Structural genomics analysis of alternative splicing and its application in modeling structures of alternatively spliced variants

Peng Wang, Bo Yan, Juntao Guo, Chindo Hicks, Ying Xu
Department of Biochemistry and Molecular Biology and Institute of Bioinformatics, University of Georgia, Athens, GA 30622, USA
peng@csbl.bmb.uga.edu, byn@csbl.bmb.uga.edu, guy@csbl.bmb.uga.edu, hicks@csbl.bmb.uga.edu, xyn@bmb.uga.edu

Alternative splicing is a sophisticated nuclear process that regulates gene expression [1] and is an important mean of enhancing protein diversity [2]. Alternatively spliced variants play important roles in development and defects in mRNA splicing are causes of many human diseases [3, 4, 5]. Information about alternatively spliced variants is mainly derived from primary transcripts and little knowledge exists regarding structures. Here we carry out a structural genomics analysis of known alternative splicing events and show that threading is a valid approach to model structures of alternatively spliced variants. We collect 3-D structures of proteins with know translated splicing products from PDB and further expand the dataset with high quality models generated with threading approach. Our analysis shows that splicing events have a strong preference for non-regular secondary structure elements and tend to avoid buried residues. Those observations suggest evolutionary constrains exist for locations of splicing in the context of 3-D environment. We then show that majority of substitutions in splicing events share high structural similarity and splicing events also tend to remove entire domain and avoid exposing hydrophobic cores when part of a domain was removed. Those observations support the notion that majority of splicing isoforms adopt same fold as full-length protein despite sequence substitutions and deletions. This principle was then utilized to generate high quality structures of splicing variants that could be a valuable resource for studying their structures and functions and may provide new insights into pathogenesis of related diseases.

References

[1] Black D.L. Mechanisms of alternative pre-messenger RNA splicing. *Annu Rev Biochem.* 2003;72:291-336.

[2] Johnson J.M., Castle J., Garrett-Engele P., Kan Z., Loerch P.M., Armour C.D., Santos R., Schadt E.E., Stoughton R., Shoemaker D.D.. Genome-wide survey of human alternative pre-mRNA splicing with exon junction microarrays. *Science.* 2003, 302(5653):2141-4.

[3] Venables J.P. Aberrant and alternative splicing in cancer. *Cancer Res.* 2004, 64(21):7647-54.

[4] Courtois S., de Fromentel C.C., Hainaut P. p53 protein variants: structural and functional similarities with p63 and p73 isoforms. *Oncogene.* 2004, 23(3):631-8.

[5]Bartel F., Taubert H., Harris L.C. Alternative and aberrant splicing of MDM2 mRNA in human cancer. *Cancer Cell.* 2002, 2(1):9-15.