Chapter IV. Conclusions and future prospects

4.1. Applications of SPLICE and further experimental characterizations

Elucidating the complex regulatory networks that control splicing represents a major challenge in the post-genomic era. Significant advances in our understanding of the mechanisms that guide splice site selection and the distribution of regulatory elements has aided the formulation of an early version of a 'splicing code'. Chapter 2 describes the development of a novel high-throughput cell-based screening strategy for ISRE function. We propose that SPLICE is a general screening platform that can be easily modified to select for a variety of *cis*-acting regulatory elements within any transcript. Although the initial 'proof-of-principle' of this screen was demonstrated with an alternative splicing pattern of exon skipping, the platform can be adapted towards other modes of alternative splicing including alternative 3' / 5' ss selection and mutually exclusive exon splicing. Using this platform, one can also screen for *cis*-acting regulators in different exonic and intronic locations.

One future application of this technology will be to select for libraries of cell-specific intronic or exonic regulatory elements that can be used to target gene expression in a tissue-specific and temporal manner. Toward this goal, our platform can be extended to develop a cell-specific screening strategy by incorporating a counter-selection step in a second cell line. This modified strategy can be utilized as well for the generation of cancer-specific and disease-specific regulators of alternative splicing. However, since we observed that the majority of ISREs generated through this screening strategy function in a second cell type, the diversity of the initial pool of oligonucleotides utilized for such

cell-specific screens may have to be significantly larger than what was examined in the study presented here. It is also likely that mutagenesis approaches will have to be applied at every round of these selections in combination with sequencing and experimental validation of selected elements to assess population enrichment and regulatory activity.

Given the role of alternative splicing in disease, the use of small molecules to target alternative splicing events has been investigated[1,2]. Robust and quantitative cell-based assays are needed to identify small molecule-based therapies targeted to specific splicing events[3]. Two screens have been recently developed, although on a modest scale, to identify small-molecule inhibitors of alternative splicing[4,5]. We have demonstrated that SPLICE is a robust cell-based screen and therefore can be used to screen for small molecule inhibitors of therapeutically-relevant spliced transcripts. The same strategy could likewise be applied to transcripts containing SPLICE selected ISREs to identify small molecules that affect their activities by interfering with the activity of corresponding trans-acting factors. Small molecules that regulate the activities of *cis*-acting regulatory elements will likely provide a means to modify alternative splicing of specific disease-relevant mRNAs.

Larger-scale experimental studies are needed to further characterize the SRNs involved in ISRE function as well as determining how large a role combinatorial control plays in the function of these elements. Several studies have applied high-throughput RNAi-based screening to probe gene expression in diseases such as cancer, where they have been used to gain insights into specific disease-associated pathways and their resulting phenotypes. A genome-wide RNAi-based screening strategy could be used to identify specific RNA-binding proteins that modulate the alternative splicing patterns of

transcripts containing selected ISRE elements. Results from these studies will aid in the elucidation of the trans-acting networks or SRNs involved in the regulation of these elements. In addition, high-throughput sequencing methods can be applied in combination with genome-wide RNAi-based screens to investigate transcriptome wide changes in alternative splicing. Bioinformatic approaches employed in our study can be applied to determine the spectrum of *cis*-elements that are conserved and likely involved in RNAi induced changes in global splicing patterns. Results from these studies will aid in the determination of a general set of rules that define the regulation and context dependence of intronic control elements.

Further experimental characterization is needed to identify how SPLICE selected ISRE sequences utilize complex SRNs to regulate splicing patterns in targeted cell types. Experimental analysis of selected ISREs demonstrated that their function is determined by combinatorial effects of multiple elements. To further characterize the combinatorial nature of these elements, SPLICE can be adapted to contain a synthetic oligonucleotide library comprising a combination of enriched ISRE elements spaced by neutral elements of similar length. This screen can be used to select for synthetic regulatory elements of enhanced function in a specific transcript or multiple transcripts towards the selection of universal or modular regulatory sequences. The experimental characterization of the SRNs involved in the regulation of these selected sequences as well as mutational studies that determine which elements play a role in regulation will provide significant insight into how combinatorial regulation is achieved.

Our studies also revealed that selected ISREs retain function when tested in a second cell type, but not in a second transcript. ISRE function was examined in a second

NMD-based reporter, based on the BRCA1 gene that contained an intron with significantly different predicted RNA secondary structure than the examined SMN1 intron. Given these differences, the context dependence of selected ISREs should be examined in several alternatively spliced mini-gene contexts. Also, the context dependence of these elements should be further explored within the introns of the endogenous genes containing these elements that were examined in our RNAi-based depletion study. Mini-genes containing the alternative spliced portions of these genes could be constructed to determine if the splicing of these transcripts depend on the presence of SPLICE generated ISREs. Overall, this work provides the first large-scale analysis of ISREs *in vivo* and our results highlight than an understanding of the complex interplay between multiple factors at a single binding site is necessary to further define the splicing code.

4.2. Applications of engineered ligand control of alternative splicing

Toward the goal of engineered regulation of alternative splicing we have created a platform to support the construction of protein-responsive alternative splicing regulatory elements based on the integration of protein-binding RNA aptamers into key intronic locations of a target alternatively spliced transcript. This protein-responsive platform was adapted to detect disease biomarkers, reprogram natural signaling pathways, and control biologically-relevant processes, such as apoptosis, in response to increased signaling through pathways associated with disease. *In vitro* splicing studies are needed to validate that the observed gene regulatory responses are directly modulating

alternative splicing patterns in response to changing levels of the input protein. Since our qRT-PCR measurements of the spliced isoforms correlate well with observed gene expression data for these RNA control systems, we anticipate that the results from *in vitro* splicing studies will be in line with the *in vivo* measurements.

This protein-responsive alternative splicing control platform can be interfaced with natural gene regulatory networks and signaling pathways to build programmable cells or cellular biosensors. As biosensors, these genetic elements can be used to interface with, investigate and perturb natural biological systems to report on the functional properties of regulatory networks. By replacing the aptamer component of the system with various protein-binding aptamers, this modular platform can be applied to early disease detection for a variety of different diseases. Therefore, as protein aptamers are selected towards desired targets, such as disease, cancer and viral biomarkers, they can be readily integrated into our ligand-regulated splicing platform. Furthermore, the ability to reprogram biological function in response to endogenous protein levels has broad applications in health and medicine, where such molecular tools can provide the basis for the design of targeted “intelligent” therapeutics.

References

1. Soret, J. *et al.* Selective modification of alternative splicing by indole derivatives that target serine-arginine-rich protein splicing factors. *Proc Natl Acad Sci U S A* 102, 8764–8769 (2005).

2. Fukuhara, T. *et al.* Utilization of host SR protein kinases and RNA-splicing machinery during viral replication. *Proc Natl Acad Sci U S A* 103, 11329–11333 (2006).

3. Cooper, T.A., Wan, L. & Dreyfuss, G. RNA and disease. *Cell* 136, 777–793 (2009).

4. Stoilov, P., Lin, C.H., Damoiseaux, R., Nikolic, J. & Black, D.L. A high-throughput screening strategy identifies cardiotonic steroids as alternative splicing modulators. *Proc Natl Acad Sci U S A* 105, 11218–11223 (2008).

5. O'Brien, K., Matlin, A.J., Lowell, A.M. & Moore, M.J. The biflavonoid isoginkgetin is a general inhibitor of Pre-mRNA splicing. *J Biol Chem* 283, 33147–33154 (2008).