Mechanistic studies of a small-molecule modulator of SMN2 splicing

Jingxin Wang, Peter G. Schultz, and Kristen A. Johnson

*California Institute for Biomedical Research, La Jolla, CA 92037; 
†Department of Chemistry, The Scripps Research Institute, La Jolla, CA 92037; and †Skaggs Institute for Chemical Biology, The Scripps Research Institute, La Jolla, CA 92037

Contributed by Peter G. Schultz, April 2, 2018 (sent for review January 9, 2018; reviewed by Peter B. Dervan and Adrian R. Krainer)

RG-7916 is a first-in-class drug candidate for the treatment of spinal muscular atrophy (SMA) that functions by modulating pre-mRNA splicing of the SMN2 gene, resulting in a 2.5-fold increase in survival of motor neuron (SMN) protein level, a key protein lacking in SMA patients. RG-7916 is currently in three interventional phase 2 clinical trials for various types of SMA. In this report, we show that SMN-C2 and -C3, close analogs of RG-7916, act as selective RNA-binding ligands that modulate pre-mRNA splicing. Chemical proteomic and genomic techniques reveal that SMN-C2 directly binds to the AGGAAG motif on exon 7 of the SMN2 pre-mRNA, and promotes a conformational change in two to three unpaired nucleotides at the junction of intron 6 and exon 7 in both in vitro and in-cell models. This change creates a new functional binding surface that increases binding of the splicing modulators, far upstream element binding protein 1 (FUBP1) and its homolog, KH-type splicing regulatory protein (KHSRP), to the SMN2-C2/3–SMN2 pre-mRNA complex and enhances SMN2 splicing. These findings underscore the potential of small-molecule drugs to selectively bind RNA and modulate pre-mRNA splicing as an approach to the treatment of human disease.

**Significance**

The development of small-molecule therapeutics that act by targeting defined DNA or RNA sequences associated with human disease remains a challenge. RG-7916, a small-molecule drug candidate for the treatment of spinal muscular atrophy (SMA), selectively regulates the alternative splicing (AS) of the SMN2 gene. Herein, we show that SMN-C2 and -C3, close analogs of RG-7916, act by binding SMN2 pre-mRNA and thereby increasing the affinity of the RNA binding proteins far upstream element binding protein 1 (FUBP1) and KH-type splicing regulatory protein (KHSRP) to the SMN2 pre-mRNA complex. These results suggest that nucleic acid targeted small molecules may have untapped potential for modulating disease processes at the level of pre-mRNA splicing.

Author contributions: J.W., P.G.S., and K.A.J. designed research; J.W. performed research; J.W., P.G.S., and K.A.J. contributed new reagents/analytic tools; J.W. analyzed data; and J.W., P.G.S., and K.A.J. wrote the paper.

Reviewers: P.B.D., California Institute of Technology; and A.R.K., Cold Spring Harbor Laboratory.

The authors declare no conflict of interest.

Published under the PNAS license.

Data deposition: Next-generation sequencing data reported for Chem-CLIP and in-cell SHAPE-MaP in this paper have been deposited in the Sequence Read Archive (SRA) database, https://www.ncbi.nlm.nih.gov/sra (accession no. SRP126430).

To whom correspondence may be addressed. Email: schultz@scripps.edu or kjohnson@calib.org.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1800260115/-/DCSupplemental.

www.pnas.org/cgi/doi/10.1073/pnas.1800260115

**RG-7916** | RNA splicing | spinal muscular atrophy | RNA binding | chemical proteomics

**SPINAL MUSCULAR ATROPHY** is one of the most common lethal genetic diseases in newborns (1). In the most severe form of SMA (type I), infants usually do not survive beyond their first 2 y of life due to progressive hypotonia, leading to respiratory failure (2). The cause of SMA in most type I patients is a recessive homozygous deletion of the survival of motor neuron (SMN) 1 gene in chromosome 5. With a reduced level of functional SMN protein, the size of motor neurons is smaller, eventually causing muscle weakness (1). The exact mechanism of SMN in motor neuron maintenance and survival has not been fully elucidated. One strategy to treat SMA is to promote increased expression of the endogenous gene SMN2, which is nearly identical to SMN1, to compensate for loss of the latter. However, the endogenous expression of SMN from SMN2 is ~85% reduced (compared with SMN1) primarily due to a pre-mRNA splicing error (3). The major transcription product of SMN2 leads to a shorter and nonfunctional SMN protein primarily because exon 7 of SMN2 pre-mRNA (abbreviated throughout as exon 7) is skipped in the process of splicing (4), as a result of a C-to-T transition at position +6 on exon 7 (4). Despite the challenges associated with the development of small-molecule therapeutics targeting nucleic acids (5–12), several exon 7 splicing regulators have been identified including the splicing activator, SRSF1 (13), the splicing repressor, hnRNP A1 (14), an SR-like protein, Tra2B1 (15), and various members in hnRNP family (16). The secondary structure or higher-order folding of SMN2 pre-mRNA is also crucial for accurate mRNA splicing: a stem-loop structure at the 5′-splice site (5′-ss) of exon 7, namely terminal stem-loop 2 (TS2L2), has been shown to be a negative regulatory element for exon 7 splicing (17).

To date, Nusinersen is the only Food and Drug Administration-approved therapeutic for SMA and has been shown to improve motor function in SMA patients after 15 mo of treatment (18). Nusinersen is an antisense oligomer (ASO) that blocks an intronic splicing silencer (ISS) binding site at intron 7, leading to an increase of SMN protein levels in motor neurons (19). Significant efforts have been made to develop an oral small molecule that can correct the splicing error of SMN2 exon 7 (12, 20, 21). The pyrido-pyrimidinone RG-7916 is currently in phase 2 clinical trials for various types of SMA. A pair of close analogs of RG-7916, SMN-C2 and SMN-C3, correct exon 7 splicing with an EC50 ~ 100 nM, presumably through the same mechanism (Fig. 1 A and B) (21). In type 2 and 3 SMA patients, RG-7916 was shown to induce a 2.5-fold increase in SMN protein levels in peripheral blood cells (22). To understand the mechanism of SMN-C2/C3 in exon 7 splicing and to guide the rational design of future splicing modulators, we undertook an effort to elucidate the cellular target of SMN-C2/C3. Chemical proteomic and genomic studies revealed that SMN-C2/C3 directly binds to the AGGAAG motif on exon 7 of the SMN2 pre-mRNA, and promotes a conformational change in two to three unpaired nucleotides at the junction of intron 6 and exon 7 in both in vitro and in-cell models. We hypothesize that this change increases binding of the splicing modulators, FUBP1 and its homolog, KHSRP, to the SMN2-C2/3–SMN2 pre-mRNA complex and enhances SMN2 splicing. SMN-C2/C3 has been documented to selectively regulate splicing of SMN2 along with only a handful other genes, such as STRN3.
Therefore, the direct target of SMN-C3 is unlikely to be the global splicing machinery. We hypothesized two potential targets for SMN-C3: (i) a sequence of RNA on or close to exon 7, or (ii) a splicing regulatory protein or protein complex that is specific to exon 7. We therefore carried out a pull-down experiment to determine whether SMN-C3 directly binds to an RNA target using the photo–cross-linking probe, SMN-C2-BD (Fig. 1C), in which a biotin-diazirine bifunctional handle replaces the ethyl group of SMN-C2. SMN-C2-BD is ~20-fold less active than its parent compound as demonstrated by Western blot analyses of SMN protein in SMNΔ7 mouse astrocytes (SI Appendix, Fig. S1).

To capture the interacting RNA fragment, 293T cells were treated with 2 μM SMN-C2-BD, in the absence or presence of 50× SMN-C3, and then cross-linked by irradiation (365 nm) of the live cells. Total RNA was extracted, pulled down by streptavidin beads, washed, and bound RNA was released under denaturing conditions. The released RNA fraction was reverse

![Fig. 1](image-url) (A) The structures of SMN-C2, SMN-C3, and SMN-C5 (the structure of RG-7916 has not been disclosed); (B) SMN-C2, SMN-C3, and SMN-C5 correct exon 7 splicing. (C) Structure of a biotin-diazirine bifunctional probe, SMN-C2-BD.

![Fig. 2](image-url) (A) A purine-rich RNA binding motif of SMN-C2 identified by MEME analysis (24) with an E value of 1.3 × 10^-151. (B) Sequences of synthetic RNA 15-mers. The coverage of each 15-mer is indicated with a bar above or below the sequences. Exon 7 is highlighted in bold with the putative binding sequence AGGAAG of SMN-C2 and a similar AAGGAG sequence at the 5′-ss in red and green, respectively. (C) Fluorescence polarization assays with SMN-C2 (200 nM) and seven RNA 15-mers that cover the whole exon 7, part of the polypyrimidine tract of intron 6, and the junction of exon 7/intron 7 (n = 2). The K_d values of SMN-C2 binding to oligo-4 and -7 were determined to be 16 ± 2 and 46 ± 3 μM, respectively.
Fig. 3. (A) Structures of SMN-C2–Phen and SMN-C2–GGH complexed with copper. (B) SMN-C2–Phen–Cu and SMN-C2–GGH–Cu exclusively cleave position +32 of exon 7. 1–3, cleavage reaction with 400, 40, and 0 nM SMN-C2–Phen–Cu; 4–6, cleavage reaction with 400, 40, and 0 nM SMN-C2–GGH–Cu; 7, RNase T1 ladder (G); 8, alkaline hydrolysis ladder; 9, purified 5′-32P-labeled RNA. The ribonuclease reaction consisted of 60 mM Hepes, pH 8.0, 20 mM MgCl2, 0.2 μg/mL yeast tRNA, 0.5 M KCl, 10 mM 5′-32P-labeled RNA, 0.1% H2O2, 1 mM sodium ascorbate, and copper-complexed chemical probe at different concentrations. See SI Appendix for the full sequence of RNA.
transcribed to cDNA by random hexamer extension and analyzed by sequencing. Using the competitively treated cells as reference control, alignment of the sequencing reads of the probe-only sample to the human genome did not reveal an enrichment of exon 7. However, we were able to identify a binding motif for SMN-C2-BD in a genome-wide analysis of the captured pre-mRNA. A de novo motif searching for common sequences was performed within the peaks of the aligned reads using MEME motif discovery algorithm (24). A purine-rich binding motif, GAGGAAGA (Fig. 2A), was discovered with a satisfying E value of 1.3 × 10^{-15}.

This putative binding sequence highly resembled the +24 to +29 (AGGAAGA) region of exon 7 (Fig. 2B). To validate the binding specificity, seven RNA 15-mers were synthesized, which cover the entire exon 7 and adjacent intron 6 and 7 regions (Fig. 2B). Each adjacent oligomer was designed to have a 5-nt overlap. An analog of SMN-C3, SMN-C2, contains a coumarin fluorophore so that fluorescence polarization could be used to measure the binding affinity between SMN-C2 and its target (Fig. 2C). As expected, the sequence containing the putative binding site (oligo-4, Fig. 2B and C), AGGAAGA, showed more than 10-fold decrease in K_d compared with sequences 2, 3, 5, and 6 (Fig. 2B and C). The lower binding affinity of oligo-3 to SMN-C2 suggests that high purine content alone is probably not an SMN-C2 recognition pattern. The pyrimidine-rich sequence, oligo-1, showed the lowest binding affinity to SMN-C2 (Fig. 2B and C). It has been shown that SMN-C5 also binds to an RNA duplex of 5'-ss of exon 7 and U1 snRNA (23). Although oligo-7 at the 5'-ss of exon 7 contains an AAGGAAG sequence that is similar to the putative binding sequence AGGAAGA, the binding affinity of SMN-C2 to oligo-7 (K_d = 46 ± 3 μM) is somewhat lower compared with binding of SMN-C2 to oligo-4 (K_d = 16 ± 2 μM).

Next, we attempted to confirm the binding site of SMN-C2 using an RNA footprinting experiment in the context of a longer RNA sequence containing the entire exon 7. Iron or copper chelates in the presence of hydrogen peroxide have previously been used as an artificial chemical ribonuclease to probe the nucleic acid binding sites of molecules (25). Conjugation of the metal ion chelate to a nucleic acid binding molecule leads to oxidative cleavage of the phosphate sugar backbone adjacent to the binding site (26). Therefore, a conjugate of SMN-C2 and 1,10-phenanthroline, SMN-C2-Phen (Fig. 3A), was synthesized with a two-carbon linker (27). To eliminate a potential bias by the phenanthroline moiety in binding to RNA, another copper ligand, Gly–Gly–His (28), conjugate was also synthesized, SMN-C2–GGH (Fig. 3A). In the presence of sodium ascorbate and H_2O_2, a 5'-32P-labeled 120-nt RNA, which spans exon 7 (54 nt) and adjacent intron 6 and 7 regions, was treated with submicromolar concentrations (40–400 nM) of the SMN-C2–Phen- and SMN-C2–GGH–copper complexes. As predicted, both SMN-C2–Phe–copper and SMN-C2–GGH–copper exclusively cleaved the RNA three nucleotides away (Fig. 3B) from the putative binding site (29), AGGAAGA. This sequence-specific cleavage data confirmed the binding site of SMN-C2.

**SMN-C2 Does Not Disrupt the Secondary Structure of SMN2 Pre-mRNA upon Binding**

After the AGGAAG binding motif was identified for SMN-C2/ C3, the next step was to determine whether any structural alterations are induced on exon 7 by SMN-C2 binding. The selective 2'-hydroxyl acylation analyzed by primer extension (SHAPE) method was used to interrogate the local nucleotide conformation of exon 7. SHAPE uses a 2'-OH alkylation reagent, 2-methylisicotinic acid imidazolide (NAI), to modify the backbone ribose of an RNA sequence. A base-paired nucleotide is less accessible than an unpaired one to NAI modification; as a consequence, reverse transcriptase stops at the ribose-NAI adduct to form a truncated product at the primer extension stage (30). The resulting band intensity of an unpaired nucleotide is higher than a paired one by PAGE analysis of the radiolabeled cDNA fragments.

A 140-nt RNA clone of SMN2 exon 7 and adjacent regions was used as a template. This 140-nt RNA was imbedded within a stabilizing cassette for in vitro SHAPE (30). The in vitro SHAPE-directed modeling revealed the following: (i) a stem-loop structure (loop 2, Fig. 4A and B) at the 5'-ss of exon 7 that was previously reported in literature as TSL2 (17); (ii) a secondary structure at the 3'-ss of exon 7 consisting of two bulges and a loop (Fig. 4A and B), designated previously as TSL1 (17); and (iii) a distal stem-loop structure on intron 7 (SI Appendix, Fig. S2 for full-length SHAPE profile). Addition of SMN-C2 (50 μM) did not change the overall secondary structure of exon 7 as determined by in vitro SHAPE (Fig. 4A, lane 2). Importantly, no band was observed for the sequence complementary to the AGGAAGA binding motif in the in vitro SHAPE analysis, suggesting that binding of SMN-C2 to AGGAAG does not displace the complementary strand. However, the intensity of the bands corresponding to three nucleotides in loop 1 and bulge 2 increased compared with DMSO control (Fig. 4A), suggesting these nucleotides adopt a more flexible conformation in the presence of SMN-C2.

The RNA secondary structure can be different from that in a cellular context primarily because of the equilibrium between naked RNA and RNA/RNA-binding protein complexes (31). Therefore, an in-cell SHAPE-mutation profiling (MaP) experiment was carried out to investigate structural differences in exon 7 upon treatment with SMN-C2. First, 293T cells were transfected with an SMN2 minigene to enrich the target RNA. The cells were subsequently treated with SMN-C2, a less active analog, SMN-C2-Ac (SI Appendix, Fig. S3) or DMSO, and NAI. Total RNA was then reverse transcribed with a gene-specific primer. Replacing Mg^{2+} with Mn^{2+} in the buffer system results in an error-prone reverse transcription at the NAI-addition position, creating mutations (32). Finally, a region of 276 nt containing exon 7 was amplified by PCR with high fidelity for sequencing analysis. Analogous to in vitro SHAPE, a high mutation rate or high SHAPE reactivity translates to an unpaired nucleotide or a more flexible conformation [see SI Appendix, Fig. S4, for SHAPE reactivity and mutation rate calculated by SHAPEmapper software (29) for pre-mRNA (Fig. S4), and full-length SHAPE (Fig. S4)].

The secondary structure within this 276-nt amplicon was predicted by matching the SHAPE reactivity for each nucleotide with a software package, SuperFold (SI Appendix, Figs. S5 and S6) (32). Specifically, the TSL2 region remains unchanged in the in-cell model but TSL1 undergoes some base pair rearrangement around bulge 2 and loop 1 (Fig. 4B). RNA-binding proteins, such as Tra2β1 (15) and hnRNP A1 (14), are known to bind to subsequences within TSL1 and may contribute to the base pair rearrangement. Despite this arrangement, the base pairing between AGGAAG and polypyrimidine tract is retained in the cellular context (Fig. 4B). Like the in vitro SHAPE experiment, treatment with 10 μM SMN-C2 results in significant SHAPE reactivity changes only in TSL1, but not TSL2 (SI Appendix, Fig. S7). An increase of SHAPE reactivity in position 34 and a decrease in position 41 near the AGGAAG binding site, compared with DMSO, indicates a more flexible conformation in position 34 and a more constrained one in position 41 (Fig. 4C). SHAPE reactivity of a less active analog, SMN-C2-Ac, resembles that of DMSO at 10 μM and does not show significant changes in positions 34 and 41 (Dataset S1). In the 276-nt sequence, there is only one other secondary structure upstream of TSL1 which showed significant SHAPE reactivity changes (SI Appendix, Fig. S7). Thus, under both in vitro and in-cell conditions, the base-paired SMN-C2–binding site AGGAAG on exon 7 and its complementary sequence in the polypyrimidine tract on intron 6 forms a stem-loop structure, TSL1, at the junction of intron
6 and exon 7. SMN-C2 treatment only affected the SHAPE reactivity of some unpaired nucleotides on TSL1 but not TSL2 or other areas on exon 7. These changes revealed by SHAPE indicated that SMN-C2 does not disrupt the secondary structure of TSL1, but tunes the conformation of the nucleotides in the bulge and loop of TSL1.

**SMN-C2, SMN2 Pre-mRNA Exon 7, and FUBP1 Form a Ternary Structure**

The interaction of SMN-C2/C3 with its binding motif, AGGAAG, does not explain the selectivity of SMN-C3 as a pre-mRNA splicing modulator. There are more than 15 sequences of AGGAAG in the SMN2 gene alone, in addition to a larger number within the whole human genome. To look for additional cofactors that may contribute to selectivity, a proteomic analysis was performed by photo-cross-linking cell lysates in the presence of 2 μM biotin-diazirine bifunctional probe, SMN-C2-BD, followed by immunoblotting for biotinylated proteins (33, 34). After ammonium sulfate fractionation, the cell lysate was further resolved by two-dimensional SDS/PAGE. Western blots with anti-biotin antibody revealed two associated proteins, FUBP1 and hnRNP A1 (SI Appendix, Fig. S8). FUBP1 and hnRNP A1 are both known to play a role in pre-mRNA splicing. To confirm specificity for SMN-C2/C3, a protein pull-down was carried out using SMN-C2-BD with or without unlabeled SMN-C3 competitor (Fig. 5A). FUBP1 was consistently competed with SMN-C3, whereas hnRNP A1 only showed a basal signal level compared with 1% of the input control (Fig. 5A). The cross-linking of hnRNP A1 likely results from the high abundance of this protein in the nucleus. This result suggests that FUBP1 is a potential protein target for SMN-C3. The pull-down of FUBP1 is dependent on endogenous RNA as evidenced by a reduced FUBP1 signal when treating the protein lysate with a mixture of RNase A/T1 (Fig. 5A).

To confirm binding of SMN-C3 to FUBP1, a cellular thermal shift assay (CETSA) (Fig. 5B) was conducted (35). In CETSA, small molecules bind and protect their specific cellular protein targets from unfolding and aggregation when the cell suspension is heated (36). THP-1 cells were treated with SMN-C3 at various concentrations or DMSO control and heated at 76 °C. After cell lysis, cell debris and aggregated protein were removed by centrifugation. The amount of FUBP1 in the supernatant increased 58% in the presence of 25 μM SMN-C3 compared with DMSO control and was dose dependent as evidenced by Western blot with an anti-FUBP1 antibody (Fig. 5B). These data suggest SMN-C3 and FUBP1 interact in live cells.

Recombinant FUBP1 was then expressed and purified to further probe its interaction with SMN-C2 in vitro. A fluorescence polarization experiment using recombinant FUBP1 and SMN-C2 (Fig. 5C) revealed no apparent binding between 120 and 160 were omitted. Significant polarization compared with FUBP1 in the presence of either a 15-mer containing the AGGAAG binding motif (oligo-4) or a 120-nt RNA covering exon 7 and adjacent regions (Fig. 5C). This result suggests a ternary complex is formed by FUBP1, SMN-C2, and SMN-C3. The pull-down of FUBP1 is dependent on endogenous RNA as evidenced by a reduced FUBP1 signal when treating the protein lysate with a mixture of RNase A/T1 (Fig. 5A).

To confirm binding of SMN-C3 to FUBP1, a cellular thermal shift assay (CETSA) (Fig. 5B) was conducted (35). In CETSA, small molecules bind and protect their specific cellular protein targets from unfolding and aggregation when the cell suspension is heated (36). THP-1 cells were treated with SMN-C3 at various concentrations or DMSO control and heated at 76 °C. After cell lysis, cell debris and aggregated protein were removed by centrifugation. The amount of FUBP1 in the supernatant increased 58% in the presence of 25 μM SMN-C3 compared with DMSO control and was dose dependent as evidenced by Western blot with an anti-FUBP1 antibody (Fig. 5B). These data suggest SMN-C3 and FUBP1 interact in live cells.

Recombinant FUBP1 was then expressed and purified to further probe its interaction with SMN-C2 in vitro. A fluorescence polarization experiment using recombinant FUBP1 and SMN-C2 (Fig. 5C) revealed no apparent binding between 120 and 160 were omitted. Significant polarization compared with FUBP1 in the presence of either a 15-mer containing the AGGAAG binding motif (oligo-4) or a 120-nt RNA covering exon 7 and adjacent regions (Fig. 5C). This result suggests a ternary complex is formed by FUBP1, SMN-C2, and SMN-C3. The pull-down of FUBP1 is dependent on endogenous RNA as evidenced by a reduced FUBP1 signal when treating the protein lysate with a mixture of RNase A/T1 (Fig. 5A).

To confirm binding of SMN-C3 to FUBP1, a cellular thermal shift assay (CETSA) (Fig. 5B) was conducted (35). In CETSA, small molecules bind and protect their specific cellular protein targets from unfolding and aggregation when the cell suspension is heated (36). THP-1 cells were treated with SMN-C3 at various concentrations or DMSO control and heated at 76 °C. After cell lysis, cell debris and aggregated protein were removed by centrifugation. The amount of FUBP1 in the supernatant increased 58% in the presence of 25 μM SMN-C3 compared with DMSO control and was dose dependent as evidenced by Western blot with an anti-FUBP1 antibody (Fig. 5B). These data suggest SMN-C3 and FUBP1 interact in live cells.

Recombinant FUBP1 was then expressed and purified to further probe its interaction with SMN-C2 in vitro. A fluorescence polarization experiment using recombinant FUBP1 and SMN-C2 (Fig. 5C) revealed no apparent binding between 120 and 160 were omitted. Significant polarization compared with FUBP1 in the presence of either a 15-mer containing the AGGAAG binding motif (oligo-4) or a 120-nt RNA covering exon 7 and adjacent regions (Fig. 5C). This result suggests a ternary complex is formed by FUBP1, SMN-C2, and SMN-C3. The pull-down of FUBP1 is dependent on endogenous RNA as evidenced by a reduced FUBP1 signal when treating the protein lysate with a mixture of RNase A/T1 (Fig. 5A).

To confirm binding of SMN-C3 to FUBP1, a cellular thermal shift assay (CETSA) (Fig. 5B) was conducted (35). In CETSA, small molecules bind and protect their specific cellular protein targets from unfolding and aggregation when the cell suspension is heated (36). THP-1 cells were treated with SMN-C3 at various concentrations or DMSO control and heated at 76 °C. After cell lysis, cell debris and aggregated protein were removed by centrifugation. The amount of FUBP1 in the supernatant increased 58% in the presence of 25 μM SMN-C3 compared with DMSO control and was dose dependent as evidenced by Western blot with an anti-FUBP1 antibody (Fig. 5B). These data suggest SMN-C3 and FUBP1 interact in live cells.

Recombinant FUBP1 was then expressed and purified to further probe its interaction with SMN-C2 in vitro. A fluorescence polarization experiment using recombinant FUBP1 and SMN-C2 (Fig. 5C) revealed no apparent binding between 120 and 160 were omitted. Significant polarization compared with FUBP1 in the presence of either a 15-mer containing the AGGAAG binding motif (oligo-4) or a 120-nt RNA covering exon 7 and adjacent regions (Fig. 5C). This result suggests a ternary complex is formed by FUBP1, SMN-C2, and SMN-C3. The pull-down of FUBP1 is dependent on endogenous RNA as evidenced by a reduced FUBP1 signal when treating the protein lysate with a mixture of RNase A/T1 (Fig. 5A).

To confirm binding of SMN-C3 to FUBP1, a cellular thermal shift assay (CETSA) (Fig. 5B) was conducted (35). In CETSA, small molecules bind and protect their specific cellular protein targets from unfolding and aggregation when the cell suspension is heated (36). THP-1 cells were treated with SMN-C3 at various concentrations or DMSO control and heated at 76 °C. After cell lysis, cell debris and aggregated protein were removed by centrifugation. The amount of FUBP1 in the supernatant increased 58% in the presence of 25 μM SMN-C3 compared with DMSO control and was dose dependent as evidenced by Western blot with an anti-FUBP1 antibody (Fig. 5B). These data suggest SMN-C3 and FUBP1 interact in live cells.

Recombinant FUBP1 was then expressed and purified to further probe its interaction with SMN-C2 in vitro. A fluorescence polarization experiment using recombinant FUBP1 and SMN-C2 (Fig. 5C) revealed no apparent binding between 120 and 160 were omitted. Significant polarization compared with FUBP1 in the presence of either a 15-mer containing the AGGAAG binding motif (oligo-4) or a 120-nt RNA covering exon 7 and adjacent regions (Fig. 5C). This result suggests a ternary complex is formed by FUBP1, SMN-C2, and SMN-C3. The pull-down of FUBP1 is dependent on endogenous RNA as evidenced by a reduced FUBP1 signal when treating the protein lysate with a mixture of RNase A/T1 (Fig. 5A).

To confirm binding of SMN-C3 to FUBP1, a cellular thermal shift assay (CETSA) (Fig. 5B) was conducted (35). In CETSA, small molecules bind and protect their specific cellular protein targets from unfolding and aggregation when the cell suspension is heated (36). THP-1 cells were treated with SMN-C3 at various concentrations or DMSO control and heated at 76 °C. After cell lysis, cell debris and aggregated protein were removed by centrifugation. The amount of FUBP1 in the supernatant increased 58% in the presence of 25 μM SMN-C3 compared with DMSO control and was dose dependent as evidenced by Western blot with an anti-FUBP1 antibody (Fig. 5B). These data suggest SMN-C3 and FUBP1 interact in live cells.
covering exon 7 was in vitro transcribed, 3′-labeled with biotin, and subjected to an electrophoretic mobility shift assay (EMSA) visualized by Northern blot with streptavidin–HRP conjugate. SMN-C3 increased the bound fraction of the RNA (10 pmol) in the presence of a low concentration of recombinant FUBP1 (120 nM) in a dose–response manner (Fig. 5D). At this concentration of FUBP1, no protein-bound RNA was observed in EMSA without SMN-C3. Taken together, we conclude that SMN-C3 increases the binding affinity of FUBP1 and exon 7 by forming a ternary structure.

FUBP1 and KHSRP Are SMN2 Splicing Activators

Next, the role of FUBP1 in SMN2 splicing was investigated. FUBP1 has four KH domains that can bind either single-stranded DNA or RNA. The optimal recognition site of multiple KH-domain protein can be long and secondary structure dependent. For example, a multi-KH domain protein, vigilin, has a 75-nt-long recognition sequence (37). In the literature, FUBP1 has been reported as both a pre-mRNA splicing activator and repressor (38–40). In the case of triadin, exon 10 splicing, FUBP1 binds to an AU-rich motif downstream of a stem-loop structure at the 3′-ss and inhibits the second step of splicing, that is, nucleophilic attack of the exon 3′-OH on the 3′-splice site to release the lariat (40). Due to the high homology (61%) of FUBP1 and KH-type splicing regulatory protein (KHSRP) (SI Appendix, Fig. S9), especially at the conserved RNA binding KH domains, we hypothesized that the two proteins likely play a similar role in modulating SMN2 splicing by SMN-C3. Indeed, photo–cross-linking pull-down experiments confirmed that SMN-C2-BD can also specifically bind to KHSRP and is competed by SMN-C3 in a dose–response manner as analyzed by Western blotting with an anti-KHSRP antibody (SI Appendix, Fig. S10). KHSRP also contains four KH domains and has been shown to regulate pre-mRNA splicing through cooperation with other splicing regulatory proteins (41, 42). The third and fourth KH domains of KHSRP have been shown to independently interact with different regions of the AU-rich elements and recognize a broad set of mRNAs (43–45).

To further explore their roles, we knocked down FUBP1 and KHSRP individually or in combination with siRNAs in SMN2 minigene-transfected 293T cells. The ratio of full-length SMN (FL SMN) to that which lacks exon 7 (∆7 SMN) was analyzed by end-point RT-PCR with a pair of minigene vector-specific primers. Although FUBP1 and KHSRP alone did not significantly affect SMN2 splicing (SI Appendix, Fig. S11), dual knockdown of both genes decreased the basal FL SMN mRNA level and shifted the dose–response curve of SMN-C3 (Fig. 5E). In summary, FUBP1 and KHSRP are both activators for exon 7 splicing.

Discussion

Recently, small-molecule RNA-targeted therapeutics have drawn considerable attention from both the pharmaceutical industry and academic laboratories (11, 46–48). One such molecule, RG-7916, was originally identified through a phenotypic high-throughput screen. An optimized analog is currently being evaluated in type I, II, and III SMA patients in multiple phase 2 clinical studies. Comprehensive proteomic, genomic, and structural studies by our laboratory and Sivaramakrishnan et al. have led to an understanding of the selective actions of RG-7916 on splicing of SMN2 pre-mRNA.
Sivaramakrishnan et al. (23) recently proposed that the RG-7916 analog, SMN-C5, binds to two distinct sites of the SMN2 pre-mRNA and stabilizes an unidentified ribonucleaseoprotein (RNP) complex that is specific to SMN-C5–SMN2 pre-mRNA complex. NMR analysis showed that SMN-C5 induces chemical shift perturbations in 7 nt in the 5′-ss of exon 7 U1 snRNA duplex and causes the broadening of one imino signal. In addition, SMN-C5 was shown to bind to immobilized Ese2 fragment of exon 7 using label-free surface plasmon resonance (SPR) technology. With this ESE2 fragment, several imino signals in H NMR were broadened in the ESE2 region with chemical shift perturbations, which indicated that the conformation of ESE2 was likely changed in the presence of SMN-C5. The authors argued that binding of SMN-C5 at the 5′-ss region is more important than ESE2 based on mutagenesis of the SMN2 minigene.

In our study, the AGGAAG motif, which is almost identical to ESE2, was identified by RNA pull-down and RNA affinity cleavage experiments using probes based on the structure of SMN-C3, a close analog of RG-7916. In-cell SHAPE-MaP analysis showed that the most significant conformational change induced by SMN-C2 is in TSL1, which contains ESE2, but not the 5′-ss. SMN-C2 does not disrupt the secondary structure of TSL1 but tunes the conformation of nucleotides in the bulge and loop of TSL1. In agreement with previous publications, mutagenesis by deleting AGGAAG on exon 7, which is a conserved sequence for the splicing activator Tra2β1 (15), or its complementary sequences at polypyrimidine tract in intron 6 in a SMN2 knockin mice, resulted in ESE2 fragment of exon 7(APendix, Fig. S12). In vitro experiments, we measured binding of SMN-C2 to oligos containing the 5′-ss, but it is weaker than that to TSL1 site.

Because of the relatively large number of AGGAAG motifs in the human genome, and the selective effect of RG-7916 on mRNA splicing, it is likely that other factors must influence the selectivity of the drug for SMN2. To identify effector proteins that may also play a role in RG-7916 action, Sivaramakrishnan et al. used immobilized RNA fragment of ESE2 and identified its interaction with hnRNP G. The binding of hnRNP G was partially blocked by SMN-C5 in an SPR-based binding assay, yet the functional role of hnRNP G in pre-mRNA splicing was not fully elucidated (23). In our study, the splicing factor FUBP1 was identified in a whole-cell photo-cross-linking experiment using SMN-C2-BD. We hypothesize that conformational changes induced by SMN-C2 or -C3 binding lead to the partial displacement of hnRNP G, and enhance pre-mRNA recognition by both FUBP1 and the highly homologous splicing factor KHSPR, as demonstrated by fluorescence polarization and EMSA. The splicing activation effect of FUBP1/KHSPR was further validated by siRNA knockdown in the presence of SMN-C3.

In conclusion, the combination of our studies and the previous work of Sivaramakrishnan et al. suggest that binding of SMN-C2/C3 to the exon 7 AGGAAG motif and the resulting effect on hnRNP G/FUBP1/KHSPR binding contribute for the specificity of SMN-C2/C3 in modulating SMN2 splicing. In the literature, there are 46 other proteins shown to be relevant for exon 7 splicing (16). It cannot be ruled out that other proteins might also recognize the binding of SMN-C2/C3 and exon 7 and regulate SMN2 splicing. Further mechanistic studies should include the elucidation of the recognition sequence of FUBP1 and KHSPR and their regulatory splicing mechanism, but are beyond the scope of this paper.

Materials and Methods

RNA Pull-Down (Chem-CLIP)-Seq. Chem-CLIP protocol was modified from published procedures (49). Briefly, a 10-cm dish of 293T cells (80% confluence) were treated with SMN-C2-BD (2 μM) or SMN-C2-BD (2 μM) plus SMN-C3 (50 μM). The cells were incubated at 37 °C for 30 min before exposure to UV (365 nm) for 20 min in a UV cross-linker (StrataLinker). The cells were immediately washed with PBS once on ice and then collected by scraping. After centrifugation, total RNA of the cell pellets was extracted by using TRIzol reagent (Jambon) and Qiagen (Qiagen) with on-column treatment of DNase (#79254; Qiagen). The cross-linked RNA (100 μL in TBE buffer) was captured by incubation with prewashed streptavidin agarose (20 μL; #51638; Sigma) in the presence of RNasin ribonuclease inhibitor (40 U/mL; #21111; Promega) for 4 h at 4 °C with rotation. The beads were washed with TBE buffer (100 μL × 3) and heated with 1× elution buffer (50 μL; 95% formamide, 10 mM EDTA, pH 8.2) at 80 °C for 5 min. The released RNA was purified by RNeasy mini kit by adding 300 μL of RLT buffer and completed by following the manufacturer’s protocol. cDNA was produced from 1 μg of RNA using a qScript cDNA Synthesis Kit (#95048; Quantabio). The resulting cDNA was desalted by NuCAway columns (#A100070; Thermo Fisher). Chem-CLIP-seq samples were prepared by using ScriptSeq, version 2, RNA-Seq library preparation kit (Epicentre) following manufacturer’s protocol. The libraries were sequenced on a NextSeq 500 using 1 × 75 single-end reads to generate ~17 million reads per sample. The sequencing result was mapped and aligned with human genome by Bowtie2 algorithm and analyzed by motif discovery function (MEME) in MEME suite (meme-suite.org) (24).

Fluorescence Polarization Assay. For measurement of the binding affinity of 15-mer RNA and SMN-C2, synthetic 15-mers were reconstituted in DEPC-treated water at 500 μM. A serial dilution of each RNA oligomer was prepared in 20 μL of DEPC-treated water. The samples were heated at 80 °C for 3 min and snap cooled on ice for at least 1 min. Binding buffer (2×; 20 μL, 100 mM Heps, 300 mM NaCl, 8 mM MgCl2) with 200 nM SMN-C2 was added into each sample, mixed well by pipetting up and down, and transferred into a clear-bottom 96-well plate. For measurement of the fluorescence anisotropy to FUBP1 and hnRNP A1, 20 μM oligo-4 or 1 μM 120-nt RNA (for sequence, see SI Appendix) were snap-cooled and mixed with 200 nM SMN-C2 and 20 μM recombinant FUBP1 (SI Appendix) or hnRNP A1 (#ab123212; Abcam) in a 96-well plate. The working solution of the proteins was exchanged into 1× binding buffer using a Zeba spin desalting column (#89882; Thermo Fisher) according to the manufacturer’s protocol before use. The plate was read after 5-min equilibration at room temperature using a plate reader with fluorescence polarization function (excitation/emission, 480/488 nm). Sequences of synthetic RNA oligomers are listed in SI Appendix.
In Vitro SHAPE. The experimental protocol from literature (30) was followed except that NAI was used for RNA 2'-O-acetylation at 37 °C for 30 min. See SI Appendix for the sequence of 140 nt RNA template flanked by stabilization cassettes and a primer binding sequence at 3′-end.

In-Cell SHAPE-MaP. The experimental protocol for in-cell SHAPE-MaP experiment was modified from the one developed by Weeks and coworkers (32). Briefly, a 10-cm dish of 293T cells (70% confluent) was transfected with SMN2 minigenome (10 µg; pC-5SMN2) using FuGENE HD (50 µL; Promega) with manufacturer’s protocol. At 24 h posttransfection, the cells were tripurinized and aliquoted into 106 cells per sample (200 µL of DEMEM with 1% FBS) in 1.5-mL Eppendorf tubes that contain compounds (in DMSO, final concentration at 2 µM) or DMSO control. The tubes were incubated at 37 °C for 20 min with gentle vortexing every 5 min. The suspension of cells in each Eppendorf tube was then immediately transferred into a new tube containing 10 µL of NAI solution (2 M in DMSO, #03:310; EMD Millipore). The resulting suspension was incubated at 37 °C for 15 min with gentle vortexing every 5 min. The samples were centrifuged at 300 × g for 2 min, and the supernatant was removed. The cells were washed twice with PBS (500 µL) to remove the residue of NAI before homogenizing with Trizol (800 µL). To each sample, CHCl3 (200 µL) was added and the tubes were vigorously vortexed and centrifuged at 12,000 × g for 5 min. The aqueous partition in each tube containing the total RNA was mixed with 1.5 vol of EtOH and extracted with RNeasy mini kit (#74104; Qiagen) following the manufacturer’s protocol. Reverse transcription was performed with MnCl2 (1 µg of RNA, and a gene-specific primer (5′-TGGTTTTACAT-TGTTTTACAT) with AccuPrime pfx DNA polymerase, and the cdNA as template. The amplicons were purified by agarose gel (1.8%) and reconstituted in deionized water. Four hundred nanograms of each PCR amplicon was concatenated for 20 °C for 2 h before EtOH precipitated at −80 °C for overnight. DNA pellets were washed with 80% vol of ice-cold EtOH and resuspended in 50 µL of buffer containing 10 mM Tris-HCl, pH 7.5, 20 mM NaCl, and 0.1 mM EDTA. DNA were then fragmented by a Covaris S2 sonicator (intensity setting, 5; duty, 10%; burst cycles, 200; 5 min with frequency sweeping mode). Ten nanograms of fragmented DNA products were then treated with NEBNext Ultra II DNA Library Prep Kit for Illumina platform following manufacturer’s protocol with 9 cycles of PCR. The libraries were cleaned up using 0.9x AmpureXP beads and then sequenced on a NextSeq 500 using 1 × 75 single-end reads to generate ~10 million reads per sample. The analysis pipeline was performed using the software packages developed by Weeks and coworkers (32).

CETSA. The experimental protocol from literature (35) was followed except that THP-1 cells were used with a denaturing temperature at 76 °C. Anti-FUBP1 antibody was purchased from EMD Millipore (#ABE1330).

FUBP1/KHSRP Knockdown. Thirty picomoles of siRNA of FUBP1 (#16967; Thermo Fisher), KHSRP (#161322; Thermo Fisher), or randomized RNA control (#4390843; Thermo Fisher) were cotransfected with 0.2 µg of SMN2 minigenome (pC-5SMN2; #72287; Addgene) and 5 µL of Lipofectamine RNAiMAX (#13778030; Thermo Fisher) into one well of 293T cells (70% confluency) in a six-well plate following manufacturer’s protocol. After 24 h, the cells were trypsinized and replated into a 48-well plate containing a serial dilution of SMN-C3. After 24-h incubation with the compound, RNA was extracted from each well by RNeasy mini kit and amplified by SuperScript III Platinum One-Step qRT-PCR Kit (#11732020, Thermo Fisher) with primers (forward, 5′-TACTTAAATGACCTAGGTCGAGTTTACATC-3′; reverse, 5′-GTATCTTTACATGTCGTCG-3′) for 35 cycles (Tc = 60 °C). The PCR product was resolved on 6% TBE-urea gel followed by staining with SYBR-Safe dye for visualization.

EMSA. LightShift Chemiluminescent EMSA kit (Thermo Fisher) was used for EMSA with 120 nM purified recombinant FUBP1, 10 µM of biotinylated RNA (500 nM), and a 1:2 serial dilution of SMN-C3 following manufacturer’s protocol. See SI Appendix for the sequence and the expression protocol of the recombinant FUBP1 and the sequence of the RNA.

Note Added in Proof. While this manuscript was in preparation, Sivaramakrishnan et al. (33) showed that another active analog of RG-7916, SMN-C5 (Fig. 1A), interacts with both exonic splicing enhancer 2 (ESE2) and the 5′-ss of exon 7. The location of ESE2 on exon 7 is almost identical to the AGGAGG motif identified in this report. These results collectively provide important insights into the mechanism of this potential SMA therapeutic.

Acknowledgments. We thank Drs. James Williamson, John Hammond, Rebecca Berlow (The Scripps Research Institute), Kevin Weeks, Steve Busan (University of Northern Carolina, Chapel Hill), Gene Yao, Stefan Aigner, Frederick Tan (University of California, San Diego), and Shooutian Zhu (California Institute for Biomedical Research) for discussions. This work was made possible by NIH Grant R01 NS094721 (to P.G.S.).