
This version is available at https://strathprints.strath.ac.uk/53862/

Strathprints is designed to allow users to access the research output of the University of Strathclyde. Unless otherwise explicitly stated on the manuscript, Copyright © and Moral Rights for the papers on this site are retained by the individual authors and/or other copyright owners. Please check the manuscript for details of any other licences that may have been applied. You may not engage in further distribution of the material for any profitmaking activities or any commercial gain. You may freely distribute both the url (https://strathprints.strath.ac.uk/) and the content of this paper for research or private study, educational, or not-for-profit purposes without prior permission or charge.

Any correspondence concerning this service should be sent to the Strathprints administrator: strathprints@strath.ac.uk
A Targeted Oligonucleotide Enhancer of SMN2 Exon 7 Splicing Forms Competing Quadruplex and Protein Complexes in Functional Conditions

Highlights
The ESE domain of a TOES oligo reduces intron 7-mediated repression of SMN2 exon 7
The ESE stimulates U2 snRNP recruitment when the oligo has annealed
The ESE forms a quadruplex and several discrete nonfunctional protein complexes
Splicing activation may require rapid exchange of proteins or ESE-protein complexes

Authors
Lindsay D. Smith, Rachel L. Dickinson, ..., Glenn A. Burley, Ian C. Eperon

Correspondence
glenn.burley@strath.ac.uk (G.A.B.), eci@leicester.ac.uk (I.C.E.)

In Brief
The use of oligonucleotides to activate selected exons is hampered by a poor understanding of the mechanisms of splicing enhancers. In this study, Smith et al. show that a bifunctional oligonucleotide enhancer of splicing of SMN2 exon 7 stimulates recruitment of U2 snRNP to the upstream intron. Surprisingly, the GGA-rich enhancer forms a quadruplex and complexes enriched in hnRNP H and other proteins. Nonetheless, these complexes do not prevent enhancer activity, suggesting that SR proteins function here via dynamic exchange.
A Targeted Oligonucleotide Enhancer of SMN2 Exon 7 Splicing Forms Competing Quadruplex and Protein Complexes in Functional Conditions

Lindsay D. Smith, Rachel L. Dickinson, Christian M. Lucas, Alex Cousins, Alexey A. Malygin, Carika Weldon, Andrew J. Perrett, Andrew R. Bottrill, Mark S. Searle, Glenn A. Burley, and Ian C. Eperon

1Department of Biochemistry, University of Leicester, Leicester LE1 9HN, UK
2Department of Chemistry, University of Leicester, Leicester LE1 7RH, UK
3School of Chemistry, Centre for Biomolecular Sciences, University of Nottingham, Nottingham NG7 2RD, UK
4Institute of Chemical Biology and Fundamental Medicine, Siberian Branch of the Russian Academy of Sciences, Novosibirsk 630090, Russia
5Protein and Nucleic Acid Chemistry Laboratory, University of Leicester, Leicester LE1 9HN, UK
6Department of Pure & Applied Chemistry, University of Strathclyde, Glasgow G1 1XL, UK

*Correspondence: glenn.burley@strath.ac.uk (G.A.B.), eci@leicester.ac.uk (I.C.E.)
http://dx.doi.org/10.1016/j.celrep.2014.08.051
This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/3.0/).

SUMMARY

The use of oligonucleotides to activate the splicing of selected exons is limited by a poor understanding of the mechanisms affected. A targeted bifunctional oligonucleotide enhancer of splicing (TOES) anneals to SMN2 exon 7 and carries an exonic splicing enhancer (ESE) sequence. We show that it stimulates splicing specifically of intron 6 in the presence of repressing sequences in intron 7. Complementarity to the 5’ end of exon 7 increases U2AF65 binding, but the ESE sequence is required for efficient recruitment of U2 snRNP. The ESE forms at least three coexisting discrete states: a quadruplex, a complex containing only hnRNP F/H, and a complex enriched in the activator SRSF1. Neither hnRNP H nor quadruplex formation contributes to ESE activity. The results suggest that splicing limited by weak signals can be rescued by rapid exchange of TOES oligonucleotides in various complexes and raise the possibility that SR proteins associate transiently with ESEs.

INTRODUCTION

Pre-mRNA splicing has the potential to be a target of considerable importance for therapeutic intervention. Most human protein-coding genes express two or more spliced isoforms of mRNA at significant levels, conferring additional diversity and flexibility to the informational capability of a limited number of genes (Djebali et al., 2012; Pan et al., 2008; Wang et al., 2008). Moreover, recent reports suggest that splicing might be stochastic, i.e., that the use of a particular exon or splice site is a matter of probability and that many minor alternative events might arise as stochastic noise (Djebali et al., 2012; Melamud and Moutt, 2009). If splicing patterns are not fixed, it may be feasible to redirect almost any splicing pattern for therapeutic purposes.

One of the most successful techniques for redirecting the splicing patterns of specific genes is to use oligonucleotides complementary to splicing signals or auxiliary motifs in the pre-mRNA (Eperon, 2012; Rigo et al., 2012; Singh and Cooper, 2013). These techniques were first designed to suppress the use of a particular pattern by blocking the binding of splicing factors to splice sites or exons (Djebali et al., 2012; Melamud et al., 2013), and were subsequently developed as potential therapies for muscular dystrophy in cases where skipping of an exon carrying a nonsense mutation would be beneficial (Cirak et al., 2011; Dunckley et al., 1998; Goemans et al., 2011). The development of oligonucleotides that had the opposite effect, stimulating exon splicing, followed from the discovery of exonic splicing enhancers (ESEs). ESE sequences in the exons of normal pre-mRNA are bound by activator proteins, the best characterized of which are the SR proteins. These proteins contain one or two RNA-binding domains and a C-terminal domain rich in RS dipeptides. The C-terminal domain of an ESE-bound SR protein was proposed to interact directly with the 3’ splice site-recognition factor, U2AF, the recruitment of which is often a limiting step in splicing, thereby increasing the level of binding of U2AF (Graveley et al., 2001; Lavigueur et al., 1993; Staknis and Reed, 1994; Wang et al., 1995; Wu and Maniatis, 1993). This led to the development of two types of oligonucleotides to stimulate usage of an exon. In one version, a PNA sequence complementary to a target exon is attached to a short RS domain peptide (Cartegni and Krainer, 2003). In the other, an oligonucleotide complementary to a target exon is extended by an ESE sequence (Skordis et al., 2003). These bipartite oligonucleotides are referred to as targeted oligonucleotide enhancers of splicing (TOES) (Eperon and Muntoni, 2003). Other sequences in or around exons have been found to act as silencers, and in such cases activation can also be achieved by using oligonucleotides to block the binding of repressor proteins (Hua et al., 2007, 2008).
One of the difficulties in designing oligonucleotides that mimic the actions of ESEs and SR proteins is that the mechanisms of activation by the latter are still poorly understood (Eperon, 2012). In addition to the recruitment of U2AF to weak 5' splice sites, the RS domains of ESE-bound SR proteins have also been shown to stabilize RNA duplexes formed between the 5' splice site and branchpoint sequences with U6 and U2 snRNA, respectively (Shen and Green, 2006). One SR protein, SRSF1, enhances U1 snRNP binding to the 5' splice site via protein interactions of its RRM domains (Cho et al., 2011). It is possible that some of these interactions may not be direct, since the introduction of a non-RNA linker between an ESE and a target 5' splice site was shown to prevent ESE activity in vitro (Lewis et al., 2012), and not all of the interactions may be involved at every ESE or made by every SR protein. Therefore, it is difficult to identify the deficiencies in an exon’s splicing signals and the best ways to compensate for them.

TOES oligonucleotides have been used to activate exons in SMN2 (Skordis et al., 2003; Marquis et al., 2007; Baughan et al., 2009), Ron (Ghigna et al., 2010), and IKBKAP (Ibrahim et al., 2007). Important determinants of the prototypical TOES oligonucleotide-activating SMN2 exon 7 include its site of annealing in the exon, the number of ESE-type motifs, and the inclusion of a non-RNA linker between the domains (Owen et al., 2011; Perrett et al., 2013). Surprisingly, previous

for designing the oligonucleotides were inadequate, which compromises our ability to apply TOES oligonucleotides to rescue other exons. Consequently, we investigated the mechanisms by which the two domains of the oligonucleotide contribute to splicing activation.

RESULTS

Identification of Sequences that Suppress Exon 7 Splicing and Are Counteracted by GGA

The most effective TOES oligonucleotides tested on SMN2 anneal to exon 6 or 7 and contain three repeats of the sequence GGAGGAC in the ESE portion (Owen et al., 2011). Here, we focus on an oligonucleotide, termed GGA, that anneals to exon 7 (Figure 1A). This stimulates inclusion of exon 7 both in vitro and in fibroblasts derived from patients (Skordis et al., 2003); in the latter case, SMN protein expression is elevated for more than 28 days after a second transfection of cells with the oligonucleotide (Owen et al., 2011). These findings suggest that the mechanistic models used as a basis

we found that it augmented splicing equally effectively even in the absence of an ESE in exon 7 (Owen et al., 2011) that is otherwise essential (Hofmann et al., 2000; Martins de Araújo et al., 2009), the sites of annealing that enabled efficient activation of SMN2 exon 7 were very restricted, and stabilizing the annealing of the oligonucleotide with modified nucleotides reduced its activity (Owen et al., 2011). These findings suggest that the mechanistic models used as a basis

Figure 1. Inhibition of Splicing of SMN2 Exon 7 by Intron 7 Sequences and Activation by GGA

(A) Sequence of the TOES oligonucleotide GGA (AD, annealing domain; ED, ESE domain) annealed to SMN2 exon 7 (lowercase, intron; uppercase, exon [with the portion omitted shown by the dotted line]; red, site of the C/U difference in SMN1 versus SMN2; blue, exonic enhancer).

(B) Effects of oligonucleotides at 250 nM on the time courses of splicing in vitro of the introns 5' and 3' of exon 7 (SMN2) in SMN2/β-globin chimeric substrates (Owen et al., 2011; Skordis et al., 2003). The substrate in the right-hand panel included 122 nt of intron 6 preceding exon 7. NO, no oligonucleotide.

(C) Effects of increasing lengths of SMN2 intron 7 on the response of a β-globin/SMN2 exon 7 chimeric substrate to GGA and AD oligonucleotides. In vitro splicing assays were done in triplicate for 2 hr with oligonucleotides at 250 nM. Error bars show the SDs of the triplicates.

The standard pre-mRNA used to test the activity of TOES oligonucleotides contains SMN2 exon 7 and adjacent intron sequences flanked by portions of introns and exons of β-globin. The effects of TOES oligonucleotides on this pre-mRNA were previously shown to correlate well with those on the endogenous SMN2 gene in patient fibroblasts (Owen et al., 2011). To identify
the splicing reactions stimulated by GGA, we tested transcripts containing single introns. The results (Figure 1B) showed that there was little stimulation of splicing. One possible explanation for this is that GGA is active when exon 7 is flanked by both introns 6 and 7. To test this, various lengths of a second intron were incorporated into the transcripts. GGA had little effect on splicing when exon 7 was preceded by the proximal 122 nt of intron 6 (Figure 1B). However, extending the length of intron 7 incorporated to the 3' side of exon 7 progressively reduced the efficiency of splicing and GGA substantially counteracted this effect (Figure 1C): an oligonucleotide containing only the annealing domain (AD) had less effect than GGA, but still enhanced splicing slightly. We conclude that GGA counteracts the inhibitory effects of intron 7 on the splicing of intron 6.

**Effects of the AD of GGA on ATP-Independent Binding of U2AF65**

These results suggest that the oligonucleotide overcomes a limiting step in spliceosome assembly imposed in part by sequences in intron 7. Early steps of assembly involve the binding of factors at the 5' and 3' splice sites, which can be studied independently of subsequent events if ATP is omitted. The first factor known to bind specifically to 5' splice sites is the U1 snRNP (Roca et al., 2013). Neither psoralen crosslinking nor immunoprecipitation revealed any effects of the oligonucleotide on the binding of U1 snRNP to the 5' splice site of intron 7 (data not shown). The binding of protein factors was analyzed by UV crosslinking to a radiolabeled transcript comprising exon 7 flanked by 41 nt of intron 6 and 58 nt of intron 7. GGA and AD increased the levels of crosslinking to a band of 50–75 kDa, detected after transfer of the proteins onto a nitrocellulose filter (Figures 2A and 2B). Comparison with PTB crosslinked to a TPM1 substrate (Cherny et al., 2010) showed that the band comigrated with U2AF65, a protein required for the assembly of early splicing complexes that binds to 3'ss polypyrimidine tracts. Moreover, the radioactivity from crosslinking coincided perfectly with the chemiluminescent signal that arose from immunodetection of U2AF65 on the same filter (Figure 2B). Finally, immunoprecipitation of U2AF65 confirmed in both of two trials that GGA and, to a lesser extent, AD increased U2AF65 crosslinking to the transcript (Figure 2C). We conclude that GGA and AD facilitate the binding of U2AF65 to the 3'ss of intron 6.

**Effects of the ESE Domain of GGA on the Formation of ATP-Dependent Complexes**

We analyzed later stages of assembly by detecting heparin-resistant complexes that formed on the same exon 7-based transcript in the presence of ATP. GGA strongly promoted the assembly of a U2-dependent complex (ED-A; Figures 3A and 3B) and base-pairing of U2 snRNA to the transcript (Figure 3C). Oligonucleotide AD had a much smaller effect. A 2'-O-methyl oligonucleotide complementary to the 5' end of U1 snRNA reduced ED-A formation, showing that U1 snRNP also contributed to assembly (data not shown). When we analyzed complex assembly on the intron 6-based transcript with 80 nt of intron 7 to the 3' side (Figure 1C), we found that GGA stimulated assembly...
of the prespliceosome (complex A) and splicesomal complexes (Figure 3D). We conclude that whereas the AD is sufficient to stimulate binding of U2AF65 to the 3’ splice site of intron 6, the ESE domain of GGA stimulates the recruitment of U2 snRNPs and the assembly of complex A and splicesomes.

Affinity Purification of GGA-Protein Complexes by Photoelution and Mass Spectrometry

To identify the proteins bound to GGA, we performed conventional affinity purification using a biotinylated transcript with the same sequence as GGA that had been incubated in nuclear extract. Proteins were separated by gel electrophoresis and bands were analyzed by mass spectrometry (MS). This revealed about eight major proteins and a number of proteins that were less abundant (data not shown). The most abundant proteins included hnRNP U, DHX36, nucleolin, hnRNP F/H, hnRNP A1, and CNBP, but not SRSF1. The high number of proteins recovered suggested that the transcript was forming heterogeneous complexes. One likely reason for this was that the concentration of the oligonucleotide was too high relative to the availability and affinity of specific proteins. Another possibility was that specific proteins, such as SRSF1, dissociated during purification and were replaced by other proteins. Native gel electrophoresis was used to test whether heterogeneous complexes formed (Figures S1A and S1B). A discrete complex formed in low concentrations of nuclear extract or in the presence of heparin, and several larger complexes formed in higher concentrations of nuclear extract and in the absence of heparin. We conclude that there is a core complex of tightly bound proteins formed on the ESE domain that is augmented or replaced by additional proteins in functional splicing conditions.

To establish that the isolated proteins came from the core complex, we sought a method that would enable intact complexes to be purified, eluted in native conditions, and analyzed by native gel electrophoresis. Moreover, it was important to identify the proteins that were associated with the full functional oligonucleotide but bound to the ESE domain. For this purpose, we performed affinity purification and photocleavage (Figure 4A). Complexes assembled on GGA were captured on avidin beads with a 3’ biotin adduct, and the ED complex was selectively eluted via a photocleavable oligonucleotide. Proteins crosslinked to the substrate are indicated. The band labeled as U2* is presumed to be U2 snRNA that was cut only at the 5’ end and was still able to associate with the substrate.

Figure 3. Effects of GGA and AD on Assembly of U2 snRNP-Containing Complexes

(A) Time courses of complex assembly in nuclear extract containing ATP on SMN2 exon 7 flanked by intron sequences as shown. As in Figure 2, GGA or AD oligonucleotides were included at 250 mM as shown; bgG is a pre-mRNA derived from β-globin. The levels of radiolabeled RNA in the absence of oligonucleotide are usually much reduced, possibly due to degradation of RNA that does not enter splicesome-related complexes. ED-A marks the exon-defined complex A.

(B) Effects on complex assembly of preincubation with DNA oligonucleotides directing RNase H degradation at two sites of U2 snRNA.

(C) Psoralen-mediated crosslinking of radiolabeled SMN2 exon 7 substrate incubated in nuclear extract with GGA and/or oligonucleotides directing degradation of U2 snRNA. Adducts crosslinked to the substrate are indicated. The band labeled as U2* is presumed to be U2 snRNA that was cut only at the 5’ end and was still able to associate with the substrate.

(D) Stimulation by GGA of splicesome assembly on pre-mRNA comprising β-globin exon 2/SMN2 exon 7 with 80 nt of intron 7 to the 3’ side.
ranked only 25th. The ESE domain was designed to recruit SRSF1, which is known to bind GGA motifs with high affinity (Tacke and Manley, 1995; Sanford et al., 2009; Cho et al., 2011; Clery et al., 2013; Pandit et al., 2013; Ray et al., 2013), and we previously showed that recombinant SRSF1 is able to bind it specifically (Owen et al., 2011). The hnRNP proteins F and H bind G-triplets, and especially GGGA (Caputi and Zahler, 2001; Dominguez et al., 2010). G-triplets are generally found in introns, where they are associated with the stimulation of splicing by hnRNP H (Chou et al., 1999; Han et al., 2005; McCullough and Berget, 1997; Xiao et al., 2009). They are generally considered to be uncommon and associated with inhibition when found in exons (Chen et al., 1999; Lim et al., 2011; Mauger et al., 2008), although other findings suggest that sites within exons can also activate inclusion by recruiting hnRNP H (Caputi and Zahler, 2002; Huelga et al., 2012). The other groups of proteins are not generally considered to be splicing factors. Nucleolin, CNBP, and DHX36 are all associated with G-quadruplexes: CNBP binds single-stranded purine-rich sequences, such as GGA (Armas et al., 2008; Liu et al., 1998b; Michelotti et al., 1995), and promotes the formation of parallel G-quadruplexes (Borgognone et al., 2010); nucleolin binds and stabilizes parallel G-quadruplexes (Bates et al., 1999; González et al., 2009); and DHX36 binds tightly to G-quadruplexes and promotes unwinding (Creacy et al., 2008; Giri et al., 2011; Vaughn et al., 2005).

Different Sets of Proteins Contact the ESE Domain

To test for the presence of a common core complex, we incubated end-labeled ED at low nanomolar concentrations in nuclear extract and proteins crosslinked to it by short-wave UV. Potential components of the complex were immunoprecipitated. Since there was no treatment with ribonuclease, we expected that all of the crosslinked proteins would be detected in constant ratios regardless of the antigen selected if there were a single discrete complex. Without immunoprecipitation,
crosslinking showed, as expected, that the ED was in contact with the proteins of ~100, 50, 35, and 20 kDa, allowing for the mass of the attached oligonucleotide (Figures 4C and 4D). Similar results were found with 5'-end labeled GGA attached to a photoreactive aryl azide and irradiated with long-wave UV (data not shown). However, immunoprecipitation with antibodies to SRSF1 and hnRNP F/H produced different patterns, demonstrating that there are different discrete complexes (Figure 4C). With anti-SRSF1, the high-Mr proteins were absent and the levels of hnRNP F/H and CNBP crosslinking were reduced 4-fold relative to SRSF1 (Figure 4C); with anti-hnRNP F/H, none of the 100 kDa, 30 kDa, or 20 kDa crosslinked proteins were recovered. This shows that in the majority of oligonucleotides where it was bound, hnRNP F/H excluded other proteins from making crosslinks to the RNA. In contrast, binding by SRSF1 was incompatible with high-Mr proteins, but not with hnRNP F/H or CNBP. The incompatibility of binding by SRSF1 with binding by DHX36 and nucleolin could be explained if the binding of the latter two proteins, unlike hnRNP F/H or CNBP, required the entire ESE domain. Alternatively, the oligonucleotide itself might exist in at least two different conformations: one bound by high-Mr proteins and one bound by hnRNP F/H, SRSF1, and CNBP in competition. The existence of competition among the high-Mr proteins, hnRNP F/H, and SRSF1 was tested by depletion of hnRNP F/H by RNAi or overexpression of GFP-hnRNP H or GFP-SRSF1 (Figure 4D). The levels of crosslinking to hnRNP F/H were inversely correlated with those of crosslinking to the high-Mr proteins and SRSF1 (Table S2).

**Formation of G-Quadruplexes by the ESE Domain**

Since nucleolin and DHX36 bind quadruplexes, whereas the binding of hnRNP F/H is prevented by quadruplex formation (Samatanga et al., 2013), quadruplex formation could account for at least part of the heterogeneity. The ESE domain of GGA contains three GGAGGAC repeats (5'-AGGAGGACGGAGGAGGACGGAGGACA) (Owen et al., 2011). We synthesized the 23-mer as RNA and its 2'-OMe and 2'-OMe/PS derivatives, together with a truncated version, as a 15-mer RNA. Analysis by 1H NMR at 800 MHz in 90% H2O solution (Figure 5A) identified imino proton resonances at high ppm values (>9 ppm) that are characteristic of hydrogen-bonded NHs. All four natural and 2'-OMe RNA sequences also revealed a set of imino-NH resonances between 11 and 12 ppm. These are consistent with the existence of G-tetrads in a quadruplex. The poor resolution of these NH resonances in the 23-mers likely arises from the existence of multiple different quadruplex folds produced by the six GG motifs. Correspondingly, the spectrum of the 15-mer ESE with four GG motifs is better resolved.

The NMR data were confirmed by far-UV circular dichroism (CD) spectra (Figure 5B), which show the characteristic features of a stable, parallel folded, quadruplex structure, with a strong positive band at 260 nm and a weak negative band at 240 nm. However, the RNA versions show an additional atypical shoulder of positive ellipticity in the CD spectrum between 290 and 305 nm. In addition, thermal stability studies by CD (at ~260 nm) show significant hysteresis between the melting and refolding curves for the RNA, but not the modified oligonucleotides (Figure 5C; Table S3). The shoulder at 290–305 nm and the hysteresis are consistent with previous reports that GGAGG-based RNA sequences form end-to-end stacks of two quadruplex motifs via tetrad-hexad arrangements in which the adenosines are recruited to form the hexad and stabilize the dimer (Liu et al., 2002; Mashima et al., 2013; Uesugi et al., 2003). Electrospray ionization (ESI)-MS analysis of the 15-mer and 23-mer ESEs provides strong support for the formation of a dimeric structure (Figure 5D). To test whether the ESE domain formed a quadruplex structure in functional conditions, we performed primer extension assays in parallel on pure GGA and on GGA in a nuclear extract in splicing reaction conditions (Figure 5E). In both cases, pause sites were seen at the nucleotide prior to the first two 3’-GGA-5’ motifs in the template, but there was much reduced pausing prior to 3’-G-GC-5’ motifs. This is consistent with a model in which AGG sequences participate in intermolecular G-quadruplexes even in nuclear extracts. The reduced intensity of the pause prior to the 5’-most AGG sequences could be accounted for by a propensity of the dimer to dissociate once the first of the four strands has been removed from the quadruplex by reverse transcription.

**Contributions of Complexes Lacking SRSF1 to Exon 7 Inclusion**

Our results showed that there are at least two major states of the oligonucleotide that are unlikely to contain SRSF1: a quadruplex dimer and the predominant hnRNP F/H complex. To test whether the quadruplex might contribute to activation of exon 7, we examined the effects of aromatic compounds that might stabilize quadruplexes. GSA-0902 and GSA-0802 are monoaand disubstituted quinolines, respectively, that bind DNA G-quadruplexes (Figure S2; Boddupally et al., 2012). CD was used to test whether these compounds stabilized the ESE RNA quadruplexes. The ligands induced changes in the quadruplex secondary structure of the 15-mer, but the spectra were consistent with the persistence of a predominantly parallel folded topology. The hysteresis in the melting and refolding curves persisted in the presence of the ligands, but the midpoints increased by 4°C–7°C, consistent with stabilization of the ligand-bound quadruplexes (data not shown). Splicing assays were done in triplicate with different concentrations of the quadruplex-binding reagents in the presence or absence of the GGA oligonucleotide (Figures 6A and 6B). In the presence of GGA, the majority of the spliced products included exon 7, but addition of the quadruplex stabilizers produced a dose-dependent decline in the level of inclusion. There was also a decline in the level of inclusion in the absence of GGA, but it was proportionately much smaller. We conclude that stabilizers, and therefore probably the quadruplexes themselves, do not contribute to and indeed compromise the activity of the bifunctional GGA oligonucleotide.

We tested the effect of hnRNP F/H on the activity of GGA after knockdown or overexpression of hnRNP F and H, using the extracts assayed for crosslinking to the ED (Figure 4D). The time courses of splicing showed very little effect on the proportion of SMN2 exon 7 inclusion, whether in the presence or absence of GGA (Figures 6C and S3), even though the reduction in levels of hnRNP F/H by RNAi was significant (Figure 6D).
DISCUSSION

Here, we have described experiments addressing the effects of oligonucleotides that stimulate exon inclusion on the processes of splicing. This is important both because our findings will facilitate future applications of bifunctional oligonucleotides and because these oligonucleotides provide tools for investigating the actions of enhancers. We have shown that GGA stimulates splicing of the upstream intron, counteracting the negative influences of intron 7. In the absence of exogenous ATP, it enhances the formation of exon-defined complexes and the recruitment of U2AF65, and in the presence of ATP, it stimulates the formation of spliceosomal complexes and the recruitment of U2 snRNP. Moreover, we have shown that the enhancer (ESE) domain forms several mutually exclusive complexes in nuclear extracts and that it forms a quadruplex.

The deficiency of splicing of SMN2 exon 7 compared with SMN1 exon 7 is known to be caused by two sequence differences. The major one, a C-to-T transition in nt +6 of exon 7 (Lorson et al., 1999; Monani et al., 1999), converts a binding site for SRSF1 (Cartegni et al., 2006; Cartegni and Krainer, 2002) into a site for hnRNP A1 and/or Sam68 (Kashima and Manley, 2003; Kashima et al., 2007a, 2007b; Pedrotti et al., 2010). This results in a 2-fold reduced recruitment of U2AF65 as well as U2 snRNP, which depends on both U2AF65 and the central Tra2β-binding enhancer in the exon (Martins de Araújo et al., 2009). The contribution of an ESE to U2 snRNP binding is still generally thought to be mediated by U2AF, which is necessary and, at least in some circumstances, sufficient for U2 snRNP recruitment via interactions with SF3B1 (Mackereth et al., 2011; Ruskin et al., 1988; Valcárcel et al., 1996; Zamore and Green, 1989; Cass and Berglund, 2006; Gozani et al., 1998). However, there is also a U2AF-independent pathway for U2 recruitment (MacMillan et al., 1997) and ESEs can function without increasing U2AF binding (Li and Blencowe, 1999; Mühlemann et al., 2000), possibly to counteract silencers (Zhu and Krainer, 2000). The effects of AD on U2AF65 binding may explain the strong dependence of TOES activity on
the site of annealing within the exon (Owen et al., 2011), even though on its own it has only a very small effect on splicing even in the presence of 250 nM GGA. The quadruplex-stabilizing molecules GSA-0820 and GSA-0902 were dissolved in DMSO and added to the reactions to produce the final concentrations shown. The pre-mRNA and inclusion and exclusion mRNA products are indicated. Charts to the right of each phosphorimage show the ratio of inclusion/exclusion of exon 7 at each of the concentrations tested. Error bars show the SDs of the ratios for the triplicates.

(C) Time courses of splicing in vitro in the nuclear extracts used in Figure 4E. The extracts were prepared from HEK293T cells transfected with plasmids expressing mEGFP-hnRNP F and mEGFP-hnRNP H (+F/H) or siRNA targeting hnRNP F and H (F/H RNAi). –, mock transfection. Reactions were done in the presence or absence of GGA and samples taken at 0, 0, 60, and 120 min. The underlined lanes were transposed in the image.

(D) Analysis by western blotting of the nuclear extracts used for the experiments in Figures 4E and 4C. Antibodies detecting hnRNP F and H and U1A were detected by fluorescent secondary antibodies. C, mock transfection; –, untransfected. See also Figures S2 and S3.

(Figure 1) suggest that the inhibitory effect of the C-to-T mutation is established in cooperation with intron 7 sequences. This might reflect a failure of an exon definition process, although exon definition currently lacks a mechanistic description and does not necessarily require more of the downstream intron than a 5′ splice site (Robberson et al., 1990; Kreivi et al., 1991). A possible explanation is that the known hnRNP A1 binding sites in intron 7 at nt +10 to +24 (Hua et al., 2008; Singh et al., 2006) and the SMN2-specific site at nt 100 (Kashima et al., 2007b), together with other likely sites in the exon (Vezain et al., 2010) and intron 7, may act in a concerted way to repress U2 snRNP recruitment. This may explain the need for additional contacts even when U2AF binding is observed.

The existence of quadruplex structures in functional conditions is generally difficult to establish, since they can be detected in short, pure RNA fragments but function within a longer RNA sequence in the presence of potential secondary structures and RNA-binding proteins. In contrast, GGA is a discrete functional unit that forms a quadruplex in isolation and also in nuclear extracts (Figure 5), where the structure antagonizes the activity of GGA (Figures 6A and 6B). The propensity to form quadruplexes may not be wholly disadvantageous. The head-to-head
dimers, in which tetrad-hexad quadruplexes are stacked together, may reduce the sensitivity of the oligonucleotide to 5’ nuclease. In addition, quadruplex formation is inimical to the binding of proteins that bind single-stranded RNA (Samatanga et al., 2013) and may reduce the sequestration of splicing activators by free GGA in the cell. This may be a generally useful property that could be exploited in the design of other functional oligonucleotides.

The other predominant form of GGA appeared to be a complex formed only with hnRNP F/H proteins, the binding of which is incompatible with quadruplex formation (Samatanga et al., 2013). SRSF1 complexes excluded nucleolin and DHX36 (Figure 4C) and thus are likely to also be in competition with quadruplexes. Overexpression of SRSF1 appeared to displace other proteins (Figure 4D), suggesting that there are multiple binding sites for SRSF1 on the ED, but that binding is not highly cooperative (Eperon et al., 2000), allowing partial occupancy and coexistence with other proteins. The coexistence of distinct complexes is likely to be a general feature of ESEs and may explain the common observation that SR proteins appear to be at best minor components of complexes when proteins bound to enhancers are analyzed after affinity purification (Staknis and Reed, 1994; Dreumont et al., 2010).

Competition between hnRNP F/H and SRSF1 may not, within limits, affect the efficacy of the oligonucleotide (Figures 4D and 6D; Table S2). Moreover, TOES activity was reduced by the use of SRSF1-binding sequences that would not bind hnRNP H (Owen et al., 2011). It is possible that the various forms of the oligonucleotide are in rapid equilibrium between pre-mRNA-bound and free states, enabling the rapid exchange of oligonucleotides in the quadruplex, hnRNP F/H, or SRSF1-containing complexes. A requirement for rapid oligonucleotide exchange would explain why increasing the strength of oligonucleotide base-pairing to exon 7 reduced TOES activity (Owen et al., 2011). Alternatively, there might be a rapid exchange of protein components bound to the oligonucleotide. Consistent with the idea that SR proteins have only a transient association with some enhancers are (1) the difficulties of demonstrating that SR proteins are the predominant proteins bound to enhancers (Staknis and Reed, 1994; Dreumont et al., 2010), (2) the inverse correlation found between SRSF2’s affinity for an enhancer sequence and the activity of the enhancer (Dreumont et al., 2010), (3) the poor correlation between the more highly occupied SRSF1 and SRSF2 binding sites transcriptome-wide and their effects on splicing (Pandit et al., 2013), and (4) the increase in TOES activity with increased numbers of binding motifs (Owen et al., 2011). Moreover, functional targets for SRSF1 are more diverse than optimal binding sites, since half of them have neither a GGA nor a GAA motif (Liu et al., 1998a; Smith et al., 2006). The strength of binding of SRSF1 is likely to be crucial in those cases where its activity requires only the second RNA-binding domain, possibly because it displaces other RNA-binding proteins (Cléry et al., 2013). For the majority of enhancers, though, it may be that their activity does not depend on the strength of their binding to SR proteins and instead is based on fleeting “kiss and tell” encounters with SR proteins.

**EXPERIMENTAL PROCEDURES**

**In Vitro Splicing and Related Assays**

Splicing assays were done in nuclear extract (Cilbiotech) as previously described (Eperon and Krainer, 1994; Skordis et al., 2003). Complex formation on exon-defined transcripts or pre-mRNA in the presence or absence of exogenous ATP was analyzed as previously described (Das and Reed, 1999). Direct crosslinking in nuclear extracts with short-wave UV light was done as previously described (Eperon et al., 2000) using RNA transcribed with [α-32P]UTP. Immunoprecipitation after crosslinking and digestion with ribonucleases A and T1 was done with an antibody to U2AF65 (Gama-Carvalho et al., 2006). Psoralen-mediated crosslinking of RNA to snRNA and ribonucleotide digestion

---

**Figure 7. Models for the Effects of Oligonucleotides on Splicing of SMN2 Exon 7**

(A) SMN1: SRSF1 bound to the 5’ end of the exon (Cartegni et al., 2006; Cartegni and Krainer, 2002), U2AF65 (Martins de Araújo et al., 2009), and a Tra2 complex (Cléry et al., 2011; Hofmann et al., 2000) contribute to recruitment of U2 snRNP. Green dotted lines show putative functional interactions and may not correspond to direct molecular contacts.

(B) SMN2: binding of hnRNP A1 and other proteins (Kashima and Manley, 2003; Kashima et al., 2007a, 2007b; Pedrotti et al., 2010) to various sites are associated with loss of binding of SRSF1 and U2AF (Martins de Araújo et al., 2009) and U2 recruitment.

(C) The AD of GGA permits some U2AF binding, but U2 snRNP recruitment is weak.

(D) The GGA tail binds SRSF1 when the quadruplex is unwound, facilitating U2 recruitment independently of Tra2(b). Nonproductive complexes of GGA with hnRNP F/H and in quadruplexes are also shown. Weak base-pairing or protein binding may enable rapid exchange of the various complexes on SMN2 exon 7 until an SRSF1 complex is stabilized by interactions with other proteins or the pre-mRNA.
Affinity Purification and Analysis of Complexes on Oligonucleotides

For conventional biotin affinity purification (Figure 4A), RNA corresponding to GGA was transcribed in the presence of biotin-11-UTP. Then 100 pmol of RNA was selected on avidin-agarose beads and incubated for 10 min in 200 μl nuclear extract under splicing reaction conditions. The beads were washed four times with 600 μl of ice-cold FSP buffer (20 mM Tris-HCl [pH 7.5], 60 mM KCl, 2.5 mM EDTA, 0.1% NP40) and once with FSP buffer without KCl, and bound proteins were eluted with a solution containing 2% SDS and 20 mM dithiothreitol. The GGA-PC-Bio oligonucleotide was synthesized by DNA Technology. For the purification of complexes, GGA-PC-Bio was incubated in nuclear extract in darkness and then the mixture was captured by incubation for 2–12 hr at 4°C. Knockdown of each hnRNP was done using two siRNAs (s230280 and s6729 for hnRNP F, and s6729 and s6730 for hnRNP H1; Ambion Life Technologies). Each knockdown was done using three successive transfections at 10 μM with Lipofectamine at intervals of 24 hr. Extracts were prepared as previously described (Lee et al., 1988).

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, three figures, and three tables and can be found with this article online at http://dx.doi.org/10.1016/j.celrep.2014.08.051.

AUTHOR CONTRIBUTIONS

L.D.S. did the experiments identifying sequences required for GGA responsiveness, assays of splicing complex assembly, and UV crosslinking. R.L.D. devised and did the affinity purification with a photocleavable linker, assays of complex formation on the oligonucleotides, and splicing assays with quadruplex binders. C.M.L. prepared depleted extracts and did the crosslinking-immunoprecipitation and splicing assays in them. A.C. did NMR, CD, and MS experiments. A.A.M. did conventional biotin affinity purifications. C.W. did the deletions by RNAi. A.J.P. synthesized the RNA and modified the oligonucleotides for physical measurements. A.R.B. did the MS and analyzed the results. M.S.S. suggested, designed, and supervised the physical measurements. G.A.B. and I.C.E. devised the project and designed and supervised the chemical and biochemical aspects, respectively. L.D.S., A.R.B., M.S.S., and I.C.E. wrote the manuscript.

ACKNOWLEDGMENTS

We thank Dr. L. Hurley (University of Arizona) for the gift of quadruplex-stabilizing reagents, Professor M. Carmo-Fonseca (IMM, Lisbon) for the anti-U2AF65 antibody, and Dr. A.R. Krainer (Cold Spring Harbor Laboratory) for the anti-SRSF1 antibody. We also thank Ms S. Ashra (University of Leicester) and Ms R. Jukes-Jones and Prof. K. Cain (MRC Toxicology Unit) for assistance with MS. This work was supported by the Leverhulme Trust (grant F/00 212Y to G.A.B. and I.C.E.), the MRC (postgraduate studentship to L.D.S.), and the Association for International Cancer Research (M.S.S. and A.C.).

Received: April 22, 2013
Revised: July 17, 2014
Accepted: August 21, 2014
Published: September 25, 2014

REFERENCES


Staknis, D., and Reed, R. (1994). SR proteins promote the first specific recognition of Pre-mRNA and are present together with the U1 small nuclear