R. Alfieri, A. Maffia, S. Sabbioneda, and C. Ghigna. 2019. 'Gene Expression Profiles Controlled by the Alternative Splicing Factor Nova2 in Endothelial Cells', *Cells*, 8.
21. Berget, S. M. 1995. 'Exon recognition in vertebrate splicing', *J Biol Chem*, 270: 2411-4.
22. Berget, S. M., C. Moore, and P. A. Sharp. 1977. 'Spliced segments at the 5' terminus of adenovirus 2 late mRNA', *Proc Natl Acad Sci U S A*, 74: 3171-5.
23. Biamonti, G., L. Infantino, D. Gaglio, and A. Amato. 2019. 'An Intricate Connection between Alternative Splicing and Phenotypic Plasticity in Development and Cancer', *Cells*, 9.
24. Biancospino, M., G. R. Buel, C. A. Nino, E. Maspero, R. Scotto di Perrotolo, A. Raimondi, L. Redlingshofer, J. Weber, F. M. Brodsky, K. J. Walters, and S. Polo. 2019. 'Clathrin light chain A drives selective myosin VI recruitment to clathrin-coated pits under membrane tension', *Nat Commun*, 10: 4974.
25. Bielli, P., R. Busa, S. M. Di Stasi, M. J. Munoz, F. Botti, A. R. Kornblihtt, and C. Sette. 2014. 'The transcription factor FBI-1 inhibits SAM68-mediated BCL-X alternative splicing and apoptosis', *EMBO Rep*, 15: 419-27.
26. Biselli-Chicote, P. M., J. M. Biselli, B. R. Cunha, R. Castro, J. V. Maniglia, D. S. Neto, E. H. Tajara, J. F. Gois Filho, E. E. Fukuyama, E. C. Pavarino, and E. M. Goloni-Bertollo. 2017. 'Overexpression of Antiangiogenic Vascular Endothelial Growth Factor Isoform and Splicing Regulatory Factors in Oral, Laryngeal and Pharyngeal Squamous Cell Carcinomas', *Asian Pac J Cancer Prev*, 18: 2171-77.
27. Black, D. L. 2003. 'Mechanisms of alternative pre-messenger RNA splicing', *Annu Rev Biochem*, 72: 291-336.
28. Black, D. L., and J. A. Steitz. 1986. 'Pre-mRNA splicing in vitro requires intact U4/U6 small nuclear ribonucleoprotein', *Cell*, 46: 697-704.
29. Bland, C. S., E. T. Wang, A. Vu, M. P. David, J. C. Castle, J. M. Johnson, C. B. Burge, and T. A. Cooper. 2010. 'Global regulation of alternative splicing during myogenic differentiation', *Nucleic Acids Res*, 38: 7651-64.
30. Boise, L. H., M. Gonzalez-Garcia, C. E. Postema, L. Ding, T. Lindsten, L. A. Turka, X. Mao, G. Nunez, and C. B. Thompson. 1993. 'bcl-x, a bcl-2-related gene that functions as a dominant regulator of apoptotic cell death', *Cell*, 74: 597-608.
31. Bond, L. M., A. A. Peden, J. Kendrick-Jones, J. R. Sellers, and F. Buss. 2011. 'Myosin VI and its binding partner optineurin are involved in secretory vesicle fusion at the plasma membrane', *Mol Biol Cell*, 22: 54-65.
32. Bonnal, S., C. Martinez, P. Forch, A. Bachi, M. Wilm, and J. Valcarcel. 2008. 'RBM5/Luca-15/H37 regulates Fas alternative splice site pairing after exon definition', *Mol Cell*, 32: 81-95.
33. Bonomi, S., A. di Matteo, E. Buratti, D. S. Cabianca, F. E. Baralle, C. Ghigna, and G. Biamonti. 2013. 'HnRNP A1 controls a splicing regulatory circuit promoting mesenchymal-to-epithelial transition', *Nucleic Acids Res*, 41: 8665-79.

34. Bonomi, S., S. Gallo, M. Catillo, D. Pignataro, G. Biamonti, and C. Ghigna. 2013. 'Oncogenic alternative splicing switches: role in cancer progression and prospects for therapy', *Int J Cell Biol*, 2013: 962038.
35. Bourgeois, C. F., F. Lejeune, and J. Stevenin. 2004. 'Broad specificity of SR (serine/arginine) proteins in the regulation of alternative splicing of pre-messenger RNA', *Prog Nucleic Acid Res Mol Biol*, 78: 37-88.
36. Boutz, P. L., P. Stoilov, Q. Li, C. H. Lin, G. Chawla, K. Ostrow, L. Shiue, M. Ares, Jr., and D. L. Black. 2007. 'A post-transcriptional regulatory switch in polypyrimidine tract-binding proteins reprograms alternative splicing in developing neurons', *Genes Dev*, 21: 1636-52.
37. Braeutigam, C., L. Rago, A. Rolke, L. Waldmeier, G. Christofori, and J. Winter. 2014. 'The RNA-binding protein Rbfox2: an essential regulator of EMT-driven alternative splicing and a mediator of cellular invasion', *Oncogene*, 33: 1082-92.
38. Briata, P., D. Bordo, M. Puppo, F. Gorlero, M. Rossi, N. Perrone-Bizzozero, and R. Gherzi. 2016. 'Diverse roles of the nucleic acid-binding protein KHSRP in cell differentiation and disease', *Wiley Interdiscip Rev RNA*, 7: 227-40.
39. Brown, R. L., L. M. Reinke, M. S. Damerow, D. Perez, L. A. Chodosh, J. Yang, and C. Cheng. 2011. 'CD44 splice isoform switching in human and mouse epithelium is essential for epithelial-mesenchymal transition and breast cancer progression', *J Clin Invest*, 121: 1064-74.
40. Buckanovich, R. J., J. B. Posner, and R. B. Darnell. 1993. 'Nova, the paraneoplastic Ri antigen, is homologous to an RNA-binding protein and is specifically expressed in the developing motor system', *Neuron*, 11: 657-72.
41. Buratti, E., C. Stuani, G. De Prato, and F. E. Baralle. 2007. 'SR protein-mediated inhibition of CFTR exon 9 inclusion: molecular characterization of the intronic splicing silencer', *Nucleic Acids Res*, 35: 4359-68.
42. Busch, A., and K. J. Hertel. 2012. 'Evolution of SR protein and hnRNP splicing regulatory factors', *Wiley Interdiscip Rev RNA*, 3: 1-12.
43. Buss, F., S. D. Arden, M. Lindsay, J. P. Luzio, and J. Kendrick-Jones. 2001. 'Myosin VI isoform localized to clathrin-coated vesicles with a role in clathrin-mediated endocytosis', *Embo j*, 20: 3676-84.
44. Buss, F., J. Kendrick-Jones, C. Lionne, A. E. Knight, G. P. Cote, and J. Paul Luzio. 1998. 'The localization of myosin VI at the golgi complex and leading edge of fibroblasts and its phosphorylation and recruitment into membrane ruffles of A431 cells after growth factor stimulation', *J Cell Biol*, 143: 1535-45.
45. Cadigan, K. M., and M. L. Waterman. 2012. 'TCF/LEFs and Wnt signaling in the nucleus', *Cold Spring Harb Perspect Biol*, 4.
46. Carstens, R. P., E. J. Wagner, and M. A. Garcia-Blanco. 2000. 'An intronic splicing silencer causes skipping of the IIIb exon of fibroblast growth factor receptor 2 through involvement of polypyrimidine tract binding protein', *Mol Cell Biol*, 20: 7388-400.
47. Castiglioni, F., E. Tagliabue, M. Campiglio, S. M. Pupa, A. Balsari, and S. Menard. 2006. 'Role of exon-16-deleted HER2 in breast carcinomas', *Endocr Relat Cancer*, 13: 221-32.
48. Cavaloc, Y., C. F. Bourgeois, L. Kister, and J. Stevenin. 1999. 'The splicing factors 9G8 and SRp20 transactivate splicing through different and specific enhancers', *Rna*, 5: 468-83.
49. Chan, R. C., and D. L. Black. 1997. 'The polypyrimidine tract binding protein binds upstream of neural cell-specific c-src exon N1 to repress the splicing of the intron downstream', *Mol Cell Biol*, 17: 4667-76.

50. Chao, T. K., T. S. Huang, Y. P. Liao, R. L. Huang, P. H. Su, H. Y. Shen, H. C. Lai, and Y. C. Wang. 2017. 'Pyruvate kinase M2 is a poor prognostic marker of and a therapeutic target in ovarian cancer', *PLoS One*, 12: e0182166.
51. Chen, C. D., R. Kobayashi, and D. M. Helfman. 1999. 'Binding of hnRNP H to an exonic splicing silencer is involved in the regulation of alternative splicing of the rat beta-tropomyosin gene', *Genes Dev*, 13: 593-606.
52. Chen, L., Y. Yao, L. Sun, J. Zhou, M. Miao, S. Luo, G. Deng, J. Li, J. Wang, and J. Tang. 2017. 'Snail Driving Alternative Splicing of CD44 by ESRP1 Enhances Invasion and Migration in Epithelial Ovarian Cancer', *Cell Physiol Biochem*, 43: 2489-504.
53. Cheng, J., T. Zhou, C. Liu, J. P. Shapiro, M. J. Brauer, M. C. Kiefer, P. J. Barr, and J. D. Mountz. 1994. 'Protection from Fas-mediated apoptosis by a soluble form of the Fas molecule', *Science*, 263: 1759-62.
54. Chibalina, M. V., A. Poliakov, J. Kendrick-Jones, and F. Buss. 2010. 'Myosin VI and optineurin are required for polarized EGFR delivery and directed migration', *Traffic*, 11: 1290-303.
55. Chou, M. Y., N. Rooke, C. W. Turck, and D. L. Black. 1999. 'hnRNP H is a component of a splicing enhancer complex that activates a c-src alternative exon in neuronal cells', *Mol Cell Biol*, 19: 69-77.
56. Chou, M. Y., J. G. Underwood, J. Nikolic, M. H. Luu, and D. L. Black. 2000. 'Multisite RNA binding and release of polypyrimidine tract binding protein during the regulation of c-src neural-specific splicing', *Mol Cell*, 5: 949-57.
57. Chow, L. T., R. E. Gelinas, T. R. Broker, and R. J. Roberts. 1977. 'An amazing sequence arrangement at the 5' ends of adenovirus 2 messenger RNA', *Cell*, 12: 1-8.
58. Cittelly, D. M., P. M. Das, V. A. Salvo, J. P. Fonseca, M. E. Burow, and F. E. Jones. 2010. 'Oncogenic HER2 Δ 16 suppresses miR-15a/16 and deregulates BCL-2 to promote endocrine resistance of breast tumors', *Carcinogenesis*, 31: 2049-57.
59. Clower, C. V., D. Chatterjee, Z. Wang, L. C. Cantley, M. G. Vander Heiden, and A. R. Krainer. 2010. 'The alternative splicing repressors hnRNP A1/A2 and PTB influence pyruvate kinase isoform expression and cell metabolism', *Proc Natl Acad Sci U S A*, 107: 1894-9.
60. Collaco, A., R. Jakab, P. Hegan, M. Mooseker, and N. Ameen. 2010. 'Alpha-AP-2 directs myosin VI-dependent endocytosis of cystic fibrosis transmembrane conductance regulator chloride channels in the intestine', *J Biol Chem*, 285: 17177-87.
61. Collisson EA, Campbell JD, Brooks AN, Berger AH, Lee W, Chmielecki J, Beer DG, Cope L, Creighton CJ, Danilova L, Ding L, Getz G, Hammerman PS, Hayes DN, Hernandez B, Herman JG, Heymach JV, Jurisica I, Kucherlapati R, Kwiatkowski D, Ladanyi M, Robertson G, Schultz N, Shen R, Sinha R, Sougnez C, Tsao MS, Travis WD, Weinstein JN, Wigle DA, Wilkerson MD, Chu A, Cherniack AD, Hadjipanayis A, Rosenberg M, Weisenberger DJ, Laird PW, Radenbaugh A, Ma S, Stuart JM, Averett Byers L, Baylin SB, Govindan R, Meyerson M, Rosenberg M, Gabriel SB, Cibulskis K, Sougnez C, Kim J, Stewart C, Lichtenstein L, Lander ES, Lawrence MS, Getz, Kandoth C, Fulton R, Fulton LL, McLellan MD, Wilson RK, Ye K, Fronick CC, Maher CA, Miller CA, Wendl MC, Cabanski C, Ding L, Mardis E, Govindan R, Creighton CJ, Wheeler D, Balasundaram M, Butterfield YS, Carlsen R, Chu A, Chuah E, Dhalla N, Guin R, Hirst C, Lee D, Li HI, Mayo M, Moore RA, Mungall AJ, Schein JE, Sipahimalani P, Tam A, Varhol R, Robertson A, Wye N, Thiessen N, Holt RA, Jones SJ, Marra MA, Campbell JD, Brooks AN, Chmielecki J, Imielinski M, Onofrio RC, Hodis E, Zack T, Sougnez C, Helman E, Sekhar Peadamallu C, Mesirov J, Cherniack AD, Saksena G, Schumacher SE, Carter SL, Hernandez B, Garraway L, Beroukheim R, Gabriel SB, Getz G, Meyerson M, Hadjipanayis A, Lee S, Mahadeshwar HS, Pantazi A, Protopopov A, Ren X, Seth S, Song X, Tang J, Yang L, Zhang J, Chen PC, Parfenov M, Wei Xu A, Santoso N, Chin L, Park PJ, Kucherlapati R, Hoadley KA, Auman JT, Meng S, Shi Y, Buda E, Waring S, Veluvolu U, Tan D, Mieczkowski PA, Jones CD, Simons JV, Soloway MG, Bodenheimer T, Jefferys SR, Roach J, Hoyle AP, Wu J, Balu S, Singh D, Prins JF, Marron JS, Parker JS, Hayes DN, Perou CM, Liu J, Cope L, Danilova L, Weisenberger DJ, Maglinte DT, Lai PH, Bootwalla MS, Van Den Berg DJ, Triche T Jr, Baylin SB, Laird PW, Rosenberg M, Chin L, Zhang J, Cho J, DiCara D,

- Heiman D, Lin P, Mallard W, Voet D, Zhang H, Zou L, Noble MS, Lawrence MS, Saksena G, Gehlenborg N, Thorvaldsdottir H, Mesirov J, Nazaire MD, Robinson J, Getz G, Lee W, Aksoy BA, Ciriello G, Taylor BS, Dresdner G, Gao J, Gross B, Seshan VE, Ladanyi M, Reva B, Sinha R, Sumer SO, Weinhold N, Schultz N, Shen R, Sander C, Ng S, Ma S, Zhu J, Radenbaugh A, Stuart JM, Benz CC, Yau C, Haussler D, Spellman PT, Wilkerson MD, Parker JS, Hoadley KA, Kimes PK, Hayes DN, Perou CM, Broom BM, Wang J, Lu Y, Kwok Shing Ng P, Diao L, Averett Byers L, Liu W, Heymach JV, Amos CI, Weinstein JN, Akbani R, Mills GB, Curley E, Paulauskis J, Lau K, Morris S, Shelton T, Mallery D, Gardner J, Penny R, Saller C, Tarvin K, Richards WG, Cerfolio R, Bryant A, Raymond DP, Pennell NA, Farver C, Czerwinski C, Huelsenbeck-Dill L, Iacocca M, Petrelli N, Rabeno B, Brown J, Bauer T, Dolzhanskiy O, Potapova O, Rotin D, Voronina O, Nemirovich-Danchenko E, Fedosenko KV, Gal A, Behera M, Ramalingam SS, Sica G, Flieder D, Boyd J, Weaver J, Kohl B, Huy Quoc Thinh D, Sandusky G, Juhl H, Duhig E, Illei P, Gabrielson E, Shin J, Lee B, Rodgers K, Trusty D, Brock MV, Williamson C, Burks E, Rieger-Christ K, Holway A, Sullivan T, Wigle DA, Asiedu MK, Kosari F, Travis WD, Rekhtman N, Zakowski M, Rusch VW, Zippile P, Suh J, Pass H, Goparaju C, Owusu-Sarpong Y, Bartlett JM, Kodeeswaran S, Parfitt J, Sekhon H, Albert M, Eckman J, Myers JB, Cheney R, Morrison C, Gaudio C, Borgia JA, Bonomi P, Pool M, Liptay MJ, Moiseenko F, Zaytseva I, Dienemann H, Meister M, Schnabel PA, Muley TR, Peifer M, Gomez-Fernandez C, Herbert L, Egea S, Huang M, Thorne LB, Boice L, Hill Salazar A, Funkhouser WK, Rathmell WK, Dhir R, Yousem SA, Dacic S, Schneider F, Siegfried JM, Hajek R, Watson MA, McDonald S, Meyers B, Clarke B, Yang IA, Fong KM, Hunter L, Windsor M, Bowman RV, Peters S, Letovanec I, Khan KZ, Jensen MA, Snyder EE, Srinivasan D, Kahn AB, Baboud J, Pot DA, Mills Shaw KR, Sheth M, Davidsen T, Demchok JA, Yang L, Wang Z, Tarnuzzer R, Zenklusen JC, Ozenberger BA, Sofia HJ, Travis WD, Cheney R, Clarke B, Dacic S, Duhig E, Funkhouser WK, Illei P, Farver C, Rekhtman N, Sica G, Suh J, Tsao MS, Travis WD, Cheney R, Clarke B, Dacic S, Duhig E, Funkhouser WK, Illei P, Farver C, Rekhtman N, Sica G, Suh J, Tsao MS. 2014. 'Comprehensive molecular profiling of lung adenocarcinoma', *Nature*, 511: 543-50.
62. Conboy, J. G. 2017. 'Developmental regulation of RNA processing by Rbfox proteins', *Wiley Interdiscip Rev RNA*, 8.
63. Crawford, J. B., and J. G. Patton. 2006. 'Activation of alpha-tropomyosin exon 2 is regulated by the SR protein 9G8 and heterogeneous nuclear ribonucleoproteins H and F', *Mol Cell Biol*, 26: 8791-802.
64. Daguene, E., G. Dujardin, and J. Valcarcel. 2015. 'The pathogenicity of splicing defects: mechanistic insights into pre-mRNA processing inform novel therapeutic approaches', *EMBO Rep*, 16: 1640-55.
65. Dance, A. L., M. Miller, S. Seragaki, P. Aryal, B. White, L. Aschenbrenner, and T. Hasson. 2004. 'Regulation of myosin-VI targeting to endocytic compartments', *Traffic*, 5: 798-813.
66. Darman, R. B., M. Seiler, A. A. Agrawal, K. H. Lim, S. Peng, D. Aird, S. L. Bailey, E. B. Bhavsar, B. Chan, S. Colla, L. Corson, J. Feala, P. Fekkes, K. Ichikawa, G. F. Keaney, L. Lee, P. Kumar, K. Kunii, C. MacKenzie, M. Matijevic, Y. Mizui, K. Myint, E. S. Park, X. Puyang, A. Selvaraj, M. P. Thomas, J. Tsai, J. Y. Wang, M. Warmuth, H. Yang, P. Zhu, G. Garcia-Manero, R. R. Furman, L. Yu, P. G. Smith, and S. Buonamici. 2015. 'Cancer-Associated SF3B1 Hotspot Mutations Induce Cryptic 3' Splice Site Selection through Use of a Different Branch Point', *Cell Rep*, 13: 1033-45.
67. Das, D., T. A. Clark, A. Schweitzer, M. Yamamoto, H. Marr, J. Arribere, S. Minovitsky, A. Poliakov, I. Dubchak, J. E. Blume, and J. G. Conboy. 2007. 'A correlation with exon expression approach to identify cis-regulatory elements for tissue-specific alternative splicing', *Nucleic Acids Res*, 35: 4845-57.
68. David, C. J., M. Chen, M. Assanah, P. Canoll, and J. L. Manley. 2010. 'HnRNP proteins controlled by c-Myc deregulate pyruvate kinase mRNA splicing in cancer', *Nature*, 463: 364-8.
69. David, C. J., and J. L. Manley. 2008. 'The search for alternative splicing regulators: new approaches offer a path to a splicing code', *Genes Dev*, 22: 279-85.
70. Davis, C. A., B. C. Hitz, C. A. Sloan, E. T. Chan, J. M. Davidson, I. Gabdank, J. A. Hilton, K. Jain, U. K. Baymuradov, A. K. Narayanan, K. C. Onate, K. Graham, S. R. Miyasato, T. R. Dreszer, J. S.

- Strattan, O. Jolanki, F. Y. Tanaka, and J. M. Cherry. 2018. 'The Encyclopedia of DNA elements (ENCODE): data portal update', *Nucleic Acids Res*, 46: D794-d801.
71. Davis-Smyth, T., R. C. Duncan, T. Zheng, G. Michelotti, and D. Levens. 1996. 'The far upstream element-binding proteins comprise an ancient family of single-strand DNA-binding transactivators', *J Biol Chem*, 271: 31679-87.
 72. Dayton, T. L., V. Gocheva, K. M. Miller, W. J. Israelsen, A. Bhutkar, C. B. Clish, S. M. Davidson, A. Luengo, R. T. Bronson, T. Jacks, and M. G. Vander Heiden. 2016. 'Germline loss of PKM2 promotes metabolic distress and hepatocellular carcinoma', *Genes Dev*, 30: 1020-33.
 73. De Conti, L., M. Baralle, and E. Buratti. 2013. 'Exon and intron definition in pre-mRNA splicing', *Wiley Interdiscip Rev RNA*, 4: 49-60.
 74. DeBoever, C., E. M. Ghia, P. J. Shepard, L. Rassenti, C. L. Barrett, K. Jepsen, C. H. Jamieson, D. Carson, T. J. Kipps, and K. A. Frazer. 2015. 'Transcriptome sequencing reveals potential mechanism of cryptic 3' splice site selection in SF3B1-mutated cancers', *PLoS Comput Biol*, 11: e1004105.
 75. Dembowski, J. A., P. An, M. Scoulos-Hanson, G. Yeo, J. Han, X. D. Fu, and P. J. Grabowski. 2012. 'Alternative Splicing of a Novel Inducible Exon Diversifies the CASK Guanylate Kinase Domain', *J Nucleic Acids*, 2012: 816237.
 76. Dewaele, M., T. Tabaglio, K. Willekens, M. Bezzi, S. X. Teo, D. H. Low, C. M. Koh, F. Rambow, M. Fiers, A. Rogiers, E. Radaelli, M. Al-Haddawi, S. Y. Tan, E. Hermans, F. Amant, H. Yan, M. Lakshmanan, R. C. Koumar, S. T. Lim, F. A. Derheimer, R. M. Campbell, Z. Bonday, V. Tergaonkar, M. Shackleton, C. Blattner, J. C. Marine, and E. Guccione. 2016. 'Antisense oligonucleotide-mediated MDM4 exon 6 skipping impairs tumor growth', *J Clin Invest*, 126: 68-84.
 77. Dole, M. G., R. Jasty, M. J. Cooper, C. B. Thompson, G. Nunez, and V. P. Castle. 1995. 'Bcl-xL is expressed in neuroblastoma cells and modulates chemotherapy-induced apoptosis', *Cancer Res*, 55: 2576-82.
 78. Duan, W., Y. P. Zhang, Z. Hou, C. Huang, H. Zhu, C. Q. Zhang, and Q. Yin. 2016. 'Novel Insights into NeuN: from Neuronal Marker to Splicing Regulator', *Mol Neurobiol*, 53: 1637-47.
 79. Dunn, T. A., S. Chen, D. A. Faith, J. L. Hicks, E. A. Platz, Y. Chen, C. M. Ewing, J. Sauvageot, W. B. Isaacs, A. M. De Marzo, and J. Luo. 2006. 'A novel role of myosin VI in human prostate cancer', *Am J Pathol*, 169: 1843-54.
 80. Dvinge, H. 2018. 'Regulation of alternative mRNA splicing: old players and new perspectives', *FEBS Lett*, 592: 2987-3006.
 81. Dvinge, H., E. Kim, O. Abdel-Wahab, and R. K. Bradley. 2016. 'RNA splicing factors as oncoproteins and tumour suppressors', *Nat Rev Cancer*, 16: 413-30.
 82. Edmond, V., E. Moysan, S. Khochbin, P. Matthias, C. Brambilla, E. Brambilla, S. Gazzeri, and B. Eymis. 2011. 'Acetylation and phosphorylation of SRSF2 control cell fate decision in response to cisplatin', *Embo j*, 30: 510-23.
 83. Effenberger, K. A., V. K. Urabe, and M. S. Jurica. 2017. 'Modulating splicing with small molecular inhibitors of the spliceosome', *Wiley Interdiscip Rev RNA*, 8.
 84. Erkelenz, S., W. F. Mueller, M. S. Evans, A. Busch, K. Schoneweis, K. J. Hertel, and H. Schaal. 2013. 'Position-dependent splicing activation and repression by SR and hnRNP proteins rely on common mechanisms', *Rna*, 19: 96-102.
 85. Eskens, F. A., F. J. Ramos, H. Burger, J. P. O'Brien, A. Piera, M. J. de Jonge, Y. Mizui, E. A. Wiemer, M. J. Carreras, J. Baselga, and J. Tabernero. 2013. 'Phase I pharmacokinetic and pharmacodynamic study of the first-in-class spliceosome inhibitor E7107 in patients with advanced solid tumors', *Clin Cancer Res*, 19: 6296-304.

86. Eswaran, J., A. Horvath, S. Godbole, S. D. Reddy, P. Mudvari, K. Ohshiro, D. Cyanam, S. Nair, S. A. Fuqua, K. Polyak, L. D. Florea, and R. Kumar. 2013. 'RNA sequencing of cancer reveals novel splicing alterations', *Sci Rep*, 3: 1689.
87. Faria, M., L. Capinha, J. Simoes-Pereira, M. J. Bugalho, and A. L. Silva. 2016. 'Extending the Impact of RAC1b Overexpression to Follicular Thyroid Carcinomas', *Int J Endocrinol*, 2016: 1972367.
88. Fiegen, D., L. C. Haeusler, L. Blumenstein, U. Herbrand, R. Dvorsky, I. R. Vetter, and M. R. Ahmadian. 2004. 'Alternative splicing of Rac1 generates Rac1b, a self-activating GTPase', *J Biol Chem*, 279: 4743-9.
89. Fili, N., Y. Hari-Gupta, Á Dos Santos, A. Cook, S. Poland, S. M. Ameer-Beg, M. Parsons, and C. P. Toseland. 2017. 'NDP52 activates nuclear myosin VI to enhance RNA polymerase II transcription', *Nat Commun*, 8: 1871.
90. Finan, D., M. A. Hartman, and J. A. Spudich. 2011. 'Proteomics approach to study the functions of Drosophila myosin VI through identification of multiple cargo-binding proteins', *Proc Natl Acad Sci U S A*, 108: 5566-71.
91. Fu, X. D. 1995. 'The superfamily of arginine/serine-rich splicing factors', *Rna*, 1: 663-80.
92. Fu, X. D., and T. Maniatis. 1990. 'Factor required for mammalian spliceosome assembly is localized to discrete regions in the nucleus', *Nature*, 343: 437-41.
93. ———. 1992. 'The 35-kDa mammalian splicing factor SC35 mediates specific interactions between U1 and U2 small nuclear ribonucleoprotein particles at the 3' splice site', *Proc Natl Acad Sci U S A*, 89: 1725-9.
94. Fukuhara, T., T. Hosoya, S. Shimizu, K. Sumi, T. Oshiro, Y. Yoshinaka, M. Suzuki, N. Yamamoto, L. A. Herzenberg, L. A. Herzenberg, and M. Hagiwara. 2006. 'Utilization of host SR protein kinases and RNA-splicing machinery during viral replication', *Proc Natl Acad Sci U S A*, 103: 11329-33.
95. Fushimi, K., P. Ray, A. Kar, L. Wang, L. C. Sutherland, and J. Y. Wu. 2008. 'Up-regulation of the proapoptotic caspase 2 splicing isoform by a candidate tumor suppressor, RBM5', *Proc Natl Acad Sci U S A*, 105: 15708-13.
96. Garcia-Blanco, M. A., S. F. Jamison, and P. A. Sharp. 1989. 'Identification and purification of a 62,000-dalton protein that binds specifically to the polypyrimidine tract of introns', *Genes Dev*, 3: 1874-86.
97. Garcia-Mayoral, M. F., I. Diaz-Moreno, D. Hollingworth, and A. Ramos. 2008. 'The sequence selectivity of KSRP explains its flexibility in the recognition of the RNA targets', *Nucleic Acids Res*, 36: 5290-6.
98. Garcia-Mayoral, M. F., D. Hollingworth, L. Masino, I. Diaz-Moreno, G. Kelly, R. Gherzi, C. F. Chou, C. Y. Chen, and A. Ramos. 2007. 'The structure of the C-terminal KH domains of KSRP reveals a noncanonical motif important for mRNA degradation', *Structure*, 15: 485-98.
99. Gautrey, H., C. Jackson, A. L. Dittrich, D. Browell, T. Lennard, and A. Tyson-Capper. 2015. 'SRSF3 and hnRNP H1 regulate a splicing hotspot of HER2 in breast cancer cells', *RNA Biol*, 12: 1139-51.
100. Geisbrecht, E. R., and D. J. Montell. 2002. 'Myosin VI is required for E-cadherin-mediated border cell migration', *Nat Cell Biol*, 4: 616-20.
101. Geuens, T., D. Bouhy, and V. Timmerman. 2016. 'The hnRNP family: insights into their role in health and disease', *Hum Genet*, 135: 851-67.
102. Gherzi, R., C. Y. Chen, A. Ramos, and P. Briata. 2014. 'KSRP controls pleiotropic cellular functions', *Semin Cell Dev Biol*, 34: 2-8.

103. Ghigna, C., S. Giordano, H. Shen, F. Benvenuto, F. Castiglioni, P. M. Comoglio, M. R. Green, S. Riva, and G. Biamonti. 2005. 'Cell motility is controlled by SF2/ASF through alternative splicing of the Ron protooncogene', *Mol Cell*, 20: 881-90.
104. Giampietro, C., G. Deflorian, S. Gallo, A. Di Matteo, D. Pradella, S. Bonomi, E. Belloni, D. Nyqvist, V. Quaranta, S. Confalonieri, G. Bertalot, F. Orsenigo, F. Pisati, E. Ferrero, G. Biamonti, E. Fredrickx, C. Taveggia, C. D. Wyatt, M. Irimia, P. P. Di Fiore, B. J. Blencowe, E. Dejana, and C. Ghigna. 2015. 'The alternative splicing factor Nova2 regulates vascular development and lumen formation', *Nat Commun*, 6: 8479.
105. Giannakis, M., X. J. Mu, S. A. Shukla, Z. R. Qian, O. Cohen, R. Nishihara, S. Bahl, Y. Cao, A. Amin-Mansour, M. Yamauchi, Y. Sukawa, C. Stewart, M. Rosenberg, K. Mima, K. Inamura, K. Noshio, J. A. Nowak, M. S. Lawrence, E. L. Giovannucci, A. T. Chan, K. Ng, J. A. Meyerhardt, E. M. Van Allen, G. Getz, S. B. Gabriel, E. S. Lander, C. J. Wu, C. S. Fuchs, S. Ogino, and L. A. Garraway. 2016. 'Genomic Correlates of Immune-Cell Infiltrates in Colorectal Carcinoma', *Cell Rep*, 15: 857-65.
106. Golan-Gerstl, R., M. Cohen, A. Shilo, S. S. Suh, A. Bakacs, L. Coppola, and R. Karni. 2011. 'Splicing factor hnRNP A2/B1 regulates tumor suppressor gene splicing and is an oncogenic driver in glioblastoma', *Cancer Res*, 71: 4464-72.
107. Goncalves, V., A. F. Henriques, J. F. Pereira, A. Neves Costa, M. P. Moyer, L. F. Moita, M. Gama-Carvalho, P. Matos, and P. Jordan. 2014. 'Phosphorylation of SRSF1 by SRPK1 regulates alternative splicing of tumor-related Rac1b in colorectal cells', *Rna*, 20: 474-82.
108. Gonçalves, V., P. Matos, and P. Jordan. 2008. 'The beta-catenin/TCF4 pathway modifies alternative splicing through modulation of SRp20 expression', *Rna*, 14: 2538-49.
109. Gong, S. G. 2014. 'Isoforms of receptors of fibroblast growth factors', *J Cell Physiol*, 229: 1887-95.
110. Gooding, C., G. C. Roberts, and C. W. Smith. 1998. 'Role of an inhibitory pyrimidine element and polypyrimidine tract binding protein in repression of a regulated alpha-tropomyosin exon', *Rna*, 4: 85-100.
111. Gout, S., E. Brambilla, A. Boudria, R. Drissi, S. Lantuejoul, S. Gazzeri, and B. Eymin. 2012. 'Abnormal expression of the pre-mRNA splicing regulators SRSF1, SRSF2, SRPK1 and SRPK2 in non small cell lung carcinoma', *PLoS One*, 7: e46539.
112. Grumati, P., and I. Dikic. 2018. 'Ubiquitin signaling and autophagy', *J Biol Chem*, 293: 5404-13.
113. Hall, M. P., S. Huang, and D. L. Black. 2004. 'Differentiation-induced colocalization of the KH-type splicing regulatory protein with polypyrimidine tract binding protein and the c-src pre-mRNA', *Mol Biol Cell*, 15: 774-86.
114. Hallay, H., N. Locker, L. Ayadi, D. Ropers, E. Guittet, and C. Branlant. 2006. 'Biochemical and NMR study on the competition between proteins SC35, SRp40, and heterogeneous nuclear ribonucleoprotein A1 at the HIV-1 Tat exon 2 splicing site', *J Biol Chem*, 281: 37159-74.
115. Han, J., J. H. Ding, C. W. Byeon, J. H. Kim, K. J. Hertel, S. Jeong, and X. D. Fu. 2011. 'SR proteins induce alternative exon skipping through their activities on the flanking constitutive exons', *Mol Cell Biol*, 31: 793-802.
116. Hayes, G. M., P. E. Carrigan, A. M. Beck, and L. J. Miller. 2006. 'Targeting the RNA splicing machinery as a novel treatment strategy for pancreatic carcinoma', *Cancer Res*, 66: 3819-27.
117. He, C., F. Zhou, Z. Zuo, H. Cheng, and R. Zhou. 2009. 'A global view of cancer-specific transcript variants by subtractive transcriptome-wide analysis', *PLoS One*, 4: e4732.
118. He, F., H. P. Wollscheid, U. Nowicka, M. Biancospino, E. Valentini, A. Ehlinger, F. Acconcia, E. Magistrati, S. Polo, and K. J. Walters. 2016. 'Myosin VI Contains a Compact Structural Motif that Binds to Ubiquitin Chains', *Cell Rep*, 14: 2683-94.

119. Hernandez, F., M. Perez, J. J. Lucas, A. M. Mata, R. Bhat, and J. Avila. 2004. 'Glycogen synthase kinase-3 plays a crucial role in tau exon 10 splicing and intranuclear distribution of SC35. Implications for Alzheimer's disease', *J Biol Chem*, 279: 3801-6.
120. Hernandez, J., E. Bechara, D. Schlesinger, J. Delgado, L. Serrano, and J. Valcarcel. 2016. 'Tumor suppressor properties of the splicing regulatory factor RBM10', *RNA Biol*, 13: 466-72.
121. Heyd, F., and K. W. Lynch. 2010. 'Phosphorylation-dependent regulation of PSF by GSK3 controls CD45 alternative splicing', *Mol Cell*, 40: 126-37.
122. Hollander, D., M. Donyo, N. Atias, K. Mekahel, Z. Melamed, S. Yannai, G. Lev-Maor, A. Shilo, S. Schwartz, I. Barshack, R. Sharan, and G. Ast. 2016. 'A network-based analysis of colon cancer splicing changes reveals a tumorigenesis-favoring regulatory pathway emanating from ELK1', *Genome Res*, 26: 541-53.
123. Horiguchi, K., K. Sakamoto, D. Koinuma, K. Semba, A. Inoue, S. Inoue, H. Fujii, A. Yamaguchi, K. Miyazawa, K. Miyazono, and M. Saitoh. 2012. 'TGF-beta drives epithelial-mesenchymal transition through deltaEF1-mediated downregulation of ESRP', *Oncogene*, 31: 3190-201.
124. Houck, K. A., N. Ferrara, J. Winer, G. Cachianes, B. Li, and D. W. Leung. 1991. 'The vascular endothelial growth factor family: identification of a fourth molecular species and characterization of alternative splicing of RNA', *Mol Endocrinol*, 5: 1806-14.
125. Hsu, T. Y., L. M. Simon, N. J. Neill, R. Marcotte, A. Sayad, C. S. Bland, G. V. Echeverria, T. Sun, S. J. Kurley, S. Tyagi, K. L. Karlin, R. Dominguez-Vidana, J. D. Hartman, A. Renwick, K. Scorsone, R. J. Bernardi, S. O. Skinner, A. Jain, M. Orellana, C. Lagiseti, I. Golding, S. Y. Jung, J. R. Neilson, X. H. Zhang, T. A. Cooper, T. R. Webb, B. G. Neel, C. A. Shaw, and T. F. Westbrook. 2015. 'The spliceosome is a therapeutic vulnerability in MYC-driven cancer', *Nature*, 525: 384-8.
126. Hu, S., Y. Guo, Y. Wang, Y. Li, T. Fu, Z. Zhou, Y. Wang, J. Liu, and L. Pan. 2019. 'Structure of Myosin VI/Tom1 complex reveals a cargo recognition mode of Myosin VI for tethering', *Nat Commun*, 10: 3459.
127. Huang, S. C., A. C. Ou, J. Park, F. Yu, B. Yu, A. Lee, G. Yang, A. Zhou, and E. J. Benz, Jr. 2012. 'RBFOX2 promotes protein 4.1R exon 16 selection via U1 snRNP recruitment', *Mol Cell Biol*, 32: 513-26.
128. Hui, J., L. H. Hung, M. Heiner, S. Schreiner, N. Neumuller, G. Reither, S. A. Haas, and A. Bindereif. 2005. 'Intronic CA-repeat and CA-rich elements: a new class of regulators of mammalian alternative splicing', *Embo j*, 24: 1988-98.
129. Hui, J., K. Stangl, W. S. Lane, and A. Bindereif. 2003. 'HnRNP L stimulates splicing of the eNOS gene by binding to variable-length CA repeats', *Nat Struct Biol*, 10: 33-7.
130. Huynh, F. C., and F. E. Jones. 2014. 'MicroRNA-7 inhibits multiple oncogenic pathways to suppress HER2Delta16 mediated breast tumorigenesis and reverse trastuzumab resistance', *PLoS One*, 9: e114419.
131. Ilagan, J. O., A. Ramakrishnan, B. Hayes, M. E. Murphy, A. S. Zebari, P. Bradley, and R. K. Bradley. 2015. 'U2AF1 mutations alter splice site recognition in hematological malignancies', *Genome Res*, 25: 14-26.
132. Imielinski, M., A. H. Berger, P. S. Hammerman, B. Hernandez, T. J. Pugh, E. Hodis, J. Cho, J. Suh, M. Capelletti, A. Sivachenko, C. Sougnez, D. Auclair, M. S. Lawrence, P. Stojanov, K. Cibulskis, K. Choi, L. de Waal, T. Sharifnia, A. Brooks, H. Greulich, S. Banerji, T. Zander, D. Seidel, F. Leenders, S. Ansen, C. Ludwig, W. Engel-Riedel, E. Stoelben, J. Wolf, C. Goparaju, K. Thompson, W. Winckler, D. Kwiatkowski, B. E. Johnson, P. A. Janne, V. A. Miller, W. Pao, W. D. Travis, H. I. Pass, S. B. Gabriel, E. S. Lander, R. K. Thomas, L. A. Garraway, G. Getz, and M. Meyerson. 2012. 'Mapping the hallmarks of lung adenocarcinoma with massively parallel sequencing', *Cell*, 150: 1107-20.
133. Inoue, D., and O. Abdel-Wahab. 2016. 'Modeling SF3B1 Mutations in Cancer: Advances, Challenges, and Opportunities', *Cancer Cell*, 30: 371-73.

134. Ishimoto, T., O. Nagano, T. Yae, M. Tamada, T. Motohara, H. Oshima, M. Oshima, T. Ikeda, R. Asaba, H. Yagi, T. Masuko, T. Shimizu, T. Ishikawa, K. Kai, E. Takahashi, Y. Imamura, Y. Baba, M. Ohmura, M. Suematsu, H. Baba, and H. Saya. 2011. 'CD44 variant regulates redox status in cancer cells by stabilizing the xCT subunit of system xc(-) and thereby promotes tumor growth', *Cancer Cell*, 19: 387-400.
135. Iwai, K., M. Yaguchi, K. Nishimura, Y. Yamamoto, T. Tamura, D. Nakata, R. Dairiki, Y. Kawakita, R. Mizojiri, Y. Ito, M. Asano, H. Maezaki, Y. Nakayama, M. Kaishima, K. Hayashi, M. Teratani, S. Miyakawa, M. Iwatani, M. Miyamoto, M. G. Klein, W. Lane, G. Snell, R. Tjhen, X. He, S. Pulukuri, and T. Nomura. 2018. 'Anti-tumor efficacy of a novel CLK inhibitor via targeting RNA splicing and MYC-dependent vulnerability', *EMBO Mol Med*, 10.
136. Izquierdo, J. M., N. Majos, S. Bonnal, C. Martinez, R. Castelo, R. Guigo, D. Bilbao, and J. Valcarcel. 2005. 'Regulation of Fas alternative splicing by antagonistic effects of TIA-1 and PTB on exon definition', *Mol Cell*, 19: 475-84.
137. Jaffe, A. B., N. Kaji, J. Durgan, and A. Hall. 2008. 'Cdc42 controls spindle orientation to position the apical surface during epithelial morphogenesis', *J Cell Biol*, 183: 625-33.
138. Jang, H. N., M. Lee, T. J. Loh, S. W. Choi, H. K. Oh, H. Moon, S. Cho, S. E. Hong, D. H. Kim, Z. Sheng, M. R. Green, D. Park, X. Zheng, and H. Shen. 2014. 'Exon 9 skipping of apoptotic caspase-2 pre-mRNA is promoted by SRSF3 through interaction with exon 8', *Biochim Biophys Acta*, 1839: 25-32.
139. Jeanes, A., C. J. Gottardi, and A. S. Yap. 2008. 'Cadherins and cancer: how does cadherin dysfunction promote tumor progression?', *Oncogene*, 27: 6920-9.
140. Jensen, K. B., K. Musunuru, H. A. Lewis, S. K. Burley, and R. B. Darnell. 2000. 'The tetranucleotide UCAY directs the specific recognition of RNA by the Nova K-homology 3 domain', *Proc Natl Acad Sci U S A*, 97: 5740-5.
141. Jeong, S. 2017. 'SR Proteins: Binders, Regulators, and Connectors of RNA', *Mol Cells*, 40: 1-9.
142. Jin, Y., H. Suzuki, S. Maegawa, H. Endo, S. Sugano, K. Hashimoto, K. Yasuda, and K. Inoue. 2003. 'A vertebrate RNA-binding protein Fox-1 regulates tissue-specific splicing via the pentanucleotide GCAUG', *Embo j*, 22: 905-12.
143. Jin, Y., J. Zhou, F. Xu, B. Jin, L. Cui, Y. Wang, X. Du, J. Li, P. Li, R. Ren, and J. Pan. 2016. 'Targeting methyltransferase PRMT5 eliminates leukemia stem cells in chronic myelogenous leukemia', *J Clin Invest*, 126: 3961-80.
144. Jung, H., D. Lee, J. Lee, D. Park, Y. J. Kim, W. Y. Park, D. Hong, P. J. Park, and E. Lee. 2015. 'Intron retention is a widespread mechanism of tumor-suppressor inactivation', *Nat Genet*, 47: 1242-8.
145. Karni, R., E. de Stanchina, S. W. Lowe, R. Sinha, D. Mu, and A. R. Krainer. 2007. 'The gene encoding the splicing factor SF2/ASF is a proto-oncogene', *Nat Struct Mol Biol*, 14: 185-93.
146. Keppetipola, N., S. Sharma, Q. Li, and D. L. Black. 2012. 'Neuronal regulation of pre-mRNA splicing by polypyrimidine tract binding proteins, PTBP1 and PTBP2', *Crit Rev Biochem Mol Biol*, 47: 360-78.
147. Kim, E., J. O. Ilagan, Y. Liang, G. M. Daubner, S. C. Lee, A. Ramakrishnan, Y. Li, Y. R. Chung, J. B. Micol, M. E. Murphy, H. Cho, M. K. Kim, A. S. Zebari, S. Aumann, C. Y. Park, S. Buonamici, P. G. Smith, H. J. Deeg, C. Lobry, I. Aifantis, Y. Modis, F. H. Allain, S. Halene, R. K. Bradley, and O. Abdel-Wahab. 2015. 'SRSF2 Mutations Contribute to Myelodysplasia by Mutant-Specific Effects on Exon Recognition', *Cancer Cell*, 27: 617-30.
148. Kim, Y. E., S. Choi, J. O. Kim, and K. K. Kim. 2017. 'RBF3X regulates Claudin-1 expression in human lung tissue via attenuation of proteasomal degradation', *Biosci Rep*, 37.

149. Kim, Y. E., J. O. Kim, K. S. Park, M. Won, K. E. Kim, and K. K. Kim. 2016. 'Transforming Growth Factor-beta-Induced RBFOX3 Inhibition Promotes Epithelial-Mesenchymal Transition of Lung Cancer Cells', *Mol Cells*, 39: 625-30.
150. Kohtz, J. D., S. F. Jamison, C. L. Will, P. Zuo, R. Luhrmann, M. A. Garcia-Blanco, and J. L. Manley. 1994. 'Protein-protein interactions and 5'-splice-site recognition in mammalian mRNA precursors', *Nature*, 368: 119-24.
151. Kruppa, A. J., C. Kishi-Itakura, T. A. Masters, J. E. Rorbach, G. L. Grice, J. Kendrick-Jones, J. A. Nathan, M. Minczuk, and F. Buss. 2018. 'Myosin VI-Dependent Actin Cages Encapsulate Parkin-Positive Damaged Mitochondria', *Dev Cell*, 44: 484-99.e6.
152. Kuroyanagi, H. 2009. 'Fox-1 family of RNA-binding proteins', *Cell Mol Life Sci*, 66: 3895-907.
153. Kwong, K. Y., and M. C. Hung. 1998. 'A novel splice variant of HER2 with increased transformation activity', *Mol Carcinog*, 23: 62-8.
154. Labourier, E., M. Blanchette, J. W. Feiger, M. D. Adams, and D. C. Rio. 2002. 'The KH-type RNA-binding protein PSI is required for Drosophila viability, male fertility, and cellular mRNA processing', *Genes Dev*, 16: 72-84.
155. Lagier-Tourenne, C., M. Polymenidou, K. R. Hutt, A. Q. Vu, M. Baughn, S. C. Huelga, K. M. Clutario, S. C. Ling, T. Y. Liang, C. Mazur, E. Wancewicz, A. S. Kim, A. Watt, S. Freier, G. G. Hicks, J. P. Donohue, L. Shiue, C. F. Bennett, J. Ravits, D. W. Cleveland, and G. W. Yeo. 2012. 'Divergent roles of ALS-linked proteins FUS/TLS and TDP-43 intersect in processing long pre-mRNAs', *Nat Neurosci*, 15: 1488-97.
156. Lapuk, A., H. Marr, L. Jakkula, H. Pedro, S. Bhattacharya, E. Purdom, Z. Hu, K. Simpson, L. Pachter, S. Durinck, N. Wang, B. Parvin, G. Fontenay, T. Speed, J. Garbe, M. Stampfer, H. Bayandorian, S. Dorton, T. A. Clark, A. Schweitzer, A. Wyrobek, H. Feiler, P. Spellman, J. Conboy, and J. W. Gray. 2010. 'Exon-level microarray analyses identify alternative splicing programs in breast cancer', *Mol Cancer Res*, 8: 961-74.
157. Lee, H. K., Y. S. Choi, Y. A. Park, and S. Jeong. 2006. 'Modulation of oncogenic transcription and alternative splicing by beta-catenin and an RNA aptamer in colon cancer cells', *Cancer Res*, 66: 10560-6.
158. Lee, S. C., and O. Abdel-Wahab. 2016. 'Therapeutic targeting of splicing in cancer', *Nat Med*, 22: 976-86.
159. Lee, S. C., H. Dvinge, E. Kim, H. Cho, J. B. Micol, Y. R. Chung, B. H. Durham, A. Yoshimi, Y. J. Kim, M. Thomas, C. Lobry, C. W. Chen, A. Pastore, J. Taylor, X. Wang, A. Krivtsov, S. A. Armstrong, J. Palacino, S. Buonamici, P. G. Smith, R. K. Bradley, and O. Abdel-Wahab. 2016. 'Modulation of splicing catalysis for therapeutic targeting of leukemia with mutations in genes encoding spliceosomal proteins', *Nat Med*, 22: 672-8.
160. Lee, Y., and D. C. Rio. 2015. 'Mechanisms and Regulation of Alternative Pre-mRNA Splicing', *Annu Rev Biochem*, 84: 291-323.
161. Lefave, C. V., M. Squatrito, S. Vorlova, G. L. Rocco, C. W. Brennan, E. C. Holland, Y. X. Pan, and L. Cartegni. 2011. 'Splicing factor hnRNPH drives an oncogenic splicing switch in gliomas', *Embo j*, 30: 4084-97.
162. Leoni, G., L. Le Pera, F. Ferre, D. Raimondo, and A. Tramontano. 2011. 'Coding potential of the products of alternative splicing in human', *Genome Biol*, 12: R9.
163. Lewis, H. A., H. Chen, C. Edo, R. J. Buckanovich, Y. Y. Yang, K. Musunuru, R. Zhong, R. B. Darnell, and S. K. Burley. 1999. 'Crystal structures of Nova-1 and Nova-2 K-homology RNA-binding domains', *Structure*, 7: 191-203.
164. Li, Y. I., L. Sanchez-Pulido, W. Haerty, and C. P. Ponting. 2015. 'RBFOX and PTBP1 proteins regulate the alternative splicing of micro-exons in human brain transcripts', *Genome Res*, 25: 1-13.

165. Li, Z., Q. Li, L. Han, N. Tian, Q. Liang, Y. Li, X. Zhao, C. Du, and Y. Tian. 2016. 'Pro-apoptotic effects of splice-switching oligonucleotides targeting Bcl-x pre-mRNA in human glioma cell lines', *Oncol Rep*, 35: 1013-9.
166. Linares, A. J., C. H. Lin, A. Damianov, K. L. Adams, B. G. Novitch, and D. L. Black. 2015. 'The splicing regulator PTBP1 controls the activity of the transcription factor Pbx1 during neuronal differentiation', *Elife*, 4: e09268.
167. Liu, C., Y. Li, M. Semenov, C. Han, G. H. Baeg, Y. Tan, Z. Zhang, X. Lin, and X. He. 2002. 'Control of beta-catenin phosphorylation/degradation by a dual-kinase mechanism', *Cell*, 108: 837-47.
168. Liu, H. X., S. L. Chew, L. Cartegni, M. Q. Zhang, and A. R. Krainer. 2000. 'Exonic splicing enhancer motif recognized by human SC35 under splicing conditions', *Mol Cell Biol*, 20: 1063-71.
169. Liu, H. X., M. Zhang, and A. R. Krainer. 1998. 'Identification of functional exonic splicing enhancer motifs recognized by individual SR proteins', *Genes Dev*, 12: 1998-2012.
170. Liu, W. R., M. X. Tian, L. X. Yang, Y. L. Lin, L. Jin, Z. B. Ding, Y. H. Shen, Y. F. Peng, D. M. Gao, J. Zhou, S. J. Qiu, Z. Dai, R. He, J. Fan, and Y. H. Shi. 2015. 'PKM2 promotes metastasis by recruiting myeloid-derived suppressor cells and indicates poor prognosis for hepatocellular carcinoma', *Oncotarget*, 6: 846-61.
171. Llorian, M., S. Schwartz, T. A. Clark, D. Hollander, L. Y. Tan, R. Spellman, A. Gordon, A. C. Schweitzer, P. de la Grange, G. Ast, and C. W. Smith. 2010. 'Position-dependent alternative splicing activity revealed by global profiling of alternative splicing events regulated by PTB', *Nat Struct Mol Biol*, 17: 1114-23.
172. Loh, T. J., H. Moon, S. Cho, H. Jang, Y. C. Liu, H. Tai, D. W. Jung, D. R. Williams, H. R. Kim, M. G. Shin, D. J. Liao, J. Zhou, W. Shi, X. Zheng, and H. Shen. 2015. 'CD44 alternative splicing and hnRNP A1 expression are associated with the metastasis of breast cancer', *Oncol Rep*, 34: 1231-8.
173. Loh, T. J., H. Moon, S. Cho, D. W. Jung, S. E. Hong, D. H. Kim, M. R. Green, X. Zheng, J. Zhou, and H. Shen. 2014. 'SC35 promotes splicing of the C5-V6-C6 isoform of CD44 pre-mRNA', *Oncol Rep*, 31: 273-9.
174. Long, J. C., and J. F. Caceres. 2009. 'The SR protein family of splicing factors: master regulators of gene expression', *Biochem J*, 417: 15-27.
175. Lovci, M. T., D. Ghanem, H. Marr, J. Arnold, S. Gee, M. Parra, T. Y. Liang, T. J. Stark, L. T. Gehman, S. Hoon, K. B. Massirer, G. A. Pratt, D. L. Black, J. W. Gray, J. G. Conboy, and G. W. Yeo. 2013. 'Rbfox proteins regulate alternative mRNA splicing through evolutionarily conserved RNA bridges', *Nat Struct Mol Biol*, 20: 1434-42.
176. Lu, Y., W. Xu, J. Ji, D. Feng, C. Sourbier, Y. Yang, J. Qu, Z. Zeng, C. Wang, X. Chang, Y. Chen, A. Mishra, M. Xu, M. J. Lee, S. Lee, J. Trepel, W. M. Linehan, X. Wang, Y. Yang, and L. Neckers. 2015. 'Alternative splicing of the cell fate determinant Numb in hepatocellular carcinoma', *Hepatology*, 62: 1122-31.
177. Luo, C., Y. Cheng, Y. Liu, L. Chen, L. Liu, N. Wei, Z. Xie, W. Wu, and Y. Feng. 2017. 'SRSF2 Regulates Alternative Splicing to Drive Hepatocellular Carcinoma Development', *Cancer Res*, 77: 1168-78.
178. Ma, P. C., T. Kijima, G. Maulik, E. A. Fox, M. Sattler, J. D. Griffin, B. E. Johnson, and R. Salgia. 2003. 'c-MET mutational analysis in small cell lung cancer: novel juxtamembrane domain mutations regulating cytoskeletal functions', *Cancer Res*, 63: 6272-81.
179. Maddugoda, M. P., M. S. Crampton, A. M. Shewan, and A. S. Yap. 2007. 'Myosin VI and vinculin cooperate during the morphogenesis of cadherin cell cell contacts in mammalian epithelial cells', *J Cell Biol*, 178: 529-40.

180. Maher, M. T., A. S. Flozak, A. M. Stocker, A. Chenn, and C. J. Gottardi. 2009. 'Activity of the beta-catenin phosphodestruction complex at cell-cell contacts is enhanced by cadherin-based adhesion', *J Cell Biol*, 186: 219-28.
181. Makeyev, E. V., J. Zhang, M. A. Carrasco, and T. Maniatis. 2007. 'The MicroRNA miR-124 promotes neuronal differentiation by triggering brain-specific alternative pre-mRNA splicing', *Mol Cell*, 27: 435-48.
182. Makishima, H., V. Visconte, H. Sakaguchi, A. M. Jankowska, S. Abu Kar, A. Jerez, B. Przychodzen, M. Bupathi, K. Guinta, M. G. Afable, M. A. Sekeres, R. A. Padgett, R. V. Tiu, and J. P. Maciejewski. 2012. 'Mutations in the spliceosome machinery, a novel and ubiquitous pathway in leukemogenesis', *Blood*, 119: 3203-10.
183. Mangold, S., S. J. Norwood, A. S. Yap, and B. M. Collins. 2012. 'The juxtamembrane domain of the E-cadherin cytoplasmic tail contributes to its interaction with Myosin VI', *Bioarchitecture*, 2: 185-8.
184. Markovtsov, V., J. M. Nikolic, J. A. Goldman, C. W. Turck, M. Y. Chou, and D. L. Black. 2000. 'Cooperative assembly of an hnRNP complex induced by a tissue-specific homolog of polypyrimidine tract binding protein', *Mol Cell Biol*, 20: 7463-79.
185. Martin, W., and E. V. Koonin. 2006. 'Introns and the origin of nucleus-cytosol compartmentalization', *Nature*, 440: 41-5.
186. Masuda, A., J. Takeda, T. Okuno, T. Okamoto, B. Ohkawara, M. Ito, S. Ishigaki, G. Sobue, and K. Ohno. 2015. 'Position-specific binding of FUS to nascent RNA regulates mRNA length', *Genes Dev*, 29: 1045-57.
187. Matera, A. G., and Z. Wang. 2014. 'A day in the life of the spliceosome', *Nat Rev Mol Cell Biol*, 15: 108-21.
188. Matos, P., J. G. Collard, and P. Jordan. 2003. 'Tumor-related alternatively spliced Rac1b is not regulated by Rho-GDP dissociation inhibitors and exhibits selective downstream signaling', *J Biol Chem*, 278: 50442-8.
189. Matter, N., P. Herrlich, and H. König. 2002. 'Signal-dependent regulation of splicing via phosphorylation of Sam68', *Nature*, 420: 691-5.
190. Mauger, D. M., C. Lin, and M. A. Garcia-Blanco. 2008. 'hnRNP H and hnRNP F complex with Fox2 to silence fibroblast growth factor receptor 2 exon IIIc', *Mol Cell Biol*, 28: 5403-19.
191. Meggendorfer, M., A. Roller, T. Haferlach, C. Eder, F. Dicker, V. Grossmann, A. Kohlmann, T. Alpermann, K. Yoshida, S. Ogawa, H. P. Koefler, W. Kern, C. Haferlach, and S. Schnittger. 2012. 'SRSF2 mutations in 275 cases with chronic myelomonocytic leukemia (CMML)', *Blood*, 120: 3080-8.
192. Melton, A. A., J. Jackson, J. Wang, and K. W. Lynch. 2007. 'Combinatorial control of signal-induced exon repression by hnRNP L and PSF', *Mol Cell Biol*, 27: 6972-84.
193. Ménétrey, J., A. Bahloul, A. L. Wells, C. M. Yengo, C. A. Morris, H. L. Sweeney, and A. Houdusse. 2005. 'The structure of the myosin VI motor reveals the mechanism of directionality reversal', *Nature*, 435: 779-85.
194. Meng, W., and M. Takeichi. 2009. 'Adherens junction: molecular architecture and regulation', *Cold Spring Harb Perspect Biol*, 1: a002899.
195. Merdzhanova, G., V. Edmond, S. De Seranno, A. Van den Broeck, L. Corcos, C. Brambilla, E. Brambilla, S. Gazzeri, and B. Eymin. 2008. 'E2F1 controls alternative splicing pattern of genes involved in apoptosis through upregulation of the splicing factor SC35', *Cell Death Differ*, 15: 1815-23.
196. Min, H., C. W. Turck, J. M. Nikolic, and D. L. Black. 1997. 'A new regulatory protein, KSRP, mediates exon inclusion through an intronic splicing enhancer', *Genes Dev*, 11: 1023-36.

197. Mitra, D., M. J. Brumlik, S. U. Okamgba, Y. Zhu, T. T. Duplessis, J. G. Parvani, S. M. Lesko, E. Brogi, and F. E. Jones. 2009. 'An oncogenic isoform of HER2 associated with locally disseminated breast cancer and trastuzumab resistance', *Mol Cancer Ther*, 8: 2152-62.
198. Morriswood, B., G. Ryzhakov, C. Puri, S. D. Arden, R. Roberts, C. Dendrou, J. Kendrick-Jones, and F. Buss. 2007. 'T6BP and NDP52 are myosin VI binding partners with potential roles in cytokine signalling and cell adhesion', *J Cell Sci*, 120: 2574-85.
199. Mozaffari-Jovin, S., K. F. Santos, H. H. Hsiao, C. L. Will, H. Urlaub, M. C. Wahl, and R. Luhrmann. 2012. 'The Prp8 RNase H-like domain inhibits Brr2-mediated U4/U6 snRNA unwinding by blocking Brr2 loading onto the U4 snRNA', *Genes Dev*, 26: 2422-34.
200. Mukherjea, M., P. Llinas, H. Kim, M. Travaglia, D. Safer, J. Ménétrey, C. Franzini-Armstrong, P. R. Selvin, A. Houdusse, and H. L. Sweeney. 2009. 'Myosin VI dimerization triggers an unfolding of a three-helix bundle in order to extend its reach', *Mol Cell*, 35: 305-15.
201. Musco, G., G. Stier, C. Joseph, M. A. Castiglione Morelli, M. Nilges, T. J. Gibson, and A. Pastore. 1996. 'Three-dimensional structure and stability of the KH domain: molecular insights into the fragile X syndrome', *Cell*, 85: 237-45.
202. Niño, C. A., S. Sala, and S. Polo. 2019. 'When ubiquitin meets E-cadherin: Plasticity of the epithelial cellular barrier', *Semin Cell Dev Biol*, 93: 136-44.
203. O'Loughlin, T., J. Kendrick-Jones, and F. Buss. 2020. 'Approaches to Identify and Characterise MYO6-Cargo Interactions', *Adv Exp Med Biol*, 1239: 355-80.
204. Obeng, E. A., R. J. Chappell, M. Seiler, M. C. Chen, D. R. Campagna, P. J. Schmidt, R. K. Schneider, A. M. Lord, L. Wang, R. G. Gambe, M. E. McConkey, A. M. Ali, A. Raza, L. Yu, S. Buonamici, P. G. Smith, A. Mullally, C. J. Wu, M. D. Fleming, and B. L. Ebert. 2016. 'Physiologic Expression of Sf3b1(K700E) Causes Impaired Erythropoiesis, Aberrant Splicing, and Sensitivity to Therapeutic Spliceosome Modulation', *Cancer Cell*, 30: 404-17.
205. Oguchi, Y., S. V. Mikhailenko, T. Ohki, A. O. Olivares, E. M. De La Cruz, and S. Ishiwata. 2008. 'Load-dependent ADP binding to myosins V and VI: implications for subunit coordination and function', *Proc Natl Acad Sci U S A*, 105: 7714-9.
206. Oh, Hk, E. Lee, H. N. Jang, J. Lee, H. Moon, Z. Sheng, Y. Jun, T. J. Loh, S. Cho, J. Zhou, M. R. Green, X. Zheng, and H. Shen. 2013. 'hnRNP A1 contacts exon 5 to promote exon 6 inclusion of apoptotic Fas gene', *Apoptosis*, 18: 825-35.
207. Pandit, S., Y. Zhou, L. Shiue, G. Coutinho-Mansfield, H. Li, J. Qiu, J. Huang, G. W. Yeo, M. Ares, Jr., and X. D. Fu. 2013. 'Genome-wide analysis reveals SR protein cooperation and competition in regulated splicing', *Mol Cell*, 50: 223-35.
208. Papasaikas, P., and J. Valcarcel. 2016. 'The Spliceosome: The Ultimate RNA Chaperone and Sculptor', *Trends Biochem Sci*, 41: 33-45.
209. Park, H., B. Ramamurthy, M. Travaglia, D. Safer, L. Q. Chen, C. Franzini-Armstrong, P. R. Selvin, and H. L. Sweeney. 2006. 'Full-length myosin VI dimerizes and moves processively along actin filaments upon monomer clustering', *Mol Cell*, 21: 331-6.
210. Paronetto, M. P., T. Achsel, A. Massiello, C. E. Chalfant, and C. Sette. 2007. 'The RNA-binding protein Sam68 modulates the alternative splicing of Bcl-x', *J Cell Biol*, 176: 929-39.
211. Paronetto, M. P., I. Bernardis, E. Volpe, E. Bechara, E. Sebestyen, E. Eyra, and J. Valcarcel. 2014. 'Regulation of FAS exon definition and apoptosis by the Ewing sarcoma protein', *Cell Rep*, 7: 1211-26.
212. Patnaik, M. M., T. L. Lasho, C. M. Finke, C. A. Hanson, J. M. Hodnefield, R. A. Knudson, R. P. Ketterling, A. Pardanani, and A. Tefferi. 2013. 'Spliceosome mutations involving SRSF2, SF3B1, and U2AF35 in chronic myelomonocytic leukemia: prevalence, clinical correlates, and prognostic relevance', *Am J Hematol*, 88: 201-6.

213. Patton, J. G., S. A. Mayer, P. Tempst, and B. Nadal-Ginard. 1991. 'Characterization and molecular cloning of polypyrimidine tract-binding protein: a component of a complex necessary for pre-mRNA splicing', *Genes Dev*, 5: 1237-51.
214. Pellarin, I., A. Dall'Acqua, A. Gambelli, I. Pellizzari, S. D'Andrea, M. Sonego, I. Lorenzon, M. Schiappacassi, B. Belletti, and G. Baldassarre. 2020. 'Splicing factor proline- and glutamine-rich (SFPQ) protein regulates platinum response in ovarian cancer-modulating SRSF2 activity', *Oncogene*, 39: 4390-403.
215. Pena, V., A. Rozov, P. Fabrizio, R. Luhrmann, and M. C. Wahl. 2008. 'Structure and function of an RNase H domain at the heart of the spliceosome', *Embo j*, 27: 2929-40.
216. Penengo, L., M. Mapelli, A. G. Murachelli, S. Confalonieri, L. Magri, A. Musacchio, P. P. Di Fiore, S. Polo, and T. R. Schneider. 2006. 'Crystal structure of the ubiquitin binding domains of rabex-5 reveals two modes of interaction with ubiquitin', *Cell*, 124: 1183-95.
217. Perez, I., J. G. McAfee, and J. G. Patton. 1997. 'Multiple RRM contribute to RNA binding specificity and affinity for polypyrimidine tract binding protein', *Biochemistry*, 36: 11881-90.
218. Pichith, D., M. Travaglia, Z. Yang, X. Liu, A. B. Zong, D. Safer, and H. L. Sweeney. 2009. 'Cargo binding induces dimerization of myosin VI', *Proc Natl Acad Sci U S A*, 106: 17320-4.
219. Poulikakos, P. I., Y. Persaud, M. Janakiraman, X. Kong, C. Ng, G. Moriceau, H. Shi, M. Atefi, B. Titz, M. T. Gabay, M. Salton, K. B. Dahlman, M. Tadi, J. A. Wargo, K. T. Flaherty, M. C. Kelley, T. Misteli, P. B. Chapman, J. A. Sosman, T. G. Graeber, A. Ribas, R. S. Lo, N. Rosen, and D. B. Solit. 2011. 'RAF inhibitor resistance is mediated by dimerization of aberrantly spliced BRAF(V600E)', *Nature*, 480: 387-90.
220. Pritchard-Jones, R. O., D. B. Dunn, Y. Qiu, A. H. Varey, A. Orlando, H. Rigby, S. J. Harper, and D. O. Bates. 2007. 'Expression of VEGF(xxx)b, the inhibitory isoforms of VEGF, in malignant melanoma', *Br J Cancer*, 97: 223-30.
221. Puente, X. S., S. Bea, R. Valdes-Mas, N. Villamor, J. Gutierrez-Abril, J. I. Martin-Subero, M. Munar, C. Rubio-Perez, P. Jares, M. Aymerich, T. Baumann, R. Beekman, L. Belver, A. Carrio, G. Castellano, G. Clot, E. Colado, D. Colomer, D. Costa, J. Delgado, A. Enjuanes, X. Estivill, A. A. Ferrando, J. L. Gelpi, B. Gonzalez, S. Gonzalez, M. Gonzalez, M. Gut, J. M. Hernandez-Rivas, M. Lopez-Guerra, D. Martin-Garcia, A. Navarro, P. Nicolas, M. Orozco, A. R. Payer, M. Pinyol, D. G. Pisano, D. A. Puente, A. C. Queiros, V. Quesada, C. M. Romeo-Casabona, C. Royo, R. Royo, M. Rozman, N. Russinol, I. Salaverria, K. Stamatopoulos, H. G. Stunnenberg, D. Tamborero, M. J. Terol, A. Valencia, N. Lopez-Bigas, D. Torrents, I. Gut, A. Lopez-Guillermo, C. Lopez-Otin, and E. Campo. 2015. 'Non-coding recurrent mutations in chronic lymphocytic leukaemia', *Nature*, 526: 519-24.
222. Puppo, M., G. Bucci, M. Rossi, M. Giovarelli, D. Bordo, A. Moshiri, F. Gorlero, R. Gherzi, and P. Briata. 2016. 'miRNA-Mediated KHSRP Silencing Rewires Distinct Post-transcriptional Programs during TGF-beta-Induced Epithelial-to-Mesenchymal Transition', *Cell Rep*, 16: 967-78.
223. Racca, C., A. Gardiol, T. Eom, J. Ule, A. Triller, and R. B. Darnell. 2010. 'The Neuronal Splicing Factor Nova Co-Localizes with Target RNAs in the Dendrite', *Front Neural Circuits*, 4: 5.
224. Ray, D., H. Kazan, E. T. Chan, L. Pena Castillo, S. Chaudhry, S. Talukder, B. J. Blencowe, Q. Morris, and T. R. Hughes. 2009. 'Rapid and systematic analysis of the RNA recognition specificities of RNA-binding proteins', *Nat Biotechnol*, 27: 667-70.
225. Reid, D. C., B. L. Chang, S. I. Gunderson, L. Alpert, W. A. Thompson, and W. G. Fairbrother. 2009. 'Next-generation SELEX identifies sequence and structural determinants of splicing factor binding in human pre-mRNA sequence', *Rna*, 15: 2385-97.
226. Reyes, J. L., E. H. Gustafson, H. R. Luo, M. J. Moore, and M. M. Konarska. 1999. 'The C-terminal region of hPrp8 interacts with the conserved GU dinucleotide at the 5' splice site', *Rna*, 5: 167-79.

227. Rogelj, B., L. E. Easton, G. K. Bogu, L. W. Stanton, G. Rot, T. Curk, B. Zupan, Y. Sugimoto, M. Modic, N. Haberman, J. Tollervy, R. Fujii, T. Takumi, C. E. Shaw, and J. Ule. 2012. 'Widespread binding of FUS along nascent RNA regulates alternative splicing in the brain', *Sci Rep*, 2: 603.
228. Rogez, B., L. Würthner, A. B. Petrova, F. B. Zierhut, D. Saczko-Brack, M. A. Huergo, C. Batters, E. Frey, and C. Veigel. 2019. 'Reconstitution reveals how myosin-VI self-organises to generate a dynamic mechanism of membrane sculpting', *Nat Commun*, 10: 3305.
229. Rothrock, C. R., A. E. House, and K. W. Lynch. 2005. 'HnRNP L represses exon splicing via a regulated exonic splicing silencer', *Embo j*, 24: 2792-802.
230. Russo, A., M. Catillo, D. Esposito, P. Briata, C. Pietropaolo, and G. Russo. 2011. 'Autoregulatory circuit of human rpL3 expression requires hnRNP H1, NPM and KHSRP', *Nucleic Acids Res*, 39: 7576-85.
231. Sahlender, D. A., R. C. Roberts, S. D. Arden, G. Spudich, M. J. Taylor, J. P. Luzio, J. Kendrick-Jones, and F. Buss. 2005. 'Optineurin links myosin VI to the Golgi complex and is involved in Golgi organization and exocytosis', *J Cell Biol*, 169: 285-95.
232. Sato, S., M. Idogawa, K. Honda, G. Fujii, H. Kawashima, K. Takekuma, A. Hoshika, S. Hirohashi, and T. Yamada. 2005. 'Beta-catenin interacts with the FUS proto-oncogene product and regulates pre-mRNA splicing', *Gastroenterology*, 129: 1225-36.
233. Schaal, T. D., and T. Maniatis. 1999. 'Selection and characterization of pre-mRNA splicing enhancers: identification of novel SR protein-specific enhancer sequences', *Mol Cell Biol*, 19: 1705-19.
234. Schaukowitch, K., J. Y. Joo, X. Liu, J. K. Watts, C. Martinez, and T. K. Kim. 2014. 'Enhancer RNA facilitates NELF release from immediate early genes', *Mol Cell*, 56: 29-42.
235. Schnelzer, A., D. Prechtel, U. Knaus, K. Dehne, M. Gerhard, H. Graeff, N. Harbeck, M. Schmitt, and E. Lengyel. 2000. 'Rac1 in human breast cancer: overexpression, mutation analysis, and characterization of a new isoform, Rac1b', *Oncogene*, 19: 3013-20.
236. Screamon, G. R., M. V. Bell, J. I. Bell, and D. G. Jackson. 1993. 'The identification of a new alternative exon with highly restricted tissue expression in transcripts encoding the mouse Pgp-1 (CD44) homing receptor. Comparison of all 10 variable exons between mouse, human, and rat', *J Biol Chem*, 268: 12235-8.
237. Seiler, M., A. Yoshimi, R. Darman, B. Chan, G. Keaney, M. Thomas, A. A. Agrawal, B. Caleb, A. Csibi, E. Sean, P. Fekkes, C. Karr, V. Klimek, G. Lai, L. Lee, P. Kumar, S. C. Lee, X. Liu, C. Mackenzie, C. Meeske, Y. Mizui, E. Padron, E. Park, E. Pazolli, S. Peng, S. Prajapati, J. Taylor, T. Teng, J. Wang, M. Warmuth, H. Yao, L. Yu, P. Zhu, O. Abdel-Wahab, P. G. Smith, and S. Buonamici. 2018. 'H3B-8800, an orally available small-molecule splicing modulator, induces lethality in spliceosome-mutant cancers', *Nat Med*, 24: 497-504.
238. Shapiro, I. M., A. W. Cheng, N. C. Flytzanis, M. Balsamo, J. S. Condeelis, M. H. Oktay, C. B. Burge, and F. B. Gertler. 2011. 'An EMT-driven alternative splicing program occurs in human breast cancer and modulates cellular phenotype', *PLoS Genet*, 7: e1002218.
239. Sharp, P. A. 1994. 'Split genes and RNA splicing', *Cell*, 77: 805-15.
240. ———. 2005. 'The discovery of split genes and RNA splicing', *Trends Biochem Sci*, 30: 279-81.
241. Shen, H., J. L. Kan, C. Ghigna, G. Biamonti, and M. R. Green. 2004. 'A single polypyrimidine tract binding protein (PTB) binding site mediates splicing inhibition at mouse IgM exons M1 and M2', *Rna*, 10: 787-94.
242. Shi, Y. 2017. 'Mechanistic insights into precursor messenger RNA splicing by the spliceosome', *Nat Rev Mol Cell Biol*, 18: 655-70.
243. Shirai, C. L., B. S. White, M. Tripathi, R. Tapia, J. N. Ley, M. Ndonwi, S. Kim, J. Shao, A. Carver, B. Saez, R. S. Fulton, C. Fronick, M. O'Laughlin, C. Lagisetti, T. R. Webb, T. A. Graubert, and M.

- J. Walter. 2017. 'Mutant U2AF1-expressing cells are sensitive to pharmacological modulation of the spliceosome', *Nat Commun*, 8: 14060.
244. Shiroki, T., M. Yokoyama, N. Tanuma, R. Maejima, K. Tamai, K. Yamaguchi, T. Oikawa, T. Noguchi, K. Miura, T. Fujiya, H. Shima, I. Sato, N. Murata-Kamiya, M. Hatakeyama, K. Iijima, T. Shimosegawa, and K. Satoh. 2017. 'Enhanced expression of the M2 isoform of pyruvate kinase is involved in gastric cancer development by regulating cancer-specific metabolism', *Cancer Sci*, 108: 931-40.
245. Shultz, J. C., R. W. Goehe, D. S. Wijesinghe, C. Murudkar, A. J. Hawkins, J. W. Shay, J. D. Minna, and C. E. Chalfant. 2010. 'Alternative splicing of caspase 9 is modulated by the phosphoinositide 3-kinase/Akt pathway via phosphorylation of SRp30a', *Cancer Res*, 70: 9185-96.
246. Siebel, C. W., A. Admon, and D. C. Rio. 1995. 'Soma-specific expression and cloning of PSI, a negative regulator of P element pre-mRNA splicing', *Genes Dev*, 9: 269-83.
247. Siebel, C. W., R. Kanaar, and D. C. Rio. 1994. 'Regulation of tissue-specific P-element pre-mRNA splicing requires the RNA-binding protein PSI', *Genes Dev*, 8: 1713-25.
248. Singh, A., A. E. Karnoub, T. R. Palmby, E. Lengyel, J. Sondek, and C. J. Der. 2004. 'Rac1b, a tumor associated, constitutively active Rac1 splice variant, promotes cellular transformation', *Oncogene*, 23: 9369-80.
249. Smith, E., W. Zhou, P. Shindiapina, S. Sif, C. Li, and R. A. Baiocchi. 2018. 'Recent advances in targeting protein arginine methyltransferase enzymes in cancer therapy', *Expert Opin Ther Targets*, 22: 527-45.
250. Spellman, R., M. Llorian, and C. W. Smith. 2007. 'Crossregulation and functional redundancy between the splicing regulator PTB and its paralogs nPTB and ROD1', *Mol Cell*, 27: 420-34.
251. Staknis, D., and R. Reed. 1994. 'SR proteins promote the first specific recognition of Pre-mRNA and are present together with the U1 small nuclear ribonucleoprotein particle in a general splicing enhancer complex', *Mol Cell Biol*, 14: 7670-82.
252. Staley, J. P., and C. Guthrie. 1999. 'An RNA switch at the 5' splice site requires ATP and the DEAD box protein Prp28p', *Mol Cell*, 3: 55-64.
253. Sun, H., E. C. Chow, S. Liu, Y. Du, and K. S. Pang. 2008. 'The Caco-2 cell monolayer: usefulness and limitations', *Expert Opin Drug Metab Toxicol*, 4: 395-411.
254. Supek, F., B. Minana, J. Valcarcel, T. Gabaldon, and B. Lehner. 2014. 'Synonymous mutations frequently act as driver mutations in human cancers', *Cell*, 156: 1324-35.
255. Sveen, A., S. Kilpinen, A. Ruusulehto, R. A. Lothe, and R. I. Skotheim. 2016. 'Aberrant RNA splicing in cancer; expression changes and driver mutations of splicing factor genes', *Oncogene*, 35: 2413-27.
256. Swiatecka-Urban, A., C. Boyd, B. Coutermarsh, K. H. Karlson, R. Barnaby, L. Aschenbrenner, G. M. Langford, T. Hasson, and B. A. Stanton. 2004. 'Myosin VI regulates endocytosis of the cystic fibrosis transmembrane conductance regulator', *J Biol Chem*, 279: 38025-31.
257. Tacke, R., and J. L. Manley. 1995. 'The human splicing factors ASF/SF2 and SC35 possess distinct, functionally significant RNA binding specificities', *Embo j*, 14: 3540-51.
258. Tan, A. Y., and J. L. Manley. 2009. 'The TET family of proteins: functions and roles in disease', *J Mol Cell Biol*, 1: 82-92.
259. Tilio, M., V. Gambini, J. Wang, C. Garulli, C. Kalogris, C. Andreani, C. Bartolacci, M. Elempuru Zabaleta, L. Pietrella, A. Hysi, M. Iezzi, B. Belletti, F. Orlando, M. Provinciali, R. Galeazzi, C. Marchini, and A. Amici. 2016. 'Irreversible inhibition of Delta16HER2 is necessary to suppress Delta16HER2-positive breast carcinomas resistant to Lapatinib', *Cancer Lett*, 381: 76-84.

260. Tomatis, V. M., A. Papadopoulos, N. T. Malintan, S. Martin, T. Wallis, R. S. Gormal, J. Kendrick-Jones, F. Buss, and F. A. Meunier. 2013. 'Myosin VI small insert isoform maintains exocytosis by tethering secretory granules to the cortical actin', *J Cell Biol*, 200: 301-20.
261. Tripathi, V., J. H. Shin, C. H. Stuelten, and Y. E. Zhang. 2019. 'TGF-beta-induced alternative splicing of TAK1 promotes EMT and drug resistance', *Oncogene*, 38: 3185-200.
262. Tu, Y., S. Renner, F. Xu, A. Fleishman, J. Taylor, J. Weisz, R. Vescio, M. Rettig, J. Berenson, S. Krajewski, J. C. Reed, and A. Lichtenstein. 1998. 'BCL-X expression in multiple myeloma: possible indicator of chemoresistance', *Cancer Res*, 58: 256-62.
263. Tumbarello, D. A., J. Kendrick-Jones, and F. Buss. 2013. 'Myosin VI and its cargo adaptors - linking endocytosis and autophagy', *J Cell Sci*, 126: 2561-70.
264. Tumbarello, D. A., B. J. Waxse, S. D. Arden, N. A. Bright, J. Kendrick-Jones, and F. Buss. 2012. 'Autophagy receptors link myosin VI to autophagosomes to mediate Tom1-dependent autophagosome maturation and fusion with the lysosome', *Nat Cell Biol*, 14: 1024-35.
265. Turunen, J. J., E. H. Niemela, B. Verma, and M. J. Frilander. 2013. 'The significant other: splicing by the minor spliceosome', *Wiley Interdiscip Rev RNA*, 4: 61-76.
266. Urbanski, L. M., N. Leclair, and O. Anczukow. 2018. 'Alternative-splicing defects in cancer: Splicing regulators and their downstream targets, guiding the way to novel cancer therapeutics', *Wiley Interdiscip Rev RNA*, 9: e1476.
267. Valacca, C., S. Bonomi, E. Buratti, S. Pedrotti, F. E. Baralle, C. Sette, C. Ghigna, and G. Biamonti. 2010. 'Sam68 regulates EMT through alternative splicing-activated nonsense-mediated mRNA decay of the SF2/ASF proto-oncogene', *J Cell Biol*, 191: 87-99.
268. Valadkhan, S., and J. L. Manley. 2001. 'Splicing-related catalysis by protein-free snRNAs', *Nature*, 413: 701-7.
269. Van Nostrand, E. L., G. A. Pratt, A. A. Shishkin, C. Gelboin-Burkhart, M. Y. Fang, B. Sundararaman, S. M. Blue, T. B. Nguyen, C. Surka, K. Elkins, R. Stanton, F. Rigo, M. Guttman, and G. W. Yeo. 2016. 'Robust transcriptome-wide discovery of RNA-binding protein binding sites with enhanced CLIP (eCLIP)', *Nat Methods*, 13: 508-14.
270. Varey, A. H., E. S. Rennel, Y. Qiu, H. S. Bevan, R. M. Perrin, S. Raffy, A. R. Dixon, C. Paraskeva, O. Zaccaro, A. B. Hassan, S. J. Harper, and D. O. Bates. 2008. 'VEGF 165 b, an antiangiogenic VEGF-A isoform, binds and inhibits bevacizumab treatment in experimental colorectal carcinoma: balance of pro- and antiangiogenic VEGF-A isoforms has implications for therapy', *Br J Cancer*, 98: 1366-79.
271. Varsano, T., M. Q. Dong, I. Niesman, H. Gacula, X. Lou, T. Ma, J. R. Testa, J. R. Yates, 3rd, and M. G. Farquhar. 2006. 'GIPC is recruited by APPL to peripheral TrkA endosomes and regulates TrkA trafficking and signaling', *Mol Cell Biol*, 26: 8942-52.
272. Venables, J. P., J. P. Brosseau, G. Gadea, R. Klinck, P. Prinos, J. F. Beaulieu, E. Lapointe, M. Durand, P. Thibault, K. Tremblay, F. Rousset, J. Tazi, S. Abou Elela, and B. Chabot. 2013. 'RBF2 is an important regulator of mesenchymal tissue-specific splicing in both normal and cancer tissues', *Mol Cell Biol*, 33: 396-405.
273. Venables, J. P., R. Klinck, A. Bramard, L. Inkel, G. Dufresne-Martin, C. Koh, J. Gervais-Bird, E. Lapointe, U. Froehlich, M. Durand, D. Gendron, J. P. Brosseau, P. Thibault, J. F. Lucier, K. Tremblay, P. Prinos, R. J. Wellinger, B. Chabot, C. Rancourt, and S. A. Elela. 2008. 'Identification of alternative splicing markers for breast cancer', *Cancer Res*, 68: 9525-31.
274. Venables, J. P., R. Klinck, C. Koh, J. Gervais-Bird, A. Bramard, L. Inkel, M. Durand, S. Couture, U. Froehlich, E. Lapointe, J. F. Lucier, P. Thibault, C. Rancourt, K. Tremblay, P. Prinos, B. Chabot, and S. A. Elela. 2009. 'Cancer-associated regulation of alternative splicing', *Nat Struct Mol Biol*, 16: 670-6.

275. Verma, B., M. V. Akinyi, A. J. Norppa, and M. J. Frilander. 2018. 'Minor spliceosome and disease', *Semin Cell Dev Biol*, 79: 103-12.
276. Vuong, C. K., D. L. Black, and S. Zheng. 2016. 'The neurogenetics of alternative splicing', *Nat Rev Neurosci*, 17: 265-81.
277. Wada, R., S. Yagihashi, and Z. Naito. 2016. 'mRNA expression of delta-HER2 and its clinicopathological correlation in HER2-overexpressing breast cancer', *Mol Med Rep*, 14: 5104-10.
278. Wagner, E. J., and M. A. Garcia-Blanco. 2001. 'Polypyrimidine tract binding protein antagonizes exon definition', *Mol Cell Biol*, 21: 3281-8.
279. ———. 2002. 'RNAi-mediated PTB depletion leads to enhanced exon definition', *Mol Cell*, 10: 943-9.
280. Wagner, W., K. Lippmann, F. F. Heisler, K. V. Gromova, F. L. Lombino, M. K. Roesler, Y. Pechmann, S. Hornig, M. Schweizer, S. Polo, J. R. Schwarz, J. Eilers, and M. Kneussel. 2019. 'Myosin VI Drives Clathrin-Mediated AMPA Receptor Endocytosis to Facilitate Cerebellar Long-Term Depression', *Cell Rep*, 28: 11-20.e9.
281. Wang, E. T., R. Sandberg, S. Luo, I. Khrebtkova, L. Zhang, C. Mayr, S. F. Kingsmore, G. P. Schroth, and C. B. Burge. 2008. 'Alternative isoform regulation in human tissue transcriptomes', *Nature*, 456: 470-6.
282. Wang, H., L. Fang, J. Jiang, Y. Kuang, B. Wang, X. Shang, P. Han, Y. Li, M. Liu, Z. Zhang, and P. Li. 2018. 'The cisplatin-induced lncRNA PANDAR dictates the chemoresistance of ovarian cancer via regulating SFRS2-mediated p53 phosphorylation', *Cell Death Dis*, 9: 1103.
283. Wang, W., Q. He, J. Sun, Z. Liu, L. Zhao, Z. Lu, X. Zhou, and A. Wang. 2017. 'Pyruvate kinase M2 deregulation enhances the metastatic potential of tongue squamous cell carcinoma', *Oncotarget*, 8: 68252-62.
284. Wang, Y., Y. Bao, S. Zhang, and Z. Wang. 2020. 'Splicing dysregulation in cancer: from mechanistic understanding to a new class of therapeutic targets', *Sci China Life Sci*.
285. Wang, Y., C. G. Cheong, T. M. Hall, and Z. Wang. 2009. 'Engineering splicing factors with designed specificities', *Nat Methods*, 6: 825-30.
286. Wassie, E. A., R. Itzykson, T. L. Lasho, O. Kosmider, C. M. Finke, C. A. Hanson, R. P. Ketterling, E. Solary, A. Tefferi, and M. M. Patnaik. 2014. 'Molecular and prognostic correlates of cytogenetic abnormalities in chronic myelomonocytic leukemia: a Mayo Clinic-French Consortium Study', *Am J Hematol*, 89: 1111-5.
287. Wells, A. L., A. W. Lin, L. Q. Chen, D. Safer, S. M. Cain, T. Hasson, B. O. Carragher, R. A. Milligan, and H. L. Sweeney. 1999. 'Myosin VI is an actin-based motor that moves backwards', *Nature*, 401: 505-8.
288. Weyn-Vanhenyryck, S. M., A. Mele, Q. Yan, S. Sun, N. Farny, Z. Zhang, C. Xue, M. Herre, P. A. Silver, M. Q. Zhang, A. R. Krainer, R. B. Darnell, and C. Zhang. 2014. 'HITS-CLIP and integrative modeling define the Rbfox splicing-regulatory network linked to brain development and autism', *Cell Rep*, 6: 1139-52.
289. Will, C. L., and R. Luhrmann. 2011. 'Spliceosome structure and function', *Cold Spring Harb Perspect Biol*, 3.
290. Witkiewicz, A. K., E. A. McMillan, U. Balaji, G. Baek, W. C. Lin, J. Mansour, M. Mollaei, K. U. Wagner, P. Koduru, A. Yopp, M. A. Choti, C. J. Yeo, P. McCue, M. A. White, and E. S. Knudsen. 2015. 'Whole-exome sequencing of pancreatic cancer defines genetic diversity and therapeutic targets', *Nat Commun*, 6: 6744.
291. Wollscheid, H. P., M. Biancospino, F. He, E. Magistrati, E. Molteni, M. Lupia, P. Soffientini, K. Rottner, U. Cavallaro, U. Pozzoli, M. Mapelli, K. J. Walters, and S. Polo. 2016. 'Diverse functions

- of myosin VI elucidated by an isoform-specific alpha-helix domain', *Nat Struct Mol Biol*, 23: 300-08.
292. Woolard, J., W. Y. Wang, H. S. Bevan, Y. Qiu, L. Morbidelli, R. O. Pritchard-Jones, T. G. Cui, M. Sugiono, E. Waite, R. Perrin, R. Foster, J. Digby-Bell, J. D. Shields, C. E. Whittles, R. E. Mushens, D. A. Gillatt, M. Ziche, S. J. Harper, and D. O. Bates. 2004. 'VEGF165b, an inhibitory vascular endothelial growth factor splice variant: mechanism of action, in vivo effect on angiogenesis and endogenous protein expression', *Cancer Res*, 64: 7822-35.
293. Xerri, L., P. Parc, P. Brousset, D. Schlaifer, J. Hassoun, J. C. Reed, S. Krajewski, and D. Birnbaum. 1996. 'Predominant expression of the long isoform of Bcl-x (Bcl-xL) in human lymphomas', *Br J Haematol*, 92: 900-6.
294. Xiao, X., Z. Wang, M. Jang, R. Nutiu, E. T. Wang, and C. B. Burge. 2009. 'Splice site strength-dependent activity and genetic buffering by poly-G runs', *Nat Struct Mol Biol*, 16: 1094-100.
295. Xu, Y., X. D. Gao, J. H. Lee, H. Huang, H. Tan, J. Ahn, L. M. Reinke, M. E. Peter, Y. Feng, D. Gius, K. P. Siziopikou, J. Peng, X. Xiao, and C. Cheng. 2014. 'Cell type-restricted activity of hnRNPM promotes breast cancer metastasis via regulating alternative splicing', *Genes Dev*, 28: 1191-203.
296. Yeo, G. W., N. G. Coufal, T. Y. Liang, G. E. Peng, X. D. Fu, and F. H. Gage. 2009. 'An RNA code for the FOX2 splicing regulator revealed by mapping RNA-protein interactions in stem cells', *Nat Struct Mol Biol*, 16: 130-7.
297. Yomoda, J., M. Muraki, N. Kataoka, T. Hosoya, M. Suzuki, M. Hagiwara, and H. Kimura. 2008. 'Combination of Clk family kinase and SRp75 modulates alternative splicing of Adenovirus E1A', *Genes Cells*, 13: 233-44.
298. Yoshida, H., W. Cheng, J. Hung, D. Montell, E. Geisbrecht, D. Rosen, J. Liu, and H. Naora. 2004. 'Lessons from border cell migration in the Drosophila ovary: A role for myosin VI in dissemination of human ovarian cancer', *Proc Natl Acad Sci U S A*, 101: 8144-9.
299. Yu, C., W. Feng, Z. Wei, Y. Miyanori, W. Wen, Y. Zhao, and M. Zhang. 2009. 'Myosin VI undergoes cargo-mediated dimerization', *Cell*, 138: 537-48.
300. Zahler, A. M., C. K. Damgaard, J. Kjems, and M. Caputi. 2004. 'SC35 and heterogeneous nuclear ribonucleoprotein A/B proteins bind to a juxtaposed exonic splicing enhancer/exonic splicing silencer element to regulate HIV-1 tat exon 2 splicing', *J Biol Chem*, 279: 10077-84.
301. Zhang, C., W. H. Li, A. R. Krainer, and M. Q. Zhang. 2008. 'RNA landscape of evolution for optimal exon and intron discrimination', *Proc Natl Acad Sci U S A*, 105: 5797-802.
302. Zhang, C., Z. Zhang, J. Castle, S. Sun, J. Johnson, A. R. Krainer, and M. Q. Zhang. 2008. 'Defining the regulatory network of the tissue-specific splicing factors Fox-1 and Fox-2', *Genes Dev*, 22: 2550-63.
303. Zhang, J., and J. L. Manley. 2013. 'Misregulation of pre-mRNA alternative splicing in cancer', *Cancer Discov*, 3: 1228-37.
304. Zhao, J., Y. Sun, Y. Huang, F. Song, Z. Huang, Y. Bao, J. Zuo, D. Saffen, Z. Shao, W. Liu, and Y. Wang. 2017. 'Functional analysis reveals that RBM10 mutations contribute to lung adenocarcinoma pathogenesis by deregulating splicing', *Sci Rep*, 7: 40488.
305. Zhong, X. Y., J. H. Ding, J. A. Adams, G. Ghosh, and X. D. Fu. 2009. 'Regulation of SR protein phosphorylation and alternative splicing by modulating kinetic interactions of SRPK1 with molecular chaperones', *Genes Dev*, 23: 482-95.
306. Zhou, A., A. C. Ou, A. Cho, E. J. Benz, Jr., and S. C. Huang. 2008. 'Novel splicing factor RBM25 modulates Bcl-x pre-mRNA 5' splice site selection', *Mol Cell Biol*, 28: 5924-36.




307. Zhou, C., S. Licciulli, J. L. Avila, M. Cho, S. Troutman, P. Jiang, A. V. Kossenkov, L. C. Showe, Q. Liu, A. Vachani, S. M. Albelda, and J. L. Kissil. 2013. 'The Rac1 splice form Rac1b promotes K-ras-induced lung tumorigenesis', *Oncogene*, 32: 903-9.
308. Zhou, H. L., A. P. Baraniak, and H. Lou. 2007. 'Role for Fox-1/Fox-2 in mediating the neuronal pathway of calcitonin/calcitonin gene-related peptide alternative RNA processing', *Mol Cell Biol*, 27: 830-41.
309. Zhou, H., M. Mangelsdorf, J. Liu, L. Zhu, and J. Y. Wu. 2014. 'RNA-binding proteins in neurological diseases', *Sci China Life Sci*, 57: 432-44.
310. Zhou, Y. Q., C. He, Y. Q. Chen, D. Wang, and M. H. Wang. 2003. 'Altered expression of the RON receptor tyrosine kinase in primary human colorectal adenocarcinomas: generation of different splicing RON variants and their oncogenic potential', *Oncogene*, 22: 186-97.
311. Zhou, Z., and X. D. Fu. 2013. 'Regulation of splicing by SR proteins and SR protein-specific kinases', *Chromosoma*, 122: 191-207.
312. Zhou, Z., J. Qiu, W. Liu, Y. Zhou, R. M. Plocinik, H. Li, Q. Hu, G. Ghosh, J. A. Adams, M. G. Rosenfeld, and X. D. Fu. 2012. 'The Akt-SRPK-SR axis constitutes a major pathway in transducing EGF signaling to regulate alternative splicing in the nucleus', *Mol Cell*, 47: 422-33.
313. Zhu, J., A. Mayeda, and A. R. Krainer. 2001. 'Exon identity established through differential antagonism between exonic splicing silencer-bound hnRNP A1 and enhancer-bound SR proteins', *Mol Cell*, 8: 1351-61.
314. Zihni, C., C. Mills, K. Matter, and M. S. Balda. 2016. 'Tight junctions: from simple barriers to multifunctional molecular gates', *Nat Rev Mol Cell Biol*, 17: 564-80.
315. Zuo, P., and T. Maniatis. 1996. 'The splicing factor U2AF35 mediates critical protein-protein interactions in constitutive and enhancer-dependent splicing', *Genes Dev*, 10: 1356-68.

ARTICLE

<https://doi.org/10.1038/s41467-019-12855-6>

OPEN

Clathrin light chain A drives selective myosin VI recruitment to clathrin-coated pits under membrane tension

Matteo Biancospino^{1,6}, Gwen R. Buel ^{2,6}, Carlos A. Niño^{1,6}, Elena Maspero¹, Rossella Scotto di Perrotolo¹, Andrea Raimondi ³, Lisa Redlingshöfer⁴, Janine Weber¹, Frances M. Brodsky^{4*}, Kylie J. Walters^{2*} & Simona Polo ^{1,5*}

Clathrin light chains (CLCa and CLCb) are major constituents of clathrin-coated vesicles. Unique functions for these evolutionary conserved paralogs remain elusive, and their role in clathrin-mediated endocytosis in mammalian cells is debated. Here, we find and structurally characterize a direct and selective interaction between CLCa and the long isoform of the actin motor protein myosin VI, which is expressed exclusively in highly polarized tissues. Using genetically-reconstituted Caco-2 cysts as proxy for polarized epithelia, we provide evidence for coordinated action of myosin VI and CLCa at the apical surface where these proteins are essential for fission of clathrin-coated pits. We further find that myosin VI and Huntingtin-interacting protein 1-related protein (Hip1R) are mutually exclusive interactors with CLCa, and suggest a model for the sequential function of myosin VI and Hip1R in actin-mediated clathrin-coated vesicle budding.

¹IFOM, Fondazione Istituto FIRC di Oncologia Molecolare, 20139 Milan, Italy. ²Structural Biophysics Laboratory, Center for Cancer Research, National Cancer Institute, Frederick, MD 21702, USA. ³Experimental Imaging Center, San Raffaele Scientific Institute, Milan, Italy. ⁴Division of Biosciences, University College London, London WC1E 6BT, UK. ⁵Dipartimento di Oncologia ed Emato-oncologia, Università degli Studi di Milano, 20122 Milan, Italy. ⁶These authors contributed equally: Matteo Biancospino, Gwen R. Buel, Carlos A. Niño. *email: f.brodsky@ucl.ac.uk; kylie.walters@nih.gov; simona.polo@ifom.eu

Clathrin-coated vesicles (CCVs) are well-characterized molecular machines, responsible for the internalization of nutrients, signaling receptors, and adhesion molecules from the plasma membrane (PM) as well as for cargo transport between the *trans*-Golgi network and endosomes^{1,2}. Clathrin is a three-legged molecule (triskelion), composed of three clathrin heavy chain (CHC) and three clathrin light chain (CLC) subunits. The clathrin triskelia assemble around a budding vesicle, creating a characteristic polyhedral coat, forming a clathrin-coated pit (CCP) that captures protein cargo and further matures into a CCV.

While CHCs are essential components of the cages, CLCs are apparently more regulatory. Functional roles for CLCs have been linked to regulation of cage rigidity, assembly and disassembly by extrapolation from *in vitro* studies¹, and to clathrin-mediated endocytosis (CME) of model cargos, including G-protein-coupled receptors^{3–5}. Importantly, CLCs appear essential at the apical surface of highly polarized tissues, where membrane tension is higher than normal and actin dynamics is critical for membrane invagination and fission⁶. This process involves the conserved interaction between CLCs and the actin-remodeling huntingtin-interacting proteins (Hip1 and Hip1R)^{7,8}, an interaction that also contributes to clathrin-actin interactions during endosomal recycling and is needed for cell migration⁹. CLCs are also required for uptake of large particles, including viruses and pathogenic bacteria, such as *Listeria*, *Yersinia*, and *Salmonella*^{10–12}. In these cases, the direct association of CLC with actin remodelers distinguishes these clathrin-dependent processes from canonical CME. Finally, CLCs, together with Hip1R and the actin motor protein myosin VI¹³, participate in cell-cell junction formation and remodeling through molecular mechanisms that have not been defined¹⁴.

In vertebrates, two genes encode the CLCa and CLCb variants, which are 40% divergent in sequence. High conservation of each protein variant in all vertebrate species suggests that they have distinct functions beyond their shared interaction with the Hip proteins. However, evidence for CLCa- and CLCb-specific roles is limited. CLCa was shown to be preferentially involved in cell spreading and migration¹⁵, whereas maturation of CCPs from flat lattices to invaginated buds requires phosphorylation of CLCb¹⁶. CLCb was also found to be overexpressed in non-small-cell lung cancer, causing altered CCP dynamics and consequently aberrant growth factor signaling¹⁷.

In a previous study, we identified clathrin as a major and specific binding partner of myosin VI isoforms that contain the $\alpha 2$ -linker (named myosin VI_{long}) and we found that myosin VI_{long} localizes to CCPs independently of the presence of the endocytic adaptor Dab-2¹⁸. These findings suggested a molecular link between clathrin and myosin VI_{long}. In this manuscript, we now demonstrate that this interaction is due to a direct and selective interaction of myosin VI_{long} with CLCa and not CLCb. We use nuclear magnetic resonance (NMR) spectroscopy to solve the structure of the myosin VI_{long}:CLCa complex and characterize the mechanism of interaction. The CLCa-binding region in myosin VI partially overlaps with the ubiquitin-binding domain MyUb¹⁹ while CLCa binds myosin VI with a surface that is in close proximity to the Hip1R-binding site⁷. Guided by the structure, we identify a CLCa point mutant that loses myosin VI-binding ability, while leaving unaltered its capacity to interact with CHC or Hip1R. We further investigate the biological implications of these molecular interactions, taking advantage of genetically reconstituted Caco-2 cells growing in three dimension (3D) as cysts. Altogether, our results indicate that the interaction between myosin VI_{long} and CLCa is needed to generate the force that leads to invagination and fission of CCPs at the apical surface of highly polarized cells and suggests a model for the sequential function of myosin VI and Hip1R in actin-mediated coated vesicle budding.

Results

CLCa is a direct and specific interactor of myosin VI_{long}. In myosin VI, the presence of helix $\alpha 2$ structurally defines a clathrin-binding domain that is unique to myosin VI_{long}^{18,19} (Fig. 1a). To identify the corresponding myosin VI_{long} interacting region within clathrin, we performed experiments with green fluorescent protein (GFP)- and glutathione S-transferase (GST)-tagged CHC truncations. The results revealed that the C-terminal part of the CHC is required for myosin VI interaction but failed to demonstrate direct binding between the two proteins (Supplementary Fig. 1a). Given that CLCs are reported to interact with the C-terminus of CHC, we hypothesized their involvement in recruiting myosin VI to CCPs. Corroborating our hypothesis, confocal microscopic analysis of HeLa cells revealed reduced colocalization of myosin VI_{long} and CHC upon CLCa transient knockdown (KD) as measured by Mander's coefficient (Supplementary Fig. 1b, c). Unexpectedly, CLCb KD had the opposite effect, increasing the co-localization levels of CHC and myosin VI_{long}, thus suggesting different roles for the two CLCs.

We next examined direct binding in a pull-down assay by using bacterially purified GST-CLCa and GST-CLCb full-length proteins and purified fragments of long and short myosin VI isoforms spanning amino acids 998–1131 (Fig. 1a). The results demonstrated direct and selective binding of CLCa to myosin VI_{long} while both CLCs showed minimal binding to myosin VI_{short} (Fig. 1b).

To formally test whether CLCa acts as a bridge between myosin VI_{long} and CHC, we exploited CLCa knockout (KO) mice that were recently generated⁵ and newly produced CLCb KO mice. Whole-brain lysates obtained from CLCa and CLCb KO mice and wild-type (WT) littermates were used in a pull-down assay with GST-myosin VI_{long}^{998–1131} from long and short myosin VI isoforms. Using WT littermates, GST-myosin VI_{long} efficiently pulled down CHC (Fig. 1c). This interaction was dependent on CLCa since clathrin was not detected using brain lysate from CLCa KO animals (Fig. 1c). Of note, brain lysate from CLCb KO mice behaved as WT littermates, similarly showing interaction between clathrin and myosin VI_{long} (Fig. 1c), despite lacking CLCb (Fig. 1d).

The lysis conditions used in these experiments and in our previous study¹⁸ favor free triskelia in solution as opposed to assembled clathrin. During clathrin cage assembly, CLCs undergo conformational changes⁸, and to investigate whether the interaction of the coat protein with myosin VI can occur in the context of a formed CCV, we exploited the ability of clathrin to self-assemble in a pH-dependent manner^{20,21}. We purified clathrin from pig brain CCVs and induced the formation of assembled cages from native clathrin containing an endogenous mixture of CLCa and CLCb, CHC with no CLCs, or CHC reconstituted with either CLCa or CLCb only (see Methods and Supplementary Fig. 2a–c for details). Purified myosin VI_{long}^{998–1131} protein co-sedimented with native clathrin cages and cages formed from clathrin with only CLCa, whereas only weak sedimentation was observed with CLCb-only cages that was similar to CHC-only cages (Fig. 1e). In summary, these results demonstrate that CLCa is the subunit of clathrin responsible for binding to myosin VI_{long}, both in the context of free triskelia and in CCVs.

CLCa and myosin VI_{long} interact with sub-micromolar affinity.

To obtain biochemical and structural insights into the myosin VI_{long}:CLCa complex, we systematically designed CLCa deletions taking into consideration the known interaction domains of the protein²² (Fig. 2a). Incubation of these constructs with bacterially purified myosin VI_{long}^{998–1131} identified a region of interaction in CLCa that spans amino acids 47–97 (Fig. 2b). This fragment of clathrin falls in a region of limited conservation between the two

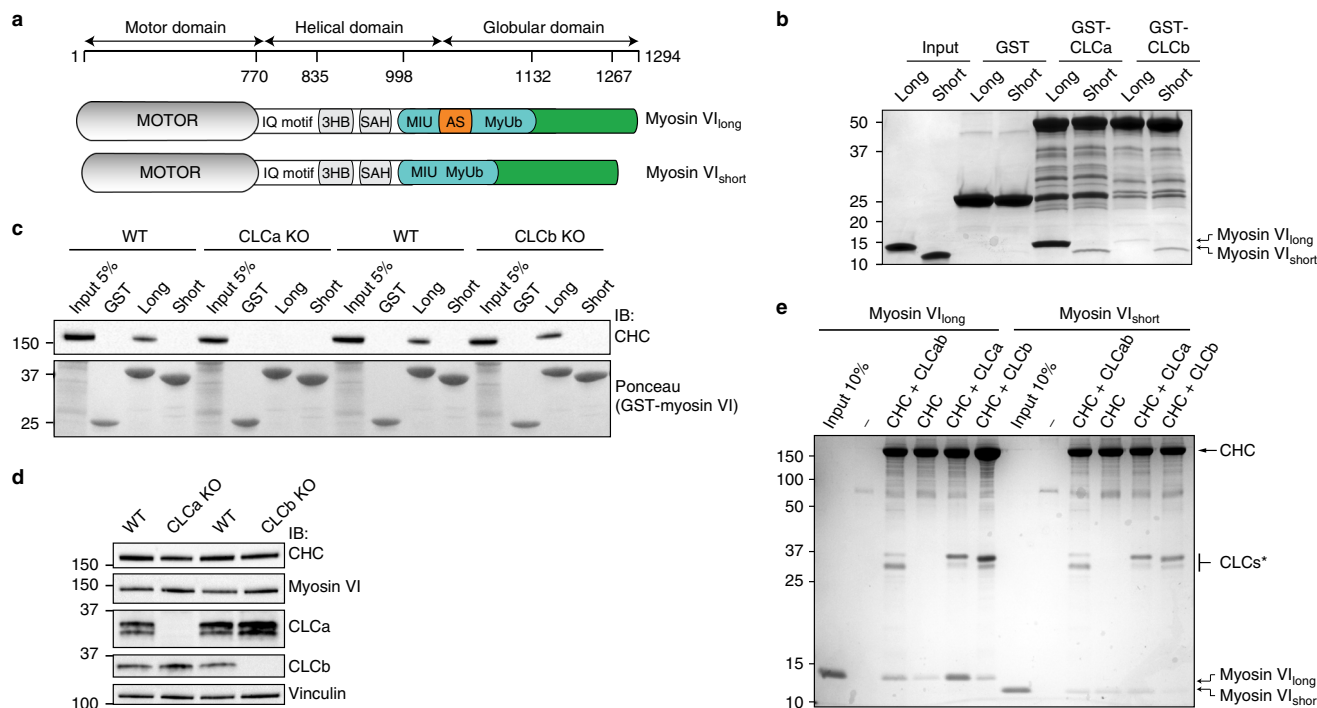


Fig. 1 CLCa is a direct and specific interactor of myosin VI_{long} in triskelia and clathrin cages. **a** Scheme of the myosin VI highlighting the region involved in clathrin binding (amino acids 998–1131 of the long isoform). Long and short isoforms are reported together with the domains and motifs previously identified, including IQ motif, 3HB (three-helix bundle), SAH (single α -helix), MIU (motif interacting with ubiquitin), AS (alternative splicing region), and MyUb (myosin VI ubiquitin-binding domain). In orange is represented the alternatively spliced region codifying for the α 2-linker¹⁸. **b** Pull-down assay with GST-CLCa and CLCb full-length and cleaved and purified fragments spanning amino acids 998–1131 of long and short myosin VI isoforms. Glutathione sepharose beads coupled to GST and GST-tagged proteins were incubated with myosin VI^{998–1131}. After washes, bound proteins were eluted in Laemmli-buffer, resolved through SDS-PAGE, and stained with Coomassie. **c** Pull-down assay using the long and short GST-myosin VI^{998–1131} constructs and brain lysates (500 μ g) obtained from the indicated mouse strains. After washes, bound proteins were eluted in Laemmli-buffer, resolved through SDS-PAGE, and transferred to a nitrocellulose membrane. Immunoblot (IB) was performed with anti-clathrin heavy-chain antibody. Ponceau detect equal loading of GST proteins. **d** IB of the brain lysates used in (c), as indicated. **e** Co-sedimentation assay. Equimolar (1.5 μ M) amount of myosin VI^{998–1131} and clathrin cages were incubated at 4 $^{\circ}$ C for 45 min in the presence of detergent (0.1% Triton X-100) and then pelleted by ultracentrifugation. Precipitated proteins were dissolved in Laemmli-buffer, resolved through SDS-PAGE, and stained with Coomassie. CLCs* indicates the various CLC proteins. Note that in the native cages CLCs (CHC-CLCab) run at different molecular weight (mw) as they are from pig brain while the human CLCs used for reconstitution are bacterially produced and cleaved from GST

paralogs (Fig. 2a), thus explaining the binding selectivity of CLCa to myosin VI_{long}. By analytical size-exclusion chromatography, the CLCa:47–97myosin VI^{998–1131} sample co-eluted at the expected molecular weight for a 1:1 complex, indicating a stable and strong interaction in solution (Supplementary Fig. 2d). Further truncation analyses narrowed down the minimal interaction surface to a short peptide, with no secondary structure predicted, extending from amino acids 51 to 61 (Fig. 2a and Supplementary Fig. 2e).

To evaluate the biochemical properties of the interacting CLCa peptide, we measured binding affinity of the myosin VI–CLCa interaction by fluorescent polarization (FP) analysis. We designed three different fluorescein-labeled CLCa peptides (Fig. 2a) that span the minimal binding surface (amino acids 51–61) exclusively or with five amino acids added at the N (amino acids 46–61) or C (amino acids 51–66) terminal end. FP measurements revealed a tenfold increase in affinity for the peptide 46–61 compared to the other two peptides (Fig. 2c), resulting in a dissociation constant (K_d) of 0.93 μ M. Thus, in addition to the minimal binding region, amino acids spanning 46–51 conserved between CLCa and CLCb contribute to binding, which may explain the residual interaction observed by both CLCs to myosin VI_{short} (Fig. 1b).

We further tested the 46–61 peptide for binding to various myosin VI constructs to better define the interaction boundaries

on myosin VI. Confirming our previous results¹⁸, the long and the short isoforms of myosin VI showed 2 log-fold difference in affinity for CLCa (0.93 μ M versus 96 μ M, respectively, Fig. 2c, d). As expected, the myosin VI^{1069–1131} construct that lacks the α 2-linker displayed a remarkably reduced K_d . Finally, the region encompassing amino acids 1050–1131 of myosin VI_{long} showed the highest affinity (0.4 μ M, Fig. 2d) and was therefore chosen for the subsequent structural experiments.

CLCa^{46–61} forms a helix in a complex with myosin VI^{1050–1131}.

To further characterize the myosin VI:CLCa binding interface, we first compared two-dimensional (2D) ¹H, ¹⁵N-HSQC spectra of myosin VI^{1050–1131} before and after addition of the unlabeled CLCa peptide spanning 46–61 (Fig. 3a). Large effects were observed for many amino acids, which were quantified by chemical shift perturbation (CSP) analysis (Supplementary Fig. 3). Residues with CSP values of one standard deviation (SD) above average cluster to three regions in the myosin VI sequence (Supplementary Fig. 3) located at the N-terminal end of α 2, N-terminal to α 3, and toward the C-terminal end of α 4. The CSPs above the mean were mapped onto the previously solved structure of free myosin VI (Fig. 3b), which revealed a localized region focused at the C-terminal half of α 4 and the N-terminal region of α 2.

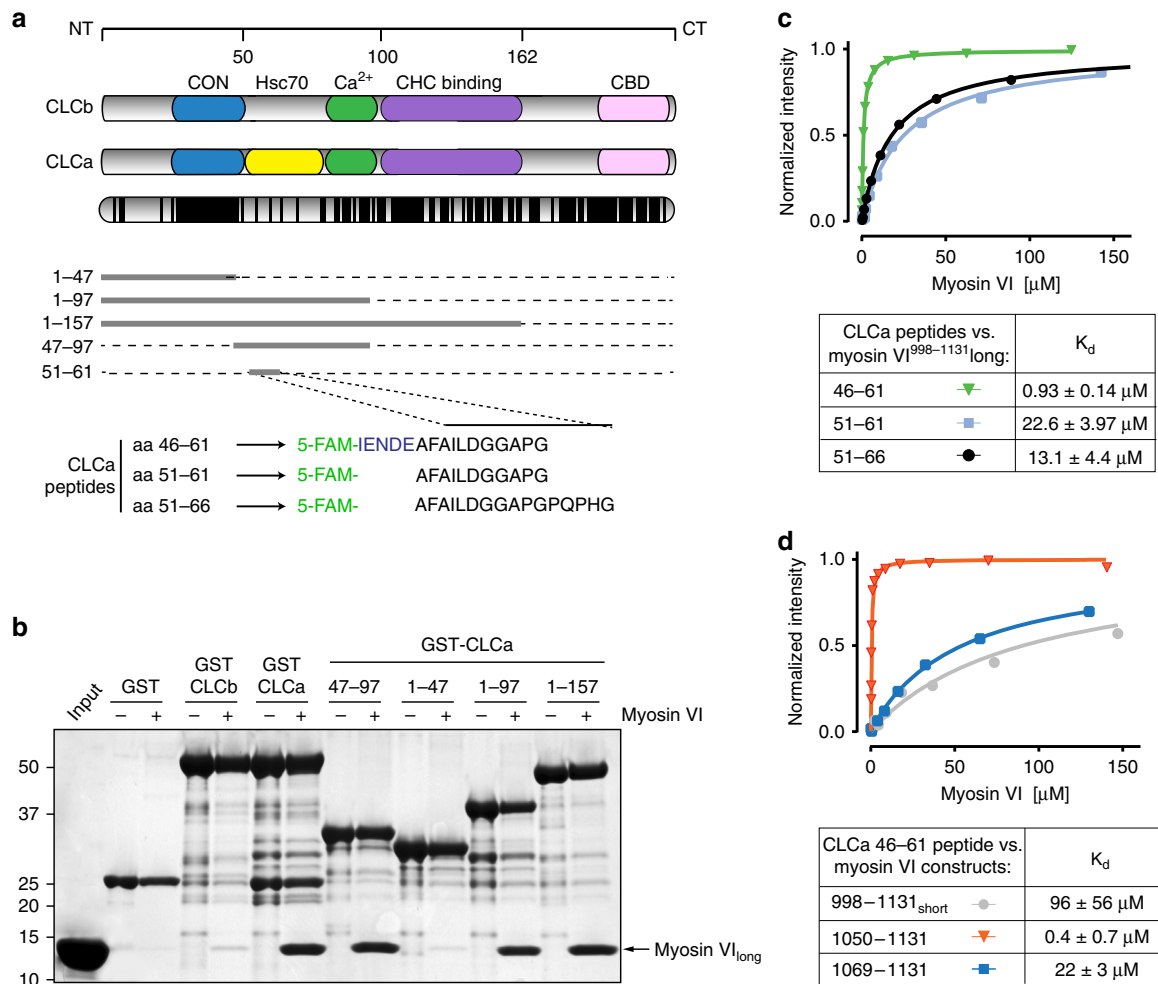


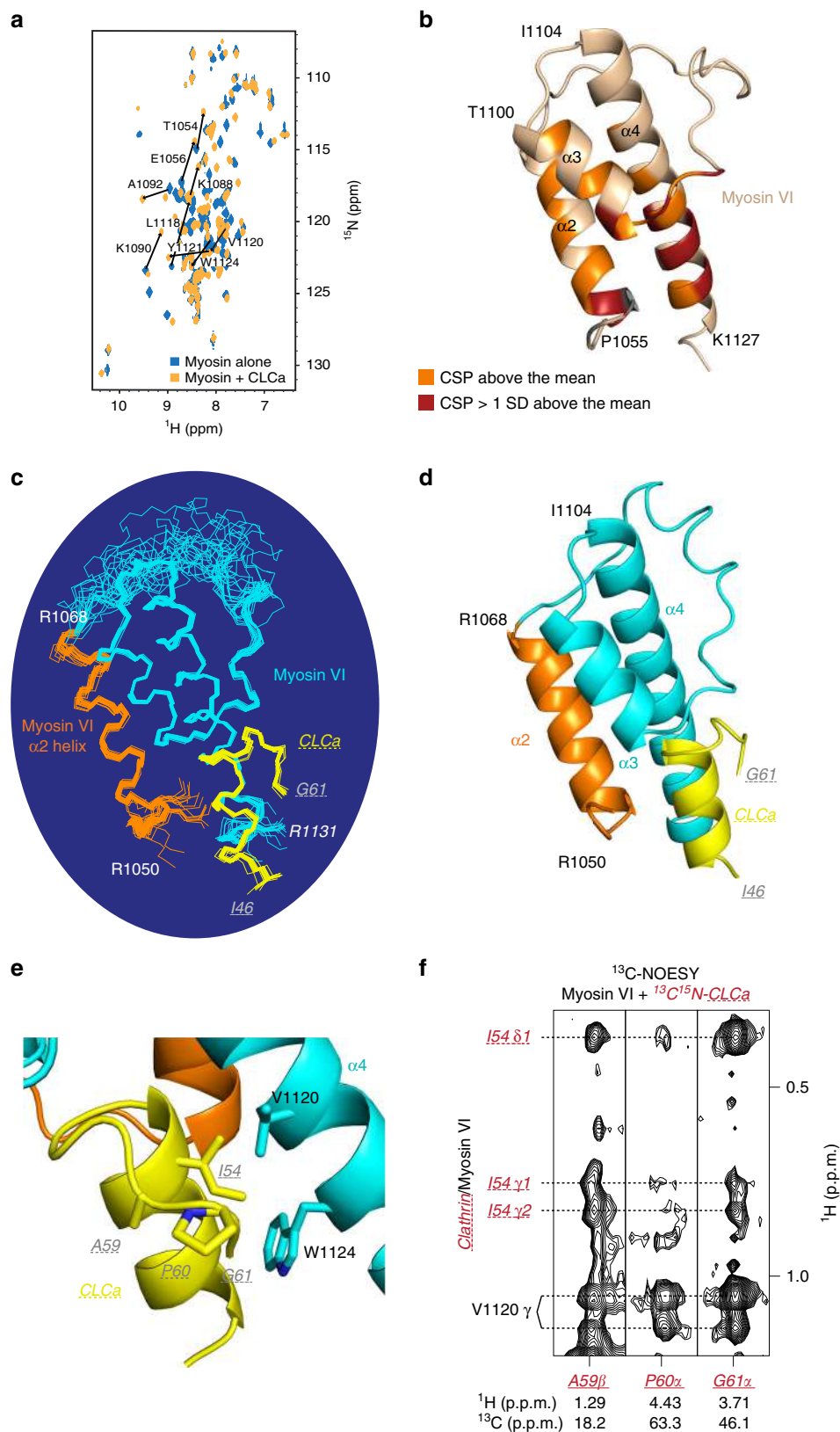
Fig. 2 CLCa:myosin VI_{long} interact with sub-micromolar affinity. **a** Domain structures of CLCa and CLCb. CON conserved Hip-binding region, Hsc70 unique region in CLCa that stimulates Hsc70 activity in vitro, Ca²⁺ EF-hand domain that binds calcium, CHC binding clathrin heavy chain binding region, CBD calmodulin-binding domain. Sequence conservation between the two proteins is reported below. Each line represents one amino acid, black line indicates identity. Lower panel, scheme of the selected constructs used in **(b)** together with the sequence of the three overlapping 5-carboxyfluorescein (5-FAM)-conjugated CLCa peptides used for FP analysis in **(c)**. **b** Pull-down assay with GST-CLCa and CLCb full length and the indicated fragments of CLCa immobilized on glutathione sepharose beads and incubated with the purified fragment spanning amino acids 998-1131 of myosin VI_{long}. After washes, bound proteins were eluted in Laemmli-buffer, resolved through SDS-PAGE, and stained with Coomassie. **c** FP assay using the three peptides shown in **(a)** and the purified fragment spanning amino acids 998-1131 of myosin VI_{long}. Dissociation constants with their respective 95% confidence interval (CI) are reported in the table at the bottom. Graph is representative of three independent experiments used to calculate K_d and CI. **d** FP assay using peptide 46-61 of CLCa and the indicated fragments of long and short myosin VI isoforms. Graph, K_d , and CI as for **(c)**

In previously published structures of clathrin lattice and pits, the myosin VI-binding region of CLCa identified here was not detectable, as it falls into a flexible region^{8,23}. As part of an effort to solve the structure of this CLCa region complexed with myosin VI¹⁰⁵⁰⁻¹¹³¹ by NMR, we assigned chemical shift values to the myosin VI-bound CLCa fragment by 3D HNCACB (Supplementary Fig. 4) and nuclear Overhauser effect spectroscopy (NOESY; examples provided in Supplementary Fig. 5) experiments recorded on equimolar ¹⁵N, ¹³C-labeled CLCa and unlabeled myosin VI¹⁰⁵⁰⁻¹¹³¹; we also assigned chemical shift values to CLCa-bound myosin VI, as described in Methods. The Ca and C' chemical shift assignments of CLCa were compared to those of randomly coiled values to generate a chemical shift index plot (Supplementary Fig. 6). This information was combined with NOE analyses obtained by NOESY experiments to define the secondary structure of CLCa (Supplementary Fig. 6) as being α -helical from residues I46-I54.

We next measured intermolecular interactions between myosin VI and CLCa directly by performing 3D ¹H, ¹³C half-filtered

NOESY experiments; resulting spectra show NMR signals between intermolecular atoms that are within 5 Å of each other²⁴. In agreement with the CSP analysis, we identified large numbers of intermolecular NOEs between CLCa and helices α_2 and α_4 of myosin VI (Supplementary Figs. 7 and 8). In particular, residues P1055, M1058, and M1062 from helix α_2 (Supplementary Figs. 7a, b and 8) and V1120, Y1121, and W1124 from helix α_4 (Supplementary Figs. 7c-f and 8) displayed numerous intermolecular NOEs to CLCa.

CLCa binds to isoform-specific α_2 and α_4 of MyoVb. NMR data (summarized in Table 1), including 221 intermolecular NOE interactions between myosin VI and CLCa, were used to calculate the myosin VI:CLCa structure. The 20 lowest energy structures (Fig. 3c) converged with a backbone root mean square deviation (r.m.s.d.) in the ordered regions of 0.36 Å (Table 1). The myosin VI structure is largely unchanged in the complex, retaining its



three helices and a disordered loop that follows the C-terminal end of isoform-specific $\alpha 2$. In agreement with our chemical shift index data (in Supplementary Fig. 6), CLCa^{46–61} forms a short α -helix (yellow, Fig. 3c–e) that appears to complete the myosin VI structural domain by abutting the N-terminal end of $\alpha 3$ and packing against $\alpha 2$ and $\alpha 4$ (Fig. 3d). At the C-terminal end, CLCa

L55–G58 forms a helical turn with A59–G61 folding back along one side of the CLCa α -helix (Fig. 3e), a structure supported by NOEs to CLCa I54 and to myosin VI $\alpha 4$ (V1120 and W1124, Fig. 3f and Supplementary Fig. 7c, e). The hydrophobicity of CLCa A59 and P60 likely drives this segment to remain in close contact with the helices (Fig. 3e), as supported by the NMR data (Fig. 3f).

Fig. 3 The myosin VI-binding region of CLCa forms an α -helix that interlocks with the helices of myosin VI. **a** 2D ^1H , ^{15}N -HSQC spectrum of 0.78 mM ^{15}N -labeled myosin VI $^{1050-1131}$ alone (blue) and in the presence of equimolar unlabeled CLCa $^{46-61}$ (orange). Notable changes in NMR signals following addition of CLCa are indicated with arrows and labeled with the assigned amino acid residue from myosin VI. **b** On the previously solved structure of myosin VI $^{1050-1131}$ (PDB 2N12), amino acids with CSPs greater than the mean or 1 SD above the mean are indicated in orange or red, respectively. CSPs were calculated according to the definition $\text{CSP} = [(0.2\Delta\delta\text{N})^2 + (\Delta\delta\text{H})^2]^{1/2}$ where $\Delta\delta\text{N}$ and $\Delta\delta\text{H}$ represent chemical shift differences for the amide nitrogen and proton of each residue, respectively. Residues without amide protons (P1051, P1055, and P1070) are indicated in gray, and those at the beginning and end of each helix are labeled. **c** Superimposition of the 20 lowest energy structures calculated for myosin VI $^{1050-1131}$ in complex with CLCa $^{46-61}$. CLCa is shown in yellow; myosin VI is colored cyan except for the isoform-specific $\alpha 2$ helix, which is highlighted in orange. The orientation of the structures is identical to **b**. **d** Ribbon representation of the structure of myosin VI bound to CLCa $^{46-61}$ using the same color scheme and orientation as in **c**. **e** Ribbon representation as in **d**, expanded and rotated about the CLCa helix to highlight the α -helical structure of CLCa $^{46-54}$ and interactions involving the C-terminal amino acids of the CLCa peptide. Sidechain heavy atoms of the highlighted residues are included. Note that the C-terminal portion of the CLCa fragment remains in close contact with myosin VI and the CLCa helix. **f** Selected regions from a 3D ^{13}C NOESY experiment acquired on 0.4 mM ^{13}C , ^{15}N -labeled CLCa $^{46-61}$ and equimolar unlabeled myosin VI $^{1050-1131}$. NOE interactions involving the C-terminal portion of CLCa (A59, P60, and G61) are shown. Intramolecular and intermolecular interactions are labeled in red and black, respectively

Table 1 NMR data collection and refinement statistics

	Total	Myosin VI $^{1050-1131}$	CLCa $^{46-61}$
NMR distance and dihedral constraints			
NOE-derived distance constraints			
Total NOE	2540		
Intramolecular		1971	348
Intra-residue		687	91
Inter-residue		1284	257
Sequential ($ i - j = 1$)		509	105
Medium-range ($2 < i - j < 4$)		545	126
Long-range ($ i - j > 5$)		230	26
Intermolecular	221		
Hydrogen bonds			
Total	60		
Intramolecular		47	9
Intermolecular	4		
Total dihedral angle constraints			
ϕ		50	9
ψ		50	9
Structure statistics			
Violations (mean and s.d.)			
Distance constraints (Å)	0		
Dihedral angle constraints ($^\circ$)	0		
Max. dihedral angle violation ($^\circ$)	<5		
Max. distance constraint violation (Å)	<0.3		
Deviations from idealized geometry			
Bond lengths (Å)	0.003 \pm 0.001		
Bond angles ($^\circ$)	0.420 \pm 0.013		
Impropers ($^\circ$)	0.302 \pm 0.009		
Average pairwise r.m.s. deviation ^a (Å)			
Heavy	0.99 \pm 0.14		
Backbone	0.36 \pm 0.09		

^aValues were calculated for the 20 lowest energy structures of myosin VI residues M1053-R1068 and L1086-N1128 and CLCa residues E47-P60

As predicted by the binding selectivity of the long isoform¹⁸, the interaction surface on myosin VI involves isoform-specific residues from $\alpha 2$, including P1055, M1058, A1059, and M1062, which surround L55 of CLCa (Fig. 4a, b). Myosin VI P1055 extends this CLCa-binding surface through its interaction with F52 (Fig. 4b, NOEs shown in Supplementary Fig. 7b). At an adjacent side of the CLCa helix, myosin VI R1117, V1120, Y1121, and W1124 surround CLCa A51 and I54 (Fig. 4a, b). In addition, the W1124 indole group forms a hydrogen bond to the sidechain carboxyl group of E50 (Fig. 4a).

The RRL motif required for binding to multiple adaptor proteins including optineurin and GIPC^{18,25,26} is embedded in $\alpha 4$ (Supplementary Fig. 9a). R1116 does not participate in the interaction and remains surface exposed, whereas both R1117 and

L1118 contribute to CLCa binding. R1117, required for myosin VI structural integrity^{18,27}, maintains its hydrogen bonds to S1087 and E1113 as in free myosin VI and forms a hydrogen bond to the backbone oxygen of CLCa D56, the sidechain of which also forms a hydrogen bond to the backbone amide of myosin VI Y1091 (Fig. 4c). Lastly, L1118 of the RRL motif contributes to binding through interactions with CLCa L55 (Fig. 4b). Notably, the CLCa amino acids critical for binding to myosin VI, including A51, I54, L55, and D56, are not conserved in CLCb (Fig. 4d), thus providing an explanation for paralog specificity.

The importance of the identified interactions is supported by GST pull-down experiments. A truncated construct confirmed that the $\alpha 4$ helix of myosin VI is involved in binding to CLCa

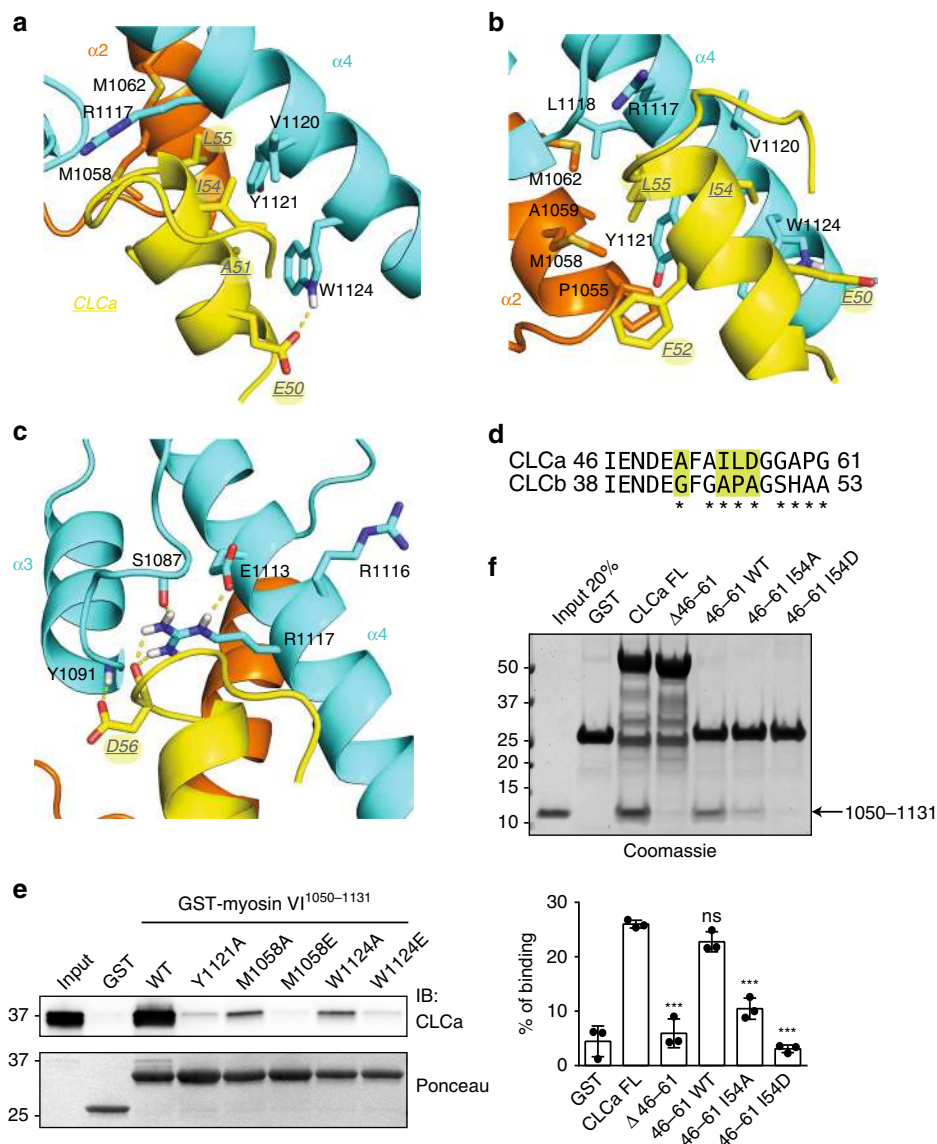


Fig. 4 A hydrophobic pocket formed by myosin VI encompasses residues I54 and L55 of CLCa. **a** Ribbon representation of the myosin VI:CLCa binding interface in the orientation of Fig. 3e with key sidechain heavy atoms displayed. Myosin VI α 4 residues R1117, V1120, Y1121, and W1124 (blue) interact with CLCa E50, A51, I54, and L55 (yellow). CLCa L55 also interacts with myosin VI α 2 residues, with M1058 and M1062 (orange) observable in this view. At the edge of the hydrophobic pocket lies a hydrogen bond between CLCa E50 and the myosin VI indole group. Sulfur and nitrogen atoms are in yellow and indigo, respectively. **b** As in **a** but rotated about the CLCa helix to highlight interactions involving the myosin VI isoform-specific α 2 helix, especially P1055 and A1059. **c** Enlarged representation of the region containing myosin VI R1117 to highlight intramolecular and intermolecular hydrogen bonds. The guanidine group of R1117 forms hydrogen bonds to the backbone and sidechain carboxyl groups of myosin VI S1087 and E1113, respectively, and to the backbone carboxyl group of CLCa D56. This view is similar to that of **a** but rotated about the sidechain of R1117. In **a**, **c**, a dashed yellow line is used to indicate a hydrogen bond with oxygen and nitrogen atoms in red and indigo, respectively. **d** Sequence alignment of CLCa 46-61 with the corresponding region of CLCb. Asterisks indicate residues not conserved between isoforms. In yellow are amino acids putatively responsible for the selective binding. **e** Pull-down assay using the indicated GST-myosin VI¹⁰⁵⁰⁻¹¹³¹ mutant constructs and lysates (1 mg) from HEK293T cells. After washes, bound proteins were eluted in Laemmli-buffer, resolved through SDS-PAGE, and IB was performed with the anti-CLCa antibody ($\times 16$). Ponceau detects equal loading of GST proteins. Representative image of three independent experiments is shown. **f** GST pull-down assay using the indicated CLCa constructs and purified fragment spanning amino acids 1050-1131 of myosin VI_{long}. After washes, bound proteins were eluted in Laemmli-buffer, resolved through SDS-PAGE, and stained with Coomassie. Bottom panel, quantitation of three independent experiments. Data are expressed as percentage of binding with respect to input and normalized for the amount of GST proteins used in each pull-down. Error bars represent s.d. *** $P < 0.001$ by two-tailed T test

(Supplementary Fig. 9a) while single substitution of myosin VI M1058, Y1121, or W1124 led to reduced binding (Fig. 4e). FP analysis revealed a 2 log-fold difference in binding affinity for the Y1121A mutant (Supplementary Fig. 9b). On the CLCa side, we tested the effect of substituting I54 with alanine or aspartic acid, using CLCa full-length protein as a control. As expected, myosin VI¹⁰⁵⁰⁻¹¹³¹ bound to CLCa WT but not Δ 46-61. I54A or I54D

significantly impairs binding to myosin VI, with aspartic acid showing the strongest defect (Fig. 4f).

Myosin VI requirement for CME in polarized cysts. While CLCa is ubiquitously expressed in animal tissues⁵, the presence of myosin VI_{long} is restricted to organs containing polarized cells of

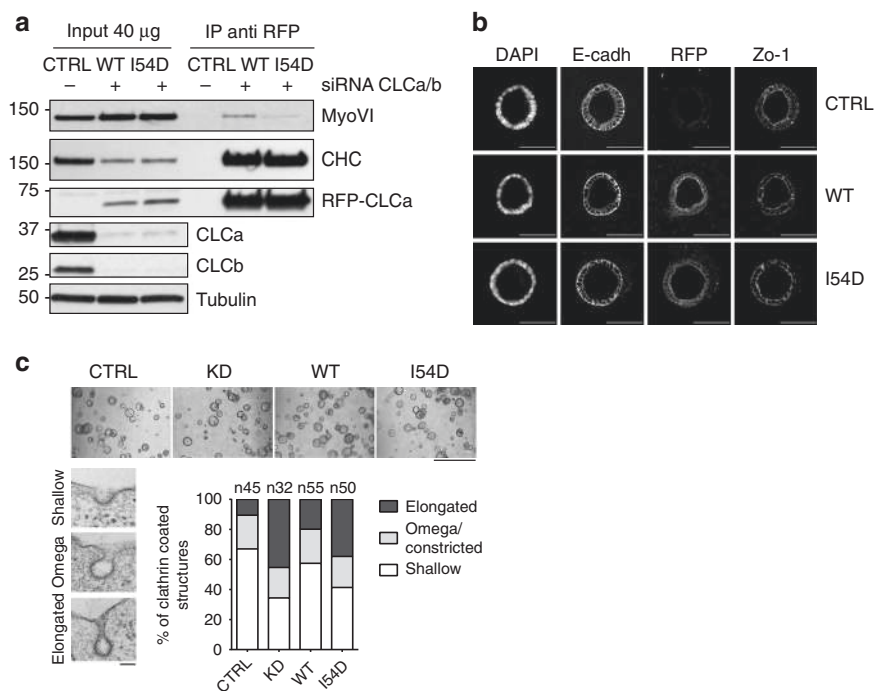


Fig. 5 CLCa I54D is a selective myosin VI-impaired mutant. **a** Co-immunoprecipitation experiment using lysates from Caco-2 reconstituted cell lines. Cells depleted of endogenous CLCs or mock treated are kept in confluency for 7 days, lysed, and RFP-CLCa wild-type and I54D were immunoprecipitated from 1 mg of lysates. IP and IB as indicated. **b** Representative confocal micrographs of CLCa wild-type and I54D localization in Caco-2 cysts. Staining with the indicated apical-basal polarity markers are shown. Scale bar, 100 μ m. **c** Upper panel, representative brightfield pictures of Caco-2 cysts generated using the indicated isogenic cell lines. Scale bar, 1 mm. Clathrin-coated structures present at the apical surface of the Caco-2 cysts were counted and classified as shallow, omega/constricted, or elongated according to their morphology. Left panel, representative EM images of the different morphology of CCPs. Scale bar, 100 nm. Right panel, distribution of clathrin-coated structures in the different isogenic cell lines expressed as percentage among the different classes and calculated on three independent experiments. Between 94 and 125 cellular profiles of well-polarized cells for each line were imaged. Number of clathrin-coated structures counted in total are reported as $n = x$. See also Supplementary Fig. 11d

epithelial origin, such as kidney and intestines, both in mice²⁸ and humans (Supplementary Fig. 10a). There, myosin VI localizes to the apical surface facing the lumen of the organs at the base of microvilli^{29,30} (Supplementary Fig. 10b). To analyze the physiological role of the CLCa:myosin VI complex in a cellular model of polarized epithelial tissue, we took advantage of the intestine-derived epithelial Caco-2 cells that form polarized cysts when plated as a single-cell suspension embedded in 3D EHS-derived matrix³¹. Notably, in this Caco-2 cellular model system, a clear switch toward the myosin VI_{long} isoform occurs during the acquisition of full polarity both in 2D and 3D systems, as measured by reverse transcriptase-polymerase chain reaction (PCR) (Supplementary Fig. 10c). Transmission electron microscopy (TEM) and confocal microscopy analysis showed that the cysts are fully formed and polarized (Supplementary Fig. 10d–f) and myosin VI_{long} is enriched in the apical terminal web region together with occludin (Supplementary Fig. 10d). We then generated Caco-2 cells stably expressing red fluorescent protein (RFP)-WT or an RFP-I54D mutant rat CLCa as these constructs are resistant to the small interfering RNA (siRNA) oligos designed on the human sequence. Upon efficient depletion of the endogenous CLCa and CLCb by siRNA oligos (Supplementary Fig. 11a–c and Fig. 5a), co-immunoprecipitation analysis performed with lysates from 2D fully polarized Caco-2 cells demonstrated that the I54D mutant was largely unable to interact with myosin VI (Fig. 5a), validating our previous *in vitro* results. Next, single Caco-2 reconstituted cells depleted of endogenous CLCs were cultured in matrigel and 7 days after cysts were counted and stained to evaluate adherens and tight junctions and the localization of RFP-CLCa WT or I54D mutant (Fig. 5b).

Exogenous RFP-CLCa I54D mutant, similarly to the WT protein, was enriched in the apical region toward the lumen and did not appear to affect the ability of the cells to form cysts (Fig. 5c) or prevent the establishment of apical basal polarity (Fig. 5b and Supplementary Fig. 10e, f).

Clues as to the molecular requirement for CLCa–myosin VI interaction were derived from EM analysis of the cysts. As CLCs are implicated in CME, we investigated the CCPs that formed at the apical surface facing the lumen in control, CLC-depleted, and reconstituted RFP-CLCa variants. CCPs stemming from the PM were classified as shallow, omega/constricted, or elongated according to their morphology (Fig. 5c, left panel). Compared to mock-treated Caco-2 cysts, cysts depleted of CLCs displayed a significantly greater proportion of CCPs with an aberrant morphology, consisting of elongated pits joined to the PM by “long necks” (Fig. 5c). These structures closely resembled those induced by the dominant-negative mutant of dynamin, DynK44A³². Strikingly, while WT RFP-CLCa fully rescued this phenotype, the I54D mutant was unable to do the same, indicating that the morphological defect scored in the absence of CLCs can be ascribed to the lack of myosin VI interaction with CLCa (Fig. 5c, Supplementary Fig. 11d).

These results are compatible with the possibility that the interaction of CLCa with myosin VI is part of the machinery that controls CCP fission. Previously, CLCs and their binding partner Hip1R were implicated in this process at the apical membrane of polarized cells⁶. Intriguingly, the myosin VI-binding region we identified here on CLCa partially overlaps with the Hip1R-binding region⁷ (Fig. 6a). Thus we evaluated the ability of myosin VI and Hip1R to bind simultaneously to CLCa, using an *in vitro*

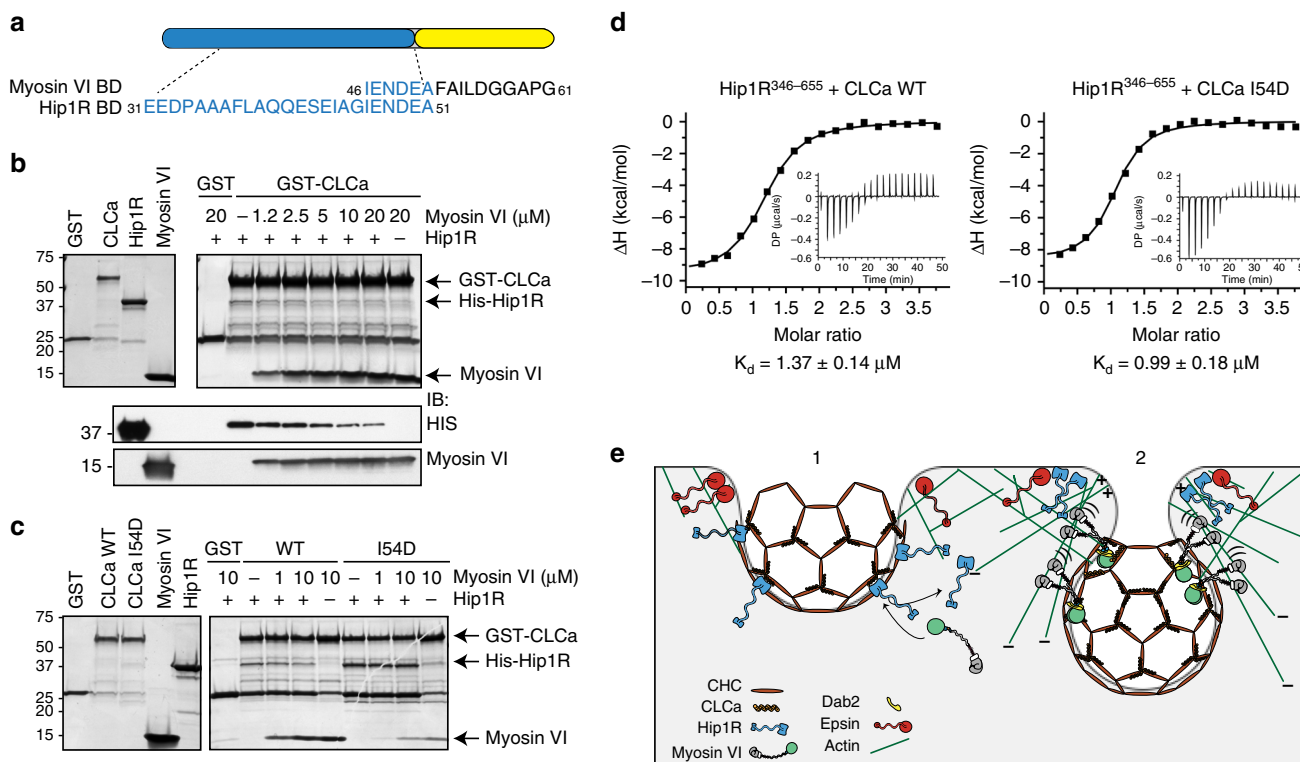


Fig. 6 Myosin VI and Hip1R are mutually exclusive binders of CLCa. **a** Binding regions of myosin VI and Hip1R on CLCa. **b** Competitive binding assay. Bacterially purified His-Hip1R coiled-coil region spanning residues 346–655 (2.5 μM) pre-incubated with GST-CLCa full-length protein (1 μM) for 1 h at 4 °C was mixed with increasing amounts of myosin VI^{1050–1131} as indicated. Bottom panel, Coomassie staining. Lower panels, 1/10 of the elutant was loaded for immunoblotting with the indicated antibodies. **c** As in **b** but using GST-CLCa I54D mutant. **d** ITC experiments with the indicated proteins. The integrated heat and raw plots are reported. Equilibrium dissociation constants (K_d) obtained by the fitting are indicated below. Relevant ITC measurements are reported in Table 2. **e** A putative model of clathrin-coated pit fission in polarized tissue. Once CCP passes the endocytic checkpoint, it undergoes a maturation process mediated by PIP2 turnover and the activity of several endocytic accessory factors (e.g., Hip1R). (1) As the bud expands, myosin VI is recruited to CCPs by CLCa and disengages Hip1R, which preferentially associates with epsin at the CCP edge. Actin anchoring and polymerization is thereby restricted to the neck of the invaginating pit where Hip1R along with epsin provide a link between actin nucleation and the CCP. Alternatively, formation of a complex between epsins and Hip1R at the neck would recruit Hip1R from the coat, exposing the myosin VI-binding site on CLCa and causing recruitment of myosin VI to the CCP. (2) Myosin monomer bound to CLCa anchors the CCP to the actin meshwork engaged in retrograde flow and facilitates movement of the vesicle into the cytoplasm. After dimerization induced by Dab2 or oligomerization, myosin VI becomes a processive motor that walk toward the minus end of actin filaments providing mechanical force to oppose membrane tension in polarized tissues and promoting the final dynamin-mediated fission step (not depicted). Future experiments are needed to test this model. In the myosin VI representation, the boxing glove and white cuff indicate the motor domain and single IQ motif, respectively, while the cargo-binding tail, which includes the clathrin-binding domain, is depicted as a green sphere; the black squiggly connecting region includes the 3HB and SAH

Table 2 Thermodynamic parameters obtained by isothermal titration calorimetry

	N (sites)	ΔH (kcal/mol)	Offset (kcal/mol)	ΔG (kcal/mol)	$-T\Delta S$ (kcal/mol)	K_d (μM)
Hip1R + CLCa FL	1.15 + 0.01	-9.70 + 0.19	2.27 + 0.06	-8.00	1.69	1.37 + 0.14
Hip1R + CLCa I54D	1.01 + 0.02	-8.72 + 0.27	1.25 + 0.08	-8.19	0.528	0.99 + 0.18
Myo6 + CLCa FL	0.79 + 0.02	-11.6 + 0.47	2.02 + 0.18	-9.00	1.69	0.09 + 0.05
Myo6 + CLCa I54D	N/A	N/A	N/A	N/A	N/A	N/A

N/A not applicable

assay with bacterially purified components. We incubated GST-tagged CLCa full-length protein with the coiled-coil region of Hip1R spanning amino acids 346–655 and added increasing amounts of myosin VI^{1050–1131} protein fragment. Analysis by Coomassie staining and immunoblotting showed that myosin VI addition caused progressive displacement of Hip1R from CLCa (Fig. 6b), indicating a direct competition between the two CLCa interactors possibly due to steric hindrance. Consistent with separate but proximal binding sites for these two CLCa-binding

proteins, Hip1R bound both the CLCa I54D mutant and WT CLCa, but its binding to the mutant was not affected by myosin VI (Fig. 6c). Notably, CLCb interaction with Hip1R was insensitive to the addition of myosin VI in the same competition assay (Supplementary Fig. 11e), confirming once again the selectivity of myosin VI for CLCa. Finally, we measured the affinity of the two interactors for CLCa full length by isothermal calorimetric assay (isothermal titration calorimetry (ITC), Table 2). Myosin VI bound CLCa with a K_d of 90 nM, while it

showed no binding to CLCa I54D (Supplementary Fig. 11f). In agreement with the competition assay, Hip1R binds WT and mutant CLCa equally well but with a K_d a log higher with respect to myosin VI (Fig. 6d).

Altogether, these data provide evidence for coordinated action of myosin VI and CLCa in the fission of CCPs at the apical surface of polarized cells and identify a direct competition between Hip1R and myosin VI.

Discussion

It is well established that actin remodeling during CME is required when the PM tension is high⁶ or when there is a need to change the mechanical properties of the clathrin lattice to engulf non-canonical bulky cargo, such as viral or bacterial pathogens^{10,12}. The functional link between CCPs and actin is represented by the CLC–Hip1R interaction in both yeast and mammals^{33–35}. Here we identified the role of a third player in this process, myosin VI, which selectively binds to CLCa and directly competes with Hip1R. While myosin VI has been previously implicated in CME^{26,36,37}, our structural definition of the CLCa–myosin VI interaction provides the molecular details of such involvement.

Myosin VI is unique among mammalian myosins in that it moves toward the minus end of F-actin³⁸. This unique motor protein supports critical functions at multiple cellular sites via interaction with different adaptors that target it to specific organelles. We recently showed that an alternatively spliced α -linker modulates the myosin VI interactome and is the causative link to its diverse functions¹⁸. In particular, the myosin VI_{long} isoform includes the exon cassette that codes for the α -linker and is endowed with clathrin-binding capacities and endocytic functions¹⁸.

Here we reveal that myosin VI is recruited to CCPs via a selective interaction with CLCa. Our biochemical and structural characterization of the myosin VI–CLCa interaction identified contacts unique to CLCa that are contributed largely from a helical structure. Key differences in the amino acid sequence of the CLCa and CLCb paralogs are likely responsible for the lack of binding between myosin VI and CLCb (Fig. 4d). Specifically, A51 and I54 in CLCa are both replaced with smaller residues (Gly and Ala, respectively, in CLCb), which likely diminishes the hydrophobic contacts with myosin VI (Fig. 4a, b). In CLCb, D56 is replaced with an alanine, which cannot form a hydrogen bond with Y1091 in myosin VI (Fig. 4c). Finally, L55, which embeds itself between helices α 2 and α 4 of myosin VI (Fig. 4a, b), is replaced with a proline in CLCb. Owing to the structural rigidity of proline, this amino acid may disrupt the helical turn, possibly preventing CLCb from fitting into the myosin VI-binding pocket that accommodates CLCa.

The CLCa peptide identified here as determinant for myosin VI interaction was invisible in the previously published structures of the clathrin lattice and the central domain of the triskelion^{8,23} and identified as a conformationally labile peptide modulated by calcium ion concentration³⁹, strongly suggesting that this flexible peptide is an important element of regulation for CLC. Calcium was also shown to mobilize myosin VI from a compact dormant state into a cargo-binding primed state⁴⁰. Interestingly, the segment of the myosin VI tail region mediating this regulation (amino acids 1060–1125) matches exactly with the clathrin-binding domain. In resting conditions, myosin VI adopts a backfolded, dormant state, with the tail domain amino acids 1060–1125 binding to apo-calmodulin; at high calcium levels, myosin VI unfolds and the release of the myosin VI tail mobilizes the motor, priming it to bind to cargo⁴⁰. In this open structural conformation of myosin VI, CLCa binding should be more

favorable. Consistently, addition of 2 mM Ca^{2+} to the lysis buffer significantly improved the co-immunoprecipitation between myosin VI and CLCa (Supplementary Fig. 12). Thus our results, together with these previous findings, suggest that release of intracellular calcium could orchestrate the timing of localized myosin VI activation and of clathrin–myosin VI functional interaction.

Like Hip1R⁴¹ and CLC⁴², myosin VI is largely dispensable for CME in 2D culture *in vitro*. However, the specific expression of myosin VI_{long} in polarized tissues (e.g., kidney and intestine), together with its features as an actin-based motor, prompted us to examine its function in Caco-2 cysts, used as proxy of the enterocytes lining the small intestine.

We show here that, in Caco-2 grown as cysts, a clear switch toward myosin VI_{long} occurs as cysts develop with this endogenous isoform enriched at the apical membrane, together with clathrin-rich structures. As further discussed below, CME in epithelial polarized cells is dependent on actin polymerization, which drives membrane invagination against the membrane tension that rigidifies the apical membrane⁶. In this process, CLCa may recruit myosin VI to CCPs to provide the force necessary to complete neck constriction, fission, and release. Indeed, using a genetic reconstitution approach, we were able to demonstrate a clear change in the distribution of CCP structures at the apical surface when we abrogated the CLCa–myosin VI interaction. The increased number of elongated CCPs present at the apical surface of CLC-depleted cysts or cysts reconstituted with I54D CLCa mutant, which does not bind myosin VI, suggests a delay in CCP progression at this surface. Our results are compatible with the idea of myosin VI acting at the neck of a mature CCP to promote fission. Consistent with this notion, recent evidence demonstrated that myosin VI, together with branched actin filaments, mediates severing by constricting the neck of budding tubules that are released from melanosomes⁴³.

The involvement of the myosin VI–CLCa interaction in the CCP fission step at the apical membrane of polarized cells is attractive, as this step is where actin has also been implicated in yeast⁴⁴ and observed in mammalian cells⁴⁵. It is tempting to speculate that the lack of CLCa–myosin VI interaction and thus a defective CME is the cause of the altered level of CTRF, megalin receptor and sodium type-IIa cotransporter (NaPi2a) observed at the apical membrane of myosin VI-depleted enterocytes and renal proximal tubular cells^{19,46–51}.

Our results unveil a direct competition between myosin VI and Hip1R for CLCa (Fig. 6a–d) that likely establishes partitioning of actin activity during the various stages of CCP formation (Fig. 6e). Current knowledge localizes Hip1R bound to CLCs throughout the CCP as well as inserted directly in the surrounding PM via its lipid-binding ANTH domain^{52,53}. The conformational change induced by CLC binding masks the actin-binding domain of Hip proteins in the coat⁸ and would exclude actin interaction at this site. At the adjacent membrane, Hip protein interaction with the epsin family proteins, both in yeast⁵² and in mammals⁵⁴, results in a multimeric heterocomplex^{52,55} that may preferentially occur at the neck of an invaginating CCP⁵⁶. Here Hip proteins dissociated from the clathrin coat could interact with actin, generating the active site required for the invagination of coated pits. Indeed, depletion of either Hip1R or epsins results in similar abnormal actin nucleation at endocytic sites where vesicle budding is inhibited^{33,54}. Thus regulation of actin activity by Hip proteins seems to be restricted to the neck of an invaginating vesicle⁸ where barbed ends of branched actin filaments are concentrated⁴⁵ and fission occurs as dynamin assembles into a helical coat^{57,58}. How might this happen? We propose here that myosin VI contributes to the temporal regulation of Hip1R action while producing the forces to overcome high membrane tension (Fig. 6e). During

vesicle invagination, myosin VI recruited by CLCa would compete out Hip1R, promoting its dissociation from the cages and recruitment by epsins at the uncoated neck. At the vesicle neck, Hip1R regulates actin filament polymerization³³ while Dab2 may tether myosin VI recruited by CLCa into a stable and active dimer⁵⁹. Finally, acting as a processive cellular motor on branched actin filaments, myosin VI may contribute to the release of the invaginating vesicle, promoting CCP fission (Fig. 6e). Notably, our findings are in line with previous studies based on total internal reflection fluorescence microscopy that revealed myosin VI to be recruited during a late stage of clathrin-mediated internalization, concomitantly with dynamin and regulators of actin polymerization, to promote the scission of the invaginated membrane⁶⁰. Consistent with the idea of myosin VI actively participating in the latest steps of CCP maturation, *in vivo* analysis of renal proximal tubule cells in mice lacking myosin VI showed reduced uptake of endocytic markers and accumulation of AP2 and Dab2 at the PM⁴⁹.

Our findings indicate that only CLCa-containing CCPs may be subjected to the regulation exerted by myosin VI. CLCa-null mice lack the myosin VI–CLCa interaction and its unique function and this deficiency may contribute to the 50% postnatal mortality rate of these animals⁵. Consistent with this idea, CLCa is ubiquitously expressed at high levels in all vertebrate tissues, whereas CLCb characteristically exhibits variable levels⁵, and compared to CLCb, CLCa is preferentially associated with peripheral clathrin-coated structures and proteins involved in cell spreading¹⁵, a process that also requires actin regulation.

We are at an early stage of appreciating the extent of the selective activity of each CLC paralog, with hints of CLCa- and CLCb-associated functions emerging in recent reports^{15–17}. The high conservation and distinction of the two CLC-encoding genes that are preserved within all vertebrates⁶¹ strongly support the notion that CLCa and CLCb proteins have non-redundant functions. In the emerging scenario, our data indicate CLCa as a key modulator of clathrin interactions with the actin cytoskeleton.

Methods

Cell lines and KO mouse strains and constructs. HEK293T cells (ICLC) were grown in Dulbecco's Modified Eagle Medium (DMEM) supplemented with 10% fetal bovine serum (FBS) S.A. and 2 mM L-glutamine. HeLa cells (ATCC) were grown in Minimum Essential Medium (MEM) supplemented with 10% FBS S.A., 0.1 mM non-essential amino acids, 2 mM L-glutamine, and 1 mM Na-Pyruvate. CaCo-2 cells (ATCC) were grown in MEM supplemented with 20% FBS S.A., 0.1 mM non-essential amino acids, and 2 mM L-glutamine.

At each batch of freezing, all cell lines were authenticated by STR profiling (StemElite ID System, Promega) and tested for mycoplasma using PCR and biochemical test (MycAlert, Lonza).

Caco-2 RFP-CLCa WT and CLCa I54D isogenic cell lines were generated by lentiviral infection of Caco-2 cells with the corresponding pLVX-RFP expressing rat-CLCa variants and bulk populations were FACS-sorted to enrich for RFP-positive cells.

Production of mice deleted for the *CLTA* gene encoding CLCa (CLCa KO) was previously described⁵. Mice deleted for the *CLTB* gene (CLCb KO) were generated in the Brodsky laboratory and loss of CLCb in the tissues studied here has been confirmed, as verified in Supplementary Fig. 2a.

Human constructs of myosin VI were previously described^{18,27}. Mutants generated here were engineered by site-directed mutagenesis and sequence verified. Details are available in Supplementary Table 1.

pRSF-duet codifying for His-Hip1R coiled coil region (amino acids 346–655) was previously described⁸.

pGEX6P1-CLCa FL construct was generated by subcloning RFP-hCLCa, kindly provided by Klemens Rottner (HZI, Braunschweig), into pGEX6P1 with XhoI enzyme. Subsequent truncated constructs were engineered by site-directed mutagenesis or recombinant PCR and sequence verified. Details are available upon request.

pGEX6P1-CLCb FL construct was generated by PCR using as template pCMV tag 2B-hCLCb, kindly provided by Elizabeth Smythe (University of Sheffield), and sequence verified.

pLVX lentiviral vector expressing rat RFP-CLCa was cloned from pMSCV-RFP-ratCLCa vector, kindly provided by Tom Kirchhausen (Harvard Medical School, Boston), using the InFusion system (Takara Clontech) and the following oligos:

forward: CTCAAGCTTCGAATTCATGGTGTCTAAGGGCGAAGAGCTG
reverse: GTCGACTGCAGAAATTCCTCAATGCACCAGGGGCGC.

pLVX RFP-CLCa I54D mutation was generated by site-directed mutagenesis and sequence verified.

pEGFP-CHC and pGEX6P1-CHC constructs described in Supplementary Fig. 1a were kindly provided by Stephen Royle (University of Liverpool).

5-Carboxyfluorescein (5-FAM)-conjugated CLCa peptides 46–61, 51–61 and 51–66 were synthesized by GenScript.

Antibodies. All antibodies used in this study are listed in Supplementary Table 2.

CLC depletion. Transient CLCa and/or CLCb KD were performed using Stealth siRNA oligonucleotides from ThermoFisher Scientific (humanCLCa-3utr target sequence: 5'-TGGAAACACTACATCTGCAATATCT-3' and humanCLCb-3utr target sequence: 5'-CGCTCCTCTCAGTCTACTCAATTG-3') or Qiagen [humanCLCa (2) target sequence: AGACAGTTATGCAGCTATT].

Caco-2 (or HeLa, Supplementary Figs. 1 and 11d) cells were transfected using RNAiMax (Invitrogen) and siRNA oligos at 8 nM final concentration. Cells were transfected twice, first in suspension and the day after in adhesion. The third day Caco-2 cells reach confluency and were kept confluent for additional 7 days in order to obtain a compact polarized monolayer. Medium was changed every 2 days.

Caco-2 cyst formation. Caco-2, Caco-2 RFP-rCLCa WT, and Caco-2 RFP-rCLCa I54D cell lines were transiently transfected with CLC oligos. Two days after siRNA transfection, single-cell suspensions were embedded in medium plus 50% Matrigel (Corning) and 1 mg/ml collagen mixture and plated on 8-well ibidi chamber (Cat. No: 80826) at a density of 9×10^3 cells/well. Approximately 150 μ l were plated in each well, allowed to solidify for 30 min at 37 °C, and overlaid with 200 μ l of medium. Medium was changed every 2 days. At day 6, cholera toxin (Sigma Cat. No.: C8052) was added at a concentration of 0.1 μ g/ml and cysts are collected the day after.

Immunofluorescence (IF) and co-localization experiments. For IF analysis, samples were fixed with 4% paraformaldehyde for 10 min and permeabilized at room temperature (RT) with 0.5% Triton-X100 in PBS for 10 min (cells) or 20 min (mice tissue). Blocking was performed in phosphate-buffered saline (PBS) containing 2% bovine serum albumin (BSA) for 1 h prior to the incubation of the primary antibodies in PBS containing 0.1% Triton X-100 and 2% BSA overnight at 4 °C. After extensive washing, sample were incubated with secondary antibodies 45 min, at RT. Coverslips were mounted in a glycerol solution (20% glycerol, 50 mM Tris pH = 8.4) to avoid mechanical deformation of the sample. Images were captured using a HXC $\times 63$ plan-apochromat objective on a Leica inverted SP5 or SP8 microscope with a laser scanning confocal system. Analysis was performed with ImageJ (<http://imagej.nih.gov/ij/>).

For co-localization experiments in HeLa cells, to remove soluble cellular proteins, cells were extracted with 0.03% saponin in cytosolic buffer (25 mM Hepes-KOH, pH 7.4, 25 mM KCl, 2.5 mM magnesium acetate, 5 mM ethylene glycol-bis(β -aminoethyl ether)-N,N,N',N'-tetraacetic acid (EGTA), 150 mM K-glutamate) for 1 min prior to fixation. For co-localization analysis, regions of interest were drawn around individual cells. The Manders' coefficient was obtained using JACoP plugin and processed for statistical analysis with Prism (GraphPad software). Statistical significance was determined by non-parametric two-tailed *t* tests. Sample size was chosen arbitrary with no inclusion and exclusion criteria.

For IF analysis of Caco-2 cysts, cysts grown in matrigel in ibidi chambers were fixed in paraformaldehyde (PFA) 4% for 25 min, washed three times in PBS–glycine 0.75%, once in PBS, and permeabilized with 0.5% Triton X-100 for 20 min. Blocking was performed with 10% goat serum in IF buffer (0.1% BSA, 0.05% Tween-20, 0.2% Triton X-100, 0.05% Na₂S₂O₃ in PBS). Samples were incubated overnight in appropriate primary antibodies dilutions in IF buffer at RT, then washed three times in IF buffer and incubated at RT for 1 h with secondary antibody dilutions 1:100 in IF buffer and 15 min with 4',6-diamidino-2-phenylindole. Finally, samples were washed in PBS and post-fixed in PFA 4% for 10 min. Images were captured using a Leica inverted SP2 microscope with a laser scanning confocal system.

Isoform detection by PCR. Expression of myosin VI isoforms in cell lines or tissues was assessed by PCR. Messenger RNA from cultured cells or cysts was isolated with Maxwell RSC instrument (Promega) according to the manufacturer's protocols. RNA from mouse kidney and intestine tissues was obtained from C57BL/6 WT mouse and from commercially available human RNA (Takara Clontech Cat.No.: 636539 and 636529 for intestine and kidney, respectively). Retro-transcription was performed with the High Capacity cDNA Reverse Transcription Kit (Applied Biosystem). cDNA obtained was used in PCR reactions with primers flanking the spliced region as described in ref. 18.

Protein expression and purification. GST fusion proteins were expressed in *Escherichia coli* Rosetta (DE3) cells (Novagen) at 30 °C for 4 h after induction with 0.8 mM IPTG at an OD₆₀₀ of 0.6. Cell pellets were resuspended in lysis buffer

(50 mM Na-HEPES pH 7.5, 150 mM NaCl, 1 mM ethylenediaminetetraacetic acid (EDTA), 5% Glycerol, 0.1% NP40, Protease Inhibitor Cocktail set III, Calbiochem). Sonicated lysates were cleared by centrifugation at $48,000 \times g$ for 30 min. Supernatants were incubated with 1 ml of glutathione-sepharose beads (GE Healthcare) per liter of bacterial culture. After 4 h at 4 °C, the beads were extensively washed with PBS/0.1% Triton and equilibrated in storage buffer (50 mM Tris pH 7.4, 100 mM NaCl, 1 mM EDTA, 1 mM DTT, 10% glycerol). When needed, proteins were incubated overnight at 4 °C with PreScission Protease to obtain samples cleaved from GST-tag. Cleaved proteins were further purified through Ion-Exchange chromatography on an ÄKTA Purifier FPLC system (GE Healthcare) and dialyzed overnight against storage buffer prior to snap-freezing in liquid nitrogen.

Biochemical experiments. For pull-down experiments, 1.5 μM of GST-fusion proteins immobilized onto GSH beads were incubated for 1 h at 4 °C in JS buffer (50 mM Hepes pH 7.5, 50 mM NaCl, 1.5 mM MgCl_2 , 5 mM EGTA, 5% glycerol, and 1% Triton X-100) with either 1 mg lysate produced in JC buffer or the indicated concentration of cleaved and purified proteins. After extensive washes with JS buffer, beads were re-suspended in Laemmli-buffer and proteins analyzed through SDS-PAGE. Detection was performed either by staining the gels with Coomassie or by immunoblotting using specific antibodies. Ponceau-stained membrane was used to show loading of GST-fusion proteins.

For competition experiments, 1.5 μM of GST-CLCa immobilized onto GSH beads were incubated with 2.5 μM of His-Hip1R³⁴⁶⁻⁶⁵⁵ for 1 hour at 4 °C in JS buffer, then increasing concentration (from 1.2 to 20 μM) of myosin VI⁹⁹⁸⁻¹¹³¹ were added and incubated for additional 30 min at 4 °C. Washes and detection as noted above.

For co-IP experiments, Caco-2 RFP-rat CLCa WT and I54D depleted of the endogenous CLCs as previously described, were lysed in JS buffer and IP was performed incubating 1 mg of lysate with RFP-Trap (Chromotek) for 120 min at 4 °C. After extensive washes with JS buffer, beads were re-suspended in Laemmli-buffer and proteins analyzed through sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and immunoblotting.

FP assays were carried out at 22 °C on a 384-well plate with Infinite 200 instrument (Tecan) using an excitation wavelength of 535 nm and an emission wavelength of 580 nm. Concentrations in the nanomolar range (20–50 nM) of 5-FAM-conjugated CLCa peptides (GenScript) were titrated with the indicated myosin VI-purified fragment starting from an initial concentration of ~150 μM . Polarization readings from three independent experiments were averaged and fitted as described in Eletr et al.⁶²

For analytical size exclusion chromatography, 50 μM CLCa⁴⁶⁻⁶¹ was incubated at 1:1 molar ratio with 50 μM of myosin VI⁹⁹⁸⁻¹¹³¹ fragment for 15 min at 4 °C and subjected to size exclusion chromatography on a Superdex75 (5/150) column using the ÄKTA microsystem (GE Healthcare). Fractions containing complex or single proteins were analyzed by SDS-PAGE and Coomassie staining.

ITC measurements were performed in a buffer containing 20 mM Tris-HCl pH 8.0, 150 mM NaCl, and 2.5% glycerol, at 25 °C on a MicroCal PEAQ-ITC (Malvern Panalytical, Malvern, UK) instrument. All protein samples were dialyzed overnight in this buffer prior to titration. In all, 380 μM CLCa WT or I54D mutant were titrated into 20 μM of either Hip1R³⁴⁶⁻⁶⁵⁵ or myosin VI¹⁰⁵⁰⁻¹¹³¹ provided in the reaction cell. For each experiment, we used 19 injections of 2 μl with 150 s spacing and 750 rpm stirring speed. Measurements were performed in duplicates or triplicates. Raw data were analyzed with the integrated Malvern analysis software, and heat production was fitted to a one-set-of-sites binding model.

CCV preparation from porcine brain. Fresh porcine brain (First Link Ltd, Birmingham) was thawed in 100 ml PBS per 50 g tissue at 37 °C. Once thawed, PBS was removed and substituted with buffer A (100 mM 2-(*N*-morpholino)ethanesulfonic acid, 1 mM EGTA, 0.5 mM MgCl_2) containing 2 mM phenylmethylsulfonyl fluoride (PMSF). Homogenization was performed at 4 °C in a mixer, at pulse mode to avoid warming of the mixture. The homogenate was then centrifuged at $8000 \times g$ for 30 min at 4 °C in a JA-17 rotor (Beckman Coulter) and supernatant filtered through two layers of gauze bandage to further remove particles. The final filtrate was pooled and centrifuged at $185,000 \times g$ for 60 min at 4 °C in a Ti-45 rotor (Beckman Coulter) to pellet CCVs. To further purify CCVs from impurities, the pellet was resuspended in a final volume of 12–15 ml buffer A containing 2 mM PMSF and 0.02% NaN_3 and homogenized on ice for 5 min using a SS30 dounce stirrer (Stuart) at 125 rpm. An equivalent volume of a 12.5% Ficoll and 12.5% sucrose solution was added to the CCV homogenate and mixed gently. The mixture was then centrifuged at $30,000 \times g$ for 45 min at 4 °C using a JA-17 rotor (Beckman Coulter). CCV-containing supernatant was diluted fivefold with buffer A + 2 mM PMSF and CCVs sedimented at $185,000 \times g$ for 60 min at 4 °C in a Ti-45 rotor (Beckman Coulter). The supernatant was discarded, vesicles resuspended in 4 ml buffer A + 2 mM PMSF and 0.02% NaN_3 , aliquoted and snap-frozen in liquid nitrogen, and stored at -80 °C until further use.⁶³

Purification of triskelia from CCVs. Tris-extraction was performed according to ref. ⁶³ with slight modifications. Purified CCV were quickly thawed at 37 °C and

kept on ice. A solution of Tris pH 7.3, EDTA, protease inhibitor cocktail, and β -mercaptoethanol was added to the CCV suspension to obtain the final concentration of the Tris-extraction buffer (0.5 M Tris pH 7.3, 2 mM EDTA, 1 mM β -mercaptoethanol and $1 \times$ protease inhibitors). The mix was left on ice for 10 min before ultracentrifugation in a TLA 100.3 rotor (Beckman Coulter) for 30 min at $220,000 \times g$ at 4 °C. Supernatant containing clathrin triskelia and adaptor proteins was further purified via size exclusion chromatography in buffer B (0.5 M Tris pH 7.4, 2 mM EDTA) using a Superose 6 increase 10/300 GL column (GE Healthcare) connected to an ÄKTA Purifier FPLC system (GE Healthcare) at a flow rate of 0.5 ml/min at 8 °C. Elution fractions of 0.5 ml volumes were collected and analyzed via SDS-PAGE. Clean clathrin fractions were pooled, and an equivalent volume of saturated NH_4SO_4 solution was carefully added. The mix was incubated on ice for 30 min to allow proteins precipitation and centrifuged for 30 min at $1600 \times g$ using an Eppendorf 5810 R centrifuge (Eppendorf) at 4 °C. The pellet was dissolved in 70–150 μl extraction buffer (depending on pellet size), dialyzed in buffer B for 1 h, and overnight at 4 °C in buffer C (50 mM Tris pH 8.0, 50 mM NaCl, 2 mM EDTA, 1 mM β -mercaptoethanol, 1 mM PMSF). Protein concentration (3–5 mg/ml) was determined by ultraviolet (UV) absorbance at 280 nm, using an extinction coefficient of 1.0 (mg/ml)/cm. Clathrin triskelia were either directly subjected to CLC dissociation (see below) or dialyzed against buffer A containing 2 mM CaCl_2 overnight at 4 °C to allow cage assembly and long-term storage as native cages.

CLC dissociation and cage re-assembly. CLC dissociation was performed as previously described (Winkler and Stanley 1983). Clathrin triskelia solution were diluted in stripping buffer D (50 mM Tris pH 8.0, 50 mM NaCl, 1.3 M NaSCN, 2 mM EDTA, 1 mM β -mercaptoethanol, 1 mM PMSF) and subjected to size exclusion chromatography on a Superose 6 increase 10/300 GL column (GE Healthcare). Elution fractions were collected and analyzed via SDS-PAGE. Clean clathrin fractions were pooled, dialyzed twice against buffer C to remove NaSCN and concentrated by spin filtration using 10 K Amicon-Ultra centrifugal filters (Merck Millipore). Protein concentrations were determined by UV absorbance at 280 nm using an extinction coefficient of 1.0 (mg/ml)/cm. To assemble CHC-only cages, CHC triskelia were dialyzed overnight in buffer A containing 2 mM CaCl_2 . To generate CHC-CLC specific cages, CHC-only cages were reconstituted with either CLCa or CLCb purified proteins in slight molar excess (CHC:CLC 1:1.2). Mix were incubated on ice for 1 h, centrifuged for 30 min at 4 °C in a TLA55 rotor at $150,000 \times g$ and washed once with buffer A to eliminate the unbound CLCs.

Co-sedimentation of clathrin cages with myosin. Native and CLCa or CLCb-only cages were diluted to 1.5 μM in buffer A + 0.1% Triton X-100 and 0.1 mg/ml BSA and mixed with 1.5 μM of purified myosin VI⁹⁹⁸⁻¹¹³¹ fragment on ice for 45 min. Mix were then centrifuged at 4 °C for 30 min in a TLS-55 rotor at $108,000 \times g$, pellets were washed once with same volume of buffer A + 0.1% Triton X-100 and 0.1 mg/ml BSA and centrifuged. Finally, pellets were dissolved in Laemmli buffer and analyzed through SDS-PAGE.

Negative staining of clathrin cages. Cages were adsorbed to freshly glow-discharged, carbon-coated formvar films on copper grids by placing grids on 10 μl droplets of samples at 0.2 mg/ml on Parafilm for 90 s. Excessive liquid was removed using filter paper. Grids were washed twice by sequentially transferring grids onto 15 μl droplets of buffer A containing 2 mM CaCl_2 . Samples were then stained with 2% uranyl acetate in water for 1 min and air-dried before subjected to EM analysis. Specimens were observed using a Tecnai G2 (FEI) electron microscope at an acceleration voltage of 120 kV. Images were obtained using a SIS Morada digital camera and TIA software (FEI).

EM of Caco-2 cysts. Cysts grown in matrigel were fixed with 2.5% glutaraldehyde in 0.1 M sodium cacodylate buffer pH 7.4 for 1 h at RT, washed three times with cacodylate buffer, and fixed in 1% osmium tetroxide and 1.5% potassium ferrocyanide in 0.1 M cacodylate for 1 h on ice. After several washes in distilled water, samples were en bloc stained with 0.5% uranyl acetate in water overnight at 4 °C. Finally, samples were dehydrated in a graded ethanol series (30%, 50%, 70%, 80%, 90%, 96%, 20 min each and 3 washes with absolute ethanol, 10 min each). Next, the samples were infiltrated in a 1:1 ethanol/Epon 812 solution for 2 h, in 100% Epoxy resin twice for 1 h each, and left overnight in resin before being embedded in resin and polymerized at 60 °C for 48 h. At the end of polymerization, the resin blocks were separated from the multiwell chamber and part of the block containing few cysts was mounted on a Leica Ultracut UCT ultramicrotome. Ultrathin (70–90 nm) sections were then collected on formvar carbon-coated slot grids and stained with uranyl acetate and Sato's lead citrate before imaged with a Talos 120 C (FEI) electron microscope; images were acquired with a 4k \times 4 K Ceta CMOS camera.

For morphometrical analysis of the distribution of clathrin-coated structures, between 94 and 125 cellular profiles of well-polarized cells for each biological sample were imaged. Images of the CCPs present on the apical surface and connected to the membrane were acquired at the nominal magnification of $\times 36,000$ and classified as shallow, omega/constricted, or elongated according to their morphology. The distribution of clathrin-coated structures has been expressed as

percentage among the different classes (Fig. 5e) or as the number per microns of apical membrane (Supplementary Fig. 10g). All statistical analyses were performed using GraphPad Prism. One-way analysis of variance was used to calculate statistical significance among different samples.

Protein expression and purification for NMR spectroscopy. GST-tagged myosin VI and CLCa were expressed separately in *E. coli* BL21 (DE3) cells (Thermo Fischer Scientific) at 17 °C overnight upon reaching an OD₆₀₀ value of 0.5–0.6 by induction with 0.4 mM IPTG. Cell pellets were resuspended in Buffer A (50 mM HEPES, pH 7.5, 300 mM NaCl, 1 mM EDTA, 5% glycerol, 2 mM DTT, and Roche Complete Mini protease inhibitor cocktail), lysed by sonication, and the lysates were cleared via centrifugation. Supernatants were incubated with glutathione sepharose beads (GE Healthcare) for 3 h at 4 °C with a fresh protease inhibitor cocktail tablet added. Myosin VI-bound beads were washed twice with Buffer A, twice with Buffer B (PBS plus 500 mM NaCl, 0.1% Triton X-100, and 2 mM DTT, pH 7.4), twice with Buffer C (PBS plus 2 mM DTT, pH 7.4), and once with Buffer D (50 mM Tris-Cl, 100 mM NaCl, 1 mM EDTA, 10% glycerol, and 2 mM DTT, pH 7.0). The washed beads were resuspended in Buffer A and incubated with PreScission Protease (GE Healthcare) overnight at 4 °C. Myosin VI was eluted with Buffer A, concentrated to 2 ml, purified via size exclusion chromatography on an FPLC system (Superdex75, GE Healthcare) with FPLC Buffer (20 mM NaPO₄, 50 mM NaCl, 2 mM DTT, pH 6.5), and concentrated to 1.5 ml. CLCa-bound glutathione sepharose beads were washed five times with Buffer A, at which point purified myosin VI (from the above purification protocol) was added to the beads. The CLCa–myosin VI mixture was incubated with PreScission Protease overnight and eluted with Buffer A. Elutions were concentrated to 2 ml, purified via size exclusion chromatography on an FPLC system with FPLC Buffer, and concentrated to ~300 µl.

NMR spectroscopy and backbone assignment. NMR data were collected at 10 °C on Bruker 700 and 850 MHz spectrometers. All experiments were conducted in 20 mM NaPO₄ pH 6.5, 50 mM NaCl, 2 mM DTT, 0.1% NaN₃, 1 mM Pefabloc, and 5% D₂O. In all, 0.78 mM ¹⁵N-labeled myosin VI and synthesized 5-FAM N-terminally tagged CLCa (GenScript) were used for the HSQC experiments acquired for CSP analysis. For structure determination, two separate samples were produced, namely, 0.4 mM ¹³C, ¹⁵N-labeled CLCa with equimolar unlabeled myosin VI, and 0.36 mM ¹³C, ¹⁵N-labeled myosin VI with equimolar unlabeled CLCa. For each sample, the following NMR experiments were performed: ¹⁵N-HSQC, ¹³C-HSQC, HNCOC, HNCACB, ¹⁵N-dispersed NOESY (120 ms mixing time), ¹³C-dispersed NOESY (100 ms mixing time), and ¹³C-dispersed-half-filtered-NOESY (100 ms mixing time). The NMR spectra were processed with NMRpipe⁶⁴ and analyzed with XEASY⁶⁵. Backbone assignments of myosin VI and CLCa in complex were done manually based on chemical shifts obtained previously on free myosin VI (BMRB 25544) and the HNCACB and HNCOC spectra; the HNCACB spectrum showed both the intra-residue and preceding residue Ca signals for most residues (Supplementary Fig. 4). Residues that were ambiguous were verified with HN-HN and α-HN NOEs in the ¹⁵NOESY (Supplementary Fig. 5). Sidechain assignments were obtained by manual analysis of HNCACB, ¹³C-dispersed NOESY, and ¹⁵N-dispersed NOESY experiments.

Structure determination and refinement. NOE assignments were done manually using ¹³C-dispersed NOESY, ¹⁵N-dispersed NOESY, and ¹³C-dispersed half-filtered NOESY experiments. Chemical shifts for backbone carbons and amide nitrogen, amide hydrogen, Ha, and Cβ from myosin VI and CLCa were used with TALOS⁶⁶ to obtain backbone φ and ψ torsion angle restraints. Structure calculations were performed with Xplor-NIH (version 2.47)⁶⁷ by using standard scripts. Twenty linear starting structures were subjected to 35,000 simulated annealing steps followed by 5000 cooling steps, both of 0.005 ps. The lowest energy structure from this set was used as the starting structure for refinement closely following the protocol described in ref. 68 but with a temperature step of 1 K used during simulated annealing. Constraints used during the structure calculations include intramolecular and intermolecular NOE restraints, dihedral angle restraints, and backbone hydrogen bonds in helical regions. A summary of the data used for the structure calculations is provided in Table 1. Hydrogen bond distances between the acceptor oxygen and donor hydrogen and nitrogen atoms were set to 1.81–2.09 and 2.71–3.13 Å, respectively. Structures and NOE assignments were checked manually to identify errors, corrections were made if necessary, and structures recalculated; this process was done iteratively. Of the 300 final refined structures calculated, the 20 lowest energy structures without distance violations >0.5 Å and dihedral angle violations >5° were selected to represent the structure of myosin VI in complex with CLCa. Pymol (Schrödinger) was used for visualization and figure generation; MOLMOL⁶⁹ was used for r.m.s.d. calculation and figure generation. Among the 20 lowest energy structures, no amino acid was disfavored by Ramachandran analyses; 1, 6, and 94% were in the generously allowed, allowed, and favored regions, respectively.

Reporting summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

Data supporting the findings of this manuscript are available in the Source Data file. All uncropped immunoblots associated with Figs. 1b, 1c, 1d, 1e, 2b, 4f, 4d, 5a, 6b, 6c and Supplementary Figs. 1a, 1c, 2d, 9a, 9b, 10a, 10c, 11c, 11d, 11e, 12a, 12b are provided as Source Data file. A reporting summary for this Article is available as a Supplementary Information file. Chemical shift assignments for the complex of myosin VI^{1050–1131} and CLCa^{46–61} have been deposited in the Biological Magnetic Resonance Bank under the ID code 30500, and the atomic coordinates for the complex have been deposited in the Protein Data Bank under the ID code 6E5N.

Received: 3 June 2019; Accepted: 27 September 2019;

Published online: 31 October 2019

References

- Brodsky, F. M. Diversity of clathrin function: new tricks for an old protein. *Annu. Rev. Cell Dev. Biol.* **28**, 309–336 (2012).
- Kirchhausen, T., Owen, D. & Harrison, S. C. Molecular structure, function, and dynamics of clathrin-mediated membrane traffic. *Cold Spring Harb. Perspect. Biol.* **6**, a016725 (2014).
- Ferreira, F. et al. Endocytosis of G protein-coupled receptors is regulated by clathrin light chain phosphorylation. *Curr. Biol.* **22**, 1361–1370 (2012).
- Poupon, V. et al. Clathrin light chains function in mannose phosphate receptor trafficking via regulation of actin assembly. *Proc. Natl Acad. Sci. USA* **105**, 168–173 (2008).
- Wu, S. et al. Clathrin light chains' role in selective endocytosis influences antibody isotype switching. *Proc. Natl Acad. Sci. USA* **113**, 9816–9821 (2016).
- Boulant, S., Kural, C., Zeeh, J. C., Ubelmann, F. & Kirchhausen, T. Actin dynamics counteract membrane tension during clathrin-mediated endocytosis. *Nat. Cell Biol.* **13**, 1124–1131 (2011).
- Chen, C. Y. & Brodsky, F. M. Huntingtin-interacting protein 1 (Hip1) and Hip1-related protein (Hip1R) bind the conserved sequence of clathrin light chains and thereby influence clathrin assembly in vitro and actin distribution in vivo. *J. Biol. Chem.* **280**, 6109–6117 (2005).
- Wilbur, J. D. et al. Conformation switching of clathrin light chain regulates clathrin lattice assembly. *Dev. Cell* **18**, 841–848 (2010).
- Majeed, S. R. et al. Clathrin light chains are required for the gyrating-clathrin recycling pathway and thereby promote cell migration. *Nat. Commun.* **5**, 3891 (2014).
- Bonazzi, M. et al. Clathrin phosphorylation is required for actin recruitment at sites of bacterial adhesion and internalization. *J. Cell Biol.* **195**, 525–536 (2011).
- Cossart, P. & Helenius, A. Endocytosis of viruses and bacteria. *Cold Spring Harb. Perspect. Biol.* **6**, a016972 (2014).
- Cureton, D. K., Massol, R. H., Whelan, S. P. & Kirchhausen, T. The length of vesicular stomatitis virus particles dictates a need for actin assembly during clathrin-dependent endocytosis. *PLoS Pathog.* **6**, e1001127 (2010).
- Buss, F., Spudich, G. & Kendrick-Jones, J. Myosin VI: cellular functions and motor properties. *Annu. Rev. Cell Dev. Biol.* **20**, 649–676 (2004).
- Bonazzi, M. et al. A common clathrin-mediated machinery co-ordinates cell-cell adhesion and bacterial internalization. *Traffic* **13**, 1653–1666 (2012).
- Tsygankova, O. M. & Keen, J. H. A unique role for clathrin light chain A in cell spreading and migration. *J. Cell Sci.* **132**, jcs224030 (2019).
- Maib, H., Ferreira, F., Vassilopoulos, S. & Smythe, E. Cargo regulates clathrin-coated pit invagination via clathrin light chain phosphorylation. *J. Cell Biol.* **217**, 4253–4266 (2018).
- Chen, P. H. et al. Crosstalk between CLCb/Dyn1-mediated adaptive clathrin-mediated endocytosis and epidermal growth factor receptor signaling increases metastasis. *Dev. Cell* **40**, 278.e5–288.e5 (2017).
- Wollscheid, H. P. et al. Diverse functions of myosin VI elucidated by an isoform-specific alpha-helix domain. *Nat. Struct. Mol. Biol.* **23**, 300–308 (2016).
- Hegan, P. S., Giral, H., Levi, M. & Mooseker, M. S. Myosin VI is required for maintenance of brush border structure, composition, and membrane trafficking functions in the intestinal epithelial cell. *Cytoskeleton (Hoboken)* **69**, 235–251 (2012).
- Liu, S. H., Wong, M. L., Craik, C. S. & Brodsky, F. M. Regulation of clathrin assembly and trimerization defined using recombinant triskelion hubs. *Cell* **83**, 257–267 (1995).
- Ybe, J. A. et al. Clathrin self-assembly is regulated by three light-chain residues controlling the formation of critical salt bridges. *EMBO J.* **17**, 1297–1303 (1998).
- Brodsky, F. et al. Clathrin light chains: arrays of protein motifs that regulate coated-vesicle dynamics. *Trends Biochem. Sci.* **16**, 208–213 (1991).
- Fotin, A. et al. Molecular model for a complete clathrin lattice from electron cryomicroscopy. *Nature* **432**, 573–579 (2004).

24. Walters, K. J. et al. Characterizing protein-protein complexes and oligomers by nuclear magnetic resonance spectroscopy. *Methods Enzymol.* **339**, 238–258 (2001).
25. Sahlender, D. A. et al. Optineurin links myosin VI to the Golgi complex and is involved in Golgi organization and exocytosis. *J. Cell Biol.* **169**, 285–295 (2005).
26. Spudich, G. et al. Myosin VI targeting to clathrin-coated structures and dimerization is mediated by binding to Disabled-2 and PtdIns(4,5)P₂. *Nat. Cell Biol.* **9**, 176–183 (2007).
27. He, F. et al. Myosin VI contains a compact structural motif that binds to ubiquitin chains. *Cell Rep.* **14**, 2683–2694 (2016).
28. Buss, F., Luzio, J. P. & Kendrick-Jones, J. Myosin VI, a new force in clathrin mediated endocytosis. *FEBS Lett.* **508**, 295–299 (2001).
29. Buss, F. et al. Myosin VI isoform localized to clathrin-coated vesicles with a role in clathrin-mediated endocytosis. *EMBO J.* **14**, 3676–3684 (2001).
30. Biemesderfer, D., Mentone, S. A., Mooseker, M. & Hasson, T. Expression of myosin VI within the early endocytic pathway in adult and developing proximal tubules. *Am. J. Physiol. Ren. Physiol.* **282**, F785–F794 (2002).
31. Jaffe, A. B., Kaji, N., Durgan, J. & Hall, A. Cdc42 controls spindle orientation to position the apical surface during epithelial morphogenesis. *J. Cell Biol.* **183**, 625–633 (2008).
32. Damke, H., Baba, T., Warnock, D. E. & Schmid, S. L. Induction of mutant dynamin specifically blocks endocytic coated vesicle formation. *J. Cell Biol.* **127**, 915–934 (1994).
33. Engqvist-Goldstein, A. E. et al. The actin-binding protein Hip1R associates with clathrin during early stages of endocytosis and promotes clathrin assembly in vitro. *J. Cell Biol.* **154**, 1209–1223 (2001).
34. Legendre-Guillemin, V. et al. HIP1 and HIP12 display differential binding to F-actin, AP2, and clathrin. Identification of a novel interaction with clathrin light chain. *J. Biol. Chem.* **277**, 19897–19904 (2002).
35. Henry, K. R. et al. Scd5p and clathrin function are important for cortical actin organization, endocytosis, and localization of sla2p in yeast. *Mol. Biol. Cell* **13**, 2607–2625 (2002).
36. Aschenbrenner, L., Lee, T. & Hasson, T. Myo6 facilitates the translocation of endocytic vesicles from cell peripheries. *Mol. Biol. Cell* **14**, 2728–2743 (2003).
37. Wagner, W. et al. Myosin VI drives clathrin-mediated AMPA receptor endocytosis to facilitate cerebellar long-term depression. *Cell Rep.* **28**, 11. e9–20.e9 (2019).
38. Wells, A. L. et al. Myosin VI is an actin-based motor that moves backwards. *Nature* **401**, 505–508 (1999).
39. DeLuca-Flaherty, C., McKay, D. B., Parham, P. & Hill, B. L. Uncoating protein (hsc70) binds a conformationally labile domain of clathrin light chain LCa to stimulate ATP hydrolysis. *Cell* **62**, 875–887 (1990).
40. Batters, C., Brack, D., Ellrich, H., Aeverbeck, B. & Veigel, C. Calcium can mobilize and activate myosin-VI. *Proc. Natl Acad. Sci. USA* **113**, E1162–E1169 (2016).
41. Engqvist-Goldstein, A. E. et al. RNAi-mediated Hip1R silencing results in stable association between the endocytic machinery and the actin assembly machinery. *Mol. Biol. Cell* **15**, 1666–1679 (2004).
42. Huang, F., Khvorova, A., Marshall, W. & Sorkin, A. Analysis of clathrin-mediated endocytosis of epidermal growth factor receptor by RNA interference. *J. Biol. Chem.* **279**, 16657–16661 (2004).
43. Ripoll, L. et al. Myosin VI and branched actin filaments mediate membrane constriction and fission of melanosomal tubule carriers. *J. Cell Biol.* **217**, 2709–2726 (2018).
44. Mund, M. et al. Systematic nanoscale analysis of endocytosis links efficient vesicle formation to patterned actin nucleation. *Cell* **174**, 884–896 e17 (2018).
45. Collins, A., Warrington, A., Taylor, K. A. & Svitkina, T. Structural organization of the actin cytoskeleton at sites of clathrin-mediated endocytosis. *Curr. Biol.* **21**, 1167–1175 (2011).
46. Swiatecka-Urban, A. et al. Myosin VI regulates endocytosis of the cystic fibrosis transmembrane conductance regulator. *J. Biol. Chem.* **279**, 38025–38031 (2004).
47. Blaine, J. et al. PTH-induced internalization of apical membrane NaPi2a: role of actin and myosin VI. *Am. J. Physiol. Cell Physiol.* **297**, C1339–C1346 (2009).
48. Ameen, N. & Apodaca, G. Defective CFTR apical endocytosis and enterocyte brush border in myosin VI-deficient mice. *Traffic* **8**, 998–1006 (2007).
49. Gotoh, N. et al. Altered renal proximal tubular endocytosis and histology in mice lacking myosin-VI. *Cytoskeleton (Hoboken)* **67**, 178–192 (2010).
50. Collaco, A., Jakab, R., Hegan, P., Mooseker, M. & Ameen, N. Alpha-AP-2 directs myosin VI-dependent endocytosis of cystic fibrosis transmembrane conductance regulator chloride channels in the intestine. *J. Biol. Chem.* **285**, 17177–17187 (2010).
51. Chen, T. et al. Myosin VI mediates the movement of NHE3 down the microvillus in intestinal epithelial cells. *J. Cell Sci.* **127**, 3535–3545 (2014).
52. Skruzny, M. et al. Molecular basis for coupling the plasma membrane to the actin cytoskeleton during clathrin-mediated endocytosis. *Proc. Natl Acad. Sci. USA* **109**, E2533–E2542 (2012).
53. Clarke, N. I. & Royle, S. J. FerriTag is a new genetically-encoded inducible tag for correlative light-electron microscopy. *Nat. Commun.* **9**, 2604 (2018).
54. Messa, M. et al. Epsin deficiency impairs endocytosis by stalling the actin-dependent invagination of endocytic clathrin-coated pits. *Elife* **3**, e03311 (2014).
55. Garcia-Alai, M. M. et al. Epsin and Sla2 form assemblies through phospholipid interfaces. *Nat. Commun.* **9**, 328 (2018).
56. Sochacki, K. A., Dickey, A. M., Strub, M. P. & Taraska, J. W. Endocytic proteins are partitioned at the edge of the clathrin lattice in mammalian cells. *Nat. Cell Biol.* **19**, 352–361 (2017).
57. Antonny, B. et al. Membrane fission by dynamin: what we know and what we need to know. *EMBO J.* **35**, 2270–2284 (2016).
58. Mettlen, M., Chen, P. H., Srinivasan, S., Danuser, G. & Schmid, S. L. Regulation of clathrin-mediated endocytosis. *Annu Rev. Biochem.* **87**, 871–896 (2018).
59. Yu, C. et al. Myosin VI undergoes cargo-mediated dimerization. *Cell* **138**, 537–548 (2009).
60. Taylor, M. J., Perrais, D. & Merrifield, C. J. A high precision survey of the molecular dynamics of mammalian clathrin-mediated endocytosis. *PLoS Biol.* **9**, e1000604 (2011).
61. Jackson, A. P. & Parham, P. Structure of human clathrin light chains. Conservation of light chain polymorphism in three mammalian species. *J. Biol. Chem.* **263**, 16688–16695 (1988).
62. Eletr, Z. M., Huang, D. T., Duda, D. M., Schulman, B. A. & Kuhlman, B. E2 conjugating enzymes must disengage from their E1 enzymes before E3-dependent ubiquitin and ubiquitin-like transfer. *Nat. Struct. Mol. Biol.* **12**, 933–934 (2005).
63. Ungewickell, E. & Ungewickell, H. Bovine brain clathrin light chains impede heavy chain assembly in vitro. *J. Biol. Chem.* **266**, 12710–12714 (1991).
64. Schaefer, A. W. et al. L1 endocytosis is controlled by a phosphorylation-dephosphorylation cycle stimulated by outside-in signaling by L1. *J. Cell Biol.* **157**, 1223–1232 (2002).
65. Bartels, C., Xia, T. H., Billeter, M., Guntert, P. & Wuthrich, K. The program XEASY for computer-supported NMR spectral analysis of biological macromolecules. *J. Biomol. NMR* **6**, 1–10 (1995).
66. Cornilescu, G., Delaglio, F. & Bax, A. Protein backbone angle restraints from searching a database for chemical shift and sequence homology. *J. Biomol. NMR* **13**, 289–302 (1999).
67. Schwieters, C. D., Kuszewski, J. J., Tjandra, N. & Clore, G. M. The Xplor-NIH NMR molecular structure determination package. *J. Magn. Reson.* **160**, 65–73 (2003).
68. Schwieters, C. D., Bermejo, G. A. & Clore, G. M. Xplor-NIH for molecular structure determination from NMR and other data sources. *Protein Sci.* **27**, 26–40 (2018).
69. Koradi, R., Billeter, M. & Wuthrich, K. MOLMOL: a program for display and analysis of macromolecular structures. *J. Mol. Graph* **14**, 51–55, 29–32 (1996).

Acknowledgements

We thank Tom Kirchhausen, Stephen Royle, Klemens Rottner, and Elizabeth Smythe for DNA constructs. We thank ALEMBIC facility at San Raffaele Scientific Institute, Milan, Italy for support in electron microscopic analysis and the Biochemistry and Structural Biology Unit at the European Institute of Oncology for support with isothermal titration calorimetric analysis. This work was supported by the Associazione Italiana per la Ricerca sul Cancro (AIRC IG19875 to S.P.); the Italian Ministry of Education, Universities and Research (PRIN 20108MXN2J to S.P.); the Intramural Research Program through the CCR, NCI, NIH (1ZIABC011627 to K.J.W.); and Wellcome Trust Investigator Award (107858/Z/15/Z to F.M.B.). C.A.N.'s work is supported by the Marie Skłodowska-Curie Actions (MSCA-IF-2016 #752553). M.B. was and R.S.d.P. is a PhD student within the European School of Molecular Medicine (SEMM). M.B. was supported by a fellowship from Fondazione Umberto Veronesi (FUV) and by an EMBO short-term travel fellowship.

Author contributions

M.B. performed most of the biochemical experiments and contributed to result interpretation; G.R.B. performed the NMR experiments and solved the structures; C.A.N. generated the genetically reconstituted Caco-2 cells and cysts and contributed to experimental design; E.M. performed the Hip1R experiments; R.S.d.P. prepared the samples for EM analysis; A.R. conducted the EM experiments; L.R. supervised the mice experiments; J.W. performed the ITC experiments. F.M.B. generated the CLC KO mice and contributed to experimental design and paper writing; K.J.W. supervised the NMR analysis and contributed to NMR interpretation and paper writing; S.P. coordinated the team, designed the experiments, and wrote the paper.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41467-019-12855-6>.

Correspondence and requests for materials should be addressed to F.M.B., K.J.W. or S.P.

Peer review information *Nature Communications* thanks the anonymous reviewers for their contribution to the peer review of this work. Peer reviewer reports are available.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2019

Theranostic Complexes

A Highly Luminescent Tetrahydrocurcumin Ir^{III} Complex with Remarkable Photoactivated Anticancer ActivityAlessia Colombo,^[a] Mattia Fontani,^[a] Claudia Dragonetti,^{*,[a]} Dominique Roberto,^[a] J. A. Gareth Williams,^{*,[b]} Rossella Scotto di Perrotolo,^[c] Francesca Casagrande,^[c] Sara Barozzi,^[c] and Simona Polo^{*,[c, d]}

Dedicated to Emeritus Professor Renato Ugo on the occasion of his 81st birthday

Abstract: Curcumin has chemopreventative properties against a variety of tumours, but has poor bioavailability. Here, two new bis-cyclometallated iridium(III) complexes have been prepared, featuring the natural product curcumin (CUR) or its reduced form, tetrahydrocurcumin (THC), as bidentate, anionic OO-binding ligands. The iridium THC complex is highly luminescent in deoxygenated solution and efficiently generates singlet oxygen under aerated conditions, whereas in the CUR analogue, other non-radiative decay

pathways are competitive. The complexes are rapidly taken up by a variety of human tumour cell lines from solutions of micromolar concentration. They show negligible cytotoxicity in the absence of irradiation. When briefly irradiated with visible light, Ir-THC becomes highly phototoxic, inducing rapid apoptosis within 2 h. The results show the high potential of such complexes as sensitizers in photodynamic therapy (PDT).

Introduction

Curcumin, 1,7-bis(4-hydroxy-3-methoxyphenyl)-1,6-heptadiene-3,5-dione (CUR, Figure 1), is a polyphenolic natural product isolated from *Curcuma longa*, a rhizomatous herbaceous plant. CUR has been a well-recognized dietary spice for centuries and its pharmacological activities have been assessed through various clinical studies, which include anti-inflammatory, anti-diabetic, anti-dementia, and anti-oxidant properties.^[1] In particular, it appears to have chemopreventative properties against a variety of human malignancies and is currently in clinical trials as an anticancer agent.^[2] In Phase I clinical trials, it was con-

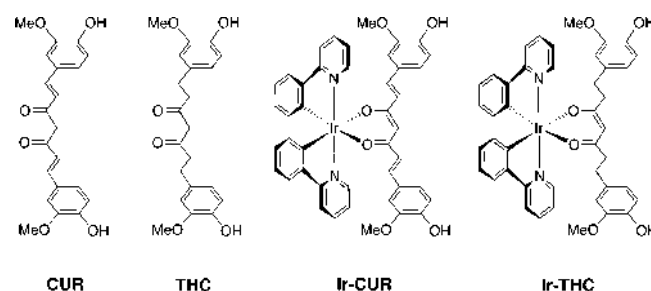


Figure 1. Structures of curcumin (CUR), tetrahydrocurcumin (THC), and their corresponding iridium(III) complexes Ir-CUR and Ir-THC featuring the ligands in their anionic (deprotonated) forms.

[a] Dr. A. Colombo, Dr. M. Fontani, Prof. C. Dragonetti, Prof. D. Roberto
Dipartimento di Chimica dell'Università degli Studi di Milano
and UdR INSTM di Milano, via Golgi 19, 20133 Milan (Italy)
E-mail: claudia.dragonetti@unimi.it

[b] Prof. J. A. G. Williams
Department of Chemistry, Durham University, Durham, DH1 3LE (UK)
E-mail: j.a.g.williams@durham.ac.uk

[c] Dr. R. Scotto di Perrotolo, Dr. F. Casagrande, Dr. S. Barozzi, Dr. S. Polo
IFOM, Fondazione Istituto FIRC di Oncologia Molecolare
Via Adamello 16, 20145 Milan (Italy)
E-mail: simona.polo@ifom.eu

[d] Dr. S. Polo
Dipartimento di Oncologia ed Emato-Oncologia
Università degli Studi di Milano, 20139 Milan (Italy)

Supporting information and the ORCID identification number(s) for the author(s) of this article can be found under:
<https://doi.org/10.1002/chem.201901527>: Details of syntheses and NMR spectra, absorption and emission spectra of the Ir^{III} complexes, details of cell lines, cellular uptake and staining results, data from time-lapse experiments, and cell viability (MTT and CCK) assays.

cluded that humans can tolerate a CUR dose as high as 8 g/day with no side effects.^[3] However, CUR shows poor oral bioavailability: its concentration in human plasma and other tissues is extremely low even after a high oral dose. This is due to its instability with respect to hydrolysis of the β -diketone unit under physiological conditions, rapidly leading to species such as ferullic acid and vanillin.^[4,5] Interestingly, binding this moiety in the deprotonated anionic form to a metal centre increases its hydrolytic stability with respect to free CUR, and the resulting complexes can show cytotoxic activity.^[4,6] Indeed, a bioreductively activated cobalt(III) carrier system for the delivery of CUR with enhanced drug stability and efficacy against colon cancer cells was recently reported.^[4b]

The high instability of CUR suggests that it may be the products of its metabolism that are responsible for its pharmacological effects in vivo.^[7] In particular, tetrahydrocurcumin (THC) is one of its active metabolites (Figure 1) and may play a cru-

cial role in CUR-induced biological effects. Its easy absorption through the gastrointestinal tract suggests that THC might even be a better candidate than CUR for the development of anticancer agents.^[8] In agreement with this observation, THC exhibits significant cell growth inhibition by inducing human breast cancer MCF-7 cells to undergo mitochondrial apoptosis and G2/M arrest.^[8]

The photobiological activity of CUR is related to its ability to generate reactive oxygen species (ROS), which induce cancer cell death by apoptosis through mitochondrial pathways.^[4] The photochemical properties of CUR are therefore important in this context.^[9] It shows a strong absorption band at around 410–430 nm and fluorescence in the region 460–560 nm in solution at room temperature, with a quantum yield in the range 0.02–0.08 depending on the solvent. While the singlet excited state decays by a non-radiative process, its triplet state reacts with ground-state $^3\text{O}_2$ to efficiently generate singlet oxygen $^1\text{O}_2$. CUR is thus of great interest for photodynamic therapy (PDT), a treatment in which the cytotoxicity of a drug is selectively activated by light only in cancerous regions, through the generation of $^1\text{O}_2$, which shows promise for several types of solid cancer.^[10,11]

Although CUR could have potential as a PDT agent, its low bioavailability would severely limit its efficacy. The greater stability of metal complexes of CUR, coupled with longer-wavelength absorption, could render them suitable for such photo-therapeutic applications.^[4] Surprisingly, despite growing interest in the use of iridium complexes as PDT agents,^[11] to the best of our knowledge only one iridium complex incorporating CUR as a ligand has hitherto been reported.^[6,12] Pettinari et al. reported that $[\text{Cp}^*\text{Ir}(\text{curc})\text{Cl}]$ (Cp^* = pentamethylcyclopentadienyl, curcH = curcumin) is moderately cytotoxic towards both human ovarian cancer cells and non-tumorigenic human embryonic kidney cells, though its potential as a PDT agent was not investigated.^[12] However, cyclometallated Ir^{III} complexes could have great potential in PDT because they may efficiently generate $^1\text{O}_2$ due to the presence of the heavy atom, which favours fast intersystem crossing to triplet states with microsecond lifetimes.^[13] Moreover, they are often luminescent, offering potential for cellular imaging, and for what is known as theranostics, that is, the combination of therapy and imaging.^[14–27] Based on these observations, we decided to prepare and study an Ir^{III} complex bearing two cyclometallated phenylpyridine (ppy) ligands and curcumin (Ir-CUR, Figure 1). We also synthesized the related complex with tetrahydrocurcumin (Ir-THC, Figure 1), metal complexes of which have apparently never been reported.

Results and Discussion

The target Ir^{III} complexes (Figure 1) are structurally similar to those of the type $[\text{Ir}(\text{NC}_2(\text{OO}))]$, where OO is an anionic β -diketonate ligand, of which acetylacetonate (acac^-) is the archetypal example.^[16] They were synthesized by reaction of $[\text{Ir}(\text{NC-ppy})_2(\mu\text{-Cl})_2]$ with the commercial curcumin or tetrahydrocurcumin proligand in methanol in the presence of sodium methoxide to deprotonate the β -diketone (Scheme S1 in the Support-

ing Information). The desired products were purified by recrystallization from a mixture of dichloromethane and diethyl ether. The identities and purities of the products were confirmed by ^1H and ^{13}C NMR spectroscopies and by elemental analysis (details are provided in the Supporting Information). The UV/Vis spectra of the two complexes show moderately intense absorption bands extending into the visible region due to charge-transfer transitions, along with π - π^* bands in the UV region (Figure S1, Supporting Information). Stability measurements under physiological conditions, both of the ligands and of the iridium complexes, have been performed by monitoring their UV/Vis spectra (Figure S2, Supporting Information). The results indicated that CUR is not stable under physiological conditions, whereas THC and both of the new iridium complexes are rather stable.

At room temperature, the Ir-CUR complex shows no detectable luminescence in deoxygenated CH_2Cl_2 solution. This lack of luminescence would be consistent with a *trans-cis* isomerization of the olefinic double bond acting as a deactivation pathway, as previously suggested for Ir^{III} complexes incorporating C=C-appended ligands.^[28,29] It may also be due to the CUR ligand being associated with low-energy π^* orbitals, such that the lowest-energy triplet excited state becomes localized on this ligand as opposed to it having the usual $d_{\text{Ir}}/\pi_{\text{ppy}} \rightarrow \pi^*_{\text{ppy}}$ charge-transfer character that would typically lead to efficient phosphorescence. Such an effect is well established for Ir^{III} complexes with OO ligands featuring more extended conjugation, such as dibenzoylmethane, which have low-energy ligand-based triplet states.^[30] A similar effect has been observed for related platinum(II) complexes with OO ligands having low-energy triplet states.^[31] At low temperature (77 K), the deactivating pathway is inhibited: our complex emits in the red region of the spectrum, displaying a highly structured profile, $\lambda_{(0,0)} = 602$ nm, and a lifetime of 5.9 μs (spectra are shown in Figure S1, Supporting Information).

In contrast, the related Ir-THC complex, which lacks olefinic bonds in the OO ligand, is an extraordinarily bright-green emitter at room temperature ($\lambda_{\text{max}} = 520$ nm). The luminescence lifetime is 1.8 μs in deoxygenated CH_2Cl_2 solution, and its photoluminescence quantum yield of 0.90 makes it one of the most brightly emitting Ir^{III} complexes reported, comparable to that of the archetypal complex *fac*- $\text{Ir}(\text{ppy})_3$.^[32] Assuming that the emitting state is formed with unitary efficiency upon light absorption, the radiative rate constant k_r and non-radiative rate constant k_{nr} can be estimated from $k_r = \Phi/\tau$ and $k_{\text{nr}} = (1-\Phi)/\tau$. The high value of k_r of $5 \times 10^5 \text{ s}^{-1}$ coupled with the relatively low value of $5.5 \times 10^4 \text{ s}^{-1}$ for k_{nr} account for the bright emission. This contrasts strikingly with THC itself, which is only very weakly emissive in ethanol solution at room temperature ($\Phi = 0.007$).^[33]

The intense green phosphorescence of Ir-THC in deoxygenated solution is strongly quenched by oxygen. The emission lifetime is reduced to 70 ns in air-equilibrated solution, from which the bimolecular rate constant for quenching by O_2 can be roughly calculated as $6.2 \times 10^9 \text{ M}^{-1} \text{ s}^{-1}$. Such quenching of transition metal complexes is often accompanied by formation of the excited $^1\Delta_g$ state of oxygen, commonly referred to as

singlet oxygen or $^1\text{O}_2$. In the present instance, the emission of the $^1\text{O}_2$ thereby generated could be readily detected in the near-infrared (NIR) region (1274 nm; Figure S1, Supporting Information). Using a previously described procedure^[34] with perinaphthenone (also known as phenalenone) as the standard ($\Phi = 0.95$ ^[35]), the quantum yield of $^1\text{O}_2$ generation was estimated to be 0.42 in CH_2Cl_2 . Further details and representative plots are provided in the Supporting Information. As noted above, some phosphorescence was still observed in air-equilibrated solution, and so the singlet oxygen quantum yield under these conditions was somewhat lower than that of phosphorescence in the deoxygenated solution. The rate of deactivation of the Ir-THC excited state by oxygen can be estimated from the lifetimes to be around $1.3 \times 10^7 \text{ s}^{-1}$ in air-equilibrated solution, which, though larger than the value of k_f given above, still allows some phosphorescence to be observed. No singlet oxygen was detected for Ir-CUR, probably because the rate of deactivation of its triplet state is too fast for it to act as a sensitizer.

The cellular uptake of the new iridium(III) complexes was probed by emission microscopy. Both complexes are characterized by high cell permeability and fast internalization kinetics in A549 (a human alveolar basal epithelial cancer cell line) and in cervical cancer (HeLa) cells (Figure 2a and Figure S3, Sup-

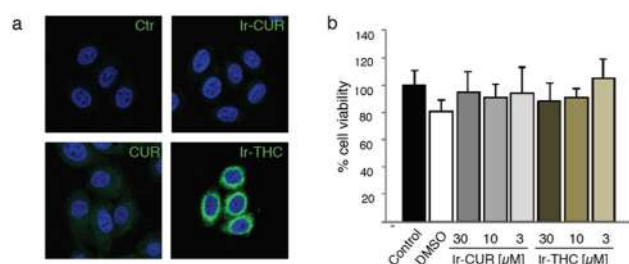


Figure 2. (a) Immunofluorescence analysis of HeLa cells incubated for 2 h with the indicated compounds at $30 \mu\text{M}$ concentration and co-stained with the nucleus-specific dye DAPI. Images were acquired by confocal microscopy with an excitation wavelength of 495 nm. Emission from the complexes is green; DAPI is blue and localized in the nuclei. (b) MTT assays of A549 cells continuously treated for 2 days with the indicated concentrations of Ir-CUR and Ir-THC without irradiation. DMSO indicates cells treated with 0.15% DMSO as for the $30 \mu\text{M}$ Ir-CUR and Ir-THC. The cell viabilities (%) of the various samples are not statistically different.

porting Information). Whereas Ir-CUR was barely detectable (as was also the case for controls treated with CUR or THC), the emission from the cells treated with the Ir-THC complex was extremely bright (Figure 2a). It was visible even after incubation at a dose of $0.3 \mu\text{M}$ and was localized inside the cells in the mitochondria, as confirmed by co-staining experiments with MitoTracker[®] dyes (Figure S3, Supporting Information).

It is notable that the cells appeared to be unaffected by the presence of the two complexes. This observation was further confirmed by an MTT assay performed at various concentrations of the complexes: no reduction in cell viability was evident, even after 48 h of continuous treatment at a concentration of $30 \mu\text{M}$ under normal cell culture conditions (Figure 2b). These results were further confirmed by a cell-counting kit

(CCK) assay (Figure S3c, Supporting Information). The photocytotoxicity of the Ir-THC complex was assessed through irradiation at different wavelengths coupled with time-lapse microscopy to view the effect of light on the cells. The cells were irradiated with light from a xenon lamp through three different excitation filters (360–370 nm, 465–495 nm, and 530–560 nm) and for different durations (0.5 s, 1 s, and 2 s) and were subsequently monitored over a period of 48 h. As can be seen in Figure S3, A549 cells treated with $30 \mu\text{M}$ Ir-THC for 2 h in the incubator, and irradiated in the wavelength range 465–495 nm for 2 s (at a power of 5.6 mW mm^{-2}), underwent apoptosis within less than 2 h. The use of light in the 530–560 nm region had an intermediate effect in initiating apoptosis, consistent with the limited absorption of the complex at such long wavelengths.

Based on previous observations, we hypothesize that $^1\text{O}_2$ or related reactive oxygen species generated upon excitation of the compound are responsible for the observed photocytotoxicity. The contrast between the light-induced cytotoxicity and the lack of toxicity in the absence of light is of crucial importance for the design of PDT agents.

To compare Ir-CUR and Ir-THC and confirm the specificity of the cytotoxicity, we performed a series of experiments in which A549 cells were treated with the two complexes and the free pro-ligands CUR and THC at a concentration of $30 \mu\text{M}$ in each case (Figure 3).

The Ir-CUR complex under irradiation showed much reduced activity compared to Ir-THC, an observation that is consistent with the earlier conclusion of rapid deactivation of the excited state. It is remarkable to note that, for cells treated with the Ir complexes, neighbouring cells in the culture plates that were not subject to irradiation were perfectly viable. Free CUR

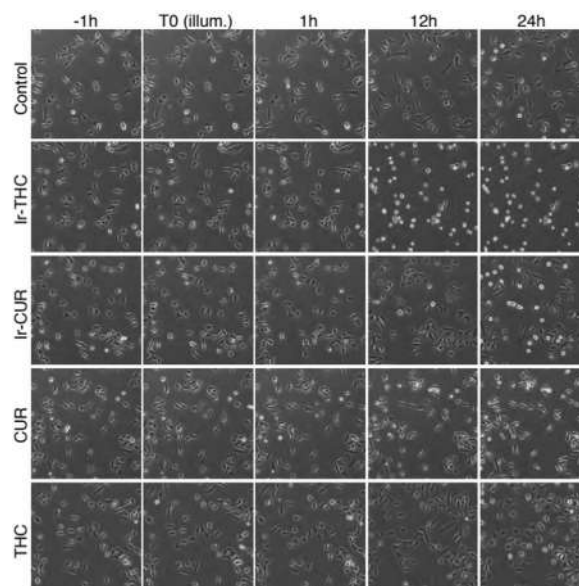


Figure 3. Sample images of the time-lapse experiment performed on A549 cells treated with Ir-THC for 2 h and the indicated compounds at $30 \mu\text{M}$ concentration before irradiation and recording. Cells were irradiated for 2 s in the wavelength range 465–495 nm (5.6 mW mm^{-2}) and this time point is considered as T0.

showed limited phototoxicity, whilst free THC did not show any such activity at all, demonstrating the fundamental role played by coordination to the iridium(III) centre for the phototoxicity.

Subsequent experiments were conducted with Ir-THC at a range of lower concentrations. Results indicated that a dose of 10 μM was sufficient to kill all of the treated cells when irradiated for 2 s (Figure 4a). In order to assess the broader applicability to other cell lines, analogous phototoxicity experiments were carried out on HeLa cells and human brain glioblastoma T89G cells, which were treated with 4 and 8 μM Ir-THC. No significant differences were discernible between the various cell lines, though the A549 cells appeared to be more resistant at 4 μM concentration (Figure 4b).

The LD₅₀ (half lethal dose concentration) of the Ir-THC complex towards HeLa and A549 cells was calculated with the set parameters (2 h pre-incubation, 465–495 nm, 2 s irradiation) using the CCK assay. Ir-THC gave LD₅₀ values of 5–7 μM . We also tested the effect of longer irradiation of cells treated with a lower concentration of complex. Strikingly, phototoxicity was discernible even with 1 μM Ir-THC when cells were irradiated for 20 s (Figure S5, Supporting Information). These results are exceptionally promising for photodynamic therapy, since the phototoxicity index (PI) of the complexes is remarkably high (>200) and the required time of light exposure is very short.

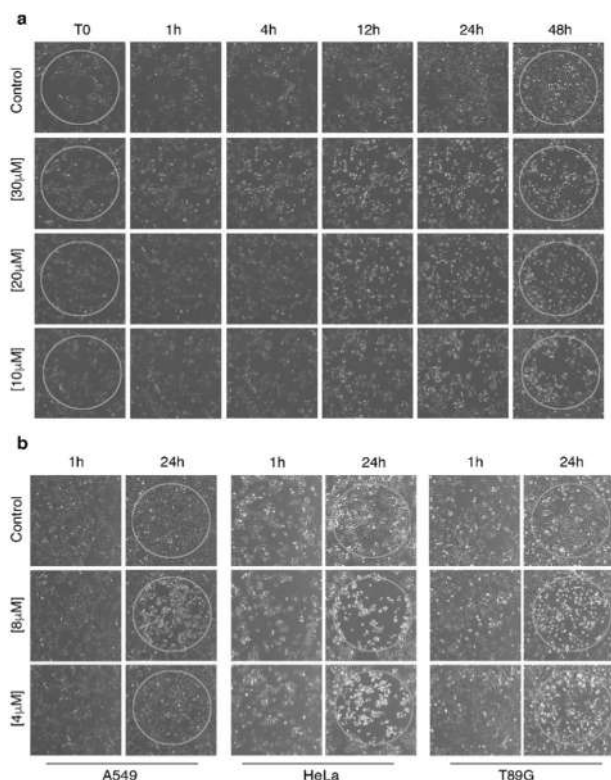


Figure 4. (a) Sample images of the time-lapse experiment performed on A549 cells treated with Ir-THC at the concentrations indicated. Circles indicate the irradiated region of the wells. (b) Comparison of images obtained after 1 and 24 h treatment of A549, HeLa, and T89G cells at 4 and 8 μM . Cells were irradiated for 2 s in the wavelength range 465–495 nm (5.6 mWmm^{-2}).

Conclusions

The new cyclometallated Ir^{III} complex with a tetrahydrocurcumin ligand, Ir-THC, appears to be a very promising tool for combined photodynamic therapy and cellular imaging. It has several desirable properties for these purposes: (i) it is highly phosphorescent and soluble; (ii) it easily enters into cells, in which it is concentrated in mitochondria, even at limited concentration; (iii) it is highly phototoxic upon irradiation with suitable light, but not cytotoxic in the dark. These results highlight how photoactive metal complexes of curcuminoids may offer interesting future potential for phototherapeutic applications in combination with emission imaging microscopy.

Acknowledgements

A.C. thanks the Università degli Studi di Milano (Piano Sostegno alla Ricerca 2015-17-LINEA 2 Azione A–Giovani Ricercatori) for financial support. Work in the laboratory of S.P. is supported by the Associazione Italiana per la Ricerca sul Cancro (AIRC, IG19875).

Conflict of interest

The authors declare no conflict of interest.

Keywords: anticancer agents • curcumin • imaging agents • iridium • natural products • photodynamic therapy • tetrahydrocurcumin

- [1] a) B. B. Aggarwal, B. Sung, *Trends Pharmacol. Sci.* **2009**, *30*, 85–94; b) P. Rinwa, B. Kaur, A. S. Jaggi, N. Singh, *Naunyn Schmiedeberg's Arch. Pharmacol.* **2010**, *381*, 529–539; c) R. A. Sharma, A. J. Gescher, W. P. Steward, *Eur. J. Cancer* **2005**, *41*, 1955–1968.
- [2] a) A. Goel, A. B. Kunnumakara, B. B. Aggarwal, *Biochem. Pharmacol.* **2008**, *75*, 787–809; b) H. Zhou, C. S. Beevers, S. Huang, *Curr. Drug Targets* **2011**, *12*, 332–347.
- [3] A. L. Cheng, C. H. Hsu, J. K. Lin, M. M. Hsu, Y. F. Ho, T. S. Shen, J. Y. Ko, J. T. Lin, B. R. Lin, W. Ming-Shiang, H. S. Yu, S. H. Jee, G. S. Chen, T. M. Chen, C. A. Chen, M. K. Lai, Y. S. Pu, M. H. Pan, Y. J. Wang, C. C. Tsai, C. Y. Hsieh, *Anticancer Res.* **2001**, *21*, 2895–2900.
- [4] a) S. Banerjee, A. R. Chakravarty, *Acc. Chem. Res.* **2015**, *48*, 2075–2083 and references therein; b) A. K. Renfrew, N. S. Bryce, T. W. Hambley, *Chem. Sci.* **2013**, *4*, 3731–3739.
- [5] a) R. A. Sharma, W. P. Steward, A. J. Gescher, *Adv. Exp. Med. Biol.* **2007**, *595*, 453–470; b) P. Anand, A. B. Kunnumakara, R. A. Newman, B. B. Aggarwal, *Mol. Pharm.* **2007**, *4*, 807–818.
- [6] F. Kühlwein, K. Polborn, W. Beck, *Z. Anorg. Allg. Chem.* **1997**, *623*, 1211–1219.
- [7] a) C. Ireson, S. Orr, D. J. Jones, R. Verschoyle, C. K. Lim, J. L. Luo, L. Howells, S. Plummer, R. Jukes, M. Williams, W. P. Steward, A. Gescher, *Cancer Res.* **2001**, *61*, 1058–1064; b) M. H. Pan, T. M. Huang, J. K. Lin, *Drug Metab. Dispos.* **1999**, *27*, 486–494.
- [8] N. Kang, M.-M. Wang, Y.-H. Wang, Z.-N. Zhang, H.-R. Cao, Y.-H. Lv, Y. Yang, P.-H. Fan, F. Qiu, X.-M. Gao, *Food Chem. Toxicol.* **2014**, *67*, 193–200.
- [9] K. I. Priyadarisni, *J. Photochem. Photobiol. C* **2009**, *10*, 81–95.
- [10] A. E. O'Connor, W. M. Gallagher, A. T. Byrne, *Photochem. Photobiol.* **2009**, *85*, 1053–1074.
- [11] a) L. K. McKenzie, I. V. Sazanovich, E. Baggaley, M. Bonneau, V. Guerschais, J. A. G. Williams, J. A. Weinstein, H. E. Bryant, *Chem. Eur. J.* **2017**, *23*,

- 234–238; b) L. K. McKenzie, H. E. Bryant, J. A. Weinstein, *Coord. Chem. Rev.* **2019**, *379*, 2–29.
- [12] R. Pettinari, F. Marchetti, C. Pettinari, F. Condello, A. Petrini, R. Scopelliti, T. Riedel, P. J. Dyson, *Dalton Trans.* **2015**, *44*, 20523–20531.
- [13] I. P. Djurovich, D. Murphy, M. E. Thompson, B. Hernandez, R. Gao, P. L. Hunt, M. Selke, *Dalton Trans.* **2007**, 3763–3770.
- [14] a) C. Dragonetti, L. Falciola, P. Mussini, S. Righetto, D. Roberto, R. Ugo, A. Valore, F. De Angelis, S. Fantacci, A. Sgamellotti, M. Ramon, M. Muccini, *Inorg. Chem.* **2007**, *46*, 8533–8547; b) C. Dragonetti, S. Righetto, D. Roberto, R. Ugo, A. Valore, F. Demartin, F. De Angelis, A. Sgamellotti, S. Fantacci, *Inorg. Chim. Acta* **2008**, *361*, 4070–4076; c) E. Margapoti, V. Shukla, A. Valore, A. Sharma, C. Dragonetti, C. C. Kitts, D. Roberto, M. Murgia, R. Ugo, M. Muccini, *J. Phys. Chem. C* **2009**, *113*, 12517–12522; d) E. Margapoti, M. Muccini, A. Sharma, A. Colombo, C. Dragonetti, D. Roberto, A. Valore, *Dalton Trans.* **2012**, *41*, 9227–9231; e) A. Colombo, F. Fiorini, D. Septiadi, C. Dragonetti, F. Nisic, A. Valore, D. Roberto, M. Mauro, L. De Cola, *Dalton Trans.* **2015**, *44*, 8478–8487.
- [15] T. Huang, Q. Yu, S. Liu, W. Huang, Q. Zhao, *Dalton Trans.* **2018**, *47*, 7628–7633.
- [16] S. Lamansky, P. Djurovich, D. Murphy, F. Abdel-Razzaq, R. Kwong, I. Tsyba, M. Bortz, B. Mui, R. Bau, M. E. Thompson, *Inorg. Chem.* **2001**, *40*, 1704–1711.
- [17] Z. Liu, P. J. Sadler, *Acc. Chem. Res.* **2014**, *47*, 1174–1185.
- [18] L. He, Y. Li, C.-P. Tan, R.-R. Ye, M.-H. Chen, J.-J. Cao, L.-N. Ji, Z.-W. Mao, *Chem. Sci.* **2015**, *6*, 5409–5418.
- [19] K. K.-W. Lo, K. K.-S. Tso, *Inorg. Chem. Front.* **2015**, *2*, 510–524.
- [20] K. K.-W. Lo, *Acc. Chem. Res.* **2015**, *48*, 2985–2995 and references therein.
- [21] Y. Li, B. Liu, X.-R. Lu, M.-F. Li, L.-N. Ji, Z.-W. Mao, *Dalton Trans.* **2017**, *46*, 11363–11371.
- [22] J. Yellol, S. A. Pérez, G. Yellol, J. Zajac, A. Donaire, G. Viguera, V. Novohradsky, C. Janiak, V. Brabec, J. Ruiz, *Chem. Commun.* **2016**, *52*, 14165–14168.
- [23] N. Wu, J.-J. Cao, X.-W. Wu, C.-P. Tan, L.-N. Ji, Z.-W. Mao, *Dalton Trans.* **2017**, *46*, 13482–13491.
- [24] T. S.-M. Tang, H.-W. Liu, K. K.-W. Lo, *Chem. Commun.* **2017**, *53*, 3299–3302.
- [25] A. M.-H. Yip, K. K.-W. Lo, *Coord. Chem. Rev.* **2018**, *361*, 138–163 and references therein.
- [26] B. Liu, S. Monro, L. Lystrom, C. G. Cameron, K. Colon, H. Yin, S. Kilina, S. A. McFarland, W. Sun, *Inorg. Chem.* **2018**, *57*, 9859–9872.
- [27] A. Zamora, G. Viguera, V. Rodriguez, M. D. Santana, J. Ruiz, *Coord. Chem. Rev.* **2018**, *360*, 34–76 and references therein.
- [28] A. Colombo, C. Dragonetti, D. Roberto, A. Valore, C. Ferrante, I. Fortunati, A. L. Picone, F. Todescato, J. A. G. Williams, *Dalton Trans.* **2015**, *44*, 15712–15720.
- [29] V. Aubert, L. Ordronneau, M. Escadeillas, J. A. G. Williams, A. Boucekine, E. Coulaud, C. Dragonetti, S. Righetto, D. Roberto, R. Ugo, A. Valore, A. Singh, J. Zyss, I. Ledoux-Rak, H. Le Bozec, V. Guerschais, *Inorg. Chem.* **2011**, *50*, 5027–5038.
- [30] a) S. Lamansky, P. Djurovich, D. Murphy, F. Abdel-Razzaq, H. E. Lee, C. Adachi, P. E. Burrows, S. R. Forrest, M. E. Thompson, *J. Am. Chem. Soc.* **2001**, *123*, 4304–4312; b) Q. Zhao, L. Fi, F. Li, M. Yu, Z. Liu, T. Yi, C. Huang, *Chem. Commun.* **2008**, 685–687.
- [31] M. Spencer, A. Santoro, G. R. Freeman, A. Díez, P. R. Murray, J. Torroba, A. C. Whitwood, L. J. Yellowlees, J. A. G. Williams, D. W. Bruce, *Dalton Trans.* **2012**, *41*, 14244–14256.
- [32] T. Sajoto, P. I. Djurovich, A. B. Tamayo, J. Oxgaard, W. A. Goddard, M. E. Thompson, *J. Am. Chem. Soc.* **2009**, *131*, 9813–9822.
- [33] A. Castellán, R. Ruggiero, L. G. da Silva, E. Portes, S. Grelier, C. Gardrat, *J. Photochem. Photobiol. A* **2007**, *190*, 110–120.
- [34] D. L. Rochester, S. Develay, S. Zális, J. A. G. Williams, *Dalton Trans.* **2009**, 1728–1741.
- [35] R. Schmidt, C. Tanielan, R. Dunsbach, C. Wolff, *J. Photochem. Photobiol. A* **1994**, *79*, 11–17.

 Manuscript received: April 2, 2019

Accepted manuscript online: April 15, 2019

Version of record online: May 13, 2019