Number of inadvertent RNA targets for morpholino knockdown in *Danio rerio* is largely underestimated: evidence from the study of Ser/Arg-rich splicing factors

Marine Joris\(^1\), Marie Schloesser\(^1\), Denis Baurain\(^2\), Marc Hanikenne\(^1\), Marc Muller\(^3\) and Patrick Motte\(^1\),\(^\ast\)

\(^1\)Laboratory of Functional Genomics and Plant Molecular Imaging, InBioS, PhytoSystems and Centre for Assistance in Technology of Microscopy (CAREm), University of Liège, 4000 Liège, Belgium, \(^2\)InBioS–PhytoSYSTEMS, Eukaryotic Phylogenomics, University of Liège, 4000 Liège, Belgium and \(^3\)Laboratory for Organogenesis and Regeneration, GIGA-Research, University of Liège, 4000 Liège, Belgium

Received February 16, 2017; Revised July 10, 2017; Editorial Decision July 11, 2017; Accepted July 13, 2017

ABSTRACT

Although the involvement of Ser/Arg-rich (SR) proteins in RNA metabolism is well documented, their role in vertebrate development remains elusive. We, therefore, elected to take advantage of the zebrafish model organism to study the SR genes’ functions using the splicing morpholino (sMO) microinjection and the programmable site-specific nucleases. Consistent with previous research, we revealed discrepancies between the mutant and morphant phenotypes and we show that these inconsistencies may result from a large number of unsuspected inadvertent morpholino RNA targets. While microinjection of MOs directed against *srsf5a* (sMO*srsf5a*) led to developmental defects, the corresponding homozygous mutants did not display any phenotypic traits. Furthermore, microinjection of sMO*srsf5a* into *srsf5a*\(^−/−\) led to the previously observed morphant phenotype. Similar findings were observed for other SR genes. sMO*srsf5a* alternative target genes were identified using deep mRNA sequencing. We uncovered that only 11 consecutive bases complementary to sMO*srsf5a* are sufficient for binding and subsequent blocking of splice sites. In addition, we observed that sMO*srsf5a* secondary targets can be reduced by increasing embryos growth temperature after microinjection. Our data contribute to the debate about MO specificity, efficacy and the number of unknown targeted sequences.

INTRODUCTION

SR proteins constitute a phylogenetically conserved family of RNA-binding proteins that are involved in many aspects of (pre-)mRNA metabolism. First described as constitutive and alternative splicing regulators, they are also implicated in transcription, non-sense-mediated decay, mRNA export, translational control as well as maintenance of genome stability (1–3). In order to clarify identification of SR splicing factors, a new definition has been proposed considering them as any protein with a modular structure consisting of one or two RNA recognition motif (RRM) at the N-terminus and a C-terminal Ser/Arg-rich domain (RS) of at least 50 amino acids with > 40% RS or SR repeats (4). Based on their architecture, SR proteins can be divided into three subfamilies in vertebrates; (i) SRSF1-like (one RRM and one pseudo-RRM), (ii) SRSF2-like (one RRM) and (iii) SRSF7-like (one RRM and one ‘zincknuckle’). SR proteins have been shown to bind exonic splicing enhancers (ESEs) sequences on pre-mRNAs via their RRM and to recruit spliceosomal components via their RS domain. SR proteins can function either as splicing activators (exon inclusion) or negative regulators (exon skipping), depending on the context (5,6). While some SR proteins seem to bind a small set of endogenous-specific targets, others share many binding sites suggesting a close collaboration between these proteins in mRNA metabolism (7,8). Whereas their functions at the molecular level are well documented, their roles in a physiological and developmental context are incompletely understood (9). The difficulties to uncover their function in a specific model may be tied to their essential role in cell viability (i.e. apoptosis and cell cycle progression) (10). Indeed, inactivation of several SR genes leads to death in early stages of embryo development in mouse and *drosophila*
Furthermore, several SR genes have redundant developmental functions in *Caenorhabditis elegans* and their inactivation had little or no effects on *Caenorhabditis* development (13). To address this issue, we evaluated the *Danio rerio* model system. Antisense morpholino oligonucleotide microinjections permitted to obtain graded phenotypes (14) and genome editing was concomitantly used to generate stable knockout lines.

Morpholinos (MOs) consist in chemically modified short oligonucleotides (25 nt) that are able to bind RNA with natural complementary base pairing, while they contain a morpholine ring instead of ribose and non-ionic phosphodiimide backbones (15). This particular MO architecture makes them nuclease-resistant with low toxicity (16). Whereas translation MOs inhibit ribosome binding to mRNAs, splicing MOs bind on pre-mRNA splice junctions and sterically block their recognition by spliceosomal components. The use of MOs was rapidly recognized as a powerful tool to investigate gene function during early embryonic development in a variety of organism including *Xenopus tropicalis*, *Mus musculus* and *Gallus gallus* (17–20). In zebrafish, MOs have been the most advocated technique to knockdown genes for the past decades (21). Recently, the emergence of new gene editing techniques like TALEN and CRISPR/Cas9 (22–25) reversed this trend and enabled researchers to rapidly and easily generate fish knockouts. The extensive generation of mutant lines has revealed discrepancies between mutant and morphant phenotypes, suggesting MOs off-target effects (26,27). It has been shown that these discrepancies can be explained by compensation mechanisms in knockout mutants (28).

Fifteen zebrafish SR genes were identified belonging to the three subfamilies of SR proteins (Supplementary Figure S1). In this study, we investigate the role of the splicing factor Srsf5a, displaying a specific expression profile and belonging to the extensively studied SRSF1 family. We observed that, while the microinjected MOs directed against srsf5a (sMOsrsf5a) led to developmental defects, the corresponding homozygous mutants did not display any phenotypic traits. Furthermore, microinjection of sMOsrsf5a into srsf5a−/− resulted in a phenotype as observed in the morphant. Similar findings were observed for two other SR genes (srsf2b and srsf9). Using deep mRNA sequencing, we found that these inconsistencies were likely due to several sMOsrsf5a alternative target genes which could not be easily predicted by straightforward blast analysis. We also uncovered that only 11 consecutive bases complementary to the 25 MO bases are sufficient for binding and subsequent blocking of splice sites. In addition, we observed that sMOsrsf5a secondary targets can be slightly reduced by increasing the growth temperature of embryos after microinjection. Our data contribute to the debate about MO specificity, efficacy and the number of unknown targeted sequences (14,29,30). Our findings also pave the way for more efficient identification of putative functional collateral MO target sites and moreover raise the question whether the experimental conditions used to perform MO, CRISPR/Cas9 or TALEN experiments could be optimized.

**MATERIALS AND METHODS**

### Zebrafish breeding

Zebrafish (*D. rerio*) were raised and staged according to standard protocols (31). Embryos were kept in E3 medium with or without 0.003% 1-phenyl-2-thiourea at 28°C until harvested. They were fixed for 2 h in 4% paraformaldehyde and dehydrated in 100% methanol before analysis. All experiments and the entire study were evaluated by the Ethical Committee of the University of Liege, Belgium and accepted under the file number 1158.

### Plasmids

Total RNA from 48 hpf embryos was extracted using TRIzol reagent (SigmaAldrich) and the RNasy mini Kit (Qiagen), and reverse-transcribed using the RevertAid First Strand cDNA Synthesis Kit (ThermoFisher Scientific). The coding sequences: srsf5a, srsf5aΔ36, srsf5a-int3 and mut-srsf5a-int3 were amplified from cDNAs (synthesized from un.injected, or sMOsrsf5a injected wt and mutant embryos) by polymerase chain reaction (PCR) and cloned into pCS2plus Vector (Addgene) or in the pGEM®-T Easy Vector (Promega) for riboprobe synthesis. Primers used were listed in the Supplementary Table S6.

### In vitro transcription

mRNA for injection were synthesized from pCS2plus after linearization by NotI using the mMESSAGEmACHINE SP6 Kit (Ambion) and were purified by LiCl precipitation. Riboprobes for *in situ* hybridization were synthesized from pGEM®-T Easy Vector using SP6 or T7 in a digoxigenin or DNP (2,4-dinitrophenol) labeling reaction. The zebrafish *isl-1* (ZDB-GENE-980526–112) and *pax6b* (ZDB-GENE-001031–1) probes were used.

### In situ hybridization, H/E blade staining and image acquisition

*In situ* hybridization were performed as described (32). Anti-digoxigenin-HRP and anti-DNP-HRP were used with tyramide-Cy3 (Red) and tyramide-FITC (green) (Perkin-Elmer TSA Kit). Embryo cell nuclei were sometimes counterstained with 0.6 μM DRAQ7®. For H/E experiments, embryos were dehydrated in ethanol and embedded in a glycol methacrylate medium (Technovit® 7100). The 8-μm sections were obtained using a microtome (Leica) and were stained with hematoxylin and eosin according to standard protocols. Pictures were taken on a Nikon® Eclipse 90i microscope controlled by NIS-Elements microscope imaging software. Fluorescent images were acquired with a Leica® SP5 inverted confocal microscope.

### Microinjections

Embryos were injected at the one cell stage into the yolk. Morpholinos were purchased from Gene Tools (Philomath, OR, USA). The sequences of the sMOsrsf5a, sMOsrsf2b; sMOsrsf9 and the CtrlMO are listed in Supplementary Table S6. MOs were injected at 3 ng for srsf5a, 2 ng for srsf2b...
and for srsf9. The severity of the observed phenotype was MO dose dependent. The dose used in our experiments was the lowest one leading to the supposedly specific phenotype in 75% of embryos. For rescue or expression experiments, 50–200 pg of coding mRNA was injected. Morpholino and mRNA were each dissolved in 1× Danieau buffer containing 0.5% Tetramethylrhodamine dextran (Invitrogen, Belgium) to be able to sort well injected embryos under a fluorescent binocular. On developmental defect severity, MOs injected embryos were divided into three main classes; weak, intermediate and strong. The intermediate phenotype was the most frequently observed (in 75% of embryos, n > 450) and was systematically used for further analysis. Each injection experiment was repeated at least three times using at least 150 individuals per experimental condition. No phenotypic differences were observed in smOsrsf5a morphants, in absence or presence of co-injection of a morpholino directed against p53 (33).

## Western blot analysis

Proteins were extracted from 80 injected embryos (24 hpf). Samples were dissolved in Laemmli buffer before gel electrophoresis. A mouse monoclonal ANTI-FLAG® M2-Peroxidase (HRP) antibody (Sigma A8592) was used for detecting Flag-tagged protein using the BM chemiluminescence Western blot kit (Roche).

## Genome editing

TALENs directed against srsf5a were designed using the TALE-NT software (https://tale-nt.cac.cornell.edu/) using the following criteria; spacer length: 15–20 bases, RVDs length: 15–20 bp. For each TALEN binding site, predicted sequences were chosen for no off-target prediction (Supplementary Table S6). TALEN assembly in the right or left genome editing was conducted using the FastTALE™ TALEN assembly kit from Sidansai Biotechnology. To produce TALEN mRNAs, left and right TALEN expression vectors were linearized by NotI and microinjected as previously described. CRISPR directed against p53 was the most frequently observed (in 75% of embryos, n > 450) and was systematically used for further analysis. Each injection experiment was repeated at least three times using at least 150 individuals per experimental condition. No phenotypic differences were observed in smOsrsf5a morphants, in absence or presence of co-injection of a morpholino directed against p53 (33).

## qPCR

Quantitative RT-PCR reactions were performed in 384-well plates with an ABI Prism 7900HT system (Applied Biosystems) using Maxima SYBR® Green qPCR Master Mix (Eurogentec) on material from three independent biological experiments (Figure 2B and Supplementary Figure S5), or from three independent mutant lines (Figure 2E) for each combination of cDNA and primer pair. Gene expression was normalized relative to ef1alpha and rpl13alpha. Supplementary Table S7 shows the primers used for these experiments.

## RNA sequencing

Total RNA from 48 hpf control and sMOsrsf5a embryos was extracted as described above. RNA sample quality and quantity were evaluated using both Bioanalyzer 2100 (Agilent) and Nanodrop. Libraries were prepared using the TruSeq kit (Illumina) and cDNA fragments ranging from 300 to 500 bp were selected and sequenced on a HiSeq2000 platform (Illumina) to generate 101-bp paired-end FASTQ sequences. Reads were mapped on the zebrafish genome (Zv9, Ensembl genome version 75, ensembl.org) using Tophat v.2.0.9 (-r 200, -p 1, -m 2, -solexa1.3-quals) (36). Gene expression was measured from the mapped reads by using HT-seq-count (intersection-strict mode) and differentially expressed (DE) genes were determined using the R package DESeq2 (37–39). MATS version 3.0.9 was used to detect differential alternative splicing events (-t paired -len 101 -a 8 -r1 178,176 -r2 183,195 -sd1 60,60 –sd2 61,67 – analysis P, a gtf file coming from the Zv9 version of the genome was also given to MATS) (40). Insert sizes were calculated using Picard Tools. Data are shown in Supplementary Tables S1 and 2. RNA-Seq data are submitted to GEO database under accession number GSE98888.
against Zebrafish GRChz10 (Genomic sequence) using the following parameters: Word size for seeding alignments: 11, Match/Mismatch scores: -1,-2. The sMOsrsf5a sequence was used as a query. Data are shown in Supplementary Tables S3–5.

RT-PCR
To visualize differential splicing due to sMOsrsf5a microinjection, total RNAs were extracted from 48 hpf injected embryos at 28°C (at 33°C the embryos were stopped at 38.5 hpf, the equivalent of the 48 hpf stage) and were reverse transcribed using oligo(dT)18. Specific primers were used to amplify a region of interest into blast predicted target genes (Supplementary Table S7).

Electrophoretic mobility shift assays
Synthetic 77 nt RNAs (see Supplementary Methods and Supplementary Table S6) were incubated at 90°C for 1 min and quickly chilled on ice for 2 min. A total of 4 pmole of sMOsrsf5a were incubated with various amount of RNAs (0.6, 1.2, 1.8, 2.4, 2.8 pmole) in 10 μl of 1× TMN buffer (20 mM Tris-Acetate, pH 7.6, 100 mM NaOAc, 5 mM Mg(OAc)2) at 28°C for 15 min. Two microliters of 6× non-denaturing RNA loading buffer were added and the samples were run on a 11% non-denaturing polyacrylamide gel in 1× TBE at 90 V for 2h. The gel was stained with GelRed™ (10 000×) for 10 min and visualized under UV.

RESULTS
Microinjection of MOs directed against srsf5a led to developmental defects
We first determined the expression profile of srsf5a by in situ hybridization (Supplementary Figure S2). Like many D. rerio SR genes (data not shown), srsf5a is maternally expressed and becomes widely expressed at later stages (24, 48 and 72 hpf) in the central nervous system and in the pharyngeal region. Noticeably, srsf5a displayed a more distinct expression in the developing eye, particularly in the ganglional cell layer and in the inner nuclear layer of the retina. These data led us to hypothesize that srsf5a could have a role in cell differentiation during eye formation.

To examine the role of Srsf5a during early embryonic development, we injected at one cell stage a splice site blocking morpholino (sMOsrsf5a) targeting the exon3-intron3 junction of the srsf5a pre-mRNA (Figure 1A). At this dose, 95% (n > 450) of srsf5a morphant embryos suffered from developmental delay and poor swimming activity. We observed opaque areas in the brain and in the eyes, indicative of necrotic zones, reduction of pigmentation (visible at 48 hpf) and often bent tails and pericardial edemas (in 30% of embryos) (Figure 1B). More thorough analysis revealed cell death in the brain and severe defects in the developing eye of morphants (Figure 1C and D). The apparent specificity of these defects for srsf5a knock-down was tested by coinjecting the sMOsrsf5a with 80 pg of Srsf5a-coding mRNA which resulted in partial rescue of the embryonic morphant defects (Figure 1B and D). Remarkably, the srsf5a expression pattern largely coincided with the MO-induced phenotype.

Analysis of the sMOsrsf5a impact on srsf5a transcripts confirmed disruption of splicing at the targeted junction, causing retention of intron3 (srsf5a-i3/153nt) and consequently leading to the introduction of a premature STOP codon within the RRM1-encoding sequence. Moreover, it revealed a cryptic splice site located within exon3, inducing a deletion of 36 bases in the srsf5a open reading frame (srsf5a/Δ36) (Figure 1A). We therefore tested whether this splicing variant lead to synthesis of a mutant protein with a 12-amino acid deletion within the RRM1 domain by injecting a srsf5a/Δ36-Flag RNA. We did not observe any phenotypic consequences following srsf5a/Δ36-Flag mRNA injection, in contrast to injection of wt srsf5a-Flag mRNA (used as a control for overexpression) which caused drastic effects on embryonic development (Supplementary Figure S3) suggesting that srsf5a/Δ36-Flag RNA is inefficient at causing a defect and at leading to the synthesis of a mutant protein able to replace the wt protein function. At this stage of the study, our data strongly suggested that MO microinjection leads to partial inactivation of Srsf5a, leading to specific developmental defects.

In addition to the sMOsrsf5a, four other MOs were used to knockdown srsf5a (trMOsrsf5a, sMOsrsf5a2, sMOsrsf5a3 and sMOsrsf5a4; see Supplementary Method) but none of these could recapitulate the sMOsrsf5a effects. No morphological defect was observed upon injection of trMOsrsf5a, sMOsrsf5a2 and sMOsrsf5a3, possibly because in all cases no difference in wt srsf5a mRNA levels could be detected. However, injection of sMOsrsf5a4 or coinjection of sMOsrsf5a3 and sMOsrsf5a4 led to brain necrosis, a high mortality rate and a decrease of wt srsf5a mRNA level (Supplementary Figure S4). Following these discrepancies between the effects of the different used MOs herein, we decided to generate stable knockout mutant lines by using TALEN (see below).

srsf5a−/− mutants failed to recapitulate the MOs induced phenotype
To concomitantly investigate the phenotype of stable srsf5a loss-of-function, we engineered two TALEN pairs and generated two srsf5a mutants in exon 2 (deletion of 11 nucleotides, Δ11) and exon 4 (deletion of 5 nucleotides, Δ5), respectively. In both cases, the mutation led to synthesis of a truncated protein theoretically missing domains critical for their function, due to premature STOP codons (Figure 2A). Analysis of the srsf5a transcript revealed an increased degradation rate in the two mutants compared to wt, confirming the efficacy of the mutation (Figure 2B, data not shown for srsf5aΔ5). Unexpectedly, these mutants failed to recapitulate the morphant phenotype (Figure 2C); no morphological defect could be observed at any embryonic stage and we obtained normal fertile homozygous adults (2-years-old). In the light of recently raised serious concerns about the use of MOs (26,27), a similar discrepancy was recently shown to result from genetic compensation in mutants through expression of a related protein (28). We therefore tested a putative complementation of the srsf5a mutation by other SR protein-encoding genes using quantitative RT-PCR (qRT-PCR). Expression levels of the 13 zebrafish SR genes were compared between offspring ob-
Injection of a splice site blocking MO targeting the srsf5a gene led to developmental defects. (A) srsf5a is composed of eight exons. The protein is encoded by exons 2–8 and consists of two RRM domains responsible for RNA binding and one RS domain, essential for protein–protein interactions. Three different transcripts are produced from srsf5a. Among them, two alternative transcripts retain intron 5 or a part of it (black up-pointing triangle and black dot). sMOsrsf5a was designed to target the exon3–intron3 junction. RT-PCR experiments to amplify srsf5a mRNA in control and morphant embryos at 48 hpf revealed intron3 retention (open hexagon mark), introducing a premature STOP codon into the RRM1 encoding part of the mRNA. Expression of the three normal srsf5a transcripts was strongly reduced in morphants, while MO injection also triggered the use by the splicing machinery of a cryptic splice site located in exon3, and leading to a deletion of 36 bases in the open reading frame of the srsf5a transcript (srsf5a/Δ36, open triangle mark). The resulting protein has a 12-amino acids deletion within the RRM1. The open square mark corresponds to a srsf5a transcript in which intron3 is retained and with a deletion of 36 bases in the exon3 (srsf5a/Δ36-intron3). All PCR products were identified by sequencing. (B) Zebrafish embryos injected with 3 ng of ctrlMO or sMOsrsf5a with or without a rescuing dose of srsf5a mRNA (80 pg) at 48 hpf. The defects in brain, eye and curved tail could be partially rescued by srsf5a mRNA injection. Pigmentation was not visible as embryos were treated with 1-phenyl-2-thiourea to increase their transparency. Bar: 200 µm. (C) Haematoxylin/eosin sections obtained from ctrlMO and sMOsrsf5a injected embryos revealed abnormal organization of cells in the retina and an increase of cell death in the eye and the entire brain at 48 and 72 hpf (data not shown). (D) Fluorescent in situ hybridization using a pax6b probe followed by nuclear staining using draq7® revealed the disorganization of the ganglional cell layer and of the inner nuclear layer in the retina at 72 hpf in morphants compared to control embryos. Rescue experiments allowed us to partially restore the control phenotype. Scale bar: 50 µm.

Evidence of MO non-specificity

Further investigations were therefore required to understand the reasons for the contradictory results observed between morphants and mutants. We injected sMOsrsf5a in srsf5a mutant embryos and strikingly, these embryos displayed the same defects as observed in wt morphants (Supplementary Figure S6a), arguing against a mechanism where MO injection simply blocks Srsf5a expression in

tained from homozygous mutants or from wt parents. We observed an increased expression of srsf1b, srsf2b, srsf3a, srsf5b and srsf6a in 24 hpf mutant larvae and of srsf3a at 48 and 72 hpf (Figure 2B and Supplementary Figure S5). In srsf5a homozygous mutant embryos born from an incross of heterozygous parents, no change in SR gene expression could be detected, likely due to maternally provided srsf5a mRNA (Supplementary Figure S5). In contrast, none of these SR gene overexpressions was observed in morphant embryos (Figure 2B). Taken together, these results suggest that compensatory mechanisms may be responsible for the lack of phenotype in srsf5a mutants. Interestingly, double srsf5a/srsf5b homozygous mutants were generated and did not display any developmental defect, suggesting that inactivation of only the closest relative would not be sufficient to overcome such a compensation effect. Comparison of SR gene expression levels in srsf5b−/− and wt at 24 hpf also revealed an overexpression of srsf3a, proposing a central role for this factor in compensation mechanism (Figure 2D and E).
wt. Similar results were obtained in otherwise unaffected srsf2b (one RRM) and srsf9 (one RRM and one pseudorRM) mutants; microinjection of the corresponding MOs produced the same effects in wt or mutant embryos (Supplementary Figure S7). Analysis of the srsf5a mRNA in sMOsrsf5a-injected mutant larvae revealed the presence of the previously observed srsf5a/Δ36 mRNA, and an unexpected accumulation of the srsf5a-i3/153nt variant (Supplementary Figure S6b). To test for the possibility that this latter mRNA could be responsible for the morphant phenotype, e.g. by acting as a non-coding RNA (41) or as a mutant protein by using an alternative start codon (42), we injected the srsf5a-i3/153nt-Flag mRNA into mutant eggs. No effect was observed during embryonic development (Supplementary Figure S6c) and no protein production could be observed (Supplementary Figure S6d).

MOs: a plethora of secondary targets

Taken together, our data strongly suggest that the effects caused by the sMOsrsf5a result from its action on one or several alternative targets. Therefore, we performed RNA sequencing analysis on mRNA extracted from control and morphant embryos. We identified 1006 DE genes and 378 differentially spliced (DS) transcripts (Supplementary Tables S1 and 2). First, we blasted the sMOsrsf5a sequence against the genomic sequences of DE or DS genes extracted using BiomaRt (43) and we obtained two lists resuming the most homologous sequences (E-value < 73) (Supplementary Tables S3 and 4). Next, we analyzed the position of these sequences within genes and we could point out many sequences localized on splice junctions, the best place for MOs to perturb splicing and consequently gene expression (Figure 3). In the DE gene list, 8 matching sequences are localized on exon-intron junctions whereas in blast results us-

Figure 2. srsf5a−/− and srsf5b−/− did not show any developmental defects, but presented an overexpression of several homologous SR genes. (A) Two
Figure 3. Process to determine secondary target RNAs for sMOsrf5a. CtrlMO and sMOsrf5a were injected at one cell stage. Total RNAs were extracted at 48 hpf and were processed according to the standard Illumina protocol, including TrueSeq mRNA library construction and sequencing on HiSeq 2000 (2 × 101 nt paired-end sequencing). A volcano plot summarizes RNAseq analysis in which 1006 genes were identified as statistically differentially expressed (DE) between control and morphant embryos. Right gray dots represent overexpressed genes (log2FoldChange > 0.5, padjust < 0.05), while left gray dots represent underexpressed genes in morphants (log2FoldChange < −0.5, padjust < 0.05). The bar plot recapitulates the number of differentially alternative splicing events (DS transcripts) detected by MATS when comparing control and MOsrf5a transcriptomes. Sequences from these DE genes and DS transcripts were extracted using BioMart and used as subjects in a blast analysis with the sMOsrf5a sequence as a query. The resulting lists (Supplementary Table S4) were scanned manually to find target regions localized on a splice junction. ES, Exon Skipping; A3SS, 3' Splice Site; A5SS, 5' Splice Site; IR, Intron Retention; MXE, Mutually exclusive exons.

Figure 4. RT-PCR analysis of eight associated putative unintended mRNAs. For 6 of them, RT-PCR analysis confirmed the splicing defect in sMOsrf5a mRNA, even though only 11–15 successive bases out of the sMOsrf5a 25 bases are complementary to the tested pre-mRNAs (Figure 4 and Supplementary Figure S8). RNA sequencing allowed us to identify new spliced transcripts in morphants and the majority of them consisted in ‘exon skipping events’. In many cases, these events imply exons located beside the junction containing a potential sMOsrf5a binding sequence suggesting that these events are due to morpholino inadvertent binding (Supplementary Table S4). Interestingly, we also found three homologous sequences localized inside skipped exon in morphants (RNA-seq data) indicating that sMOsrf5a can influence the splicing independently of its splice junction binding (Supplementary Table S4). To make sure that these altered splicing events are not due to the partial decrease of Srsf5a expression, we analyzed mRNAs from the sMOsrf5a3 or coinjected sMOsrf5a3/ sMOsrf5a4 that were previously shown to decrease the levels of wt srsf5a mRNAs (Supplementary Figure S4A). In three randomly selected RNAs from the above list, no altered splicing was detected (Supplementary Figure S4B).

When sMOsrf5a-injected eggs were incubated at 33°C instead of the recommended 28°C (31), this slight increase in temperature resulted in some cases in a nearly complete absence of splicing defects (Figure 4 and Supplementary Figure S8) and less pronounced developmental defects in morphants (not shown).

The binding of sMOsrf5a to these inadvertent targets was also confirmed in vitro using electrophoretic mobility shift assay (EMSA). In this test, we mix the sMOsrf5a with its putative target RNA sequences and analyze the resulting duplex formation on a polyacrylamide gel (Figure 5 and Supplementary Figure S9). We could observe duplex formation of sMOsrf5a with the unspliced srsf5a sequence, but not with a sequence mutated at every second position in the sequence. Duplexes were also observed with the wt unspliced sequences of three randomly selected inadvertent targets, and with the wt spliced srsf5a sequence.

Taken together, these observations indicate that the defects observed in sMOsrf5a morphants are mainly due
Figure 4. Injection of sMOsrsf5a disturbed splicing of secondary target genes. (A) Schematic of the ifrd1 splicing junction presenting 15 contiguous bases complementary to the morpholino. RT-PCR analysis using primers targeting exons 6 and 11 of ifrd1 confirmed intron 9 retention due to sMOsrsf5a binding. When the injected embryos were incubated at 33°C, retention of intron 9 was abolished. In srsf5a−/−, splicing is disturbed as observed for wt embryos in response to sMOsrsf5a injection. (B) Splicing of the g6pca.1 gene was also affected by morpholino binding on 12 contiguous bases. Primers used to perform RT-PCR targeted the exon 2 and 6. (C) Amplification of tfpia mRNA (from exon1 to exon5) in control and morphant embryos revealed the existence of an exon skipping event in sMOsrsf5a injected embryos showing 11 bases are sufficient for MO binding. sMOsrsf5a*, embryos were injected with 2 ng of sMOsrsf5a; sMOsrsf5a, embryos were injected with 3 ng of sMOsrsf5a.

to the binding of the 25bp MO to RNA sequences of a plethora of unrelated pre-mRNAs at 28°C rather than to the inactivation of srsf3a.

**DISCUSSION**

**SR genes regulatory network**

Herein, we attempt to study SR protein functions during zebrafish embryonic development. For this purpose, we used MO microinjection and TALENs/CRISPR gene editing techniques to generate stable knockouts. Our data revealed that injection of MOs against srsf5a, srsf2b or srsf9 led to developmental defects that could not be recapitulated in the corresponding mutants. These observations are somehow consistent with recent data showing poor correlation between morphant and mutant phenotypes (26). Similar data have been previously reported in mouse, *Drosophila* and *Arabidopsis thaliana* when comparing knockout and knockdown results (44–46). Recently, Rossi *et al.* demonstrated the implementation of a compensatory network in response to deleterious mutations performed in the zebrafish genome (28). This compensatory network was not activated upon gene knockdown, explaining why MOs led to many overt phenotypes while mutants did not.

We reasonably asked whether at least one of the fifteen *D. rerio* SR proteins could take over the Srsf5a developmental function. Quantification of SR gene expression levels in srsf5a−/− compared to control and morphant embryos revealed an induced expression of srsf1b, srsf2b, srsf3a, srsf5b and srsf6a. We could similarly observe an overexpression of srsf3a in srsf5b−/−. These findings prove the existence of a regulatory signaling cascade triggered by SRSF loss-of-function and converging on expression of other SR protein genes during embryonic development. Although we could not provide the insight into SR protein functions during early vertebrate development, our results reveal the compensatory mechanisms in mutants which have not been previously described for SR proteins. Interestingly, a recent RNA-mapping study revealed that mouse SRSF1 and SRSF2 exhibit extensive overlap in their RNA targets and that loss of RNA binding by one SR protein induces compensatory changes in RNA binding by another SR protein (7). Similarly, recent works in *Drosophila* showed that SR proteins act in a combinatorial manner to regulate splicing (i.e. exon inclusion and skipping) (6). Despite its different structural composition (one RRM versus two RRMs), Srsf3a seems to be a central regulator in compensation mechanism. Further studies will be required to identify the RNA targets of individual SR protein and to understand how different SR proteins can play cooperative (or redundant) roles in splicing in living organisms. Our findings call for a thorough investigation of the SR protein-specific interaction network regulating mRNA splicing during vertebrate development.

**MO RNA inadvertent target**

Despite possible compensatory mechanisms, we could recapitulate the morphant phenotype in stable mutant embryos (therefore lacking MO binding sites) suggesting other explanations for phenotypic discrepancies. The same observa-
Figure 5. Detection of sMOsrzf5a-RNAs interactions by EMSA. Various amounts of RNAs (from 0.6 to 2.8 pmoles) were incubated with (+sMOsrzf5a) or without (−sMOsrzf5a) 4 pmoles of morpholino. (A) Two 77-nt RNAs containing either the morpholino binding site of srzf5a (srzf5a RNA) or a mutated binding site (srzf5amut RNA) were run on an 11% native polyacrylamide gel. In case of sMOsrzf5a binding to the RNA, a shift was observed due to duplex formation (see arrow). The srzf5a RNA was bound by the morpholino (positive control) while the mutated srzf5a RNA was not (negative control). (B) sMOsrzf5a was able to bind ifrd1 RNA via its 15 bases complementary sequence. Interaction between sMOsrzf5a and the 12 contiguous bases of the g6pca.1 RNA was also assayed.

sions have been reported for the megamind mutant (26,47). Many causes for such inconsistencies have been proposed including hypomorphic alleles, maternal contribution or off-target effects. In our study, the first two hypotheses were unlikely: (i) qPCR data confirmed an increased degradation rate of srzf5a−/− transcripts suggesting the absence of Srsf5a proteins, and (ii) we used splicing MOs that only act on zygotic transcripts. MO-induced defects were likely due to off-target effects.

A major concern about the use of sequence-specific knockout and knockdown techniques is unwanted sequence binding. With the expansion of genomic and transcriptomic data, many algorithms were designed to predict associated MO or sgRNA inadvertent targets. A method called Genome-wide Unbiased Identification of Double Strand Breaks Enabled by Sequencing showed recently that blast-based methods were unable to efficiently predict real off-targets for CRISPR RNA-guided nucleases (48). Indeed, the majority of identified binding sites for sMOsrzf5a using mRNA sequencing were not (easily) detected by a simple straightforward blast analysis.

MOs are supposed to be virtually free of off-target effects achieving an exquisite sequence specificity (15). They are neutral and unable to electrostatically interact with proteins. Moreover, their calculated minimum inhibitory lengths (MIL) are supposed to contain sufficient sequence information to ensure targeting of a unique transcript. Using deep mRNA sequencing, we identified DE genes and DS transcripts in sMOsrzf5a injected embryos. Their corresponding sequences were used for downstream blast analysis to identify several potential srzf5a MO-binding sites spanning a splice junction. RT-PCR analysis detected new splice variants for DE genes, thus confirming inadvertent binding of sMOsrzf5a. New alternative splicing events due to MO inadvertent binding were also validated by MATS analysis of transcriptomic data. Although previous work already theorized that MOs have a MIL value of 14–15 contiguous bases (at 37°C) (15), we show here that, at 28°C, the MIL is decreased down to 11 nt. Experiments in sea urchin embryo culture are performed at 14°C, suggesting an even much lower MIL value which may explain the already reported off target effects in sea urchin embryos injected with MOs (49). Despite the so-called ‘exquisite’ sequence specificity of MOs, we have established its effect on secondary target mRNAs in the zebrafish genome. The number of these unintended RNA targets may be largely underestimated, especially because MOs act at a rather low temperature in the zebrafish model. The influence of MOs on non-coding RNA (i.e. miRNAs, IncRNAs) or on regulatory exonic and intronic sequences (in which similar homology stretches of 11 nt might exist) has never been evaluated. Our study suggests a new mechanism of MOs action by which they influence splicing regardless of their binding to splice junctions. Indeed, by binding within an exon, they may affect the splicing process by hiding ESEs or ESSs (exonic splicing silencers) and inhibiting trans splicing factor binding. At last, the relatively high number of conserved positions at the 5′-splice sites increases the likelihood of finding imperfect MO-binding sites at splicing junctions in other RNAs. Similarly, translation MOs are systematically designed to target the conserved Kozak and AUG sequences. By computational analysis and based on the observed MIL value of 11, we determined the probable number of sMOsrzf5a’s secondary
targets located on introns, exons and exon-intron junctions in the zebrafish genome. To this end, we searched perfect matches between the reverse complemented MO sequence (5'-TGCAAGTGAGGTCAGATCC-3') and all pre-mRNA sequences (see Supplementary Methods). We were able to find 279,182 and 489 matches in exons, introns and junctions respectively (Supplementary Table S5). For statistical assessment of sMOsrsf5a preference for junctions over exons and introns, the same search was performed on 200 sets of random background sequences controlled for either the mononucleotide or dinucleotide composition. When using these simulated pre-mRNA sequences, the number of matches found on introns and exons did not significantly decrease (1.35- to 2.56-fold), in contrast to the number of matches found on junctions (44.66- to 59.09-fold) (Supplementary Table S5). These results clearly indicate that the presence of conserved junction sequences should be taken into account in MO sequence design. From these data, we propose two ways for reducing MO binding to secondary targets: (i) by promoting the use of a relatively high incubation temperature in order to increase stringency and indeed the k-value (Supplementary Table S5) and (ii) by preventing the presence of residues complementary to the highly conserved splice sequences (AG-GUNAG at the donor site and C/UAG-G/A at the acceptor site) by targeting (if at all possible) known regulatory intronic/exonic sequences (50) in the MO.

Does rescue mean 'no off-target'? 

RNA rescue experiments have been thought to be a robust control to approve MO specificity. However, morphant embryos have been rescued in many studies reporting phenotype inconsistencies (27,30). In most of these studies using a splice site morpholino, the phenotype was rescued using a mature spliced mRNA missing a considerable part of the MO target sequence. Our work suggests that sMOsrsf5a binding to 13 consecutive bases of the co-injected mRNA can provide an explanation for any rescue by decreasing knockdown efficiency due to sMOsrsf5a titration (26). This hypothesis was confirmed in vitro using EMSA (Supplementary Figure S9). Sequential rescue experiments were performed, intending to decrease this titration phenomenon, by injecting the mRNA at the one-cell stage and the MO at the 4-cells stage. However, partial rescue still occurred, suggesting that the injection delay is not sufficient per se for preventing MO binding to the injected RNA (Supplementary Figure S10). Given the high copy number of injected MOs and RNAs, we hypothesize that they can still hybridize in the embryo cells. The more severe phenotype observed in srsf5a−/− injected embryos could be explained by a more important sMOsrsf5a binding to inadvertent RNAs due to the absence of srsf5a mRNAs.

Many other standard controls were recommended to check MO specificity (51). Several different MOs should be tested, although functional analysis of SpRun in sea urchin provided a cautionary example of the insufficiency of two different morpholinos as a control for specificity (49). The use of a five-base mismatch control MO was also suggested. With only 11 nt being sufficient for effective binding, such controls appear somehow problematic. Mismatches should be designed along the entire MO sequence to interrupt any 11 nt inadvertent complementarity stretch; however this does not avoid appearance of new unintended target sites. Therefore, we also think that the only valid control of MO specificity might be the confirmation that its effects are lost in a null background (29). On the other hand, studies using mutants may suffer from masking of the correct phenotype by compensatory mechanisms, as also illustrated here. In such a case, MOs may be useful as they seem unable to trigger the compensatory response, if their specificity is sufficiently proven, and if the 11-nt MIL value is considered. With all these precautions, mutants combined with antisense MOs represent a valid toolbox to elucidate the true function of a gene.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We thank Steven Fanara for helpful discussions. We also want to thank the zebrafish facility of the GIGA-R center. Author contributions: P.M. and M.M. conceived the project and secured funding. M.J., P.M. and M.M. coordinated the study and designed experiments. M.J. and M.S. performed the experiments. D.B. performed the statistical analysis about the number of sMOsrsf5a RNA secondary targets within the zebrafish transcriptome. The paper was written and figures prepared by M.J., P.M. and M.M. All authors interpreted and discussed the results, and commented on the manuscript.

FUNDING

‘Fonds National de la Recherche Scientifique—FNRS’ [FRFC n°2.4631.11F, CDR J.0080.15 to D.B.]; ‘Fonds Spéciaux du Conseil de la Recherche’ from the University of Liège; FRIA (Fonds de la Recherche pour l’Industrie et l’Agriculture, Belgium Doctoralship (to M.J.)); University of Liège ‘Crédit de démarrage 2012’ SFRD-12/04 (to D.B.); Chercheur Qualifié du F.N.R.S. (Fonds National de la Recherche Scientifique, Belgium) (to M.H., M.M.). Funding for open access charge: Fonds De La Recherche Scientifique—FNRS and University of Liège.

Conflict of interest statement. None declared.

REFERENCES

1. Zhong, X.Y., Wang, P., Han, J., Rosenfeld, M.G. and Fu, X.D. (2009) SR proteins in vertical integration of gene expression from transcription to RNA processing to translation. Mol. Cell, 35, 1–10.
2. Long, J.C. and Caceres, J.F. (2009) The SR protein family of splicing factors: master regulators of gene expression. Biochem. J., 417, 15–27.
3. Howard, J.M. and Sanford, J.R. (2015) The RNAissance family: SR proteins as multifaceted regulators of gene expression. Wiley Interdiscip. Rev. RNA, 6, 93–110.
4. Manley, J.L. and Krainer, A.R. (2010) A rational nomenclature for serine/arginine-rich protein splicing factors (SR proteins). Genes Dev., 24, 1073–1074.
5. Fu, X.D. and Ares, M.Jr (2014) Context-dependent control of alternative splicing by RNA-binding proteins. Nat. Rev. Genet., 15, 689–701.
