Steroid receptor RNA activator protein binds to and counteracts SRA RNA-mediated activation of MyoD and muscle differentiation

Florent Hubé[1],*, Guillaume Velasco[1], Jérôme Rollin[2], Denis Furling[3] and Claire Francastel[1]

[1]Université Paris Diderot-Paris7, CNRS UMR7216, Epigenetics and Cell Fate, Paris, [2]Département d'Hématologie-Hémostase, Université François Rabelais, Hôpital Trousseau, Tours and [3]Université Pierre et Marie Curie-Paris6, Institut de Myologie UM76, INSERM U974 and CNRS UMR7215, Paris, France

ABSTRACT

The steroid receptor RNA activator (SRA) has the unusual property to function as both a non-coding RNA (ncRNA) and a protein SRAP. SRA ncRNA is known to increase the activity of a range of nuclear receptors as well as the master regulator of muscle differentiation MyoD. The contribution of SRA to either a ncRNA or a protein is influenced by alternative splicing of the first intron, the retention of which disrupts the SRAP open reading frame. We reported here that the ratio between non-coding and coding SRA isoforms increased during myogenic differentiation of human satellite cells but not myotonic dystrophy patient satellite cells, in which differentiation capacity is affected. Using constructs that exclusively produce SRA ncRNA or SRAP, we demonstrated that whereas SRA ncRNA was indeed an enhancer of myogenic differentiation and myogenic conversion of non-muscle cells through the co-activation of MyoD activity, SRAP prevented this SRA RNA-dependant co-activation. Interestingly, the SRAP inhibitory effect is mediated through the interaction of SRAP with its RNA counterpart via its RRM-like domain interacting with the functional sub-structure of SRA RNA, STR7. This study thus provides a new model for SRA-mediated regulation of MyoD transcriptional activity in the promotion of normal muscle differentiation, which takes into account the nature of SRA molecules present.

INTRODUCTION

Until recently, the central dogma of biology held that genetic information, stored on DNA, through RNA as intermediate molecules, was translated into the final protein product that fulfils most structural, catalytic and regulatory functions. However, the growing number of non-protein-coding RNA (ncRNA) discovered, and the variety of genetic and epigenetic phenomena in which they have been implicated, now suggest that this traditional assumption needs to be revised. In humans, ncRNA account for 98% of the transcriptional output, and have been implicated in a large range of cellular processes [for a review, see (1)]. While the function of classical rRNA, tRNA and microRNA in protein translation, or small nuclear RNA in mRNA splicing, has now been well established, the regulation of transcription itself appears to involve new classes of ncRNA. Several ncRNA have been implicated in the control of transcription by mediating changes in the structure of chromatin at genes involved in imprinting, dosage compensation or development (1–3). Other ncRNA modulate the activity of transcriptional activators or co-activators, directly or through the regulation of their sub-cellular partitioning. Examples include B2-SINE that can directly bind RNA polymerase II leading to the inhibition of gene expression under heat-shock conditions (4), 7SK that represses transcriptional elongation through its interaction with the basal transcription factor P-TEFb (5), or NRON [non-coding repressor of NFAT] that binds to the transcriptional activator NFAT and prevents its nuclear localization (6).

In 1999, a very peculiar co-activator called steroid receptor RNA activator or SRA (7) was added to an already long list of co-regulators of steroid receptors [for a review, see (8)]. Indeed, SRA differs from all previously described co-activators since it functions as a ncRNA molecule. It has been now established that SRA transcripts co-activate numerous nuclear receptors (7,9–14) as well as the activity of MyoD, a transcription factor involved in skeletal myogenesis (15,16). SRA RNA

*To whom correspondence should be addressed. Tel: +33 1 57 27 89 32; Fax: +33 1 57 27 89 10; Email: florent.hube@univ-paris-diderot.fr

may therefore have a function wider than previously thought and may be important in regulating proliferation and/or differentiation in various cell types.

The first characterization of SRA isoforms demonstrated that they share a common core region, characterized by discrete stem–loop structures, required for the full co-activator function of SRA (7). However, and as for all other so-called ncRNA, no evidence indicated the existence of an SRA protein. Further investigations identified additional SRA RNA isoforms produced by alternative splicing or multiple transcription start sites. One of these isoforms, with a deletion within the SRA core sequence resulting from splicing of exon 3, was observed in breast and ovarian tissues (17,18). High levels of expression of this isoform in breast tumours relative to normal tissue, presumably impaired in its co-activator function, correlated with a higher tumour grade (18). Other SRA isoforms exhibited an additional exon upstream of the core exons, containing two initiating methionines and a predicted open reading frame (ORF) of 236/224 amino acids (19–21). Consistent with these hallmarks associated with coding sequences, two SRAP proteins have been detected (19,22). These data provided the demonstration that SRA, primarily classified as an ncRNA, was the founding member of a new family of ncRNA exhibiting the ability to encode for proteins (19).

Given the existence of both coding and non-coding SRA transcripts, we proposed that differential splicing of SRA transcripts might regulate the balance between each type of molecule and influence the overall effect of SRA expression. This study was intended to exhaustively identify and assess the levels of SRA molecules in human primary muscle satellite cells, isolated from both healthy and myotonic dystrophy muscles, taking into account their coding or non-coding features, and test their function on both MyoD activity and myogenic differentiation.

MATERIALS AND METHODS

Cell lines and primary cells

Human myogenic LHCN-M2 derived from primary human satellite cells and their *in vitro* differentiated myotubes (MT) counterpart, murine myoblast C2C12 (C2C12), murine mesenchymal stem C3H10T1/2 (10T1/2), Cos6 (COS), HEK-293 and MCF-7 cells were grown according to supplier's recommendations or as described earlier (21,23,24). Primary murine satellite cells (MB and MT) were isolated from adult murine skeletal muscles as described (25) and were a generous gift from Drs S. Charrin and E. Rubinstein (André Lwoff Institute, Villejuif, France).

Briefly, C2C12, primary murine MB and undifferentiated 10T1/2 cells were grown in growth media (GM; DMEM, 20% foetal bovine serum, 100 U/ml penicillin, 100 µg/ml streptomycin, 2 mM glutamine, all from Invitrogen Life Technologies). Muscle differentiation of these cells was induced by switching to serum-deprived media (DM). Likewise, primary human MB were cultured in growth media (F-12 Nutrient mixture/Ham from Invitrogen, 20% foetal bovine serum, 100 U/ml penicillin, 100 µg/ml streptomycin, 2 mM glutamine) and *in vitro* differentiation was induced by switching to serum-deprived media. COS, HEK-293 and MCF-7 were cultured in DMEM, supplemented by 10% foetal bovine serum, 100 U/ml penicillin, 100 µg/ml streptomycin, 2 mM glutamine.

RNA from primary human satellite cells isolated from muscle biopsies of three foetuses showing clinical symptoms of the congenital myotonic dystrophy type 1 (DM1), carrying more than 2000 CTG repeats, as well as RNA from satellite cells isolated from non-affected age-matched muscle, were included in this study (26).

Antibodies

Primary antibodies were directed against SRAP (7H1G1, Abcam), MyoD (sc-304, Santa Cruz Biotechnology), V5-tag (sc-81594, Santa Cruz), Flag-tag (OctA-probe D8, sc-807, Santa Cruz) and β-tubulin (T4026, Sigma-Aldrich).

Plasmids and constructions

SRA plasmid constructs are depicted in Supplementary Figure S1. Briefly, a fragment between nucleotides 418 and 1314 relative to the sequence reference NM_001035235 was subcloned into pcDNA3.1/V5-His© TOPO® TA (Invitrogen) (27) and used as a template to produce all other SRA constructs, except the 'SRAP only' plasmid, by mutagenesis using QuikChange® Site-Directed Mutagenesis Kit (Stratagene) following the manufacturer's instructions with primers described in the Supplementary data. 'SRAP only' and 'noSRA/noP' were a generous gift from Dr E. Leygue (University of Manitoba, Canada) (27). MyoD-3HA expression vector was a gift of Dr S. Ait-Si-Ali (UMR7216, University of Paris Diderot, Paris, France). Promoter-driven luciferase reporter vector containing E-box responsive elements (pE-Luc), GATA-3 and ETS-2 binding sites (pGE-Luc, containing the promoter region of the GpIIb gene), ETS2 expression plasmid and pRL-TK were gifts of Dr Z. Kadri (CEA Fontenay-aux-roses, Paris, France) and Dr I. Dussanter (Institut Cochin, Paris, France). PPARγ-Flag (pcDNA flag PPAR gamma), pP-Luc (PPRE X3-TK-luc), p53-Flag (pcDNA3 flag-p53), p53-Luc (PG13-luc wt p53 binding sites) plasmids were requested from Addgene (www.addgene.org).

RNA extraction, RT–PCR and triple primer PCR

Total RNA was isolated with Trizol reagent (Invitrogen) according to the manufacturer's instructions and as described earlier (28,29). Primers used for amplification are described in the Supplementary data and PCR were performed as described earlier (21,28,29). triple primer PCR (TP–PCR) was performed as described earlier (21) with minor modifications. Briefly, PCR was achieved using non-radioactive dNTP and PCR products were analysed on agarose gels. Quantitative PCR was carried out using 50 ng of cDNA and primers listed in the Supplementary data, using the LightCycler® 480 SYBR Green I master on the Light cycler 480 real-time PCR system (Roche Diagnostics). The ratio between the

expression of non-coding (NC) and coding (C) isoforms was obtained by calculating the $2^{-\Delta\Delta Ct}$ for each sample and using MCF7 cells as the reference, i.e. $2^{-((NC\ sample\ -\ C\ sample)\ -\ (NC\ MCF7\ -\ C\ MCF7))}$.

Rapid amplification of the 5′-cDNA end

We performed 5′-RACE PCR using the GenRacer kit (Invitrogen Life Technologies) as published earlier (21,30,31).

Short hairpin RNA production

Short hairpin RNA (shRNA) directed against human SRA transcripts (shI1, located in intron 1; shE1-2, overlapping the junction between exons 1 and 2) and luciferase gene (shLuc) were produced using MessageMuter™ shRNA production kit (Epicentre Biotechnologies) following manufacturer's instructions.

In vitro transcription

SRA transcripts were *in vitro* transcribed using the T7 RiboMax large-scale production system (Promega) following the manufacturer's instructions.

Transfection experiments

Plasmid constructs and shRNA were transiently transfected using jetPEI™ and jetSI™-endo reagents (PolyPlus transfection), respectively, following the manufacturer's instructions. COS and human primary myoblast cells were grown on P12 wells at 80% confluence. Expression vectors were transfected at $100\,ng/cm^2$, firefly luciferase reporter plasmids at $37.5\,ng$ per cm^2, the pRL-Tk renilla luciferase reporter vector at $12.5\,ng/cm^2$, co-transfected with firefly reporter vectors to normalize for transfection efficiency, and shRNA at $2\,pmol/cm^2$.

Luciferase assay

Firefly and renilla luciferase activities were measured 24 h post-transfection using the Dual-Luciferase® Reporter Assay System (Promega) and the GloMax®-Multi Detection System Luminometer (Promega). Luciferase activities were acquired as relative luciferase unit (RLU), corrected by renilla luciferase activity and expressed as the fold activity over the signal obtained from the luciferase reporter vector alone. Significant differences were assessed using Student's *t*-test. Error bars represent standard error at the mean (SEM).

Cell lysis and western blot

Proteins were extracted from cells using a single-detergent lysis buffer (0.3M NaCl, 10 mM $MgCl_2$, 50 mM Tris pH 7.5, NP40 0.4%) and the protein concentration for each sample was determined using Bio-Rad Protein Assay Kit (Bio-Rad). Equal amounts of total protein were either used in Co-IP/NChIP assay or directly analysed by SDS–PAGE and immunoblotting as described earlier (28).

RNA and protein immunoprecipitation

Cells were lysed and quantified as described earlier and used for native RNA ChIP (NChIP) or protein immunoprecipitation/co-immunoprecipitation (co-IP) experiments. Protein extracts ($100\,\mu g$ for NChIP and $1\,mg$ for co-IP) were incubated with the appropriate antibody ($1\,\mu g/mg$ of proteins) for 2 h at 4°C as described earlier (28,29). The precipitated material was then re-suspended in Trizol reagent for further analysis of associated RNA as described earlier, or Laemmli buffer to analyse protein partners after separation on SDS–PAGE and immunoblotting.

To analyse the effects of STR deletions, immunoprecipitation using HEK-293 protein extract was first performed in the presence of $50\,\mu g$ of RNAse A to remove all traces of endogenous RNA. After five extensive washes with the lysis buffer adjusted to 300 mM NaCl, $10\,\mu g$ of *in vitro* transcribed SRA RNA, in which mutations in the different STR were introduced as described earlier, was added for an additional 1 h followed by three washes in the lysis buffer. RNA was then extracted using the Trizol reagent as described earlier.

Validation of the putative SRAP RNA binding domain (RNP) was performed using NChIP following transient transfections of V5 tagged-SRA expression constructs in COS cells. COS cells were transfected with 'FL', 'mRNP' or 'ΔATG mRNP' expression plasmids, and tagged SRAP was immunoprecipitated using V5 antibody. Associated RNA was then extracted using the Trizol method, reverse transcribed and PCR amplified as described earlier.

Database analysis and RNA structure prediction

Database analysis was performed using freely available on-line ASmodeler web-based tool (http://genome.ewha.ac.kr). ASmodeler combines genome-based EST clustering and transcript assembly procedures in a coherent and consistent fashion to produce gene models (32) and currently supports human genome version hg16. In this version, input SRA1 sequences corresponded to UniGene ID number Hs.32587 and contained 262 EST sequences and 9 already published mRNA sequences. These sequences were genomically aligned and clustered according to shared exon–intron boundaries using ASmodeler web-based tool, and led to five alternatively spliced clusters of gene models (AS5C1 to 5). Clusters 2–5, as well as five gene models in AS5C1, with isoforms containing two or fewer exons, probably representing degraded or truncated species, were not further considered. In addition, two clusters showed redundancy (AS5C1.16 and AS5C1.17). Altogether, this analysis led to the identification of 20 distinct clusters of gene models (Supplementary Figure S2A).

Mfold server (http://frontend.bioinfo.rpi.edu) is an algorithm widely used to predict nucleic acid folding and hybridization by free energy minimization using empirically derived thermodynamic parameters (33). Parameters used were as follows: without constraint information; percent sub-optimality of 5%; maximum number of folding limited to 50; windows parameters setup at 57.

RESULTS

Novel classes of SRA transcripts uncovered by the analysis of public databases

Multiple SRA isoforms have been described in various systems, yet a systematic analysis of these molecules in terms of sequence, expression and function is still missing. Therefore, we first performed a thorough analysis of information available from public databases to search for EST associated with the human SRA1 gene as described in the 'Materials and Methods' section. Genomic alignment of these sequences led to 20 possible gene models (Supplementary Figure S2A).

This data mining led to the identification of additional classes of SRA isoforms. We assessed the occurrence of individual exons and intron 1 (Supplementary Figure S2B) and identified exons 4 and 5 as constitutive exons, being included in >99.7% of isoforms, although the 3'-extremity of exon 5 appeared to vary between isoforms. Surprisingly, only 59% of SRA RNA isoforms appeared to include exons 2 through 5, previously defined as the 'core element', and reported as necessary for the SRA RNA function as a transcriptional co-activator (7,11). In agreement with our previous findings (19,20), an additional exon, exon 1, was included in 72% of the isoforms, with variable 5'-extremities between transcripts. Of note, a sixth exon was identified in 2 of the 20 clustered transcripts (#6 and #7), representing only 2% of identified isoforms.

To ease further functional analysis of these different RNA molecules, we further grouped the 20 categories of SRA transcripts into two potentially functionally distinct categories of SRA transcripts, based on whether they displayed the recently identified ORF spanning exons 1 to 5 (19,20,34) or not (Figure 1). The first group of isoforms represents true non-protein coding RNA and accounts for 61% of the identified SRA isoforms. These transcripts are characterized by the absence of the 236/224 amino acids long ORF, due in part to the retention of intron 1, or by the absence of exon 3. They also show highly variable length of the 5'-end (Figure 1). On the contrary, nine transcript clusters exhibit the ability to encode the SRAP protein and account for 39% of all SRA isoforms. The presence of the SRAP ORF in this group of transcripts resulted from both the splicing of intron 1 and the retention of exon 3.

In essence, this analysis broadened our knowledge of the spectrum of SRA RNA molecules and stressed the

need to further analyse their individual expression and function.

Evidence of the diversity of SRA transcripts *in vivo*

In order to systematically assess the *in vivo* existence of the SRA isoforms identified above, the total RNA was extracted from human myogenic cells and analysed by RACE PCR (Figure 2 and Supplementary Figure S3). Because of the constitutive occurrence of exon 4 in SRA isoforms, we designed the reverse primer used for cDNA amplification in this exon and identified six distinct amplification products (Figure 2A–F) that were further cloned into pCR4-TOPO vector. A total of 40 clones were subjected to sequencing, which allowed us to identify three different groups of transcriptional initiation (Supplementary Figure S3). Band A identified the first cluster that resides at the furthest 5'-end of exon 1, probably matching the initiation site of the longest SRA isoform experimentally observed to date (positions −288 and −250 relative to the first nucleotide of exon 2). As represented in Figure 2, the sequence analysis of this band revealed that RNA initiated at this position do not retain intron 1 and conserved exon 3, suggesting that these RNA, initiated upstream of the ATG, potentially encode the SRAP protein. The second cluster (Supplementary Figure S3), between −68 and −27, surrounds the two ATG from which SRAP can be translated, assuming the resulting transcript retains the ORF. Sequence analysis of the transcripts initiated at this position indeed indicated that part of them, mainly identified in band D, display the correct sequence to encode SRAP, characterized by the absence of intron 1 and retention of exon 3. Other transcripts initiated at this position, contained in bands B and C, retain the intron 1 and therefore correspond to non-coding SRA RNA. The last cluster (Figure 2 and Supplementary Figure S3) lies downstream of the second ATG (from −24 to +127). Transcripts initiated at this position should not produce the SRAP protein product. Consistent with the above database analysis, we could conclude that most of the SRA isoforms identified by RACE–PCR retain the core sequence, exon 2 to 4, whereas about one-third of these isoforms potentially encode SRAP. The remaining and marginal RNA species do not display the core elements or the ability to encode SRAP.

Altogether, these data confirm that most of the EST identified in the public databases are actually expressed in human myogenic cells.
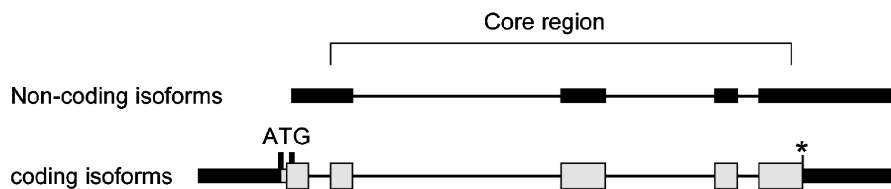

Figure 1. Archetypical coding and non-coding isoforms of SRA transcripts. Coding sequences corresponding to the longest open reading frame are indicated in pale grey and untranslated sequences are shown in black. The two ATG are indicated by vertical lines, the stop codon by a star and the core region described by Lanz *et al.* (7) is indicated above.
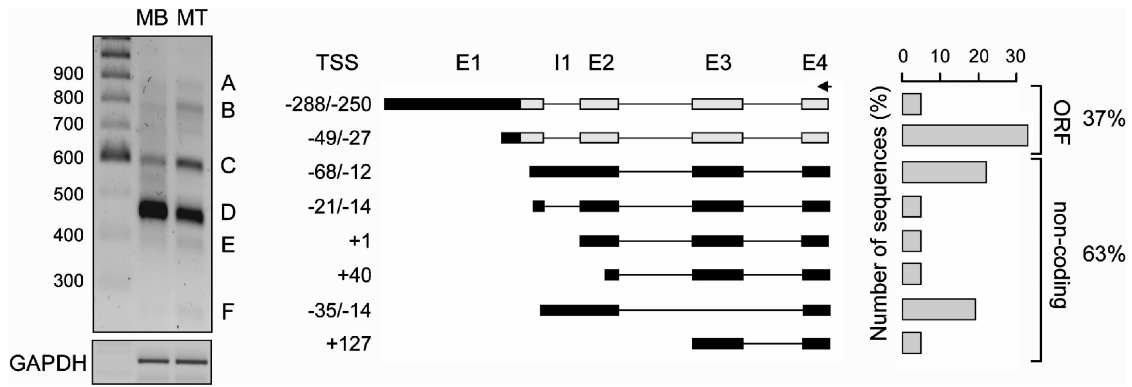
Figure 2. Diversity of SRA RNA transcripts in human primary myoblasts and myotubes. Total RNA was extracted from human primary myoblasts (MB) or *in vitro* differentiated myotubes (MT) and analysed by RACE PCR. The reverse primer used for cDNA amplification is indicated by the black arrow head located in exon 4 on the schematic SRA1 gene represented on the top. Exonic regions are indicated on the top and numbered as annotated in GenBank. Coding sequences from the longest open reading frame are indicated in pale grey and untranslated sequences are shown in black. GAPDH amplification was used as an internal control. The six bands identified on the gel were cut, cloned and sequenced. The schematic representation of SRA amplicons illustrates the identification of their transcriptional start site combined to the analysis of their sequence. Positioning of TSS is arbitrarily indicated relative to the beginning of exon 2. The histograms on the right represent the number of sequenced clones identified in each category with their corresponding coding capacity determined after sequencing.

The ratio between coding and non-coding SRA isoforms varies during muscle differentiation

Since SRA was considered as a co-activator of MyoD (15), a master regulator of muscle differentiation, we performed the RACE–PCR analysis of RNA isolated from undifferentiated myoblasts (MB) or their *in vitro* differentiated myotubes (MT). As shown in Figure 2, the two major bands detected, C and D, showed an inverse correlation with the differentiated state. B and C, mainly containing non-coding isoforms, were more intense in MT than in MB. In contrast, band D, mainly containing coding isoforms, was fainter in MT than in MB. To confirm this variation in the ratio between isoforms retaining or not an ORF during muscle differentiation, we used the TP–PCR method that we previously used (21,35) to amplify reverse transcribed RNA isolated from undifferentiated or differentiated muscle cells as described in Supplementary Figure S4. It is noteworthy that levels of isoforms containing the core region (E2/E4) present in most of SRA isoforms remained stable during differentiation (Supplementary Figure S4B). However, the analysis of the expression of RNA species identified above showed that muscle differentiation, monitored by the appearance of muscle-specific differentiation markers Myogenin, MCK and Myosin heavy chain, was accompanied by a decrease in coding isoforms concomitant with an increase in non-protein-coding isoforms. Because of the simultaneous increase in ncRNA and decrease in SRAP coding isoforms, the ratio between non-coding and coding SRA RNA in differentiated muscle cells is largely in favour of non-coding transcripts (Supplementary Figure 4C). This ratio is two to five times higher in human myogenic differentiated MT than in MB. Interestingly, this variation in SRA RNA isoforms seems to be conserved in primary murine MT and in differentiated cells from the muscle line C2C12 (Supplementary Figure S4D). As shown in Figure 3, we further confirmed this change in the ratio between non-coding and coding SRA isoforms associated
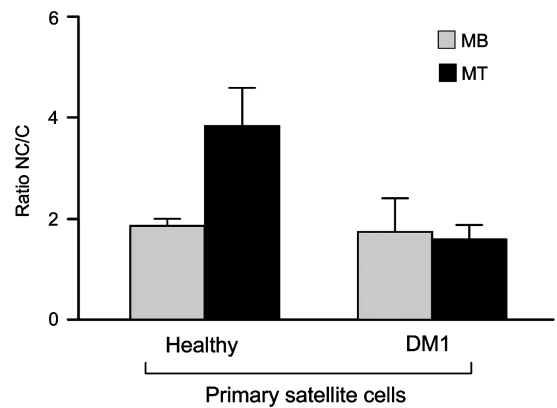

Figure 3. Real-time PCR amplification of coding and non-coding SRA RNA during myogenic differentiation of primary human satellite cells. RNA extracted from primary satellite cells (grey bars), isolated from healthy or myotonic dystrophy muscles (DM1), and their *in vitro* differentiated counterpart (black bars), was reverse transcribed and amplified by real-time PCR as described in the 'Materials and Methods' section, using primers that distinguish between non-coding and coding SRA isoforms. Signals obtained from three individual healthy and three individual DM1 samples done in triplicate were quantified, and the average proportion of non-coding versus coding SRA RNA expressed as relative units is represented on the histogram for each sample. Error bars represent the standard error at the mean.

with muscle differentiation of primary human satellite cells by real-time PCR, using primers that distinguish between these isoforms. Delayed differentiation of skeletal muscle cells has been reported in the severe congenital form of myogenic dystrophy type 1 (DM1) (26). We therefore examined the levels of non-coding and coding SRA RNA in primary muscle satellite cells isolated from DM1 muscles, maintained for up to 6 days in normal or serum-deprived culture conditions. As shown by real-time PCR experiments (Figure 3), induction of muscle differentiation was not accompanied by a change in the ratio between non-coding and coding SRA isoforms in DM1 cells compared to non-affected muscle cells.

Our examination of the variety of SRA RNA species combined to RACE PCR (Figure 2), TP–PCR (Supplementary Figure S4) and real-time PCR (Figure 3) analysis of their expression allowed us to reveal an increase in the ratio between non-coding and coding SRA RNA isoforms, associated with the capacity of the cells to differentiate into muscle fibres.

Distinct functions of SRA molecules on the activity of transcription factors *in vitro*

The role of SRA ncRNA as a co-activator of many transcription factors has been the subject of several reports (7,10–12,15,36–38). Yet, since our present work describes SRA isoforms with potentially distinct functions, the activity of individual SRA molecules has to be clarified. To this aim, we performed luciferase gene reporter assays as described earlier (21,27,30,39). We chose to transfect COS cells since they express very low levels of endogenous SRA that would interfere with the co-activation assay (40). COS cells were co-transfected with a MyoD expression vector and a promoter-driven luciferase reporter vector containing E-box responsive elements, in the presence of increasing amounts of SRA expression vectors that we previously used (27). Transcription factor activities were expressed as fold-change over the activity of the reporter construct alone and set at 1, in order to only evaluate the co-activation by SRA molecules. Surprisingly, as shown in Supplementary Figure S5A, a full-length SRA molecule ('FL', nucleotides 418–1314 of RefSeq NM_001035235) did not reproduce the known 2.5-fold activation of MyoD activity by SRA (15). For each transfection, we verified the expression of SRA molecules at the level of RNA and protein. The 'FL' construct was indeed able to drive both SRA RNA (Supplementary Figure S5B) and SRAP (Supplementary Figure S5C) expression while 'noSRA/noP' plasmid did not allow the synthesis of RNA and protein related to SRA.

Because the 'FL' construct could contribute to both SRA RNA and SRAP, and in order to elucidate the individual activity of SRA molecules, we used constructs designed to exclusively produce either a SRA ncRNA molecule ('RNA only') or a protein SRAP ('SRAP only'), but not both (27). The 'RNA only' construct indeed enhanced MyoD transcriptional activity in a dose-dependent manner. In contrast, the 'SRAP only' or the 'noSRA/noP' construct, used as a negative control, had no effect on the MyoD activity, regardless of the amount of construct used (Figure 4).

The SRA RNA-only molecule ('RNA only') was also able to activate the activity of two other unrelated transcription factors, GATA-3 and PPARγ, by an ∼2- to 8-fold depending on the transcription factor considered, in a dose-dependent manner (Supplementary Figure S6). This is not a general phenomenon, since two other transcription factors were not activated by SRA RNA, regardless the amount of expression vector used (Supplementary Figure S6). In contrast, the activity of the ETS2 transcription factor was increased by up to 3-fold by the 'SRAP only' construct but not in the presence of 'RNA only',
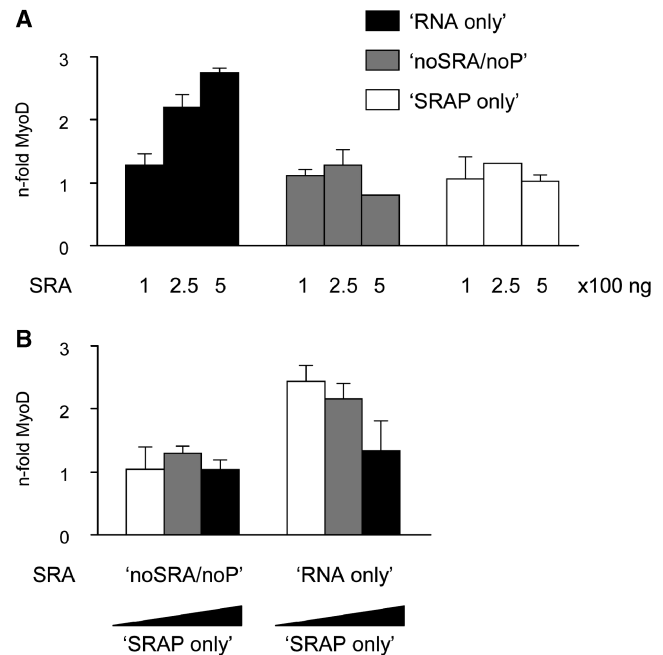

Figure 4. SRA RNA and SRAP have distinct effects on MyoD-activated transcription. (**A**) COS cells were co-transfected with the pE-Luc plasmid and the MyoD expression vector, in combination or not with increasing amounts of SRA expression vectors. SRA expression vectors used are 'RNA only' (black histograms), 'SRAP only' (white histograms) and 'noSRA/noP' (grey histograms). Luciferase activity, shown as the *n*-fold value compared to cells transfected with the MyoD expression vector alone, was measured as indicated in the 'Materials and Methods' section. Histograms are the mean of at least five independent experiments; error bars represent the standard error at the mean. (**B**) COS cells were co-transfected with constant amounts of the pE-Luc plasmid, MyoD expression vector and SRA RNA expression vector ('RNA only' or the control 'noSRA/noP'), together with increasing amounts of the SRAP expression vector, as described in the 'Materials and Methods' section. White, grey and black bars correspond to 100, 250 and 500 ng of the SRAP expression vector, respectively. Luciferase activity from at least three independent experiments is represented as the *n*-fold value relative to cells transfected with MyoD expression vector alone. Error bars represent the standard error at the mean.

while p53 was not affected by either SRA RNA or SRAP (Supplementary Figure S6).

These data led to the important conclusion that although SRAP does not have an intrinsic co-regulatory function on MyoD activity, SRA RNA is able to clearly enhance MyoD transcriptional activity only when SRAP coding features are disrupted.

SRAP prevents SRA RNA-activation of the MyoD activity

In order to verify that SRA RNA was able to enhance MyoD transcriptional activity only in the absence of SRAP, we co-transfected cells using constant amounts of luciferase reporter construct with constant amounts of MyoD and SRA 'RNA only' expression vectors, in conditions that produce the known 2.5-fold activation of MyoD transcriptional activity, and increasing amount of 'SRAP only' expression plasmid. The 'noSRA/noP'

plasmid served as a negative control and did not affect MyoD's activity (Figure 4, white bar). However, the SRA-enhanced transcriptional activity of MyoD was repressed when SRAP was co-expressed, in a dose-dependant manner (Figure 4B, grey and dark bars).

Therefore, although SRAP does not have an intrinsic co-regulatory function on the MyoD activity, it can prevent SRA RNA-mediated activation of MyoD.

SRAP interacts with its SRA RNA counterpart

Because SRAP could prevent SRA RNA-mediated activation of the MyoD activity (Figure 4B), we assessed whether they could belong to a same complex. We performed NChIP (28) using an antibody directed against the endogenous SRAP followed by the analysis of the associated RNA, using RT–PCR to specifically amplify the endogenous SRA RNA core sequences. NChIP performed using human myogenic cellular extracts (Figure 5) or murine myoblats (Supplementary Figure S7A), showed that SRA RNA was associated with SRAP, whereas no antibody or an irrelevant antibody directed against Flag-tagged proteins, absent in cell extracts, did not co-precipitate SRA RNA. Likewise, β-actin RNA was not immunoprecipitated with anti-SRAP or -Flag antibodies (Figure 5 and Supplementary Figure S7A). In addition, SRA RNA was not immunoprecipitated with an anti-MyoD antibody at high-to-moderate stringency (300 mM NaCl buffer) whereas at lower stringency (150 mM), SRA transcripts were indeed detected. Noteworthy, the SRA RNA and SRAP interaction was
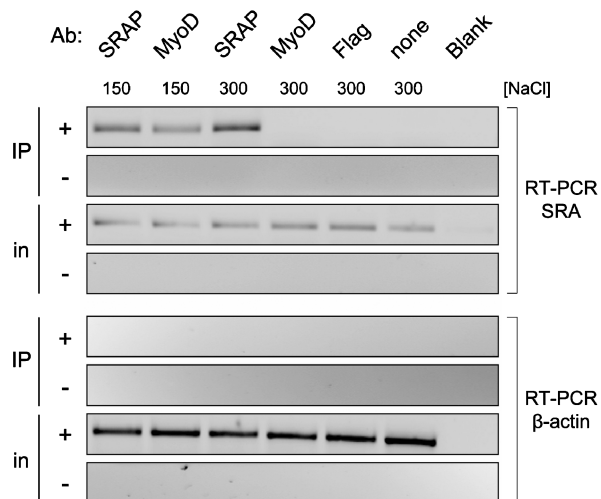

Figure 5. Interaction between SRAP and SRA RNA. NChIP performed on human myogenic cellular extracts followed by analysis of the associated RNA using RT–PCR to specifically detect the endogenous SRA RNA core sequence or β-actin mRNA. Antibodies used were directed against endogenous SRAP, MyoD or against Flag-tagged proteins, absent in cell extracts. Control NChIP was performed in which no antibody was added to the reaction. After precipitation, high (300 mM NaCl buffer) or low (150 mM NaCl buffer) stringency washes were performed. Input (in) sample represented 5% of the material used for precipitation. Immunoprecipitated (IP) RNA was analysed by RT–PCR, in reactions containing (+) or not (−) reverse transcriptase (RT).

unmodified in myoblast and myotube cells (Supplementary Figure S7B) and we did not detect interaction between SRAP and MyoD protein even at low stringency (Supplementary Figure S7C).

The interaction between SRA RNA and SRAP requires the STR7 stem-loop domain

In order to distinguish which sub-structural domains of SRA transcripts were involved in the interaction between SRA RNA and SRAP, we constructed four SRA mutants, each of them being deleted of one STR domain (Figure 6A). The STR domains considered were selected based on the similarities exhibited by STR sequences among rat, mouse and human SRA species (Supplementary Figure S8), i.e. exhibiting >85% of sequence identity. Thus, STR1 (91.7 and 94.4% similarity between human and mouse and rat, respectively), STR9 (87.5 and 85.3%), STR5 (both at 86.7%) or STR7 (87.8 and 91.8%) were deleted and the secondary structure of the resulting SRA RNA constructs ('ΔSTR1', 'ΔSTR9', 'ΔSTR5', 'ΔSTR7') was established using Mfold (33). Each construct was designed to allow for the transcription of SRA RNA mutants deleted of one STR but leaving the other STR domains unaltered (Figure 6A). We therefore used these constructs in *in vitro* RNA-ChIP assays to evaluate the consequences of STR deletions on SRA RNA interaction with SRAP (Figure 6B). Endogenous SRAP was immunoprecipitated from HEK-293 cell extracts (Figure 6C), treated with RNAse A to remove all endogenous RNA materials and after extensive washes, incubated with *in vitro* transcribed SRA RNA mutants. The analysis of the precipitated material revealed that SRA mutants deleted in STR1, STR9 or STR5 domains were still precipitated with SRAP. However, the SRA RNA construct deleted in STR7 sub-structure ('ΔSTR7') had lost its ability to interact with SRAP (Figure 6B, lane 5).

The interaction between SRAP and SRA RNA is RRM-dependent

Considering that SRA RNA could interact with SRAP *in vitro* and *in vivo*, we hypothesized that the interaction between STR7 sub-domain of the transcript and SRAP was achieved through a specific RNA-binding sequence. The most abundant type of eukaryotic RNA-binding motifs are the RNA recognition motif (RRM) and RRM-like domains that both contain an eight and six amino acids long RNP-1 and RNP-2 consensus sequences (41), respectively. While no RNP-1 was identified in the SRAP sequence, an atypical RNP-2 motif (LLVQEL) was clearly predictable between amino acids 163 and 168, this sequence being indeed exceptionally well-conserved in SRAP sequences among eukaryotic species (Figure 7A). We therefore mutated 3/6 codons of the RNP motif (Figure 7A). Corresponding 'mRNP' and 'ΔATG mRNP' constructs were transiently transfected in HEK-293 cells and NChIP experiments were performed as described earlier. As shown in Figure 7B, SRA RNA was co-precipitated with SRAP-V5 tagged protein,
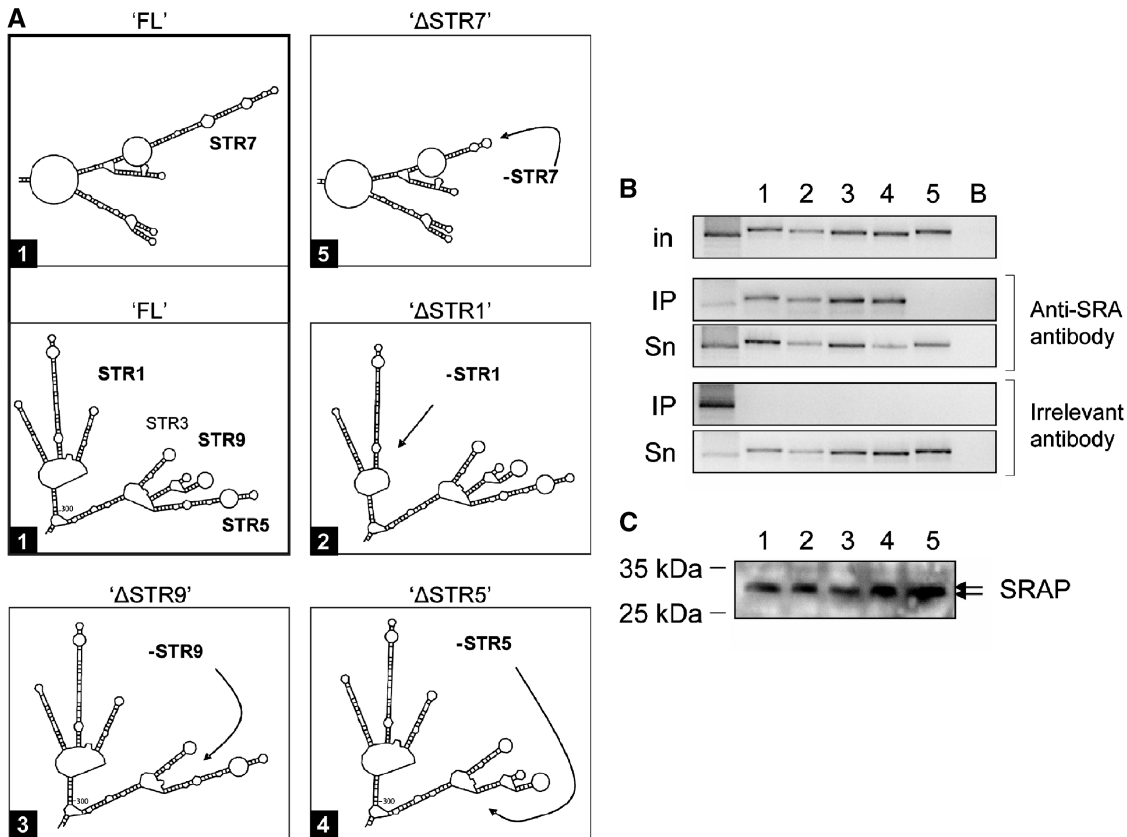
Figure 6. SRA RNA interacts with SRAP through its STR7 sub-domain. (**A**) Details of secondary structures of SRA deletion mutants used in this study. Each construct was designed to allow the *in vitro* transcription of SRA RNA mutants deleted in one given STR while leaving the other STR domains intact. (**B**) NChIP was performed as described in the 'Materials and Methods' section, using an antibody directed against the endogenous SRAP or an irrelevant antibody directed against Flag-tagged proteins, absent in cell extracts, and followed by analysis of the associated RNA using RT–PCR to specifically amplify the endogenous SRA RNA core sequence. SRA constructs were *in vitro* translated and used in NChIP experiments. 1, 'FL'; 2, 'ΔSTR1'; 3, 'ΔSTR9'; 4, 'ΔSTR5'; 5, 'ΔSTR7'; (**B**) no *in vitro* transcribed RNA control. in, 5% input; IP, immunoprecipitation; Sn, supernatant. (**C**) Western blot analysis of immunoprecipitated material using an antibody directed against the endogenous SRAP. The characteristic doublet related to SRAP is observed between 25 and 35 kDa.

whereas RNP-mutated SRAP ('mRNP') or the non-protein coding negative control ('ΔATG mRNP') did not lead to the immunoprecipitation of SRA RNA. An irrelevant RNA, GAPDH mRNA, was not precipitated with the wild-type SRAP nor with the mutated version of the protein.

Together, these data strongly suggest that SRA RNA and SRAP belong to a same complex, interacting through a specific SRAP domain (RNP-2) and an RNA stem-loop sub-structure (STR7).

MyoD-forced myogenic conversion of non-muscle cells is enhanced by SRA RNA *in vivo* but not by SRAP

In order to functionally confirm *in vivo* the ability of individual SRA molecules to enhance MyoD transcriptional activity observed *in vitro*, we performed gain and loss of function experiments.

Mesenchymal stem cell line C3H10T1/2 (10T1/2) was first transfected with combinations of MyoD and SRA expression vectors, and their myogenic conversion assessed by the appearance of the muscle-specific marker Muscle Creatine Kinase (MCK). The SRA expression

vectors used allowed the exclusive production of SRA ncRNA ('RNA only' plasmid), SRAP ('SRAP only' plasmid) or none of them ('noSRA/noP' plasmid), as described earlier. Cells were transfected and cultured in serum-containing media for 24 h, and then switched to a serum-deprived media in order to promote myogenic conversion. The efficiency of myogenic conversion was assessed by measuring levels of MCK transcripts 24 h, and 72 h after serum withdrawal. As shown in Figure 8A, when 10T1/2 cells were co-transfected with MyoD and control noSRA/noP vectors, transcription of MCK mRNA was detected as early as 24 h post-induction and further increased during the MyoD-forced myogenic conversion. Results obtained in cells co-transfected with MyoD and 'SRAP only' constructs were rigorously identical. As negative controls, in the absence of MyoD, SRA 'RNA only' and 'SRAP only' constructs did not alter the time course of conversion promoted by serum deprivation (Figure 8A). In contrast, a stronger and faster increase in MCK levels was observed in cells transfected with MyoD in combination with the SRA 'RNA only' construct 24 h post-induction.

Selective knockdown of intron 1-containing SRA ncRNA prevents the activation of MyoD target genes during the differentiation of human myogenic cells

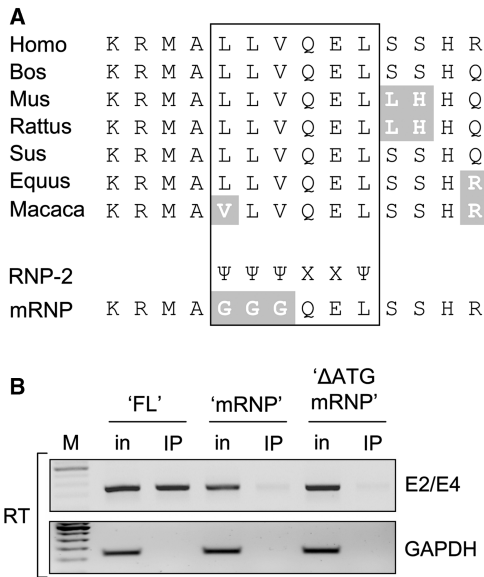We showed above that an 'RNA-only' construct could activate MyoD transcriptional activity *in vitro* whereas SRAP had no intrinsic activity on MyoD nor interacted with MyoD, but prevented SRA RNA-mediated activation of MyoD through binding to SRA RNA. Rather, the balance between non-coding and coding isoforms seemed to be critical in the promotion of normal myogenic differentiation. We therefore performed RNA interference designed to selectively knockdown non-coding or coding SRA isoforms. Since ncRNA isoforms predominate in differentiated cells, we predicted that an RNA interference approach designed to selectively knockdown intron 1-containing ncRNA SRA isoforms will have an impact on the MyoD activity and subsequently on muscle differentiation, while knockdown of SRAP coding isoforms will have only minor effects on differentiation. Human myogenic cells were transfected with shRNA directed against SRA intron 1 to target intron 1-containing ncRNA (shIn1), SRA junction between exons 1 and 2 to target coding isoforms deleted of intron 1 (shE1-2) or against luciferase mRNA as a control (shLuc), and induced towards muscle differentiation by serum deprivation for 24 h. SRA RNA levels were analysed by RT–PCR using primers that specifically amplify noncoding or coding isoforms, or the constitutive core elements. While a substantial reduction of coding isoforms was observed using shE1-2, no detectable effect was observed on the expression of the early marker of differentiation and MyoD target gene Myogenin (Supplementary Figure S9). Levels of SRA transcripts retaining intron 1 were strongly reduced using shIn1 compared to control shRNA (Figure 8B). Expression of isoforms that contain the core region, observed in most of SRA isoforms, was slightly decreased as a consequence of the diminution of isoforms containing intron 1 (Figure 8B). In contrast to shRNA directed against the exon 1–2 junction, shRNA directed against intron 1 led to a dramatic decrease in the expression of several MyoD target genes, normally activated upon induction of myogenic differentiation (Figure 8B). We observed that the expression of Myogenin, MCK and, to a lower
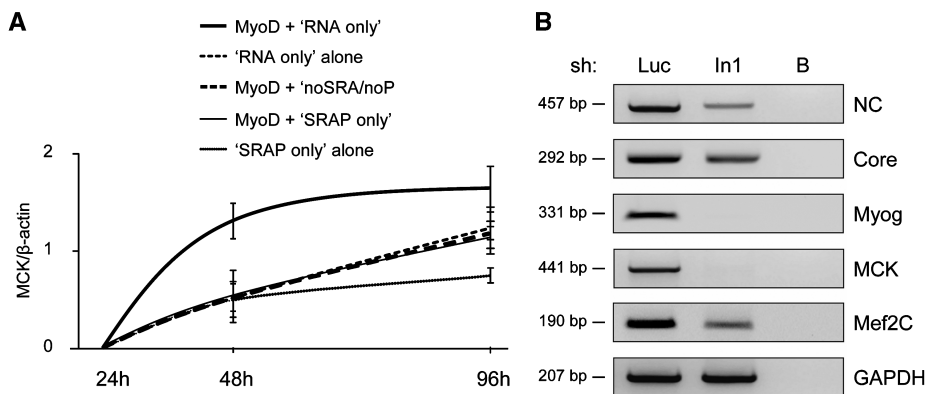

Figure 7. SRAP interacts with SRA RNA through its RRM structure. (**A**) Putative SRA RNP-2 sequence conservation between species. Homo, *H. sapiens*; Bos, *B. taurus*; Mus, *M. musculus*; Rattus, *R. norvegicus*; Sus, *S. scrofa*; Equus, *E. caballus*; Macaca, *M. mulata*; RNP, RRM RNA binding domain; mRNP, mutation performed in the putative SRA RNP motif; Ψ, aliphatic residues; X, any residue. (**B**) NChIP was carried out using V5 antibody on transiently transfected HEK-293 with 'FL' (allowing SRA RNA and SRAP synthesis), mRNP (allowing translation of a RNP-mutated version of SRAP) and 'ΔATG mRNP' (no SRAP-V5 is translated, used as a negative control). PCR using specific primers for the SRA core region (E2/E4) and control GAPDH were performed after reverse transcription of precipitated materials. IP, immunoprecipitation; in, 5% input.

Figure 8. SRA ncRNA and not SRAP enhance myogenic differentiation or conversion. (**A**) SRA ncRNA enhances the ability of MyoD to induce differentiation of C3H/10T1/2 cells. C3H/10T1/2 cells were transiently transfected with MyoD expression vector in combination with the SRA 'RNA only', the 'SRAP only' or with the 'noSRA/noP' expression vectors. As a negative control, cells were also transfected with SRA expression vectors alone. MCK and β-actin mRNA relative expression from at least two independent experiments were measured 0, 24 and 72 h after serum withdrawal as described in the 'Materials and Methods' section. (**B**) Human myogenic cells were transfected with shRNA (sh) targeting SRA intron 1-containing ncRNA isoforms (In1) or luciferase transcripts absent from cells (Luc), and switched to serum-deprived media to induce myogenic differentiation during 24 h. Total RNA was extracted by the Trizol method, reverse transcribed and amplified by PCR as described in the 'Materials and Methods' section. NC, non-coding SRA; core, core SRA elements; Myog, Myogenin; MCK, muscle creatine kinase; B, blank PCR.

extent, Mef2C was detected as early as 24 h after muscle differentiation was induced by serum withdrawal in cells transfected with control shRNA. In contrast, their expression was not induced in cells depleted of SRA ncRNA isoforms containing intron 1 (Figure 8B).

Taken as a whole, SRA RNA isoforms retaining intron 1, i.e. true SRA ncRNA, were able to promote myogenic differentiation or conversion, in contrast to SRAP-coding isoforms. In addition, in agreement with opposing functions ascribed to SRA coding or non-coding isoforms, shRNA directed against distinct portions of SRA RNA have distinct effects on the expression of MyoD target genes and myogenic differentiation.

DISCUSSION

The steroid receptor RNA activator (SRA) was originally described as a functional ncRNA involved in the regulation of gene expression by steroid receptors (7,9–16). Subsequently, several transcripts, one of which encodes a protein, have been characterized (18–21). Thus, any attempt to decipher SRA function as a transcriptional co-regulator must take into account this remarkable bifunctional characteristic consisting of mRNA that encodes a protein and potentially several ncRNA. Because SRA RNA is highly expressed in the skeletal muscle (7) and was reported to co-activate the master regulator of muscle differentiation MyoD, we focused on this tissue to perform an exhaustive analysis of SRA RNA isoforms present in immortal and primary human myogenic cells, from both healthy and myotonic dystrophy patients cells. We reported here opposing functions for the two broad families of SRA isoforms regarded as ORF-containing or non-protein-coding SRA transcripts, on both MyoD transcriptional activity and myogenic differentiation. The finding that SRA RNA acted as an enhancer of myogenic differentiation through the co-activation of MyoD activity was legitimate only in the case of true non-protein coding SRA RNA, i.e. in which hallmarks associated with coding sequences were disrupted. Indeed, we provide evidence that the protein SRAP prevents SRA RNA-dependant co-activation through interaction with its ncRNA counterpart. In addition, we identified an RRM-like RNA binding domain in SRAP required for its interaction with the functional substructure of SRA RNA, STR7. Together, these data suggest that the overall effect of SRA on both MyoD activity and myogenic differentiation results from the correct balance between coding and non-coding SRA molecules.

Multiple SRA isoforms classified into two broad families with regard to their coding capacity

The original report described multiple SRA isoforms, differing in their 5′- and 3′-ends, but sharing a common 'core' region required for the function of these ncRNA as nuclear receptors co-activators (7). Subsequently, we highlighted the fact that the range of SRA isoforms was larger than previously thought. New isoforms containing an additional exon 1 were described, with the peculiarity that these isoforms exhibited a long predicted ORF and

ATG within the new exon, compatible with their competence to encode a protein SRAP (19,20). In addition, we identified isoforms that displayed retention of intron 1 or splicing of exon 3, leading to the disruption of the ORF (21,35). We now have broadened the scope of SRA molecules. Most of the isoforms we identified in human myogenic cells matched our analysis of public EST databases. Indeed, identification of 5′-ends combined to the sequence analysis of systematically cloned SRA transcripts showed that about two-thirds of them contained a disrupted ORF, and therefore belong to the class of true non-protein coding RNA, while the other one-third preserved the features associated with coding sequences, i.e. initiating methionines and a long ORF. This analysis also represents the first experimental evidence for the existence of long SRA RNA isoforms, initiated ∼250–300 bp upstream of the two ATG, and therefore containing the longest exon 1 included in an SRA RNA isoform. While isoforms initiated around the 2 ATG contained or not the SRA ORF, all the longest SRA RNA isoforms presented the features associated with the SRAP coding sequence. Likewise, this is the first report of SRA isoforms deleted of exon 3 present in a normal tissue, since it was primarily thought to be a characteristic feature of breast cancer tissues (18).

The balance between coding and non-coding SRA isoforms varies during myogenic differentiation of primary human cells

We described herein a variation in the ratio between non-coding and coding SRA isoforms, specifically associated with myogenic differentiation capacity of primary human muscle cells and concomitant with the appearance of early markers of differentiation. In contrast, we showed that this ratio was unchanged in congenital myotonic dystrophy muscle satellite cells, in which muscle differentiation and maturation is impaired (26). While the global amount of SRA transcripts, measured through the presence of the core sequence exhibited by the majority of SRA transcripts, was unchanged during the differentiation process [this work and (15)], we refined this observation and showed that an increase in non-protein-coding RNA species paralleled a decrease in SRAP coding isoforms during the course of myogenic differentiation of healthy muscle satellite cells but not of cells isolated from DM1 patients. We previously proposed that alternative splicing of intron 1 was one mechanism used to regulate the balance between coding and non-coding SRA isoforms (21,35). Since SRA was shown to enhance cellular proliferation or differentiation processes (15,42,43), deregulation of distinct SRA isoforms may have significant implications in regulating the proliferation/differentiation balance in primary cells. For example, several reports have described an overexpression of SRA RNA in breast, uterus and ovarian tumours (17,18,42,44) and proposed its participation in tumourigenesis and tumour progression. In contrast, the expression of SRAP in a small subset of patients with primary breast cancer subsequently treated with Tamoxifen correlated with an overall better survival of the patients (22). In satellite muscle cells isolated from

myotonic dystrophy patients, in which disease phenotype has been directly linked to disrupted regulation of alternative splicing [reviewed in (45,46)], we found that the levels of intron 1-retaining non-coding SRA isoforms remained unchanged under myogenic differentiation conditions. These data are in correlation with their reported poor capacity of maturation and fusion (26) and suggest that impaired splicing of SRA may represent one of the splicing events that are misregulated in this pathology.

Alternative splicing is widely used to generate protein diversity and to control gene expression in many biological processes (47), including cell fate determination. In the light of these data, it is tempting to conclude that alternative splicing may be used to regulate RNA diversity in the case of a bifunctional transcript, to affect the balance between non-coding and coding RNA species, which may be important in the regulation of cellular processes. As mentioned earlier, this balance is disrupted in diseased situations.

SRAP contains an RNA recognition motif and interacts with SRA RNA

SRA RNA has been originally described as a co-regulator of nuclear receptor (NR), and then shown to activate MyoD, whereas it failed to potentiate other transcription factors, suggesting that SRA is not a general transcriptional co-activator (7,38). Here we added GATA-3 and PPARγ to the list of transcription factors potentiated by SRA, whereas the activity of Ets2 was unchanged in the presence of SRA RNA. Interestingly, the activity of Ets2, and not that of GATA-3 and PPARγ, was potentiated by SRAP, stressing again the need to precisely determine which SRA molecule is considered.

Most of these factors, except for thyroid receptors (TR) and the orphan receptor Dax-1, do not contain an RNA recognition motif. Therefore, their interaction with SRA RNA must be indirect through interaction with RNA binding partners like the RNA helicases p68/p72 in the case of MyoD (15). At the RNA level, a set of discrete stem-loop structures, forming sub-structural domains (STR), were described within SRA RNA (11), and subsequently implicated in the SRA function and formation of SRA-containing ribonucleoprotein complexes. For example, STR1 in combination with STR7 are required for co-activation of steroid hormone receptors (11). The observation that pseudouridylation of the STR5 sub-structure affects SRA function (14,38) strongly suggested that this post-transcriptional modification might also participate in the establishment of functional interfaces between SRA and its different partners.

We demonstrated here that SRAP was able to interact with its SRA RNA counterpart through the RRM-like domain that we described in SRAP protein and the functional substructure STR7 of the RNA. The STR7 sub-structure specifically mediates the interaction of SRA RNA with SRAP (this study) as well as with SHARP (SMRT/HDAC1 associated repressor protein) and SLIRP (SRA stem-loop interacting RNA binding protein). SHARP and SLIRP contain RRM domains and were identified as co-repressors of nuclear receptor activity through binding to SRA RNA (10,12). Interestingly, these three cases represent examples where transcriptional repressors bind to SRA RNA through STR7 to prevent its co-activator function.

Differential function of SRA RNA and SRAP

SRA was shown to operate as a ncRNA activator of NR and MyoD in experiments using the core SRA devoid of protein coding features (7,9–16). When longer RNA were employed, a clear reduced activity was observed [this study, (22,27,39)] suggesting that the bifunctionality of the RNA prevented the activity of one or the other molecules. Here we demonstrated that disrupting coding features of SRA, i.e. by mutating ATG or disrupting the coding sequence while preserving SRA RNA substructures, led to the known 2.5-fold activation of MyoD activity. In addition, forcing expression of a bifunctional SRA RNA had little effect on MyoD-dependant myogenic conversion of non-muscle cell line, whereas its non-protein coding counterpart accelerated the process with the appearance of muscle markers as early as 24 h after transfection. These findings are consistent with studies using the core SRA in murine muscle cells (15) or a SRA ncRNA construct on the activity of NR (22,27).

Along the same lines, down-regulation of all SRA transcripts using interference RNA against the core led to the inhibition of muscle differentiation, although these experiments did not discriminate between the two classes of SRA molecules. We therefore reasoned that if SRAP prevents SRA RNA-mediated activation of MyoD, changing the balance between isoforms producing either the protein or an ncRNA is predicted to have an impact on MyoD activity and subsequently on muscle differentiation. Using interference to specifically target SRA intron 1 containing isoforms, i.e. isoforms in which the ORF is disrupted, to favour SRAP-producing isoforms, we indeed delayed the appearance of muscle-specific markers.

In essence, SRAP does not display intrinsic activities on myogenic differentiation nor interacts with MyoD. In addition, forced expression of SRAP in non-muscle cell lines did not influence MyoD-driven trans-differentiation. However, SRAP has indirect negative effects on both ncRNA SRA-mediated activation of MyoD and myogenic differentiation or conversion, through binding to SRA RNA and prevention of its co-activator function. Contribution of both SRA RNA and SRAP in the same pathway is not a unique case. Additional examples of bifunctional RNA have recently emerged, such as VegT in *Xenopus*, Oskar in *Drosophila*, SgrS in *Escherichia coli* and likely some others that will be identified in a near future (48). To date, all these molecules appear to be involved in the same pathway at both ncRNA and protein level.

It is noteworthy that studying one or the other of the various aspects of a bifunctional RNA necessitates being able to discriminate between each of its components. Without the discrimination between distinct functions of SRA molecules, we could have missed the important message that the protein translated from SRA RNA prevents the co-activator function of its ncRNA

counterpart. Another recent striking example is given by p53 that interacts with Mdm2 at both protein and RNA level with opposing effects but through domains encoded by the same genomic region (49). Therefore, some previous conclusions may have to be revised in the light of the reported differing functions of SRA molecules in this study. SRA is the first example of a mammalian bifunctional RNA and we believe that our thorough analysis of SRA isoforms with distinct molecular and cellular functions will ease the functional characterization of SRA molecules and other bifunctional RNA in various normal or pathological conditions.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

FUNDING

Conflict of interest statement. None declared.

REFERENCES

1. Mattick,J.S. and Makunin,I.V. (2006) Non-coding RNA. *Hum. Mol. Genet.*, **15**, R17–R29.
2. Mattick,J.S. (2001) Non-coding RNAs: the architects of eukaryotic complexity. *EMBO Rep.*, **2**, 986–991.
3. Mercer,T.R., Dinger,M.E. and Mattick,J.S. (2009) Long non-coding RNAs: insights into functions. *Nat. Rev. Genet.*, **10**, 155–159.
4. Espinoza,C.A., Allen,T.A., Hieb,A.R., Kugel,J.F. and Goodrich,J.A. (2004) B2 RNA binds directly to RNA polymerase II to repress transcript synthesis. *Nat. Struct. Mol. Bio.*, **11**, 822–829.
5. Yik,J.H., Chen,R., Nishimura,R., Jennings,J.L., Link,A.J. and Zhou,Q. (2003) Inhibition of P-TEFb (CDK9/Cyclin T) kinase and RNA polymerase II transcription by the coordinated actions of HEXIM1 and 7SK snRNA. *Mol. Cell*, **12**, 971–982.
6. Willingham,A.T., Orth,A.P., Batalov,S., Peters,E.C., Wen,B.G., za-Blanc,P., Hogenesch,J.B. and Schultz,P.G. (2005) A strategy for probing the function of noncoding RNAs finds a repressor of NFAT. *Science*, **309**, 1570–1573.
7. Lanz,R.B., McKenna,N.J., Onate,S.A., Albrecht,U., Wong,J., Tsai,S.Y., Tsai,M.J. and O'Malley,B.W. (1999) A steroid receptor coactivator, SRA, functions as an RNA and is present in an SRC-1 complex. *Cell*, **97**, 17–27.
8. Leygue,E. (2007) Steroid receptor RNA activator (SRA1): unusual bifaceted gene products with suspected relevance to breast cancer. *Nucl. Recept. Signal*, **5**, e006.
9. Charette,M. and Gray,M.W. (2000) Pseudouridine in RNA: what, where, how, and why. IUBMB. *Life*, **49**, 341–351.
10. Hatchell,E.C., Colley,S.M., Beveridge,D.J., Epis,M.R., Stuart,L.M., Giles,K.M., Redfern,A.D., Miles,L.E., Barker,A., MacDonald,L.M. *et al.* (2006) SLIRP, a small SRA binding protein, is a nuclear receptor corepressor. *Mol. Cell*, **22**, 657–668.
11. Lanz,R.B., Razani,B., Goldberg,A.D. and O'Malley,B.W. (2002) Distinct RNA motifs are important for coactivation of steroid hormone receptors by steroid receptor RNA activator (SRA). *Proc. Natl Acad. Sci. USA*, **99**, 16081–16086.
12. Shi,Y., Downes,M., Xie,W., Kao,H.Y., Ordentlich,P., Tsai,C.C., Hon,M. and Evans,R.M. (2001) Sharp, an inducible cofactor that integrates nuclear receptor repression and activation. *Genes Dev.*, **15**, 1140–1151.
13. Watanabe,M., Yanagisawa,J., Kitagawa,H., Takeyama,K., Ogawa,S., Arao,Y., Suzawa,M., Kobayashi,Y., Yano,T., Yoshikawa,H. *et al.* (2001) A subfamily of RNA-binding DEAD-box proteins acts as an estrogen receptor alpha coactivator through the N-terminal activation domain (AF-1) with an RNA coactivator, SRA. *EMBO J.*, **20**, 1341–1352.
14. Zhao,X., Patton,J.R., Ghosh,S.K., Fischel-Ghodsian,N., Shen,L. and Spanjaard,R.A. (2007) Pus3p- and Pus1p-dependent pseudouridylation of steroid receptor RNA activator controls a functional switch that regulates nuclear receptor signaling. *Mol. Endocrinol.*, **21**, 686–699.
15. Caretti,G., Schiltz,R.L., Dilworth,F.J., Di,P.M., Zhao,P., Ogryzko,V., Fuller-Pace,F.V., Hoffman,E.P., Tapscott,S.J. and Sartorelli,V. (2006) The RNA helicases p68/p72 and the noncoding RNA SRA are coregulators of MyoD and skeletal muscle differentiation. *Dev. Cell*, **11**, 547–560.
16. Caretti,G., Lei,E.P. and Sartorelli,V. (2007) The DEAD-box p68/p72 proteins and the noncoding RNA steroid receptor activator SRA: eclectic regulators of disparate biological functions. *Cell Cycle*, **6**, 1172–1176.
17. Hussein-Fikret,S. and Fuller,P.J. (2005) Expression of nuclear receptor coregulators in ovarian stromal and epithelial tumours. *Mol. Cell Endocrinol.*, **229**, 149–160.
18. Leygue,E., Dotzlaw,H., Watson,P.H. and Murphy,L.C. (1999) Expression of the steroid receptor RNA activator in human breast tumors. *Cancer Res.*, **59**, 4190–4193.
19. Chooniedass-Kothari,S., Emberley,E., Hamedani,M.K., Troup,S., Wang,X., Czosnek,A., Hube,F., Mutawe,M., Watson,P.H. and Leygue,E. (2004) The steroid receptor RNA activator is the first functional RNA encoding a protein. *FEBS Lett.*, **566**, 43–47.
20. Emberley,E., Huang,G.J., Hamedani,M.K., Czosnek,A., Ali,D., Grolla,A., Lu,B., Watson,P.H., Murphy,L.C. and Leygue,E. (2003) Identification of new human coding steroid receptor RNA activator isoforms. *Biochem. Biophys. Res. Commun.*, **301**, 509–515.
21. Hube,F., Guo,J., Chooniedass-Kothari,S., Cooper,C., Hamedani,M.K., Dibrov,A.A., Blanchard,A.A., Wang,X., Deng,G., Myal,Y. *et al.* (2006) Alternative splicing of the first intron of the steroid receptor RNA activator (SRA) participates in the generation of coding and noncoding RNA isoforms in breast cancer cell lines. *DNA Cell Biol.*, **25**, 418–428.
22. Chooniedass-Kothari,S., Hamedani,M.K., Troup,S., Hube,F. and Leygue,E. (2006) The steroid receptor RNA activator protein is expressed in breast tumor tissues. *Int. J. Cancer*, **118**, 1054–1059.
23. Bouzinba-Segard,H., Guais,A. and Francastel,C. (2006) Accumulation of small murine minor satellite transcripts leads to impaired centromeric architecture and function. *Proc. Natl Acad. Sci. USA*, **103**, 8709–8714.
24. Zhu,C.H., Mouly,V., Cooper,R.N., Mamchaoui,K., Bigot,A., Shay,J.W., Di Santo,J.P., Butler-Browne,G.S. and Wright,W.E. (2007) Cellular senescence in human myoblasts is overcome by

human telomerase reverse transcriptase and cyclin-dependent kinase 4: consequences in aging muscle and therapeutic strategies for muscular dystrophies. *Aging Cell*, **6**, 515–523.

25. Montarras,D., Morgan,J., Collins,C., Relaix,F., Zaffran,S., Cumano,A., Partridge,T. and Buckingham,M. (2005) Direct isolation of satellite cells for skeletal muscle regeneration. *Science*, **309**, 2064–2067.

26. Furling,D., Coiffier,L., Mouly,V., Barbet,J.P., St Guily,J.L., Taneja,K., Gourdon,G., Junien,C. and Butler-Browne,G.S. (2001) Defective satellite cells in congenital myotonic dystrophy. *Hum. Mol. Genet.*, **10**, 2079–2087.

27. Chooniedass-Kothari,S., Vincett,D., Yan,Y., Cooper,C., Hamedani,M.K., Myal,Y. and Leygue,E. (2010) The protein encoded by the functional steroid receptor RNA activator is a new modulator of ER alpha transcriptional activity. *FEBS Lett.*, **584**, 1174–1180.

28. Ferri,F., Bouzinba-Segard,H., Velasco,G., Hube,F. and Francastel,C. (2009) Non-coding murine centromeric transcripts associate with and potentiate Aurora B kinase. *Nucleic Acids Res.*, **37**, 5071–5080.

29. Velasco,G., Hube,F., Rollin,J., Neuillet,D., Philippe,C., Bouzinba-Segard,H., Galvani,A., Viegas-Pequignot,E. and Francastel,C. (2010) Dnmt3b recruitment through E2F6 transcriptional repressor mediates germ-line gene silencing in murine somatic tissues. *Proc. Natl Acad. Sci. USA*., **107**, 9281–9286.

30. Hube,F., Reverdiau,P., Iochmann,S., Cherpi-Antar,C. and Gruel,Y. (2003) Characterization and functional analysis of TFPI-2 gene promoter in a human choriocarcinoma cell line. *Thromb. Res.*, **109**, 207–215.

31. Hube,F., Chooniedass-Kothari,S., Hamedani,M.K., Miksicek,R.J., Leygue,E. and Myal,Y. (2006) Identification of an octamer-binding site controlling the activity of the small breast epithelial mucin gene promoter. *Front. Biosci.*, **11**, 2483–2495.

32. Kim,N., Shin,S. and Lee,S. (2004) ASmodeler: gene modeling of alternative splicing from genomic alignment of mRNA, EST and protein sequences. *Nucleic Acids Res.*, **32**, W181–W186.

33. Zuker,M. (2003) Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res.*, **31**, 3406–3415.

34. Kawashima,H., Takano,H., Sugita,S., Takahara,Y., Sugimura,K. and Nakatani,T. (2003) A novel steroid receptor co-activator protein (SRAP) as an alternative form of steroid receptor RNA-activator gene: expression in prostate cancer cells and enhancement of androgen receptor activity. *Biochem. J.*, **369**, 163–171.

35. Cooper,C., Guo,J., Yan,Y., Chooniedass-Kothari,S., Hube,F., Hamedani,M.K., Murphy,L.C., Myal,Y. and Leygue,E. (2009) Increasing the relative expression of endogenous non-coding steroid receptor RNA activator (SRA) in human breast cancer cells using modified oligonucleotides. *Nucleic Acids Res.*, **37**, 4518–4531.

36. Klinge,C.M., Jernigan,S.C., Mattingly,K.A., Risinger,K.E. and Zhang,J. (2004) Estrogen response element-dependent regulation of transcriptional activation of estrogen receptors alpha and beta by coactivators and corepressors. *J. Mol. Endocrinol.*, **33**, 387–410.

37. Xu,B., Yang,W.H., Gerin,I., Hu,C.D., Hammer,G.D. and Koenig,R.J. (2009) Dax-1 and steroid receptor RNA activator (SRA) function as transcriptional coactivators for steroidogenic factor 1 in steroidogenesis. *Mol. Cell. Biol.*, **29**, 1719–1734.

38. Zhao,X., Patton,J.R., Davis,S.L., Florence,B., Ames,S.J. and Spanjaard,R.A. (2004) Regulation of nuclear receptor activity by a pseudouridine synthase through posttranscriptional modification of steroid receptor RNA activator. *Mol. Cell*, **15**, 549–558.

39. Chooniedass-Kothari,S., Hamedani,M.K., Auge,C., Wang,X., Carascossa,S., Yan,Y., Cooper,C., Vincett,D., Myal,Y., Jalaguier,S. *et al.* (2010) The steroid receptor RNA activator protein is recruited to promoter regions and acts as a transcriptional repressor. *FEBS Lett.*, **584**, 2218–2224.

40. Kurisu,T., Tanaka,T., Ishii,J., Matsumura,K., Sugimura,K., Nakatani,T. and Kawashima,H. (2006) Expression and function of human steroid receptor RNA activator in prostate cancer cells: role of endogenous hSRA protein in androgen receptor-mediated transcription. *Prostate Cancer Prostatic Dis.*, **9**, 173–178.

41. Kielkopf,C.L., Lucke,S. and Green,M.R. (2004) U2AF homology motifs: protein recognition in the RRM world. *Genes Dev.*, **18**, 1513–1526.

42. Lanz,R.B., Chua,S.S., Barron,N., Soder,B.M., DeMayo,F. and O'Malley,B.W. (2003) Steroid receptor RNA activator stimulates proliferation as well as apoptosis in vivo. *Mol. Cell. Biol.*, **23**, 7163–7176.

43. Xu,B. and Koenig,R.J. (2004) An RNA-binding domain in the thyroid hormone receptor enhances transcriptional activation. *J. Biol. Chem.*, **279**, 33051–33056.

44. Murphy,L.C., Simon,S.L., Parkes,A., Leygue,E., Dotzlaw,H., Snell,L., Troup,S., Adeyinka,A. and Watson,P.H. (2000) Altered expression of estrogen receptor coregulators during human breast tumorigenesis. *Cancer Res.*, **60**, 6266–6271.

45. Amack,J.D. and Mahadevan,M.S. (2004) Myogenic defects in myotonic dystrophy. *Dev. Biol.*, **265**, 294–301.

46. Lee,J.E. and Cooper,T.A. (2009) Pathogenic mechanisms of myotonic dystrophy. *Biochem. Soc. Trans.*, **37**, 1281–1286.

47. Mollet,I.G., Ben-Dov,C., Felicio-Silva,D., Grosso,A.R., Eleuterio,P., Alves,R., Staller,R., Silva,T.S. and Carmo-Fonseca,M. (2010) Unconstrained mining of transcript data reveals increased alternative splicing complexity in the human transcriptome. *Nucleic Acids Res.*, **38**, 4740–4754.

48. Dinger,M.E., Pang,K.C., Mercer,T.R. and Mattick,J.S. (2008) Differentiating protein-coding and noncoding RNA: challenges and ambiguities. *PLoS Comput. Biol.*, **4**, e1000176.

49. Candeias,M.M., Malbert-Colas,L., Powell,D.J., Daskalogianni,C., Maslon,M.M., Naski,N., Bourougaa,K., Calvo,F. and Fahraeus,R. (2008) P53 mRNA controls p53 activity by managing Mdm2 functions. *Nat. Cell Biol.*, **10**, 1098–1105.