Serine Rich Splicing Factor 2 (SRSF2) Flips for RNA Binding

M. Bohn, M. Cina, M. Flanagan, L. Hautala, A. Jaeke, C. Kennedy, B. Lane, A. O'Bryon, J. Pepin, E. Rohloff,
F. Rrahmani, S. Schmitt,, S. Staus, M. Tomashek
School: Hartford Union High School, 805 Cedar St. Hartford, WI 53027
Advisor: Mark Arnholt
Mentor: Mark McNally, PhD, Medical College of Wisconsin, Milwaukee, WI, 53226

I. Introduction

Mistakes in alternative splicing of RNA alter proteins, causing diseases such as acute myeloid leukemia. Serine Rich Splicing Factor 2 (SRSF2), a member of the SR protein family, is a splicing factor that controls alternative splicing by promoting inclusion of exon sequences (Fig. 1). SR proteins are composed of two domains, the RNA recognition motif (RRM) located at the N-terminus that binds RNA, and the RS domain that recruits the spliceosome. Therefore, it's not surprising that mutations to SRSF2 RRM are linked to cancers, since the mutated protein causes misregulation of RNA splicing. SRSF2 has the unusual ability to bind to both pyrimidine and purine rich RNA sequences by flipping two cytosine or guanine nucleotides on the mRNA into anti or syn position to allow binding to UCCAGU and UGGAGU sequences. The mutation Pro95His allows for stronger binding to UCCAGU and has been linked to leukemia. The Hartford Union High School SMART (Students Modeling A Research Topic) Team has designed a model of SRSF2 using 3D printing technology to highlight structural characteristics involved in RNA binding.

Alternative RNA Splicing

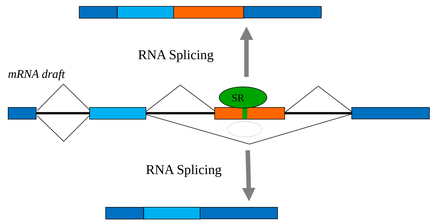

Fig. 1. Alternative splicing. In the middle is a pre-mRNA showing exons (boxes) and introns (lines). Splicing removes introns and the exons are pasted together. In alternative splicing, an SR protein like SRSF2 (green oval) binds and exon recruits the spliceosome, which causes the exon to be included (top); a lack of binding (the 'shadow' oval) causes the exon to be skipped (bottom). Changes in RNA binding of SRSF2 can alter splicing and can cause deleterious splicing events that lead to cancer.

Structure of a Typical RRM

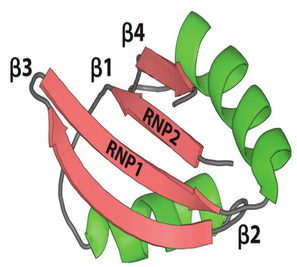

Fig. 2. Diagram of a typical RRM. Two alpha helices (green) underlie four antiparallel beta sheets (red), which contain signature amino acids in the RNP1 and RNP2 regions that are generally involved in RNA binding. The RNA typically lies across the four beta sheets. Most RRMs can only bind one optimal sequence, but SRSF2 can bind two.

II. One Protein, Two Target Sequences

SRSF2 has the unique ability to bind two different sequences. As seen in the structure below, SRSF2 has the general RRM structure shown in Fig. 2, with alpha helixes (rosy brown) and beta sheets (light salmon). Six nucleotides stretch over the sheet surface, but only C2, C3, and G5 (medium spring green) are specifically recognized by the RRM. On the beta sheets, Arg 61 (goldenrod) forms a hydrogen bond to N3 on C3, and Phe59 (yellow) forms a hydrogen bond with C2. Tyr92 (goldenrod) is not on the beta sheet, but is important as it binds with C2 and forms a hydrogen bond with C3. Additionally, evidence suggests that Lys17 (goldenrod) is involved in flipping the bases into syn or anti conformation due to the way it binds to G2 verses C2.

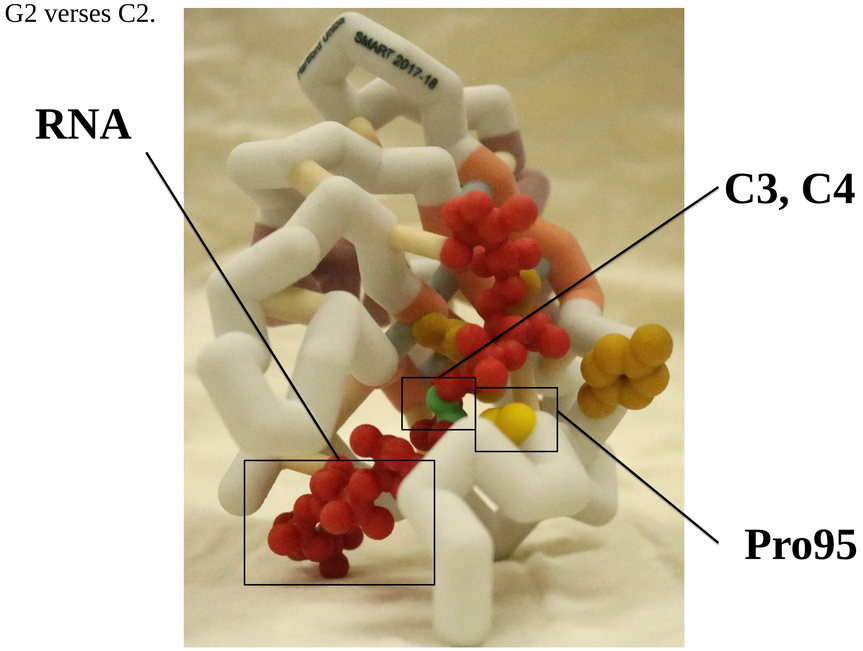

Fig. 3. Structure of the RRM of SRSF2, PDB 2LEB

III. An SRSF2 Mutation is Associated with Cancer

A Pro95His mutation in SRSF2 is associated with leukemia and shifts the binding preference from UGGAG to UCCAG because it can form hydrogen bonds with cytosine better than guanine (Zhang et al, 2012). The protein binds to cancer promoting genes due to this mutation. *Mutated* SRSF2 binds to different exons, causing the splicing factor to include these cancer promoting genes. The inclusion of these genes results in the synthesis of cancer promoting proteins. This new understanding of how SRSF2 RNA binding has changed may lead to new ways to correct aberrant splicing in cancer and other diseases.

References
Primary Citation:
Daubner GM, Cléry A, Jayne S, Stevenin J, Allain F H-T (2012). A syn-anti conformational difference allows SRSF2 to recognize guanines and cytosines equally well. EMBO J. 31:162-74

Secondary Citation:
Jian Zhang, Yen K. Lieu, Abdullah M Ali, Alex Penson, Kathryn S. Reggio, Raul Rabadan, Azra Raza, Siddhartha Mukherjee, James L Manley (2015). Disease-associated mutation in SRSF2 misregulates splicing by altering RNA-binding affinities. PNAS

The 5'-SSNG-3' Consensus is Optimal for SRSF2 Splicing In Vitro

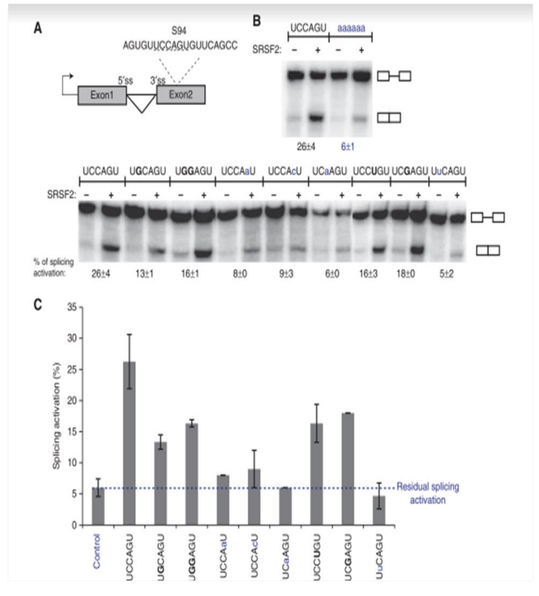

Fig. 4. SRSF2 promotes splicing of RNAs with Cs or Gs. A) A diagram of the RNA substrate that was used for the splicing experiment; it only splicing if an SRSF2 binding site is present. B) Radioactive RNAs with the indicated sequences were incubated in a nuclear extract, + or – SRSF2 to allow splicing, and the products were revealed on an acrylamide gel. Percent splicing is shown below the lanes. C) This figure shows the percent splicing from (B) in graphical form. The results show that RNAs with Cs or Gs at positions 1 and 2, and a G at position 5, splice best with SRSF2 present. These results show that the 5'-SSNG-3' sequence found from the structure mirrors the best sequence required for SRSF2 splicing.

IV. Conclusion

While it was shown that SRSF2 binds two sequences by flipping two cytosines or guanines into syn or anti position, it is currently unknown how SRSF2 induces the base flips. Mutated SRSF2 (Pro95His) misregulates splicing of numerous cancer-related genes and occurs frequently in certain leukemia and myelodysplastic syndrome patients. SRSF2 (Pro95His) has stronger binding to cytosine nucleotides than guanine nucleotides because the histidine residue hydrogen bonds to the second cytosine, which explains the splicing change. Current research is looking into which genes that are misspliced are important for cancer. This can be used to identify other protein imbalances in induced by this mutation. Antisense oligonucleotides are small sequences of RNA complementary to predetermined sequences. They are a promising fix to splicing errors because they prevent proteins like mutant SRSF2 from binding to the mRNA, which could reverse disease-related splicing. Further research into how mutations affect other splicing factors could lead to the possible treatment of other diseases as well.

The MSOE Center for BioMolecular Modeling would like to acknowledge and thank the National Institutes of Health Science Education Partnership Award (NIH-SEPA 1R25OD010505-01) and the National Institutes of Health Clinical and Translational Science Award (NIH-CTSA UL1RR031973) for their support in funding the 2017-2018 SMART Team program.