Characterization of regulatory mechanisms for BIM exon 3 alternative splicing

Bhadra, Malini

2017

Bhadra, M. (2017). Characterization of regulatory mechanisms for BIM exon 3 alternative splicing. Doctoral thesis, Nanyang Technological University, Singapore.

<http://hdl.handle.net/10356/69548>

<https://doi.org/10.32657/10356/69548>

NANYANG
TECHNOLOGICAL
UNIVERSITY

CHARACTERIZATION OF
REGULATORY MECHANISMS FOR *BIM*
EXON 3 ALTERNATIVE SPLICING

MALINI BHADRA
SCHOOL OF BIOLOGICAL SCIENCES
2017

CHARACTERIZATION OF
REGULATORY MECHANISMS FOR *BIM*
EXON 3 ALTERNATIVE SPLICING

MALINI BHADRA

School of Biological Sciences

A thesis submitted to the Nanyang Technological
University in partial fulfillment of the requirement
for the degree of Doctor of Philosophy

2017

ACKNOWLEDGEMENTS

During the course of my Doctoral research, I have had the opportunity of working with some of the most ignited minds in their respective fields of study.

I express my heartfelt gratitude to my supervisor, Asst. Professor Xavier Roca, for his invaluable guidance, support and critical insights which gave me a larger understanding of the scope of my research. His discipline and eye for detail inspired me to work to the best of my ability and grow as a researcher. I shall remain grateful to him for giving me the opportunity to explore the wonderful arena of RNA splicing and instilling the scientific temper in me.

I heartily thank each and every member of Dr Roca's lab for all the help and support I received from them. First and foremost, I am thankful to my postdoctoral fellow, Dr Liu Jun who is an intimate/intrinsic part of the BIM project. She was extremely supportive and without her mentorship and resourcefulness, my project would not have taken shape. Her technical expertise and scientific knowledge were instrumental in producing the results I obtained in these four years of my PhD work. I am grateful to my students, Ms Loh Yan Ping for her help in performing the knockdown/overexpression assays and Ms Lue Wan Lin for her work on *BIM* intronic *cis*-acting elements.

To Dr Tang Sze Jing and Dr Tan Jiazi, I am sincerely appreciative of their continuous experimental guidance and invaluable suggestions throughout my research work. Their valuable inputs and the stimulating scientific sessions that we always had helped to enrich my thesis. I really appreciate their patience and kindness. I also express my gratitude to Ms Luo Shufang and Ms Ho Jia Xin Jessie for their help in some experiments and for being supportive in all possible ways. Mr Lorenzini Paolo and Dr Liu Hongfei need particular mention for their immense contribution to my personal and professional growth. Indeed, my lab group has been a great source of succor and sustenance throughout my journey in NTU.

My sincere thanks to Nitin Loganathan, Guo Zhijun and Divyanshu Mahajan for all their help with this thesis completion. I am also indebted to NTU SBS for giving me the opportunity to pursue my research in the field. I would like to acknowledge the great

support I received from our collaborators, Professor Ong Sin Tiong and his team at Duke-Nus. I take this opportunity to thank Professor Mark Featherstone, Asst. Professor Zhang Lifeng and Asst. Professor Eugene Makeyev for being a part of my thesis committee and providing valuable suggestions.

I extend my warmest thanks to all my friends Geeta, Nitin, Anuradha, Alolika, Ajit, Mufeeda, Areetha, Anjali, Ramya, Manisha, Prabha, Shraddha, Candace, Toby, Alma, Harshyaa, Priya and my cousin Avantika Deb for their sustained companionship, help and moral support. We have shared many memorable and happy times together.

Finally, I would like to thank my family for all their love and encouragement and for instilling in me a passion for science. I am greatly indebted to my father, Dr Amit Bhadra (ex Senior Regional Director of Pfizer, Singapore) whose passion for knowledge and pursuit of excellence inspired me to move forward in my studies. He was knowledgeable, inspiring, humble and kind and touched many lives in significant ways. He had been diagnosed with brain cancer during the first year of my PhD program. Seeing him suffer and fade away has been one of the most difficult things I had to struggle with. His spirit guided me like a kindly light through very trying times. It would have been almost impossible without the unconditional love and support of my friends, and lab mates who held my hand and helped me carry on with my work. Thanks to Prof Roca for being there with me like the shade of a banyan tree at those trying moments.

I thank my mother, Dr Nandini Bhadra for being a true source of inspiration and my most intimate companion in moments of joy and sorrow. Both my mother and brother, Arindam have been pillars of strength during extremely tough times and were always there for me whenever I needed them. Arindam almost assumed the role of a loving guardian and sent me endearing messages and gifts to liven up my spirits.

I dedicate this work to my beloved father who passed away in the second year of my doctoral studies. I hope I will fulfill his dreams through this dissertation.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS.....	I
LIST OF FIGURES	VII
LIST OF TABLES	IX
ABBREVIATIONS	X
ABSTRACT.....	XV
1 INTRODUCTION	1
1.1 Spliceosome assembly and splicing	1
1.2 Splicing regulators and exon definition.....	4
1.3 Alternative splicing.....	7
1.3.1 Alternative splicing types	7
1.3.2 Regulatory mechanisms of alternative splicing.....	8
1.4 Splicing and disease.....	13
1.4.1 <i>Cis</i> -acting mutations	13
1.4.2 <i>Trans</i> -acting mutations and alterations in expression.....	15
1.4.3 Single nucleotide polymorphisms.....	16
1.5 Splicing and therapy	17
1.6 Apoptosis	19
1.7 BCL-2 protein family	21
1.8 BIM.....	22
1.8.1 Gene structure and isoforms	23
1.8.2 Physiological roles and implication in cancer-based therapy	24
1.9 Chronic myeloid leukemia (CML)	26
1.9.1 BCR-ABL protein isoforms.....	27
1.9.2 BCR-ABL signaling	28

1.9.3	Clinical features of CML	30
1.9.4	CML treatment.....	31
1.10	EGFR NSCLC	32
1.11	Primary TKI resistance in East Asian CML and EGFR NSCLC patients	33
1.12	Rationale of project	35
1.13	Objectives	35
2	MATERIALS AND METHODS	36
2.1	Cell culture	36
2.2	Plasmids.....	36
2.2.1	Generation of mutant plasmids for <i>BIM</i> $\Delta 10$ and $\Delta 11$ minigenes	36
2.2.2	pSXN heterologous plasmid	47
2.2.3	Overexpression plasmids	49
2.3	Bacterial transformation and plasmid extraction.....	49
2.4	Transfection with plasmids.....	50
2.5	Knockdown studies.....	50
2.6	Rescue studies.....	51
2.7	RNA Extraction and DNase I treatment	52
2.8	Reverse Transcription and gel electrophoresis	52
2.9	Real time PCR	53
2.10	Semi-quantitative PCR	55
2.11	Protein extraction and western blotting	56
2.11.1	Sample preparation	56
2.11.2	SDS-PAGE and transfer	56
2.11.3	Detection	56

2.12	RNA pull-down	57
2.13	Silver staining.....	58
2.14	CRISPR guide RNA Vector construction and ssODN design	58
2.15	Nucleofection.....	61
2.16	Fluorescence-activated cell sorting (FACS).....	61
2.17	Genomic DNA extraction and genotyping CRISPR-edited <i>BIM</i>	61
2.18	<i>BIM</i> expression plasmid construction.....	63
3	RESULTS.....	67
3.1	<i>BIM</i> minigene plasmids	67
3.2	$\Delta 10$ and $\Delta 11$ mimic splicing patterns of full-length and deletion minigenes.....	68
3.3	Multiple <i>cis</i> -acting elements regulate <i>BIM</i> E3 splicing.....	69
3.3.1	<i>BIM</i> E3 is rich in ESEs and ESSs	69
3.3.2	E3 PolyA signal and flanking regions exhibit enhancer-like effects	72
3.3.3	Intronic sequences upstream of <i>BIM</i> E3 3'ss contain splicing elements	74
3.3.4	Point mutations within ESEs of <i>BIM</i> E3 reduced enhancer effects.. ..	78
3.3.5	Validation of ESE sequences in a heterologous context.....	81
3.4	<i>BIM</i> E3 splicing is regulated by length of PolyU tract at 3'ss	83
3.5	Identification of <i>trans</i> -acting factors regulating <i>BIM</i> E3 splicing.....	84
3.5.1	RNA pull-down approach	84
3.5.2	Candidate approach.....	87
3.5.2.1	SRSF1 promotes <i>BIM</i> E3 inclusion.....	88
3.5.2.2	SRSF3 is likely a repressor of <i>BIM</i> E3 inclusion	92
3.5.2.3	SRSF5 mildly promotes endogenous <i>BIM</i> E3 inclusion	95
3.6	SNPs affect <i>BIM</i> E3 splicing in context of minigenes	97

3.7	Testing the effects of <i>BIM</i> E3 SNPs in K562 cell line	103
3.7.1	Introduction of SNP minor alleles in endogenous <i>BIM</i> by CRISPR editing	103
3.7.2	<i>BIM</i> knockout by CRISPR/Cas9 and rescue by <i>BIM</i> expression plasmid	108
4	DISCUSSION.....	110
4.1	<i>BIM</i> E3 is regulated by multiple regulatory <i>cis</i> -acting splicing elements	110
4.2	Polyadenylation regulatory sequences favor splicing of <i>BIM</i> E3.....	112
4.3	Validation of ESEs in a heterologous context.....	113
4.4	Length of PolyU tract at 3'ss regulates <i>BIM</i> E3 splicing.....	116
4.5	Identification of <i>trans</i> -acting factors regulating <i>BIM</i> E3 splicing.....	118
4.5.1	Effects of hnRNP A1, hnRNP A2/B1 and PTBP1 on <i>BIM</i> E3 inclusion.....	118
4.5.2	SRSF1 might protect cancer cells from apoptosis by promoting E3 inclusion.....	119
4.5.3	SRSF3 may enhance apoptosis via <i>BIM</i> splicing changes.....	120
4.6	Effect of SNPs on <i>BIM</i> E3 splicing in minigene context	121
4.7	Testing the effect of <i>BIM</i> E3 SNPs by CRISPR genome editing in K562 cells	124
5	CONCLUSION AND FUTURE DIRECTIONS	127
6	REFERENCES	129

LIST OF FIGURES

Figure 1.1 Schematic representation of spliceosome assembly and splicing.	4
Figure 1.2 Splicing regulation and exon definition	6
Figure 1.3 Schematic of the seven general types of alternative splicing	8
Figure 1.4 Regulation of alternative splicing by <i>cis</i> -acting elements and <i>trans</i> -acting factors	10
Figure 1.5 Alternative splicing regulation by transcription and chromatin	12
Figure 1.6 Schematic of apoptotic signaling pathways	20
Figure 1.7 Schematic of <i>BIM</i> gene structure and transcripts.	24
Figure 1.8 Schematic of Philadelphia chromosome	27
Figure 1.9 <i>ABL</i> and <i>BCR</i> breakpoint locations and respective fusion mRNA isoforms.....	28
Figure 1.10 Schematic showing imatinib mechanism of action	32
Figure 2.1 pSXN plasmid	48
Figure 2.2 Generation of <i>BIM</i> expression plasmid by overlap PCR.....	66
Figure 3.1 Schematic of the <i>BIM</i> minigenes.....	67
Figure 3.2 E3 versus E4 splicing pattern in <i>BIM</i> minigenes.....	68
Figure 3.3 Schematic map of serial deletions (D) and staggered deletions (SD) in <i>BIM</i> E3	70
Figure 3.4 Numerous <i>cis</i> -acting splicing elements are present in <i>BIM</i> E3	72
Figure 3.5 Polyadenylation signal and flanking region within <i>BIM</i> E3 promote E3 splicing	73
Figure 3.6 Intronic sequences upstream E3 contain ISEs and ISSs.....	75
Figure 3.7 E3/E4 correlation for deletion mutants across $\Delta 10$ and $\Delta 11$	76
Figure 3.8 Map of enhancers and silencers for <i>BIM</i> E3 splicing by deletion analysis.....	77
Figure 3.9: <i>In silico</i> predictions of <i>trans</i> -acting factors binding to the E3 ESEs and ESSs	78
Figure 3.10 List of <i>BIM</i> E3 WT ESE sequences and the corresponding mutants ..	79
Figure 3.11 Point mutations reduced enhancer effects	80
Figure 3.12 Testing ESEs in heterologous context.....	82

Figure 3.13 <i>BIM</i> E3 splicing is regulated by length of PolyU tract.....	84
Figure 3.14 RNA pull-down assay for ESE2 and ESE3	86
Figure 3.15 Western blot confirmed protein binding to ESE2 and ESE3	86
Figure 3.16 SRSF1 overexpression promotes endogenous <i>BIM</i> E3 inclusion	89
Figure 3.17 SRSF1 overexpression promotes <i>BIM</i> E3 inclusion in minigenes.....	90
Figure 3.18 SRSF1 knockdown promotes E3 exclusion and rescue partially reverses it	91
Figure 3.19 SRSF3 knockdown promotes endogenous <i>BIM</i> E3 inclusion.....	93
Figure 3.20 SRSF3 rescue does not reverse the knockdown effects of SRSF3.....	94
Figure 3.21 SRSF5 overexpression mildly promotes endogenous <i>BIM</i> E3 inclusion.....	95
Figure 3.22 SRSF5 overexpression does not change <i>BIM</i> E3 splicing in minigenes	96
Figure 3.23 SNPs affect <i>BIM</i> E3 splicing.....	98
Figure 3.24 Combinations of major and minor alleles of the four SNPs show additive effect.....	100
Figure 3.25 Comparison of 16 SNP haplotypes to reciprocal deletion or nondeletion alleles	102
Figure 3.26 Approaches for testing effect of <i>BIM</i> SNPS in K562 cell line.....	103
Figure 3.27 Introduction of SNP 3c and SNP 4a in endogenous K562 <i>BIM</i> by CRISPR.....	107
Figure 3.28 Agarose electrophoresis of PCRs after targeting E2A with two CRISPR gRNAs.....	109
Figure 4.1 Schematic of mutual stimulation between splicing and polyadenylation.....	113
Figure 4.2 Hypotheses for ESE1 as a mild enhancer in heterologous context	115
Figure 4.3 Schematic showing AG-dependent 3'ss and AG independent 3'ss	118
Figure 4.4 Schematic showing SNPs positions in map of enhancers and silencers.....	124

LIST OF TABLES

Table 2.1 Primers for generation of serial deletions in <i>BIM</i> E3	37
Table 2.2 Primers for generation of serial staggered deletions in <i>BIM</i> E3	40
Table 2.3 Primers for generation of serial deletions in <i>BIM</i> Intron 2	42
Table 2.4 Primers for generation of serial staggered deletions in <i>BIM</i> Intron 2.....	43
Table 2.5 Primers for varying the PolyU tract length at 3'ss of <i>BIM</i> E3	44
Table 2.6 Primers for introducing point mutations in <i>BIM</i> E3 enhancer sequences	45
Table 2.7 Primers for introducing SNPs in <i>BIM</i> E3 and upstream intronic region	46
Table 2.8 Conditions for PCR mutagenesis.....	47
Table 2.9 Oligonucleotide inserts for pSXN plasmid.....	49
Table 2.10 List of siRNA and dsRNA sequences.....	51
Table 2.11 Primers for detecting U-E3 and U-E4 transcripts (RT-PCR)	52
Table 2.12 Conditions for PCR.....	53
Table 2.13 Conditions for Real time PCR	54
Table 2.14 Primers for detecting different transcripts (Real time PCR).....	54
Table 2.15 List of radioactive primers.....	55
Table 2.16 List of antibodies	57
Table 2.17 List of RNA oligonucleotides for RNA pull-down	58
Table 2.18 Design of CRISPR guide sequences.....	60
Table 2.19 Design of ssODNs	60
Table 2.20 Conditions for PCR.....	63
Table 2.21 Primer sequences for generating <i>BIM</i> expression plasmid.....	65

ABBREVIATIONS

3'ss	3' splice site
5'ss	5' splice site
ABL	Abelson murine leukemia viral oncogene homolog
APAF-1	Apoptosis protease-activating factor-1
ASO	Antisense oligonucleotide
ATP	Adenosine Triphosphate
BAD	BCL-2-associated death promoter
BAK	BCL-2 homologous antagonist/ killer
BAX	BCL-2-associated X protein
BCL2	B-cell lymphoma 2
BCL2L11	B-cell lymphoma 2 like 11
BCL-xL	B-cell lymphoma-extra-large
BCR	Breakpoint cluster region
BH	BCL-2 homology
BID	BH3 interacting-domain death agonist
BIK	BCL-2-interacting killer
BIM	BCL-2-interacting mediator of cell death
BPS	Branch-point sequence
BRCA1	Breast cancer 1
Cas	CRISPR associated protein
Cas9	CRISPR associated protein 9
cDNA	Complementary DNA
CELF	CUG-BP and ETR-3 Like factors
CFIm	Cleavage factor Im
CFIIm	Cleavage factor IIm
CFTR	Cystic fibrosis transmembrane conductance regulator

CML	Chronic myeloid leukemia
CPSF	Cleavage and polyadenylation specificity factor
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats
CrRNA	CRISPR RNA
CstF	Cleavage stimulation factor
CTD	Carboxy terminal domain
CUGBP1	CUG-binding protein 1
CCND1	Cyclin D1
DNA	Deoxyribonucleic acid
DMD	Duchenne muscular dystrophy
DM1	Myotonic dystrophy type 1
DSE	Downstream element
DMPK	Myotonic dystrophy protein kinase
DSB	Double-stranded break
dsiRNA	Dicer substrate siRNA
E3	Exon 3
E4	Exon 4
<i>E.coli</i>	<i>Escherichia coli</i>
EGFR NSCLC	Epidermal growth factor receptor mutated non-small cell lung cancer
Elav	Embryonic lethal abnormal visual
EMSA	Electrophoretic mobility shift assay
ERAP2	Endoplasmic reticulum aminopeptidase 2
ERK	Extracellular-signal-regulated kinase
ESE/ESS	Exonic splicing enhancer/silencer
FACS	Fluorescence-activated cell sorting
FBS	Fetal bovine serum
FD	Familial dysautonomia
FISH	Fluorescent <i>in situ</i> hybridization

FOXO3A	Forkhead box O3
FTDP-17	Frontotemporal dementia and parkinsonism linked to Chromosome 17
gRNA	Guide RNA
GAB2	Grb2 associated binding protein 2
GFP	Green fluorescent protein
GRB2	Growth factor receptor-bound protein 2
GTP	Guanosine triphosphate
HDR	Homology directed repair
hnRNP	Heterogeneous nuclear ribonucleoprotein
IKAP	I κ B kinase complex associated protein
IKBKAP	Inhibitor of kappa light polypeptide gene enhancer in B-cells, kinase complex-associated protein
ISE/ISS	Intronic splicing enhancer/silencer
JAK	Janus kinase
KLF6	Kruppel-like factor 6
MAPK	Mitogen-activated protein kinase
MBNL1	Muscleblind like splicing regulator 1
MHC	Major histocompatibility complex
MNK2	Mitogen-activated protein kinase interacting serine/threonine kinase 2
MCL -1	Myeloid cell leukemia sequence 1
M-MuLV	Moloney Murine Leukemia Virus
mRNA	Messenger RNA
MLH1	MutL homolog 1
mTOR	Mammalian target of rapamycin
NMD	Nonsense mediated decay
NOVA	Neuro-oncological ventral antigen
NOXA	Phorbol-12-myristate-13-acetate-induced protein 1
Nt	Nucleotide

OAS1	2'-5'oligoadenylate synthetase 1
PABII	polyA-binding protein II
PAP	PolyA polymerase
PAS	Polyadenylation signal
PCR	Polymerase chain reaction
Pen/Strep	Penicillin/Streptomycin mixture
PI3K	Phosphoinositide 3-kinase
PH	Pleckstrin homology
Pol II	Polymerase II
PRPF3	Pre-mRNA processing factor 3
PRPF8	Pre-mRNA processing factor 8
PRPF31	Pre-mRNA processing factor 31
PTBP1	Polypyrimidine tract binding protein 1
PUMA	P53 upregulated modulator of apoptosis
qRT-PCR	Quantitative Reverse transcription PCR
R	Purine
RBM5	RNA-binding protein 5
RBM10	RNA-binding protein 10
Rho-GEF	Rho guanine nucleotide exchange factor
RNA	Ribonucleic acid
RPMI	Roswell Park Memorial Institute
RT-PCR	Reverse transcription PCR
SDS-PAGE	Sodium dodecyl sulfate-Polyacrylamide gel electrophoresis
SF1	Splicing factor 1
sgRNA	Single guide RNA
SH1	Src Homology 1
SH2	Src Homology 2
SH3	Src Homology 3

---

si-RNA	Small interfering RNA
SOS	Son of sevenless
SSB	Single-stranded break
S6K1	Ribosomal S6 kinase 1
SMA	Spinal muscular atrophy
SMN	Survival of motor neuron
SNP	Single nucleotide polymorphism
snRNP	Small nuclear ribonucleoproteins
snRNA	Small nuclear ribonucleic acid
ss	Splice site
SSO	Splice switching oligonucleotide
ssODN	Single-stranded DNA oligonucleotide
STAT	Signal Transducer and Activator of Transcription
TracrRNA	Trans-activating crRNA
TKI	Tyrosine kinase inhibitor
TNF	Tumor necrosis factor
U2AF	U2 auxiliary factor
Y	Pyrimidine

ABSTRACT

BCL-2-interacting mediator of cell death (BIM) is a proapoptotic member of the BCL-2 protein family. BIM upregulation is required for tyrosine kinase inhibitor (TKI)-mediated cell death in many kinase-driven cancers including chronic myeloid leukemia (CML) and epidermal growth factor receptor mutated non-small cell lung cancer (EGFR NSCLC). Previously, a 2.9 kb deletion polymorphism in *BIM* intron 2 was found to contribute to TKI resistance in CML and EGFR NSCLC patients. The deletion allele switched splicing from *BIM* exon 4 (E4) to exon 3 (E3) in a mutually exclusive manner, generating isoforms lacking the proapoptotic BCL2-homology domain 3 (BH3) encoded by E4. This suggests that modulating *BIM* splicing to exclude E3 might resensitize the polymorphism containing cancer cells to TKI. In this project, the mechanisms of *BIM* E3 alternative splicing regulation were studied. To identify the *cis*-acting splicing elements regulating *BIM* E3, a series of sequential 10 nucleotide deletions throughout E3 and upstream intronic region were generated in two *BIM* minigenes with and without the polymorphic fragment. By comparing the E3/E4 ratio of the deletion mutants to that of the full-length minigenes by quantitative real-time PCR, putative splicing enhancers and silencers regulating *BIM* exon 3 were defined. The exonic enhancers were subsequently verified by point mutations and heterologous minigene analysis. The length of the *BIM* E3 polypyrimidine tract, which is an essential sequence within 3' splice sites was varied, and it was found that an unusually long minimum of 16 uridines was required for maximal E3 inclusion. The splicing factor, SRSF1 was confirmed to promote *BIM* E3 inclusion by knockdown/overexpression study. In addition minor alleles for four annotated single nucleotide polymorphisms (SNPs) were found to potentially alter E3 splicing in *BIM* minigenes and the allele combinations showed additive effects. These annotated SNPs will be subsequently tested for their effects on TKI responses. The splicing regulatory elements elucidated in our study may reveal therapeutic targets for sensitization of TKI responses in cancers dependent on BIM expression for drug sensitivity.

1 INTRODUCTION

Subtle changes in gene expression can lead to phenotypic variability and even disease. Hence understanding the regulation of gene expression becomes very important. The eukaryotic gene expression pathway involves processes such as transcription, precursor mRNA processing, translation, and post-translational modification of a protein. Pre-mRNA processing involves capping, editing, splicing and polyadenylation. The focus of our lab is pre-mRNA splicing which is an extremely intricate process governed by many factors and is intimately coordinated with other gene expression processes such as mRNA transcription, transport and translation. The extensive integration of splicing with other layers of gene regulation stresses the need to explore splicing regulation. Splicing is further complicated by the process of alternative splicing which contributes to the generation of multiple mRNA and protein isoforms from one gene. Perturbations in constitutive splicing and alternative splicing could directly lead to a disease, increase susceptibility to a disease or affect drug responses. Hence studying the link between splicing and disease provides opportunities for diagnostics as well as treating the cause of the disease. In my PhD project, I have studied the regulatory mechanisms of *B cell lymphoma-2 (BCL-2)-interacting mediator of cell death (BIM)* exon 3 (E3) splicing. The importance of studying splicing regulation of *BIM* E3 lies in the fact that BIM is a proapoptotic protein and the increased splicing of exon 3 (E3) due to a genomic polymorphism is associated with tyrosine kinase inhibitor (TKI) resistance in chronic myeloid leukemia (CML) and epidermal growth factor receptor mutated non-small cell lung cancer (EGFR NSCLC) patients. This suggests that modulating *BIM* splicing to exclude E3 might restore sensitivity to TKIs.

In this section, I will be providing general background on splicing, apoptosis and BIM, and its implications in treatment of kinase-driven cancers such as CML and EGFR NSCLC.

1.1 Spliceosome assembly and splicing

A crucial step in ensuring accurate gene expression in eukaryotes is the process of precursor mRNA (pre-mRNA) splicing whereby introns are removed as a lariat and exons

are ligated in a pre-mRNA resulting in formation of a mRNA transcript (House and Lynch, 2008; Keren et al., 2010; Kim et al., 2008). The discovery of splicing as a post-transcriptional process was made in 1977 when Roberts and Sharp found that discontinuous regions of adenoviral genomic DNA generated a continuous mRNA transcript (Berget et al., 1977; Chow et al., 1977). Apart from producing different mRNA and proteins, splicing also regulates other gene expression pathways such as mRNA transcription, turnover, transport, and translation (Braunschweig et al., 2013).

Splicing is accomplished in a dual-step *trans*-esterification reaction which is facilitated by a dynamic macromolecular complex known as the spliceosome. Small nuclear ribonucleoproteins (snRNPs) - U1, U2, U4, U5, and U6 and more than hundred proteins are constituents of the spliceosome. A single snRNP is formed of a small nuclear RNA (snRNA) and associated proteins (Wahl et al., 2009). There are three essential *cis*-acting splicing elements, which define the core splice signals necessary for recognition by the spliceosome. They are: the 5' splice site (5'ss) and the 3' splice site (3'ss) at the 5' and 3' end of introns, respectively, and the intronic branch-point sequence (BPS) that forms the branched lariat (Keren et al., 2010; Kim et al., 2008).

The spliceosomal components are loaded onto a pre-mRNA in an orderly fashion forming distinct spliceosomal complexes (Figure 1.1). Splicing is initiated upon association of U1 snRNP with the 5'ss and the recruitment of splicing factor 1 (SF1) at the BPS. This generates the E' complex. Next, the 3' terminal AG dinucleotide and upstream polypyrimidine tract are recognized by the protein heterodimer, U2 auxiliary factor (U2AF). The 3' terminal AG is bound by U2AF35 (small subunit) while the polypyrimidine tract is bound by U2AF65 (large subunit), generating the E complex. The generation of E' complex and E complex does not require ATP. Subsequently in an ATP dependent manner, U2 snRNP replaces SF1 to associate with the branch point, generating complex A (Figure 1.1). During this complex A formation, there is base-pairing between the BPS and the U2 snRNA which leads to bulging out of branch point adenosine.

Subsequent loading of U4/U6–U5 pre-assembled tri-snRNP gives rise to the B complex. Further changes in conformation and composition give rise to the activated spliceosome,

B^{act} complex. These changes include disruption of U4 and U6 snRNAs base-pairing followed by U6 and 5'ss base-pairing and discharge of U1 and U4 from the complex. Subsequently U2-U6 base-pairing forms and brings together the 5'ss and BPS for the first catalytic splicing reaction. Other conformational rearrangements include formation of U6 intramolecular stem-loop structure (U6-ISL) and interaction of U5 snRNA with nucleotides of the 5' exon. Subsequent remodeling by the ATPase Prp2, generates the catalytically active B* complex. The first trans-esterification reaction of splicing then occurs by adenosine 2'-hydroxyl group at the BPS attacking the 5'ss phosphodiester bond. This liberates the 5' exon from the intron lariat-3' exon generating the C complex. Additional rearrangements of the C complex enable it to conduct the second trans-esterification reaction of splicing. During this second catalytic step, the 3'ss phosphodiester bond is attacked by the 3' hydroxyl group of the liberated 5'exon. As a result, the exons are linked to form a mature mRNA while the intron-lariats are usually degraded (Figure 1.1). The spliceosome subunits are recycled for subsequent splicing reactions (Chen and Manley, 2009c; Kornblihtt et al., 2013; Matera and Wang, 2014; Wahl et al., 2009; Will and Luhrmann, 2011). Throughout the splicing process, RNA-protein and RNA-RNA remodeling is mediated by different RNA-dependent helicases/ATPases at various stages. The recognition of the essential *cis*-acting elements multiple times by different snRNAs and proteins ensures the precision of splicing (Wahl et al., 2009).

Figure 1.1 Schematic representation of spliceosome assembly and splicing.

U1 snRNP interacts with the 5'ss while SF1 binds BPS. U2AF heterodimer interacts with 3'ss. SF1 at the branch point is replaced by U2 snRNP. Subsequently, the tri-snRNP U4/U5/U6 is recruited, and this is followed by conformational and compositional rearrangements of the spliceosome including loss of U1 and U4. The two transesterification reactions of splicing then take place. Boxes represent exons; Lines represent introns. For simplicity, E' and B^{act} complex are not shown in figure.

1.2 Splicing regulators and exon definition

The essential *cis*-acting elements in higher eukaryotes are poorly conserved (Sheth et al., 2006). Due to this degenerate nature of splice-site consensus sequences, auxiliary *cis*-acting elements are required for splice-site recognition (Braunschweig et al., 2013; Wahl et al., 2009). Regulatory elements promoting the recognition of splice-sites are known as exonic or intronic splicing enhancers (ESEs/ISEs) (Figure 1.2A). Conversely, regulatory elements inhibiting the recognition of splice sites are known as exonic or intronic splicing silencers (ESSs/ISSs) (Figure 1.2A) (House and Lynch, 2008). The auxiliary *cis*-acting

elements exert their functions by providing sites for binding of *trans*-acting regulatory factors (Figure 1.2A) (Wahl et al., 2009). The members of SR protein family (enriched in Arg-Ser dipeptides) typically bind to ESEs/ISEs and promote splicing while the heterogeneous nuclear ribonucleoprotein (hnRNP) protein family members typically bind to ESSs/ISSs, inhibiting splicing (Figure 1.2A) (House and Lynch, 2008).

As human pre-mRNAs are comprised of shorter exons among very long introns, the recognition of a 5'ss by U1 snRNP can be facilitated by the binding of U2AF at the 3'ss of the same exon and vice-versa. This process is referred to as exon definition (Figure 1.2B) (Berget, 1995; De Conti et al., 2013). For the first exon, 5' end capping of the pre-mRNA enables the binding of U1 snRNP at the 5'ss (Figure 1.2B) (Lewis et al., 1996). For the last exon, the recruitment of U2AF at the 3'ss is augmented by polyadenylation signals at the 3' end (Figure 1.2B). In fact, splicing of last exon and polyadenylation are mutually coordinated and are achieved by the cooperation amongst the splicing machineries and cleavage or polyadenylation factors (Kyburz et al., 2006; Millevoi et al., 2006).

Polyadenylation of nearly all eukaryotic mRNAs takes place in two consecutive steps. First, pre-mRNA endonucleolytic cleavage occurs and then polyadenylate tail addition ensues downstream of the cleavage site. The processing is achieved by protein factors binding to different *cis*-acting elements. The cleavage and polyadenylation specificity factor (CPSF) binds to the highly conserved polyadenylation (polyA) site, AAUAAA, which lies upstream (10–30 nucleotides) of the cleavage site. The cleavage stimulatory factor (CstF) recognizes a variable G/U enriched sequence located downstream of the cleavage site. Following this, cleavage occurs with the presence of cleavage factors I and II (CFI and CFII). PolyA polymerase (PAP) adds around 10 adenosine nucleotides at the 3' end. Further elongation of the polyA tail by PAP depends on the interaction between polyA-binding protein II (PABII), PAP and CPSF (Li et al., 2001; Murthy and Manley, 1995).

Figure 1.2 Splicing regulation and exon definition

(A) SR proteins promote splicing by binding to ESEs or ISEs while hnRNPs inhibit splicing by binding to ESSs or ISSs. (B) The recruitment of U1 snRNP at the 5'ss is enhanced by the binding of U2AF at the 3'ss and vice-versa of the exon by exon definition. In addition, capping enhances the binding of U1 snRNP at the 5' splice site of first exon while polyadenylation enhances recruitment of U2AF at the 3'ss of last exon.

1.3 Alternative splicing

Pre-mRNA can be spliced in multiple ways to give rise to different mRNA isoforms that are translated to different proteins in a process called alternative splicing. Alternative splicing results in the generation of the multitude of distinct proteins in multicellular eukaryotic organisms (Nilsen and Graveley, 2010). In fact, the production of >90,000 different proteins from the ~21000 protein-coding genes in the humans can be attributed to the process of alternative splicing (Ast, 2004). A recent examination of transcriptome wide high-throughput RNA sequencing data has revealed that in humans alternative splicing occurs in > 90% of pre-mRNAs (Wang et al., 2008).

Alternative splicing is differentially regulated in tissues, at different developmental or differentiation stages, or in response to various signals (Keren et al., 2010; Nilsen and Graveley, 2010). At cellular level, alternative splicing plays a crucial role in diverse processes including apoptosis, cell-cycle progression, pluripotency, circadian rhythms, responses to environmental stimuli and pathogen invasion (Braunschweig et al., 2013; Chen and Manley, 2009).

1.3.1 Alternative splicing types

Seven different types of alternative splicing have been defined to date: cassette-exon, mutually exclusive exons, intron retention, alternative 5' or 3' splice sites, and alternative promoters or polyA sites (Black, 2003; Keren et al., 2010) (Figure 1.3). Cassette exons comprise exons which can be retained or skipped from the mature mRNA. Mutually exclusive exons involve inclusion of only one of the two or more exons in the mature mRNA. Intron retention leads to intron-containing mature transcripts. Alternative 5' or 3' splice sites gives rise to exons having different sizes and boundaries. Alternative promoters or polyA sites generate different terminal exons due to use of different 5'ss or 3'ss.

Figure 1.3 Schematic of the seven general types of alternative splicing

See text for details. Image is adapted from Zahler, 2005.

1.3.2 Regulatory mechanisms of alternative splicing

Alternative splicing is regulated by *cis*-acting elements and *trans*-acting factors, pre-mRNA secondary structure, transcription and chromatin structure (Chen and Manley, 2009; Nilsen and Graveley, 2010; Smith and Valcarcel, 2000).

Auxillary *cis*-acting elements include ESEs/ISEs that promote the use of splice sites and ISSs/ESSs that repress the use of splice sites. *Trans*-acting factors include splicing activators such as SR proteins which bind to ESEs and ISEs and promote exon inclusion and splicing repressors such as hnRNPs which bind to ESSs/ISSs and promote exon exclusion. The common SR proteins include SRSF1, SRSF2, SRSF3, SRSF4, SRSF5 and SRSF6. Some common hnRNP proteins include hnRNP A1, hnRNP A2/B1, PTBP1 (hnRNP I), PTBP2, hnRNP L and hnRNP LL (Chen and Manley, 2009). Splicing

activators can favor the interaction between U1 snRNP and U2AF across an exon, promoting their recruitment to the 5'ss and 3'ss by exon definition (Figure 1.4A) (Smith and Valcarcel, 2000). Alternatively, splicing activators can antagonize the binding of splicing repressors or cooperate with other positive regulatory factors to facilitate the recognition of splice sites (Figure 1.4A) (Cartegni et al., 2002; Chen and Manley, 2009). Splicing repressors can sterically block binding of snRNPs or splicing activators to the pre-mRNA. Splicing inhibition could also be achieved by multimerization of repressors along the pre-mRNA, thus preventing splice-site recognition by the spliceosome (Figure 1.4B). Alternatively, some repressors could engage in a protein–protein interaction across an alternative exon causing 'looping out' of the exon. This loop formation may sterically hinder snRNP recruitment (Figure 1.4B) (Chen and Manley, 2009; House and Lynch, 2008). Ultimately, it is the sum of positive and negative factors that leads to inclusion or exclusion of an exon (Nilsen and Graveley, 2010; Wahl et al., 2009).

Apart from the widely expressed SR and hnRNP group of splicing factors, tissue-specific expression of splicing regulators that contributes to the differential expression of alternative splicing isoforms in the specific tissues (Wang et al., 2008). The most abundant occurrence of tissue-specific alternative splicing isoforms is found in the brain which is the most diverse tissue functionally. Examples of these neural-specific factors include neuro-oncological ventral antigen (NOVA) (Ule et al., 2005) and embryonic lethal abnormal visual (Elav) (Perrone-Bizzozero and Bolognani, 2002).

Figure 1.4 Regulation of alternative splicing by *cis*-acting elements and *trans*-acting factors

(A) Splicing activators such as SR proteins bind to ESEs or ISEs and facilitate exon inclusion by recruiting the spliceosome or inhibiting the binding of nearby repressors. (B) Splicing repressors such as hnRNPs bind to ESSs or ISSs and promote exon skipping. The repression is attained by (i) multimerization of splicing repressors which blocks the access of the spliceosome to the splice sites or activator to enhancer sequences; (ii) looping out of the alternative exon due to interaction between repressors located at the two flanking introns; (iii) inhibiting the binding of spliceosome components or splicing activators thereby preventing exon definition.

Pre-mRNA secondary structure formation can influence alternative splicing by blocking core splicing signals or enhancer or silencer regions (Chen and Manley, 2009). The regulation of splicing by transcription has been explained by two models: recruitment model and kinetic model (Figure 1.5A and B respectively). In the recruitment model, the RNA polymerase II carboxy-terminal domain (CTD) (Pol II) directly recruits the splicing factors to pre-mRNA or structurally different promoters cause recruitment of different splicing factors leading to exon inclusion or exclusion (Figure 1.5A). In the kinetic model, the RNA Pol II transcription rate influences splicing. For example, if splice sites are weak, rapid elongation will favor recognition of distal stronger splice sites by the spliceosome causing exclusion of the exon, while slow elongation will favor recruitment of the spliceosomal complex at weak splice sites, promoting exon inclusion (Figure 1.5B) (Braunschweig et al., 2013; Chen and Manley, 2009; Kornblihtt et al., 2013).

Alternatively, slowing the rate of transcription can also promote exon skipping by favoring the recognition of silencer elements by splicing repressors (Dujardin 2014). Splicing regulation by chromatin structure encompasses the influence of histone modifications and nucleosome positioning. Histone post-translational modifications can indirectly affect splicing by causing chromatin structure to become loose or compact thus causing a faster or slower rate of RNA Pol II transcription, respectively (Figure 1.5C). Histone marks can also regulate splicing by binding to chromatin adaptors which in turn can interact with splicing factors (Figure 1.5D) (Montes et al., 2012). The favorable positioning of nucleosomes at exons compared to introns, also tag the beginning of an exon. This in turn causes transient RNA Pol II elongation pause, thus providing a time-window for the recognition of 3'ss by U2AF (Kornblihtt et al., 2013).

Thus splicing is a highly complex and regulated process and a splicing disruption occurring at any level of regulation can lead to disease.

Figure 1.5 Alternative splicing regulation by transcription and chromatin

(A) Recruitment model: The CTD of RNA Pol II directly recruits the splicing repressor to pre-mRNA thereby leading to exon exclusion (left) while lack of interaction between Pol

II CTD and splicing repressor leads to exon inclusion (right). (B) Kinetic model: Fast Pol II elongation favors recognition of distal stronger splice sites instead of the upstream weak 3'ss causing exon exclusion (left). Slow Pol II elongation favors recognition of the spliceosomal complex at the upstream weak 3'ss splice sites, promoting exon inclusion (right). (A, B) Images adapted from Hall and Georgel, 2011. (C) Histone post-translational modifications can cause chromatin structure to be loose causing faster rate of Pol II elongation leading to exon exclusion (left). Other histone modifications can also cause chromatin structure to be more compact causing a slower rate of pol II elongation leading to exon inclusion (right). (D) Some histone modifications can recruit a splicing repressor through the interaction with adaptor protein, thereby leading to exon exclusion (left). In contrast, lack of the histone modification inhibits the recruitment of repressor thereby leading to exon inclusion. (C, D) Images adapted from Kornblihtt et al., 2013. The green circle indicates splicing repressor; the brown circle indicates spliceosome and the blue square indicates an adaptor protein.

1.4 Splicing and disease

Mutations in splicing regulatory elements and factors account for a high fraction of human hereditary diseases and cancers (Kornblihtt et al., 2013). Between 15 to 50% of mutations cause disease by disrupting splicing (Krawczak et al., 1992; López-Bigas et al., 2005). These mutations affect *cis*-acting elements while *trans*-acting factors could be affected by mutations as well as alterations in their levels.

1.4.1 *Cis*-acting mutations

The essential or auxiliary *cis*-acting elements can have mutations which may impair recognition of constitutive or alternative splice sites.

Disruptions in the recognition of constitutive splice sites can cause skipping of constitutive exons, intron retention or cryptic splice site activation. As a consequence, there could be an abnormal protein production or mRNA nonsense mediated decay (NMD) (Singh and Cooper, 2012). An example of a mutation affecting constitutive exon splicing

is familial dysautonomia (FD) which is an autosomal recessive condition affecting the autonomic as well as sensory neurons. FD is caused by mutations occurring in the *inhibitor of kappa light polypeptide gene enhancer in B-cells, kinase complex-associated protein (IKBKAP)* gene that hampers the production of IκB kinase complex associated protein (IKAP), an important constituent of the transcriptional elongation complex (Rubin and Anderson, 2008). In 98 % of FD cases, a substitution in intron 20 (a T > C transition in 5'ss) causes exon 20 skipping. The resultant mRNA has a premature termination codon due to shift in the open reading frame. This leads to decreased production of functional IKAP due to NMD (Anderson et al., 2001; Rubin and Anderson, 2008; Slaugenhaupt et al., 2001), affecting the transcription of many genes involved in cell motility. The consequential neuronal dysfunction is possibly due to impaired neuronal cellular migration (Close et al., 2006).

Disruptions in the alternative splice sites recognition can change proportion of alternatively spliced mRNA transcripts resulting in undesired biological effects (Faustino and Cooper, 2003; Singh and Cooper, 2012). Examples include frasier syndrome, atypical cystic fibrosis and frontotemporal dementia and parkinsonism linked to Chromosome17 (FTDP-17). Atypical cystic fibrosis is a milder form of cystic fibrosis and occurs due to mutations in the gene encoding cystic fibrosis transmembrane conductance regulator (CFTR) protein. These mutations reduce production of functional CFTR protein, affecting the pancreas, lungs and testes (Noone and Knowles, 2001). However, humans comprising the same *CFTR* mutations exhibit variations in the severity of disease. One such factor accounting for the difference in severity is polymorphisms affecting UG and U repeats located at the 3'ss of *CFTR* exon 9 that can affect exon 9 skipping. There is always slight exon 9 skipping even in unaffected individuals, but in individuals with atypical cystic fibrosis, increased skipping of exon 9 further compromises the production of full-length functional protein (Steiner et al., 2004).

Mutations in certain splicing regulatory elements can also cause cancers. Mutations in tumor-suppressor genes like *breast cancer 1 (BRCA1)* and *mutL homolog 1 (MLH1)* have been shown to affect splicing thereby inactivating their tumor-suppressor function (Tazi et al., 2009). For example, a nonsense mutation in the *BRCA1* constitutive exon 18 gene

abrogates binding of SRSF1 to an ESE thereby causing exclusion of the exon (Liu et al., 2001), while activation of a cryptic 3'ss due to AA to AG mutation adds a few nucleotides to *BRCA1* mRNA, generating a truncated protein in a type of breast cancer (Hoffman et al., 1998).

Thus mutations in *cis*-acting elements can affect splicing of a single gene and contribute to disease or cause variations in disease severity among individuals.

1.4.2 *Trans*-acting mutations and alterations in expression

Mutations occurring in genes encoding *trans*-acting splicing factors or alterations in their levels have a wide impact on the cell because they affect expression of many other genes. Such mutations affecting the expression or availability of splicing factors are linked with diseases such as retinitis pigmentosa, myotonic dystrophy, autism spectrum disorder and amyotrophic lateral sclerosis (Singh and Cooper, 2012; Ward and Cooper, 2010).

Retinitis pigmentosa involves the progressive damage of retinal cells resulting in visual impairment. The autosomal dominant form of retinitis pigmentosa can occur due to mutations affecting general splicing factors such as pre-mRNA processing factor 3 (PRPF3), pre-mRNA processing factor 8 (PRPF8) and pre-mRNA processing factor 31 (PRPF31) that are components of the tri-snRNP U4/U6.U5 (Mordes et al., 2006).

Myotonic dystrophy type 1 (DM1) which is the most prevalent form of myotonic dystrophy, is caused by expansion of a 3' untranslated region CUG repeat of *myotonic dystrophy protein kinase* (*DMPK*) gene (Mahadevan et al., 1992). The long CUG repeats provide strong binding site for the splicing factor, muscleblind like splicing regulator 1 (MBNL1). This causes its nuclear foci colocalization, exhausting nucleoplasmic MBNL1 (Lin et al., 2006). The splicing regulator, CUG binding protein 1 (CUGBP1), is also up-regulated due to DMPK CUG-repeat mediated protein kinase C-mediated hyperphosphorylation and stabilization of CUGBP1 (Kuyumcu-Martinez et al., 2007). Thus, disease pathogenesis such as myotonia and insulin resistance results from functional loss of MBNL1 and upregulation of CUGBP1, resulting in the perturbations in alternative splicing events in muscle and other tissues (Charlet et al., 2002; Cho and Tapscott, 2007).

Interestingly, MBNL1 and CUGBP1 have antagonistic effects on splicing which further adds to the alterations in splicing events (Cho and Tapscott, 2007).

Most cancer-related splicing changes are caused due to alterations in levels of *trans*-acting splicing factors. Altered expression of splicing factors such as SRSF1, hnRNP A1, hnRNP H, RNA-binding protein 5 (RBM5), and RNA-binding protein 10 (RBM10) is known to cause aberrant splicing of pre-mRNA targets resulting in malignant phenotypes (Singh and Cooper, 2012). A well-known example is the up-regulation of SRSF1 that causes production of tumorigenic isoforms (Anczukow et al., 2012; Anczuków et al., 2015; Ghigna et al., 2005; Karni et al., 2007). For example, SRSF1 overexpression promotes *Recepteur d'Origine Nantaïs* (*RON*) tyrosine kinase receptor pre-mRNA alternative splicing to produce a constitutively active *RON Δ II* isoform, which enhances cell metastasis, a phenotype of aggressive tumor cells (Ghigna et al., 2005). SRSF1 overexpression also favors production of the isoforms of the ribosomal S6 kinase 1 (S6K1) and mitogen-activated protein kinase interacting serine/threonine kinase 2 (MNK2) kinases that have antiapoptotic and proliferative effects. Further SRSF1 overexpression promotes the generation of *BINI*+12a isoform that impairs Myc binding and hence eliminates its tumor suppressor function (Karni et al., 2007). Thus even moderate variations in expression or levels of *trans*-acting factors can have significant effects on alternative splicing.

1.4.3 Single nucleotide polymorphisms

Similar to splicing *cis*-acting and *trans*-acting factor mutations that affect splicing, single nucleotide polymorphisms (SNPs) can also change splicing phenotype. A SNP can be defined as a variation in a single nucleotide of any gene, occurring in >1 % of at least one ethnic group (Cartegni et al., 2002). Any SNP located at the splice sites or positive or negative splicing regulatory elements can modulate splicing and have detrimental effects (Cartegni et al., 2002; Lu et al., 2012). It has been revealed that 21% of alternatively spliced transcripts emerge from allele-specific splicing (Nembaware et al., 2004). Lately several studies have demonstrated SNPs to affect pre-mRNA splicing. For example, *endoplasmic reticulum aminopeptidase 2* (*ERAP2*), which is a gene required for major

histocompatibility complex (MHC) class I antigen presentation contains a SNP (A to G) located at exon 10 5'ss. This activates a downstream cryptic splice site resulting in NMD of the mRNA and subsequent reduction in functional protein levels (Andrés et al., 2010). In individuals that are biallelic for the A to G SNP, there is less cleavage of antigenic peptides and reduced expression of MHC Class I at the surface of B cell. Similarly, a G to A SNP occurs in *2'-5' oligoadenylate synthetase 1(OAS1)*, a gene responsible for innate immune reaction to viral infections. This SNP present at exon 7 3'ss promotes use of an internal 3'ss generating a protein with compromised activity (Bonnevie-Nielsen et al., 2005). Hepatitis C virus patients are also affected by this SNP because patients harbouring the biallelic SNP respond sub-optimally to interferon treatment and develop liver fibrosis (El Awady et al., 2011).

SNPs altering splicing have been also shown to be associated with cancer. For example, a G to A SNP in the *cyclin D1 (CCND1)* gene results in production of an alternative splice variant, which increases the risk of developing lung cancer (Gautschi et al., 2007).

Furthermore, an intronic G to A SNP in a tumor suppressor gene encoding Kruppel-like factor 6 (KLF6), a member of Kruppel-like family of zinc finger transcription factors, results in the generation of three alternative splicing forms of KLF6. This SNP was found to be associated with increased prostate cancer risk in men (Narla et al., 2005).

Thus SNPs affecting splicing could contribute to phenotypic heterogeneity in terms of disease susceptibility, disease severity or response to drug treatment (Wang and Cooper, 2007) .

1.5 Splicing and therapy

Given the plethora of diseases caused due to improper splicing, studying the mechanisms of splicing regulation and the subsequent management of the diseases is of key relevance. Indeed, substantial progress has been made in therapies used to correct aberrant splicing defects, which include use of splice-switching anti-sense oligonucleotides (ASOs), small molecule compounds, modified snRNAs and *trans*-splicing (Havens et al., 2013).

ASOs which are of particular relevance in modulating *BIM* E3 splicing, are short oligonucleotides (~15-25 nt) that can hybridize to specific complementary regions of the pre-mRNA by Watson-Crick base-pairing. Splice-switching ASOs can redirect splicing by sterically blocking interactions of the splicing machinery with target pre-mRNA. Splice-switching ASOs binding splice sites can induce exon skipping. Splice-switching ASOs binding splicing enhancers or silencers can prevent association of *trans*-acting splicing factors and cause exon skipping or exon inclusion respectively (Bennett and Swayze, 2010). Clinical trial studies have established the success rates of splice-switching ASOs for Duchenne muscular dystrophy (DMD) (Mendell et al., 2013; Voit et al., 2014) and spinal muscular atrophy (SMA) (<https://clinicaltrials.gov/ct2/show/NCT02386553>).

DMD is a neuromuscular disorder that results in muscle loss and death (Douglas and Wood, 2013). It occurs due to deletions and mutations within the dystrophin gene. Splicing of remaining exons disrupts the translational reading frame leading to loss of dystrophin function. ASOs binding specific exons have shown to promote skipping of those exons thereby rescuing the reading frame and enhancing expression of dystrophin protein (Hammond and Wood, 2011).

Spinal muscular atrophy (SMA) is an autosomal recessive neurodegenerative disease causing muscle loss and is one of the major hereditary causes of deaths in infants. It occurs due to mutations/deletions in the survival of motor neuron 1 (*SMN1*) alleles causing its loss of function (Lefebvre et al., 1995). *SMN2*, paralog of *SMN1*, fails to rescue *SMN1* loss. This is because *SMN2* contains a C to T substitution in the 6th nucleotide of exon 7 leading to its skipping and subsequent generation of a nonfunctional shortened protein (Lorson et al., 1999; Wirth, 2000). *SMN1* is essential for snRNP assembly ubiquitously, however its deficiency mainly causes impairment of motor neurons (Monani, 2005).

In SMA, various ASOs targeting *SMN2* exon 7 and nearby intronic regions have been tested to promote exon 7 inclusion (Hua et al., 2007; Passini et al., 2011; Porensky et al., 2011). The splice switching ASO ISIS-SMN-Rx which binds and blocks an ISS in *SMN2* intron 7 has caused increase in expression of full-length protein in cell culture and mouse

models and has shown promising improvements in reversing the SMA defects (Havens et al., 2013; Hua et al., 2008; Hua et al., 2010; Hua et al., 2011). Moreover, splice-switching ASOs for cancer against targets like B-cell lymphoma-X (BCL-X) and signal transducer and activator of transcription 3 (STAT3 β) are in clinical trials (Bauman et al., 2010; Zammarchi et al., 2011).

In my project, I studied the mechanisms of alternative splicing regulation of the proapoptotic factor *BIM* which is clinically relevant in the context of CML and EGFR NSCLC. In the below sections each of these concepts are introduced in detail.

1.6 Apoptosis

Apoptosis is a physiological process of programmed cell death that is critical for development, tissue homeostasis and defense mechanism in immune responses (Adams and Cory, 2007; Elmore, 2007). Therefore its deregulation underlies many diseases, including cancer, autoimmune and neurodegenerative disorders (Elmore, 2007; Strasser et al., 2000). The characteristic morphological changes associated with apoptosis are shrinking of cells, condensation of chromatin, blebbing of cell membrane, fragmentation of nucleus, and separation of cellular fragments into apoptotic bodies. Eventually, the apoptotic bodies are cleared by phagocytosis by macrophages or surrounding cells (Elmore, 2007; Yuan et al., 2003).

The intrinsic or mitochondrial and the extrinsic pathway constitute the two key apoptotic pathways in mammals (Figure 1.6) (Adams and Cory, 2007; Klener et al., 2006), which are executed by a family of cysteine-dependent aspartate-directed proteases known as caspases (Adams and Cory, 2007). Cellular stress like radiation, growth factor deprivation, and cytotoxic drugs stimulate the intrinsic pathway (Strasser et al., 1995). The regulation of this pathway is achieved by the B-cell lymphoma 2 (BCL-2) protein family members. The activation of proapoptotic proteins BCL-2-associated X protein (BAX) and BCL-2 homologous antagonist/killer (BAK) causes the discharge of mitochondrial apoptogenic molecules like cytochrome *c* due to mitochondrial outer membrane permeabilization. Cytochrome *c* binds to apoptosis protease-activating factor-1, and procaspase-9 (initiator caspase) into a multiprotein complex, called apoptosome, thus activating caspase-9

(Chipuk et al., 2010). Caspase-9 in turn activates the effector caspases (i.e. caspases-3, 6 and 7). The effector caspases ultimately cleave target substrates, activating DNases and mediating cell death. The extrinsic pathway begins when extracellular ligands of the tumor necrosis factor (TNF) receptor family bind to their receptors. This causes a final activation of the cascade of “effector caspases”, mediated by caspase-8 activation (Elmore, 2007). Though the two pathways are independent, the intersection of the two pathways is found in few cell types (e.g. hepatocytes) whereby caspase-8 mediates cleavage and activation of BH3 interacting-domain death agonist (BID) to truncated BID (tBID). This triggers effector caspases to be activated through the mitochondrial pathway (Figure 1.6) (Adams and Cory, 2007; Youle and Strasser, 2008).

Figure 1.6 Schematic of apoptotic signaling pathways

Apoptosis is triggered by ligands of the tumor necrosis factor family binding to their respective receptors (extrinsic pathway) or by cellular stress (intrinsic pathway). In the extrinsic pathway, there is direct activation of caspase-8. The intrinsic pathway involves cytochrome *c* release and caspase-9 activation upon BAK/BAX activation. Both pathways lead to activation of the effector caspases (caspase-3,6,7). In few cell types such as hepatocytes, intersection of the intrinsic and extrinsic pathways is observed. Image is adapted from Adams and Cory, 2007.

1.7 BCL-2 protein family

The BCL-2 protein family members contain up to four BCL-2 homology (BH) domains (BH1 to BH4) which are conserved motifs (Adams and Cory, 1998). Based on the differences in structure and function, they can be grouped into three subfamilies and to date, up to 30 members of this family have been isolated (Chipuk et al., 2010; Tzifi et al., 2011). The pro-survival or antiapoptotic subfamily includes BCL-2, B-cell lymphoma-extra-large (BCL-xL), and myeloid cell leukemia sequence 1 (MCL-1) proteins, which contain all four conserved BH domains. The proapoptotic proteins include: (1) the multidomain members, such as BAX and BAK protein containing BH 1–3 domains and (2) the “BH3-only” proteins, such as BCL-2-interacting mediator of cell death (BIM) and, p53 upregulated modulator of apoptosis (PUMA) which contain only the BH3 domain (Packham and Stevenson, 2005). The BH3 domain is the essential domain responsible for cell death in the proapoptotic members as observed by mutagenesis and deletion studies (Gross et al., 1999).

There are two proposed models that describe the role of BH3-only proteins in the intrinsic apoptotic pathway. The “direct activation model”, involves cooperation between activators and sensitizers of the BH3-only proteins. Activators include BH3-only proteins (BIM, BID) that upon binding to BAK/BAX, cause their activation. While the sensitizers include other BH3-only proteins such as BCL-2-associated death promoter (BAD), BCL-2-interacting killer (BIK) and Phorbol-12-myristate-13-acetate-induced protein 1 (NOXA), that bind to the antiapoptotic BCL-2 proteins. This releases any activator proteins bound to the antiapoptotic BCL-2 proteins, permitting them to directly activate BAX and BAK. The “indirect activation model” suggests that all BH3-only proteins engage with the antiapoptotic BCL-2 proteins, thus preventing them from inhibiting BAK/BAX activation (Adams and Cory, 2007; Kang and Reynolds, 2009). Once activated, BAX and BAK oligomerize causing mitochondrial outer membrane permeabilization (Chipuk et al., 2010). This leads to release of cytochrome *c* from the mitochondria and subsequent activation of caspases resulting in apoptotic cell death (Figure 1.6) (Chipuk et al., 2010; Tzifi et al., 2011). Thus tissue homeostasis is maintained by the delicate balance between the levels of proapoptotic and antiapoptotic BCL-2 proteins and either upregulation of antiapoptotic

members or downregulation of proapoptotic members can be oncogenic (Adams and Cory, 2007). For example, BCL-2 was first found to be upregulated in follicular B-cell lymphoma (Bakhshi et al., 1985). Subsequently, over-expressed BCL-2 and MCL-1 have been implicated in a number of other cancers (Dewson and Kluck, 2010), rendering tumor cells resistant to apoptotic stimuli, including most chemotherapeutics (Reed, 2003). Recently, drugs specifically targeting the BCL-2 protein family for the treatment of malignancies are being developed as they have shown considerable potential (Dewson and Kluck, 2010). Among the BCL-2 family members, the proapoptotic protein, BIM has garnered special attention as a potent target for cancer-based therapy, due to its tumor suppressor role (Akiyama et al., 2009).

1.8 BIM

BIM also known as BCL2-like 11 or BCL2L1 is a BH3-only protein that was first identified by O'Connor et al in 1998, while screening a bacteriophage λ cDNA expression library for BCL-2 binding proteins (O'Connor et al., 1998). Later in the same year mouse BIM was discovered while screening an ovarian cDNA library using MCL1 as bait via a yeast two-hybrid approach (Hsu et al., 1998).

BIM contains the short (nine amino acid) BH3 motif present in all Bcl-2 family members. The BH3 domain in BIM has the proapoptotic activity, as lack of BH3 domain was found to prevent cell death (O'Connor et al., 1998). In mammalian cells, BIM binds to all the antiapoptotic proteins of the BCL-2 family by inserting its BH3 domain into the hydrophobic groove of the antiapoptotic members, making it highly cytotoxic (Willis and Adams, 2005). Moreover, BIM can directly interact with BAK and BAX, to cause mitochondrial outer membrane permeabilization (Youle and Strasser, 2008). Like BCL-2, a hydrophobic C-terminus in BIM targets it to intracytoplasmic membranes (O'Connor et al., 1998). The expression of BIM has been found in tissues such as brain, lung, trachea, heart, liver, kidney, thymus, spleen, testis and ovary but most highly in cells of hematopoietic origin (O'Reilly et al., 2000).

1.8.1 Gene structure and isoforms

In humans, *BIM* is located at chromosome 2q13 (Bouillet et al., 2001). The *BIM* gene contains five exons (Figure 1.7A). Exon 2A (E2A) contains the start codon while exon 3 (E3) and exon 5 (E5) both contain an in-frame stop codon and polyadenylation signal (PAS) for the 3'-end formation of mature mRNA (Figure 1.7A) (Ng et al., 2012).

Alternative splicing of the *BIM* pre-mRNA generates different mRNA isoforms, and to date, at least 12 splice variants have been identified in humans in different cell types (Adachi et al., 2005). However, the three major splice variants are *BIMEL*, *BIML* and *BIMS* (O'Connor et al., 1998) which contain the BH3 domain encoding E4 (Figure 1.7B). *BIMEL* and *BIML* are more prevalent in tissues, whereas *BIMS* is less prevalent (Bouillet et al., 2001). The three isoforms differ in their proapoptotic activity, the shortest *BIMS* being the most potent (O'Connor et al., 1998). This is because of association of the less potent forms, *BIML* and *BIMEL* but not *BIMS* to the cytoskeletal dynein light chain, LC8. The dynein binding domain (DBD) is encoded by Exon 2C which is present in *BIMEL* and *BIML*, but not *BIMS* (Figure 1.7B). Only when triggered by an apoptotic signal, there is release of *BIML* and *BIMEL* from microtubules to inhibit BCL-2 (Puthalakath et al., 1999) while *BIMS* directly interacts with the proapoptotic BAX (Marani et al., 2002).

Apart from these transcripts, two other transcripts that have gained attention are *BIM γ 1* and *BIM γ 2*. These transcripts are produced as a result of inclusion of the alternative E3. E3 is a terminal exon because it lacks a 5'ss to undergo splicing with E4 and has its own polyA signal. As a result, *BIM γ 1* and *BIM γ 2* lack the proapoptotic BH3 domain (encoded by E4) and hence cannot activate the intrinsic apoptotic-pathway (Anczukow et al., 2012; Ng et al., 2012).

Figure 1.7 Schematic of *BIM* gene structure and transcripts.

(A) *BIM* comprises five exons (E1, E2, E3, E4, and 5). E2A contains the translational start codon, while E3 and E5, both contain PolyA signal and stop codon. A dynein binding domain (DBD) is encoded by E2C. E4 contains the proapoptotic BH3 domain while E5 contains a hydrophobic domain. (B) Alternative splicing can give rise to different *BIM* isoforms (*BIM_{EL}*, *BIM_L*, *BIM_s* and *BIM_{γ1}* and *BIM_{γ2}*). Abbreviations: E, exon; PAS, polyadenylation signal (Ng et al., 2012). Image is adapted from Ng et al., 2012.

1.8.2 Physiological roles and implication in cancer-based therapy

Most of the physiological functions of BIM have been elucidated in mice by gene-targeting experiments. Firstly, BIM plays a critical role in hematopoietic homeostasis as BIM was found to provoke apoptosis of lymphocytes upon stimuli such as calcium ion flux and deprivation of cytokines. Moreover, BIM deficient mice showed accumulation of lymphoid and myeloid cells (Bouillet et al., 1999). Homeostasis of naive and memory T cells is regulated by BIM (Wojciechowski et al., 2007) and apoptosis of antibody-forming cells and memory B cells (Fischer et al., 2007) is governed by BIM. Besides, BIM contributes to prevention of autoimmunity by mediating death of autoreactive T cells

(Bouillet et al., 2002) and B cells (Enders et al., 2003) as well as terminating immune reactions after clearance of infections (Hildeman et al., 2002; Pellegrini et al., 2003) .

Given the functional significance of BIM, its deficiency could lead to autoimmune disorders or malignancies (Pinon et al., 2008). Indeed BIM deficient mice developed a fatal autoimmune kidney disease, resembling systemic lupus erythematosus in human (Bouillet et al., 1999). BIM's role as a tumor suppressor is beginning to be understood and was first demonstrated in *Eu-myc* mice, whereby deletion of a single *Bim* allele augmented the progression to acute B-cell-leukemia (Egle et al., 2004). In Burkitt's lymphomas the *BIM* gene was found to be inactivated by enhanced methylation of promoter regions (Mestre-Escorihuela et al., 2007) and in mantle cell lymphomas biallelic loss of *BIM* gene was observed (Tagawa et al., 2004).

Given its role as a tumor suppressor, BIM has been in the spotlight as a plausible target in many cancer-based therapies (Akiyama et al., 2009). For instance, BIM expression is essential for inducing apoptosis by glucocorticoids in acute lymphocytic leukemia (Bouillet et al., 1999) and paclitaxel in lymphocytes and epithelial tumor cells (Bouillet et al., 1999; Tan et al.). Proteasome inhibitors such as the boronic anhydride bortezomib (Velcade) activate BIM indirectly by upregulating NOXA which in turn causes displacement of BIM from MCL1. Free BIM then activates BAK/BAX to cause apoptosis in multiple myeloma (Gomez-Bougie et al., 2007). Bortezomib also causes apoptosis in prostate and colon cancer cells by death receptor ligand Trail mediated BIM activation (Nikrad et al., 2005). Histone deacetylase inhibitors also mediate their lethal effects by upregulating BIM. For example the epigenetic inactivation of BIM was reversed by histone deacetylase inhibitor which in turn sensitized Burkitt's lymphoma cells to many chemotherapeutic agents (Richter-Larrea et al., 2010). Vorinostat, a histone deacetylase inhibitor was shown to enhance BIM expression, among many other BH3-only proteins by facilitating acetylation of the promoter sequences in mantle cell lymphomas (Xargay-Torrent et al., 2011).

Importantly, BIM has a crucial function in mediating apoptosis in many tyrosine kinase inhibitor (TKI) based therapies. For example, the TKI, imatinib upregulates BIM in CML

(Aichberger et al., 2005; Kuribara et al., 2004; Kuroda et al., 2006) while the TKI, gefitinib upregulates BIM in EGFR NSCLC (Costa et al., 2007). The upregulation of BIM is achieved by transcriptional as well as post-translational mechanisms (Essafi et al., 2005; Ley et al., 2003; Luciano et al., 2003). Transcriptional upregulation of *BIM* is driven by the Forkhead box O3 (FOXO3A) transcription factor which is activated due to shutdown of Phosphatidylinositol 3-kinase (PI3K) Akt signaling (Essafi et al., 2005). The post-translational upregulation of BIM is mediated by inhibition of extracellular-signal-regulated kinase (ERK), which phosphorylates BIM, thus targeting it for ubiquitination and proteasomal degradation (Ley et al., 2003; Luciano et al., 2003).

1.9 Chronic myeloid leukemia (CML)

CML is a myeloproliferative disorder resulting from the neoplasm of hematopoietic stem cells (Garcia-Manero et al., 2003; Ren, 2005). Most CML patients (more than 90%) exhibit the characteristic Philadelphia chromosome (Figure 1.8) (An et al., 2010; Epstein et al., 1999). Discovered in 1960 by Nowell and Hungerford, the Philadelphia chromosome was the first ever chromosomal aberration to be linked to a malignant disease (Nowell, 1960). Subsequently Rowley and colleagues showed that the Philadelphia chromosome resulted from a reciprocal translocation between the long arms of chromosomes 9 and 22 t(9;22)(q34;q11) (Rowley, 1973) and a decade later this translocation was found to involve Abelson murine leukemia viral oncogene homolog (ABL) on chromosome 9 (Bartram et al., 1983) and the breakpoint cluster region (BCR) gene on chromosome 22 (Groffen et al., 1984) (Figure 1.8). This fusion gene was first demonstrated to be a potent oncogene when murine recipients of BCR-ABL retrovirus-transduced bone marrow manifested a CML-like disease (Daley et al., 1990).

Figure 1.8 Schematic of Philadelphia chromosome

Reciprocal translocation between parts of chromosome 9 and chromosome 22 generates *BCR-ABL* fusion gene encoding the BCR-ABL oncoprotein. Image is modified from Lydon et al., 2009.

1.9.1 BCR-ABL protein isoforms

Patients can have various BCR-ABL isoforms, depending on the translocation breakpoints and alternative pre-mRNA splicing (Ren, 2005). In the *ABL* gene, breakpoints can occur either upstream of alternative exon Ib, downstream of the alternative exon Ia, or, more commonly between the two. However due to splicing, the final hybrid mRNA contains exon a2 (Figure 1.9). Three different breakpoints in the *BCR* gene: major (M-bcr), minor (m-BCR) and micro (μ -bcr) generate BCR-ABL fusion proteins with different molecular weights (Deininger et al., 2000; Quintás-Cardama and Cortes, 2009).

The breakpoints within *BCR* occurring between exons e12-e16 (also called b1-b5) are referred to as the major breakpoint cluster region (M-bcr). The breakpoint at this region gives rise to fusion transcripts with either b2a2 (e13a2) or b3a2 (e14a2) junctions that are translated to a 210 kDa protein (p210BCR-ABL). The breakpoint in the minor breakpoint cluster region (m-bcr) between the alternative *BCR* exons e2' and e2, results in a fusion transcript named e1a2, which is translated to a 190 kDa protein, (p190BCR-ABL). The third breakpoint cluster region (μ -bcr) downstream of exon 19, results in fusion transcript

e19a2 that gives rise to a 230 kDa protein (p230BCR-ABL) (Figure 1.9) (Deininger et al., 2000; Quintás-Cardama and Cortes, 2009). Out of these fusion proteins, the p210 is typical of CML (95 % of patients) while the p190 and p230 are more commonly associated with acute lymphoblastic leukemia (ALL) and chronic neutrophilic leukemia (CNL) respectively.

Figure 1.9 *ABL* and *BCR* breakpoint locations and respective fusion mRNA isoforms

The breakpoints in *ABL* can occur upstream of exon Ib, downstream of exon Ia, or, more frequently between the two. Nevertheless, the chimeric mRNA transcripts contain exon a2 due to splicing. The breakpoints within *BCR* can occur between i) alternative exons e2' and e2 (m-bcr), resulting in a fusion transcript e1a2 ii) exons e12-e16 (M-bcr) generating either b2a2 (e13a2) or b3a2 (e14a2) fusion transcripts (iii) downstream of exon 19 (μ-bcr) generating fusion transcript e19a2. Image is adapted from Deininger et al, 2000.

1.9.2 BCR-ABL signaling

ABL proteins are nonreceptor tyrosine kinases with two isoforms that can travel between the nucleus and cytoplasm. In the nucleus, ABL is involved in regulation of the cell cycle and in genotoxicity and in the cytoplasm it is involved in signal transduction and

cytoskeletal molding (Kurzrock et al., 2003). The N-terminus of ABL contains three SH domains: SH3, SH2 and SH1. The SH1 domain harbors the tyrosine kinase function, whereas the SH2 domain binds to phosphotyrosine residues and the SH3 domain binds to proline-rich (PxxP) motifs and negatively regulates ABL tyrosine kinase activity. Downstream there are three proline-rich sequences which can interact with SH3-containing proteins such as Crk. The C-terminal part of ABL contains a DNA-binding domain, nuclear localization signals, and a binding site for actin (Deininger et al., 2000; Epstein et al., 1999; Quintás-Cardama and Cortes, 2009). The BCR protein also is found in both, the cytoplasm and nucleus and is involved in two major signaling pathways: (i) phosphorylation and (ii) guanosine triphosphate (GTP) binding. The N-terminus of Bcr contains a coiled-coil oligomerization domain, followed by a SH2 binding domain. Other downstream domains include a serine/threonine kinase domain, a Rho guanine nucleotide exchange factor (Rho-GEF) domain, a pleckstrin homology domain (PH domain), a calcium dependent lipid binding (CaLB) domain and a Rac GTPase activating protein (Rac-GAP) domain (Deininger et al., 2000; Kurzrock et al., 2003; Quintás-Cardama and Cortes, 2009) .

The usually controlled tyrosine kinase activity of the ABL protein is constitutively activated when ABL is fused with BCR regions. This is because (i) BCR inhibits the adjacent ABL SH3 kinase regulatory domain, and (ii) a coiled coil domain in the BCR promotes dimerization and trans-autophosphorylation. The subsequent increase in the phosphorylated tyrosine residues on BCR-ABL provides binding regions for the SH2 domains of other proteins (Cilloni and Saglio, 2012; Epstein et al., 1999). The BCR–ABL oncoprotein activates signal transduction pathways such as RAS/mitogen-activated protein kinase (MAPK), phosphatidylinositol 3-kinase (PI-3K), janus kinase-signal transducers and activators of transcription (JAK-STAT) and Src (Salesse and Verfaillie, 2002). The activation of these pathways promotes cell proliferation, inhibits apoptosis and reduces cellular adhesion (Deininger et al., 2000; Thielen et al., 2011). Importantly, a tyrosine at position 177 in BCR is essential for the binding of adaptor proteins such as growth factor receptor-bound protein (Goldman and Melo, 2003). Growth factor receptor-bound protein 2 (GRB2) binds to son of sevenless (SOS) and Grb2 associated binding protein 2 (GAB2) and the formation of GRB2/GAB2/SOS complex causes activation of

RAS and recruitment of PI3K. RAS in turn activates downstream MAPK which causes uncontrolled cell proliferation. PI3K activates the serine-threonine kinase AKT, which in turn prevents apoptosis by inhibiting forkhead O (FOXO) transcription factors and stimulates cell proliferation by inducing p27 proteasomal degradation and mammalian target of rapamycin (mTOR) activation. Additionally direct phosphorylation of signal transducer and activator of transcription 5 (STAT5) by or indirect phosphorylation by HCK or janus kinase 2 (JAK2) inhibits apoptosis by causing up-regulation of antiapoptotic BCL-xL. In addition, CML cells show reduced adhesion to their bone marrow microenvironment resulting in enhanced proliferation (Cilloni and Saglio, 2012; O'Hare et al., 2011).

1.9.3 Clinical features of CML

CML accounts for 15% of all cases of newly diagnosed leukemia in adults and has an annual incidence of 1–2 cases per 100,000 adults (Jabbour and Kantarjian, 2014). CML is a triphasic disease that begins with a chronic phase and progresses to a terminal blast phase, through an intermediate accelerated phase (Rana et al., 2011). In the chronic phase (CP) of CML which lasts for 3–4 years (Ren, 2005), the massive proliferation of myeloid progenitors causes an increase in granulocytes in peripheral blood (Vigneri and Wang, 2001). Most CML patients are diagnosed in the chronic phase. Symptoms include splenomegaly, anemia, lethargy and weight loss. However, half of the patients do not manifest the symptoms and are diagnosed by a blood test for other reasons (Perrotti et al., 2010). CML is diagnosed by the visualization of Philadelphia chromosome by cytogenetics or *BCR-ABL* fusion oncogene by fluorescent *in situ* hybridization (FISH), or detection of *BCR-ABL* mRNA with quantitative reverse-transcription polymerase chain reaction (qRT-PCR) (Deininger and Druker, 2003; Jabbour and Kantarjian, 2014). The disease inevitably progresses into an accelerated phase and eventually to a fatal blast crisis in several weeks or months if left untreated or due to treatment resistance (Smith et al., 2003; Thielen et al., 2011). In the blast phase, there is rapid overproduction of immature myeloid blast cells (in two thirds of patients) and immature lymphoid blast cells (in one third of patients) which overcome the bone marrow space and lose the ability to differentiate (Kurzrock et al., 2003; Savona and Talpaz, 2008; Thielen et al., 2011). Most

patients acquire additional cytogenetic abnormalities while progressing into blast phase and the median survival is six months (Perrotti et al., 2010). Hence treatment is necessary in the chronic phase.

1.9.4 CML treatment

The standard treatment for CML includes imatinib (also known as “Gleevec” or “Glivec”), a tyrosine kinase inhibitor, which has revolutionized the treatment of CML by converting a fatal cancer into a chronic illness. Approved in 2001, imatinib has markedly improved response rates, overall survival, and patient outcomes in CML patients, compared to prior therapeutic options (An et al., 2010) such as spleen irradiation, hydroxycarbamide and busulfan or interferon-alpha. Before the advent of imatinib, only Interferon-alpha was able to induce cytogenetic responses ($\leq 35\%$ Ph-positive metaphases) in around 20% of patients. Allogenic stem cell transplantation was another potential alternative, but only in few patients due to age limits and donor availability (Thielen et al., 2011). The eight-year IRIS study (International Randomized Study of Interferon and STI571) of newly diagnosed patients with chronic phase CML treated with imatinib reported a complete cytogenetic response rate of 83% and an estimated overall survival (OS) of 93% when only CML-related deaths were considered (Deininger, 2009).

Imatinib competes with ATP for its binding site in the tyrosine kinase domain of BCR-ABL, locking it in a non-ATP-binding closed state (Figure 1.10) (Savage and Antman, 2002). This prevents tyrosine trans-autophosphorylation and activation of downstream pathways that trigger oncogenesis (Melo and Chuah, 2008). Imatinib only induces apoptosis of hematopoietic cells expressing BCR-ABL without affecting normal cells due to its specificity for BCR-ABL. Side effects related to the use of imatinib are relatively well tolerated which include muscle cramps, edema, and diarrhoea (von Bubnoff and Duyster, 2010).

Although imatinib has revolutionized the treatment of CML, primary/intrinsic resistance (inability to achieve hematological or cytogenetic response within a given time period) or secondary resistance (loss of previously achieved hematological or cytogenetic responses) could arise (von Bubnoff and Duyster, 2010). Secondary resistance can be caused due to

BCR-ABL gene amplification or point mutations in the ABL kinase domain. Point mutations can occur in the kinase domain or can change the BCR-ABL conformation to the active form, preventing imatinib binding. More than 100 different BCR-ABL point mutations have been identified of which the T315I mutations occurring in the P-loop are the most common ones (Bhamidipati et al., 2013). Indeed second-generation TKIs such as dasatinib and nilotinib have been approved for patients showing resistance to imatinib due to point mutations. The effects of BCR-ABL gene amplification can be overcome by dose escalation. However the reasons underlying intrinsic resistance were not well defined until very recently and they still demand intensive characterization (Zhang et al., 2009).

Figure 1.10 Schematic showing imatinib mechanism of action

Imatinib binds the ATP binding site of the BCR-ABL tyrosine kinase domain thereby preventing activation of downstream pathways. Image is adapted from Savage et al., 2002.

1.10 EGFR NSCLC

Lung cancer leads to the largest number of cancer deaths in the world and more than 85% of lung cancer cases comprise of NSCLC (Chen et al., 2014). In a subset of NSCLC patients (30% East Asians and 10% North Americans and Western Europeans), activating

EGFR mutations within the kinase domain of EGFR are observed (Bell et al., 2005; Shigematsu et al., 2005). These mutations occur more commonly in females compared to males and in non-smokers compared to smokers (Midha et al., 2015; Rosell et al., 2009).

Most of the activating EGFR mutations comprise small in-frame deletions in exon 19 and exon 21 point mutations (L858R) that occur around the ATP-binding pocket of the EGFR tyrosine kinase domain. (Rosell et al., 2009). These “activating mutations” cause ligand-independent increased tyrosine kinase activation by destabilizing the normally autoinhibited conformation (in absence of ligand) of the domain (Yun et al., 2007). This leads to activation of downstream AKT and STAT pathways, which confer oncogenic properties (Sordella et al., 2004).

The activating EGFR mutations are sensitive to the TKIs gefitinib and erlotinib (Paez et al., 2004) (Lynch et al., 2004) (Pao et al., 2004). These TKIs reversibly bind to the tyrosine kinase domain of EGFR (Antonicelli et al., 2013; Soria et al., 2012). Around 25 %- 30% of EGFR-NSCLC patients show primary or intrinsic resistance to EGFR-TKIs i.e. they fail to respond to TKIs for which the mechanisms are not well characterized (Mitsudomi and Yatabe, 2007). In many patients who are initially sensitive to TKI treatment, at some point develop EGFR-TKI acquired resistance. For example, in more than 50% of patients developing acquired resistance, a gatekeeper second-site mutation (T790M) occurs in *EGFR* (Pao et al., 2005). Other causes of acquired resistance include: amplification of *MET* proto-oncogene amplification, epithelial-to-mesenchymal transition and transformation to small-cell lung cancer (Sequist et al., 2011; Yano et al., 2008). In patients developing acquired resistance to EGFR-TKIs, new generation EGFR-TKIs and EGFR-TKI in combination with MET-TKI have been tested (Pao and Chmielecki, 2010).

1.11 Primary TKI resistance in East Asian CML and EGFR NSCLC patients

Despite the high success rates of targeted therapies in cancers, such as TKIs, there is a high degree of heterogeneity in patient responses such as variations in initial response to treatment, response durations and overall survival (Gazdar, 2009; Sequist et al., 2007; Sharma et al., 2007). The importance of BIM activation in sensitizing these cancers to TKI-based therapies, suggests that the expression levels of BIM is a major contributor to

heterogeneity in patient responses. However the reasons underlying the differences in BIM expression remain poorly understood (Faber et al., 2011). Genomic polymorphisms have been known to contribute to differences in patient responses to cancer therapeutics, but have not been researched in the field of TKI-based cancer therapies (Feero et al., 2011).

In an attempt to identify the basis for sub-optimal response to TKIs, our collaborators at Duke-Nus, Dr. Ong Sin Tiong and his team recently found that a 2.9 kb deletion in *BIM* gene intron 2 was associated to intrinsic TKI resistance, in CML and EGFR NSCLC which are cancers caused by kinase activation. They found a strong correlation between this deletion polymorphism and TKI resistance not only in CML patients, but intrinsic TKI resistance was also observed when the polymorphism was introduced into CML and EGFR NSCLC cell lines that are TKI-sensitive. Interestingly, the deletion polymorphism was found to occur at a frequency of 12.3% in healthy East Asian populations and was completely absent in African and European populations. Analysis of *BIM* splicing from minigene and endogenous transcripts was done using CML and EGFR NSCLC cell lines. The analyses revealed that the 2.9 kb deletion in intron 2 altered the splicing pattern of the *BIM* pre-mRNA such that it favored inclusion of E3 over E4. This gave rise to proapoptotic BH3 domain (encoded by E4) lacking BIM isoforms. Consequently, individuals with the polymorphism had primary resistance to TKI as polymorphism-containing cells were unable to fully induce apoptosis following TKI treatment. Thus Dr. Ong Sin Tiong and his team speculated that the 2.9 kb intronic fragment contained silencer elements that repress *BIM* E3 inclusion (Ng et al., 2012). Indeed the silencer elements have been narrowed down to a 322-nt region of which a 23-nt ISS at the 3' end was found to be imperative (Juan et al., 2014). Subsequent studies by other groups also found a correlation between the *BIM* 2.9 kb deletion polymorphism and sub-optimal responses to TKI in TKI treated cancer patients (Isobe et al., 2014; Ying et al., 2015; Zhao et al., 2014). However two South Korean centers did not show a correlation between the polymorphism and TKI resistance (Lee et al., 2013; Lee et al., 2015), suggesting that genomic variations apart from the 2.9kb polymorphism could also contribute to TKI resistance.

1.12 Rationale of project

The discovery of the *BIM* deletion polymorphism has opened possibilities of identifying the patients at risk of inferior TKI responses, thus personalizing therapy to circumvent TKI resistance in these patients. A promising therapeutic would be the use of splice-switching ASOs that switch splicing from E3 to E4 and restore production of functional BIM. In this project I intended to characterize *cis*-acting elements regulating *BIM* E3 splicing. This included identification of enhancers and silencers in E3 and upstream intronic region after the 2.9 kb fragment with a focus on enhancers as they are potential targets for splice-switching ASOs. Identification of potential *trans*-acting factors promoting *BIM* Exon 3 inclusion was another area of research as these factors could be oncogenic biomarkers. Furthermore, in an attempt to find sequence polymorphisms apart from the 2.9 Kb fragment that could affect TKI resistance, I studied the effect of annotated SNPs on E3 splicing. Lately a growing number of studies have shown SNPs to modulate pre-mRNA splicing. A recent study found a SNP in E4 that led to drug resistance and shorter overall survival in acute lymphoblastic leukemia patients (Gagné et al., 2013). I looked for SNPs in E3 and upstream intronic region that could alter E3 inclusion possibly affecting TKI responses. Identification of SNPs associated with TKI resistance could be beneficial as these SNPs could be used as biomarkers for predicting risk of TKI resistance. Accordingly treatment could be personalized for these patients.

1.13 Objectives

The objectives of my project are to (i) identify *cis*-acting splicing elements in *BIM* E3 and upstream intronic region as mapping of splicing enhancers could reveal potential targets for antisense oligonucleotides (ASOs), (ii) identify potential *trans*-acting factors acting through enhancers, and (iii) study the effects of SNPs on E3 splicing and TKI response in order to identify biomarkers for TKI resistance.

2 MATERIALS AND METHODS

2.1 Cell culture

K562 CML cell line, lacking the *BIM* deletion polymorphism was maintained in Roswell Park Memorial Institute 1640 (RPMI-1640) medium (Hyclone, USA) supplemented with 10% Fetal Bovine Serum (FBS; Gibco, USA), and 100 U/ml Penicillin and 100 µg/ml Streptomycin (Gibco, USA). The cells were incubated at 37°C, and 5% CO₂.

CRISPR/Cas9-edited K562 cells were generated and maintained in RPMI-1640 medium supplemented with 20% FBS and 100 U/ml Penicillin and 100 µg/ml Streptomycin and incubated at 37°C, and 5% CO₂.

2.2 Plasmids

2.2.1 Generation of mutant plasmids for *BIM* Δ10 and Δ11 minigenes

Mutant plasmids were made in the context of both *BIM* Δ10 and Δ11 minigenes, using specific primers and KAPA HiFi DNA polymerase (KAPA Biosystems, USA), in accordance with manufacturer's instructions. The Δ10 and Δ11 minigenes were kindly provided to us by our collaborator, Dr. Ong Sin Tiong. These are chimeric constructs containing adenovirus first exon (U) followed by E3 and flanking intronic sequences, and E4 and upstream intronic sequence. E4 is fused with adenovirus exonic sequence D, which has a PAS. The Δ10 minigene contains only the 322 nt fragment at the 3' end of the 2.9 kb polymorphic region, which is sufficient to suppress E3 splicing while Δ11 has the 322-nt fragment removed (Juan et al., 2014; Ng et al., 2012). Serial deletions and staggered deletions of 10 nt each were generated in *BIM* E3 (220 nt) and upstream 106 nt intronic region (downstream of the 2.9 kb polymorphic region). Specific primers were used which contained the flanking regions of deleted nucleotides, as shown in Table 2.1, Table 2.2, Table 2.3, and Table 2.4 respectively. PolyU tract lengths at 3'ss of *BIM* E3 were varied using specific primers, as stated in Table 2.5. Point mutations in enhancer sequences of *BIM* E3 were introduced using specific primers, as shown in Table 2.6. SNPs in *BIM* E3 and upstream intronic regions were introduced using specific primers, shown in Table 2.7. All the DNA primers were ordered from IDT, USA. Reactions were

performed using the PCR mutagenesis conditions as depicted in Table 2.8. Total PCR products were treated with 20 U DpnI (New England Biolabs, USA) to remove methylated PCR templates followed by transformation of chemically competent *E. coli* DH5 α with 4 μ l PCR products.

Table 2.1 Primers for generation of serial deletions in *BIM E3*

Plasmid	Forward Primer	Reverse Primer
D1	CCAACACTTTTTTTTTTTT ACAGTAAGGATATAGGTGATCT TTCAC	AGTGAAAGATCACCTATATCCT TACTGTTAAAAAAAAAAAAAAG TGTGG
D2	TAACAGTAGTCATCCTAGTGAT CTTTCAC	GCACAGTGAAAGATCACTAGGA TGACTACTGTTAAAAAAAAAAAA AAAGTG
D3	GTAGTCATCCTAGAGGATATAG GCTGTGCTTTGGATTATATTT ACTG	CAGTAAATATAAATCCAAAGCA CAGCCTATATCCTCTAGGATGA CTAC
D4	GGATATAGGTGATCTTTCAGAT TTATATTTACTGGCTTAGATTT GTATGG	CTAAGCCAGTAAATATAAATCT GAAAGATCACCTATATCCTCTA GG
D5	GGTGATCTTTCAC	CCATACAAATCTAAGCCAGTAC AAAGCACAGTGAAAGATCACCT ATATC
D6	CTTTCAC	GGTGGTGGCCATACAAATCAAT ATAAATCCAAAGCACAGTGAAA G
D7	GTGCTTTGGATTATATTTACT GGCTTAGCCACCACCATAGTCA AGATAC	GTATCTTGACTATGGTGGTGGC TAAGCCAGTAAATATAAATCCA AAGCAC
D8	CTGGCTTAGATTTGTATGTAGT CAAGATACAGAACAAC	GGATTTATATTTACTGGCTTAG ATTTGTATGTAGTCAAGATACA GAAC

D9	GATTTGTATGGCCACCACCAAC AGAACAACCTCAACCACAAGG	CCTTGTGGTTGAGTTGTTCTGT TGGTGGTGGCCATACAAATC
D10	GCCACCACCATAGTCAAGATTC AACCACAAGGATTTCTCATG	CATGAGAAATCCTTGTGGTTGA ATCTTGACTATGGTGGTGGC
D11	CCATAGTCAAGATACAGAACAA CGGATTTCTCATGATACCTTTT TATAGC	GCTATAAAAAGGTATCATGAGA AATCCGTTGTTCTGTATCTTGA CTATGG
D12	CAGAACAACTCAACCACAAATGA TACCTTTTTTATAGCCACAG	CTGTGGCTATAAAAAGGTATCA TTGTGGTTGAGTTGTTCTG
D13	CTCAACCACAAGGATTTCTCAT TATAGCCACAGCCACC	GGTGGCTGTGGCTATAATGAGA AATCCTTGTGGTTGAG
D14	CCACAAGGATTTCTCATGATAC CTTTAGCCACCTCTCTCCCTC	GAGGGAGAGAGGTGGCTAAAGG TATCATGAGAAATCCTTGTGG
D15	CTCATGATACCTTTTTTATAGCC ACCTCCCTCTTCCTTGAGC	GCTCAAGGAAGAGGGAGGTGGC TATAAAAAGGTATCATGAG
D16	GCCACAGCCACCTCTCTTGAGC ATTTTGTTCATATGGTCATTG	CAATGACCATATGACAAAATGC TCAAGAGAGGTGGCTGTGGC
D17	CACCTCTCTCCCTCTTCTTGTC ATATGGTCATTGGTG	CACCAATGACCATATGACAAGA AGAGGGAGAGAGGTG
D18	CCTCTTCCTTGAGCATTGTCAT TGGTGATTAAATAAAATGT	ACATTTTATTTAATCACCAATG ACAATGCTCAAGGAAGAGG
D19	GAGCATTTTGTTCATATGATTAA ATAAAATGTATTTTAATATTGA CTTTC	CCTTGAGCATTTTGTTCATATGA TTAAATAAAATGTATTTTAATA TTG
D20	GTCATATGGTCATTGGTGATGT ATTTTAATATTGACTTTCTCTG	CAGAGAAAGTCAATATTAAAT ACATCACCAATGACCATATGAC
D21	CATTGGTGATTAAATAAAATAT TGACTTTCTCTGTTTCTTTCTA CCTTTTTTAACATGG	CCATGTTTAAAAAGGTAGAAAG AAACAGAGAAAGTCAATATTTT ATTTAATCACCAATG
D22	GGTCATTGGTGATTAAATAAAA TGTATTTTATCTCTGTTTCTTT	GTTTAAAAAGGTAGAAAGAAC AGAGATAAAATACATTTTATTT

	CTACCTTTTAAAC	AATCACCAATGACC
D23	GGTGATTAAATAAAATGTATT TAATATTGACTTTTCTACCTT TTTAAACATGGCTAC	GTAGCCATGTTTAAAAGGTAG AAAAGTCAATATTTAAATACA TTTATTTAATCACC

Note: Sequences upstream of the respective deletions are shown in red, while sequences downstream of the deletions are shown in blue. All primer sequences are in 5' to 3' direction.

Table 2.2 Primers for generation of serial staggered deletions in *BIM* E3

Plasmid	Forward Primer	Reverse Primer
SD1	CCAACACTTTTTTTTTTTT ACAGTAGTCATTAGGTGATCTT TCACTGTGC	GCACAGTGAAAGATCACCTATA TGACTACTGTTAAAAAAAAAAAA AAAGTGTGG
SD2	GTAGTCATCCTAGAGGATTTTC ACTGTGCTTTGG	CCAAAGCACAGTGAAAATCCTC TAGGATGACTAC
SD3	CATCCTAGAGGATATAGGTGAT CCTTTGGATTATATTTACTGG CTTAG	CTAAGCCAGTAAATATAAATCC AAAGGATCACCTATATCCTCTA GGATG
SD4	GGATATAGGTGATCTTTCCTG TGATATTTACTGGCTTAGATT GTATGG	CCATACAAATCTAAGCCAGTAA ATATCACAGTGAAAGATCACCT ATATCC
SD5	CTTTCCTGTGCTTTGGATTG CTTAGATTGTATGGCCAC	GTGGCCATACAAATCTAAGCAA ATCCAAAGCACAGTGAAAG
SD6	CTGTGCTTTGGATTATATTTA CTGGTATGGCCACCACCATAG	CTATGGTGGTGGCCATACCAGT AAATATAAATCCAAAGCACAG
SD7	GGATTTATATTTACTGGCTTAG ATTTCAACCATAGTCAAGATACA GAAC	GTTCTGTATCTTGACTATGGTG AAATCTAAGCCAGTAAATATAA ATCC
SD8	GCTTAGATTTGTATGGCCACAA GATACAGAACAACCTCAACC	GGTTGAGTTGTTCTGTATCTTG TGGCCATACAAATCTAAGC
SD9	GCCACCACCATAGTCACACTC AACCACAAGG	CCTTGTGGTTGAGTTGTGACTA TGGTGGTGGC
SD10	CCACCATAGTCAAGATACAGAC ACAAGGATTTCTCATGATACC	GGTATCATGAGAAATCCTTGTG TCTGTATCTTGACTATGGTGG
SD11	GTCAAGATACAGAACAACCTCAA CTCTCATGATACCTTTTATAG CCAC	GTGGCTATAAAAAGGTATCATG AGAGTTGAGTTGTTCTGTATCT TGAC
SD12	CAACTCAACCACAAGGATTCCT TTTTATAGCCACAGCC	GGCTGTGGCTATAAAAAGGAAT CCTTGTGGTTGAGTTG
SD13	CAACTCAACCACAAGGATTTCT CATGATAGCCACAGCCACCTCT C	GAGAGGTGGCTGTGGCTATCAT GAGAAATCCTTGTGGTTGAGTT G
SD14	CACAAGGATTTCTCATGATACC TTTTTATACCTCTCTCCCTCTT CCTTG	CAAGGAAGAGGGAGAGAGGTAT AAAAAGGTATCATGAGAAATCC TTGTG
SD15	GATACCTTTTATAGCCACAGC CATCTTCCTTGAGCATTGTC ATATGG	CCATATGACAAAATGCTCAAGG AAGATGGCTGTGGCTATAAAAA GGTATC
SD16	CAGCCACCTCTCTCCCGCATTT TGTCATATGGTCATTGGTG	CACCAATGACCATATGACAAAA TGC GGGAGAGAGGTGGCTG
SD17	CTCTCTCCCTCTTCCTTGAATA TGGTCATTGGTGATTAAATAAA ATGTATTTTAATATTG	AATATTAAAATACATTTTATTT AATCACCAATGACCATATCAA GGAAGAGGGAGAGAGG

SD18	CTTCCTTGAGCATT TTTGTC TGG TGATTAAATAAAATGTATTTTA ATATTG	CAATATTAAAATACATTTTATT TAATCACCA GACAAAATGCTCA AGGAAG
SD19	GAGCATTTTGT CATATGGTCAT ATAAAATGTATTTTAATATTGA CTTTCTCTG	CAGAGAAAGTCAATATTAAAAT ACATTTTAT ATGACCATATGAC AAAATGCTC
SD20	GTCATATGGTCATTGGTGATTA ATTTTAATATTGACTTTCTCTG TTTCTTTC	GAAAGAAACAGAGAAAGTCAAT ATTAAAATTAATCACCAATGAC CATATGAC
SD21	GGTCATTGGTGATTAAATAAAA TGTAGACTTTCTCTGTTTCTTT CTACC	GGTAGAAAGAAACAGAGAAAGT CTACATTTTATTTAATCACCAA TGACC
SD22	CATTGGTGATTAAATAAAATGT ATTTTAATATTGTTTCTTTCTA CCTTTTAAACATG	CATGTTTAAAAAGGTAGAAAGA AACAAATATTAAAATACATTTTA TTTAATCACCAATG
SD23	GATTAAATAAAATGTATTTTAA TATTGACTTTCTCTACCTTTT AAACATGGCTACTAG	CTAGTAGCCATGTTTAAAAAGG TAGAGAAAGTCAATATTAAAAT ACATTTTATTTAATC

Note: Sequences upstream of the staggered deletions are shown in red, while sequences downstream of the staggered deletions are shown in blue. All primer sequences are in 5' to 3' direction.

Table 2.3 Primers for generation of serial deletions in *BIM* Intron 2

Plasmid	Forward Primer	Reverse Primer
-D1	GGTCCAGTTTCTCCACATACT TTTTTTTTTTTAAACAGTAGT CATCCTAGAGG	CCTCTAGGATGACTACTGTAA AAAAAAAAAAAGTATGTGGAG AAACTGGAACC
-D2	GCACAAGGGTTCCAGTTTCTAC CAACACTTTTTTTTTTTTAA CAGTAGTC	GACTACTGTAAAAA AAGTGTTGGTAGAACTGGAAC CCTTGTGC
-D3	CAACAGGGCACAAGGGTCCACA TACTTACCAACACTTTTTTTT TTTTAAC	GTAAAAAAGTGTT GGTAAGTATGTGGACCCTGTG CCCTGTTG
-D4	CATTTTACATTCCCAACAG GTCCAGTTTCTCCACATACTTA CC	GGTAAGTATGTGGAGAACTGG ACCTGTTGGTGGGAATGTAAAA TG
-D5	GGCTGTGCCATTTTACATTCC GCACAAGGGTTCCAGTTTC	GAAACTGGAACCCTTGTGC GGAATGTAAAATGGCACAGCC
-D6	CCATAGAGGCTGTGCCAT CACCAACAGGGCACAAG	CTTGTGCCCTGTTGGTG ATGGCACAGCCTCTATGG
-D7 (Δ10)	CCATCATGCTGTTCTCCATAGA GTTTACATTCCCACCAACAGGG	CCATCATGCTGTTCTCCATAGA GTTTACATTCCCACCAACAGGG
-D7 (Δ11)	CGCGATATCCTGTTCTCCATAG AGTTTACATTCCCACCAACAGG G	CGCGATATCCTGTTCTCCATAG AGTTTACATTCCCACCAACAGG G
-D8 (Δ10)	GGGGAACCATCATGCTGTT GCTGTGCCATTTTACATTCCC	GGGGAACCATCATGCTGTT GCTGTGCCATTTTACATTCCC
-D8 (Δ11)	GGCCGCGATATCCTGTT GCTGTGCCATTTTACATTCCC	GGCCGCGATATCCTGTT GCTGTGCCATTTTACATTCCC

Note: Sequences upstream of the respective deletions are shown in red, while sequences downstream of the deletions are shown in blue. All primer sequences are in 5' to 3' direction.

Table 2.4 Primers for generation of serial staggered deletions in *BIM* Intron 2

Plasmid	Forward Primer	Reverse Primer
-SD1	GGGTTCCAGTTTCTCCACACAC TTTTTTTTTTTTTTAACAGTAG TCATCC	GGATGACTACTGTTAAAAAAA AAAAAAGTGTGTGGAGAACTG GAACCC
-SD2	GGCACAAGGGTTCAGTACTTA CCAACACTTTTTTTTTTTTTTA ACAG	CTGTTAAAAAAGTGTG TTGGTAAGTACTGGAACCTTG TGCC
-SD3	CCACCAACAGGGCACATTTCTC CACATACTTACCAACAC	GTGTTGGTAAGTATGTGGAGAA ATGTGCCCTGTTGGTGG
-SD4	GCCATTTTACATTCCCACCA AGGGTTCAGTTTCTCCAC	GTGGAGAACTGGAACCT TGGTGGGAATTTAAAATGGC
-SD5	CATAGAGGCTGTGCCATTTTAC ACAGGGCACAAGGGTTC	GAACCTTGTGCCCTGT GTAAAATGGCACAGCCTCTATG
-SD6	CTGTTCTCCATAGAGGCTGT ATTCCCACCAACAGGGC	GCCCTGTTGGTGGGAAT ACAGCCTCTATGGAGAACAG
-SD7 (Δ10)	GGAACCATCATGCTGTTCTCCA GCCATTTTACATTCCCACCAAC AGG	CCTGTTGGTGGGAATGTAAAAT GGCTGGAGAACAGCATGATGGT TCC
-SD7 (Δ11)	CCGCGATATCCTGTTCTCCA GCCATTTTACATTCCCACCAAC AG	CTGTTGGTGGGAATGTAAAATG GCTGGAGAACAGGATATCGCGG
-SD8 (Δ10)	TTTTTGGGGAACCATCATGTAG AGCTGGCCATTTTACATTCC	GTAAAATGGCACAGCCTCTACA TGATGTTCCCCAAAAAATTAAA AATAGAATTACC
-SD8 (Δ11)	CGCGGCCGCGATATCTAGAGGC TGTGCCATTTTACATTCC	GGAATGTAAAATGGCACAGCCT CTAGATATCGCGGCCGCG

Note: Sequences upstream of the respective deletions are shown in red, while sequences downstream of the deletions are shown in blue. All primer sequences are in 5' to 3' direction.

The identification of intronic *cis*-acting elements regulating *BIM* E3 splicing was performed by the CN Yang Scholars program, Year 1 undergraduate student Yue Wan Lin.

Table 2.5 Primers for varying the PolyU tract length at 3'ss of *BIM* E3

Plasmid	Forward primer	Reverse primer
6 T	CCACATACTTACCAACAC TTT TT AACAGTAGTCATCCTAGAGG ATATAGGTGATC	CCTCTAGGATGACTACTGTT AA AAA AGTGTGGTAAGTATGTGG AGAAACTGG
8 T	CCACATACTTACCAACAC TTTT TTTT AACAGTAGTCATCCTAGA GGATATAGGTGATC	CCTCTAGGATGACTACTGTT AA AAAAA AGTGTGGTAAGTATGT GGAGAACTGG
10 T	CCACATACTTACCAACAC TTTTT TTTTTT AACAGTAGTCATCCTA GAGGATATAGGTGATC	CCTCTAGGATGACTACTGTT AA AAAAAAA AGTGTGGTAAGTAT GTGGAGAACTGG
12 T	CATACTTACCAACAC TTTTTTTT TTTTTTT AACAGTAGTCATCCTAG AGGATATAGGTGATC	CCTCTAGGATGACTACTGTT AA AAAAAAAAA AGTGTGGTAAGT ATGTGGAGAACTGG
16 T	CATACTTACCAACAC TTTTTTTTT TTTTTTTTT AACAGTAGTCATC CTAGAGGATATAGG	AAAAAAAAAAAAA AGTGTGGT AAGTATGTGGAGAACTGGAAC C
18 T	CTTACCAACAC TTTTTTTTTTTT TTTTTTTTT AACAGTAGTCATCCT AGAGGATATAGG	AAAAAAAAAAAAA AGTGTGGT AAGTATGTGGAGAACTGGAAC C
20 T	CTTACCAACAC TTTTTTTTTTTTT TTTTTTTTTTT AACAGTAGTCATC CTAGAGGATATAGG	AAAAAAAAAAAAA AGTGTGGT AAGTATGTGGAGAACTGGAAC C

Note: The poly T tract in the forward primer and its complementary polyA tract in the reverse primer are shown in red. All primer sequences are in 5' to 3' direction.

Table 2.6 Primers for introducing point mutations in *BIM* E3 enhancer sequences

Plasmid	Forward primer	Reverse primer
<u>Enhancer 1 region (D1+ SD1)</u>		
Mut 1	CAGTAGT AAT C A TAG C GGATAT AGGTGATC	CATACTTACCAACACTTTTTTT TTTTTTTAAACAGTA
Mut 2	CAGTAGTCAG C A T AG C GGATAT AGGTGATC	GATCACCTATATCC G CTA TGCT GACTACTG
Mut 3	CAGTAGTCAT AAT AG C GGATAT AGGTGATC	GATCACCTATATCC G CTA TTAT GACTACTG
<u>Enhancer 2 region (SD9+SD10)</u>		
Mut 1	GGCCACCACCATAGTCAAC C ATA A TTAACC A TTCAACCACAAGGA TTTCTCATGATACC	GGTATCATGAGAAATCCTTGTG GTTGA A T G TT AAT TAT G TTGA CTATGGTGGTGGCC
Mut 2	GGCCACCACCATAGTCAAGATA A TTAA T C A TTCAACCACAAGGA TTTCTCATGATACC	GGTATCATGAGAAATCCTTGTG GTTGA A T G ATT AAT TATCTTGA CTATGGTGGTGGCC
Mut 3	GGCCACCACCATAGTCAAGATA T CTAA T AACT C CACCACAAGGA TTTCTCATGATACC	GGTATCATGAGAAATCCTTGTG GT G GAGTT A TT AG A T ATCTTGA CTATGGTGGTGGCC
Mut 4	GGCCACCACCATAGT C ATA C T CA T AA T AACT C CACCACAAGGA TTTCTCATGATACC	GGTATCATGAGAAATCCTTGTG GT G GAGTT A TT AT G AG T AT G GA CTATGGTGGTGGCC
<u>Enhancer 3 region (SD11+D12)</u>		
Mut 1	GATACAGAACAAC T CAACCA T A AG C ATTT C CA T GATACCTTTT TATAGCC	GGCTATAAAAAGGTATCAT G G AAAT G CTT A TGGTTGAGTTGTT CTGTATC
Mut 2	GATACAGAACAAC T CAACCA T A A CCATTTCTCATGATACCTTTT TATAGCC	GGCTATAAAAAGGTATCATGAG AAAT G GTT A TGGTTGAGTTGTT CTGTATC
Mut 3	GATACAGAACAAC T CAACCA T AG C ATTTCTCATGATACCTTTT TATAGCC	GGCTATAAAAAGGTATCATGAG AAAT G CT A GTGGTTGAGTTGTT CTGTATC
Mut 4	GATACAGAACAAC T CAACCA T A T C ATTTCTCATGATACCTTTT TATAGCC	GGCTATAAAAAGGTATCATGAG AAAT G AT A GTGGTTGAGTTGTT CTGTATC

Note: The mutations in the forward and reverse primer are shown in red. All primer sequences are in 5' to 3' direction.

Table 2.7 Primers for introducing SNPs in *BIM* E3 and upstream intronic region

Plasmid	Forward primer	Reverse primer
SNP 1	GCTGTGCCATTTTACATTCCCA CCA G CAGGGCACAAGGGTTCCA G	CTGGAACCCTTGTGCCCTG C TG GTGGGAATGTAAATGGCACAG C
SNP2	CATTCCCACCAACAGGGCACA G GGGTTCCAGTTTCTCCACATAC TTACC	GGTAAGTATGTGGAGAACTGG AACCC C TGTGCCCTGTTGGTGG GAATG
SNP3	GGATATAGGTGATCTTTCCTG TGCTT C GGATTTATATTTACTG GCTTAGATTTGTATGG	CCATACAAATCTAAGCCAGTAA ATATAAATCC G AAGCACAGTGA AAGATCACCTATATCC
SNP 4	GGTGATCTTTCCTGTGCTTTG A ATTTATATTTACTGGCTTAGA TTTGTATGGC	CCATACAAATCTAAGCCAGTAA ATATAAAT T CAAAGCACAGTGA AAGATCACC
SNP 5	CACTGTGCTTTGGATTTATATT TACTGG T TTAGATTTGTATGGC CACCACC	GGTGGTGGCCATACAAATCTAA A CCAGTAAATATAAATCCAAAG CACAGTG
SNP 6	GCTTTGGATTTATATTTACTGG CTTAGAT G TGTATGGCCACCAC CATAG	CTATGGTGGTGGCCATAC A T CTAAGCCAGTAAATATAAATCC AAAGC
SNP 7	CCACAAGGATTTCTCATGATA C TTTTTTATAGCCACAGCCACC	GGTGGCTGTGGCTATAAAAA A GTATCATGAGAAATCCTTGTGG
SNP 3+4+5+6	ATTTATATTTACTGG T TTAGAT G TGTATGGCCACCACCATAG	CCAGTAAATATAAAT T C GAAGC ACAGTGAAAGATCACCTATATC C
SNP 3+4	ATTTATATTTACTGGCTTAGAT TTGTATGGCCACCACCATAG	CCAGTAAATATAAAT T C GAAGC ACAGTGAAAGATCACCTATATC C
SNP 5+6	ATTTATATTTACTGG T TTAGAT GT G TATGGCCACCACCATAG	CCAGTAAATATAAATCCAAAGC ACAGTGAAAGATCACCTATATC C
SNP 3+5	ATTTATATTTACTGG T TTAGAT TTGTATGGCCACCACCATAG	CCAGTAAATATAAATCC G AAGC ACAGTGAAAGATCACCTATATC C
SNP 3+6	ATTTATATTTACTGGCTTAGAT G TGTATGGCCACCACCATAG	CCAGTAAATATAAATCC G AAGC ACAGTGAAAGATCACCTATATC C
SNP 4+5	ATTTATATTTACTGG T TTAGAT TTGTATGGCCACCACCATAG	CCAGTAAATATAAAT T CAAAGC ACAGTGAAAGATCACCTATATC C
SNP 4+6	ATTTATATTTACTGGCTTAGAT G TGTATGGCCACCACCATAG	CCAGTAAATATAAAT T CAAAGC ACAGTGAAAGATCACCTATATC C
SNP	CACTGTGCTTTGGATTTATATT	CCAGTAAATATAAAT T C GAAGC

3+4+5	TACTGGTTTAGATTTGTATGGC CACCACC	ACAGTGAAAGATCACCTATATC C
SNP 3+4+6	GCTTTGGATTTATATTTACTGG CTTAGATGTGTATGGCCACCAC CATAG	CCAGTAAATATAAATTCGAAGC ACAGTGAAAGATCACCTATATC C
SNP 3+5+6	ATTTATATTTACTGGTTTAGAT GTGTATGGCCACCACCATAG	CCATACAAATCTAAGCCAGTAA ATATAAATCCGAAGCACAGTGA AAGATCACCTATATCC
SNP 4+5+6	ATTTATATTTACTGGTTTAGAT GTGTATGGCCACCACCATAG	CCATACAAATCTAAGCCAGTAA ATATAAATCAAAGCACAGTGA AAGATCACC

Note: The mutations in the forward and reverse primer have been shown in red. All primer sequences are in 5' to 3' direction.

Table 2.8 Conditions for PCR mutagenesis

Cycling step	Temperature (°C)	Time	
Initial denaturation	95	2 min	
Denaturation	98	30 sec	} 18 cycles
Annealing	50	1 min	
Extension	72	5 min	
Final extension	72	5 min	
Cooling	4	∞	

2.2.2 pSXN heterologous plasmid

pSXN plasmid (kindly provided by Prof. Thomas A. Cooper from Baylor College of Medicine, USA) contains a small weakly spliced alternative exon placed between β -globin exon 1 and exon 2 as well as a downstream exon 18 from *cTNT* (Figure 2.1) (Coulter et al., 1997). To test the *BIM* E3 ESE sequences in context of the pSXN minigene, equimolar amounts of a pair of complementary DNA oligonucleotides containing designed test sequences and restriction overhangs (Table 2.9) were first annealed by heating in at 95°C for 5 minutes followed by gradual cooling to room temperature for 1 h. The annealed oligonucleotides were then 5'-end phosphorylated by

T4 polynucleotide kinase (T4PNK; New England Biolabs, USA) in 1x ligation buffer [50 mM Tris-HCl, 10 mM MgCl₂, 1 mM ATP and 10 mM DTT] and ligated to the BamHI and SalI sites in the alternative exon within pSXN13 vector.

Figure 2.1 pSXN plasmid

The pSXN plasmid contains a weak exon located between β globin exons 1 and 2. Tested sequences (Table 2.9) were cloned into the weak exon using the SalI and BamHI sites and their effects on exon inclusion were examined by RT-PCR using the depicted primers. Primer sequences have been shown in Table 2.15.

Table 2.9 Oligonucleotide inserts for pSXN plasmid

BIM.Ex3.ESE1.WT.F	TCGACTAGTCATCCTAGAGGATG
BIM.Ex3.ESE1.WT.R	GATCCATCCTCTAGGATGACTAG
BIM.Ex3.ESE1.Mut 3.F	TCGACTAGTCATAA TAGCGGATG
BIM.Ex3.ESE1.Mut 3.R	GATCCATCCGCTATTATGACTAG
BIM.Ex3.ESE2.WT.F	TCGACTCAAGATACAGAACAACCTCAACCAG
BIM.Ex3.ESE2.WT.R	GATCCTGGTTGAATGATTAATTATCTTGAG
BIM.Ex3.ESE2.Mut 1.F	TCGACTCAAGAUAA TTAATCATTCAACCAG
BIM.Ex3.ESE2.Mut 1.R	GATCCTGGTTGAATGATTAA TTATCTTGAG
BIM.Ex3.ESE3.WT.F	TCGACACCACAAGGATTTCTCATGG
BIM.Ex3.ESE3.WT.R	GATCCCATGAGAAATCCTTGTTGGTG
BIM.Ex3.ESE3.Mut 1.F	TCGACACCAUAAGCAUUUC CATGG
BIM.Ex3.ESE3.Mut 1.R	GATCCCATGGGAAATGCTTATGGTG
pSXN.Linker.F	TCGACATTTCGGCTAG
pSXN.Linker.R	GATCCTAGCCGAATG

Insert sequence is in black font, SalI overhang is in blue font and BamHI overhang is in green font. Mutation is labeled in red color. Note: All oligonucleotide inserts are in 5' to 3' direction.

2.2.3 Overexpression plasmids

Overexpression plasmid, pCGT7-SRSF1 (SRSF1 cDNA plasmid) was a gift from Javier F Cáceres from MRC Human Genetics Unit at Edinburgh, UK. SRSF3 overexpression plasmid in pcDNA3.1 vector was purchased from GenScript, USA. 1.2 µg overexpression plasmids were transfected with or without *BIM* minigene plasmids. The transfected cells were incubated for 48 h before RNA and/or protein extraction unless otherwise stated.

2.3 Bacterial transformation and plasmid extraction

All bacterial transformations were carried out as follows: 50 µl of competent *E. coli* DH5α was transformed with 4 µl PCR or ligation products and incubated on ice for 30

min followed by a heat shock at 42 °C for 45 s. Cells were then incubated on ice for another 5 min and then recovered in 946 µl of pure lysogeny broth (LB) at 37° C for 1 h with agitation. Cells were spun down at 10,000 g for 1 min and 900 µl of the supernatant was discarded. The transformed cells were re-suspended in the remaining supernatant by gentle pipetting, before being plated on ampicillin-supplemented LB agar plate (70 µg/ml). Plates were incubated overnight at 37°C. Colonies were grown in 3 ml of LB overnight at 37°C. The plasmids were then extracted using MiniPrep kit (Omega Bio-tek, USA). Purified plasmids were sequenced (1st Base, Singapore) to confirm the successful introduction of designed mutations. For most of the ligation products, plasmids were digested with the respective restriction enzymes to confirm the insertion of specific fragments before sequencing.

2.4 Transfection with plasmids

Cells were seeded in 12-well plates at a density of 200,000 cells/ml 24 h before transfection to achieve a confluency of 60-70%. 1.2 µg plasmid DNA was mixed with 100 µl OptiMEM (Life Technologies, USA) and 1 µl Roche Xtreme GENE HP transfection reagent (Roche Applied Science, Switzerland). The transfection mixture was incubated for 30 min before adding to cells. The transfected cells were incubated at 37°C with 5% CO₂ for 48 hs before RNA extraction.

2.5 Knockdown studies

Depletion of SRSF1 was performed using Silencer® Select SRSF1 small interfering RNA (siRNA) s12725 (Ambion, United States). Depletion of SRSF3 was performed using Dicer substrate interfering RNA (dsiRNA) which were chemically synthesized by (IDT, USA). DsiRNAs are 27-mer duplex RNAs that have increased potency of up to 100-fold in RNA interference compared to traditional 21mers siRNAs. Additionally, dsiRNAs are able to effectively target some sites that 21mers siRNAs are not able to silence (Kim et al., 2005). The siRNA and dsiRNA sequences have been shown in (Table 2.10).

Knockdowns were performed using siRNA or dsiRNA at 50µM concentration in 1×10^6 K562. Negative controls for knockdown studies were carried out with Silencer® Select

Negative Control No. 1 siRNA (Ambion, USA). Nucleofection was performed according to manufacturer's instruction by using the FF-120 program with Amaxa 4D-Nucleofector (Lonza, Switzerland). The nucleofected cells were then incubated for 48 hs before RNA and/or protein extraction unless otherwise stated.

Table 2.10 List of siRNA and dsRNA sequences

Target gene	Sense	Antisense
SRSF1	GACCUAUGCAGUUCGAAAAtt	UUUUCGAACUGCAUAGGUCat
SRSF3	AGAGCUAGAUGGAAGAACACUAUgt	ACAUAGUGUUCUCCAUCUAGCUCU

*Lowercase indicates DNA nucleotides.

2.6 Rescue studies

Rescue studies were carried out for SRSF1 and SRSF3. Prior to rescue studies, the overexpression plasmids were mutated at the siRNA or dsRNA binding sites to make them resistant to the effects of knockdown studies. Mutations were introduced at every third nucleotide of each amino acid codon within the siRNA or dsRNA binding site. Mutations were generated by a PCR site-directed mutagenesis using complementary mutant primers. SRSF1 siRNA resistant overexpression plasmids were previously obtained by Dr. Tang Sze Jing while SRSF3 dsRNA resistant overexpression plasmids were obtained using the following primers (forward primer- 5'-ACAGCTGATGCAGTCCGCGAACTTGACGGGAGGACTCTATGTGGCTGCCGTGT AAG-3' and reverse primer - 5'-CTTACACGGCAGCCACATAGAGTCCTCCCGTCAAGTTCGCGGACTGCATCAGC TG -3'. 24 h after knockdown, K562 cells were transfected with its corresponding siRNA or dsRNA resistant overexpression plasmids. Cells were incubated for either 24 h (for SRSF1) or 48 h (for SRSF3) before RNA and/or protein extraction.

2.7 RNA Extraction and DNase I treatment

The transfected cells were lysed 48 h or 24 h (when transfected with SRSF1 rescue plasmid) post transfection/ nucleofection, and total cellular RNA was extracted using Purelink RNA mini kit (Ambion, USA). RNA concentration was measured using Nanodrop 2000 (Thermo Scientific, USA). 4 µg RNA was treated with 2U RQ1-RNase free DNase I (Promega, USA) at 37°C for 1 h to eliminate DNA contamination. RNA was then subjected to ethanol precipitation by first topping up the reaction to 100 µl with RNase-free water, then adding 0.1x volume of 3 M sodium acetate (pH 5.2) and 3x volume of 100% ethanol, and finally incubating at -80 °C overnight. RNA was centrifuged at 13500 rpm for 45 min and then washed with 70% ethanol once before resuspending the pellets in 10µl RNase-free water.

2.8 Reverse Transcription and gel electrophoresis

1µg of DNase I-treated RNA was used as template for complementary DNA (cDNA) synthesis using Moloney Murine Leukemia Virus (M-MuLV) reverse transcriptase (New England Biolabs, USA) and oligo-dT, at 42°C for 1 h. cDNA was then amplified with specific primers, as depicted in Table 2.11 using PCR conditions as shown in Table 2.12. PCR products were resolved to determine the splicing patterns of *BIM* E3 versus E4 in full length, deletion, Δ10 and Δ11 minigenes by 1.5% (w/v) agarose gel electrophoresis and viewed under UV light.

Table 2.11 Primers for detecting U-E3 and U-E4 transcripts (RT-PCR)

Transcript	Forward primer	Reverse primer
U-E3 	CGAGCTCACTCTCTTCCG	GTCAATATTAAAATACATTTTATT TAATCACCAATGACC
U-E4 	CGAGCTCACTCTCTTCCG	CCTCCTTGCATAGTAAGCGT

Table 2.12 Conditions for PCR

Cycling step	Temperature (°C)	Time	
Initial denaturation	95	2 min	
Denaturation	95	30 sec	} 35 cycles
Annealing	54	40 sec	
Extension	72	50 sec	
Final extension	72	5 min	
Cooling	4	∞	

2.9 Real time PCR

Quantification of *trans*-acting factors (*SRSF1* and *SRSF3*), endogenous *BIM* and *BIM* minigenes transcripts was done using Applied Biosystems 7500/ CFX96 Touch™ (Bio-Rad, USA) Real-Time PCR detection system and SYBR Select Mastermix (Applied Biosystems). PCR was set up in 96-well microplates containing 10 µL of Master Mix, 1 µL of cDNA, and 4 pmol each of forward and reverse primer in a 20 µL reaction volume. The conditions for PCR are indicated in Table 2.13.

Transcript levels of β -actin was used to normalize *SRSF1*, *SRSF3* and *BIM* endogenous transcripts. Transcript levels of adenovirus exonic sequence (U) were used to normalize *BIM* minigene transcripts. Primer sequences used to amplify β -actin, *SRSF1*, *SRSF3*, endogenous *BIM* transcripts and *BIM* minigenes transcripts are shown in Table 2.14. For amplifying transcripts from minigenes, we used a common forward primer that matched the sequence in U (adenovirus exon); the reverse primer for detecting U-E3 transcripts matched the junction between U and E3 while the reverse primer for detecting U-E4 transcripts matched the junction between U and E4 (Table 2.14). Ct values were used to calculate relative mRNA expression by the $2^{-\Delta\Delta C_t}$ relative quantification method (Livak and Schmittgen, 2001). Three technical replicates for each of three biological replicates was done.

Table 2.13 Conditions for Real time PCR

	Temperature (°C)	Time	
Initial denaturation	95	3 min	
Denaturation	95	10 sec	} 40 cycles
Annealing	58	30 sec	
Extension	72	20 sec	

Table 2.14 Primers for detecting different transcripts (Real time PCR)

Transcript	Forward primer	Reverse primer
Primers for detecting β-actin transcripts		
β -actin	CCAGAGGCGTACAGGGATAG	CCAACCGCGAGAAGATGA
Primers for determining the expression level of <i>trans</i>-acting factors		
SRSF1	TTCTACAAATACGGCGCTATCC	GTACCCATCGTAATCATAGCCG
SRSF3	AGCTGATGCAGTCCGAGAG	GGTGGGCCACGATTTCTAC
Primers for detection of <i>BIM</i> minigene transcripts		
U-E3 	CGAGCTCACTCTCTTCCGC	CTCTAGGATGACTACTGGTAGG GT
U-E4 	CGAGCTCACTCTCTTCCGC	CCTCATGGAAGCTGGTAGGGT
Primers for detection of <i>BIM</i> endogenous transcripts		
Exon 2A	TTCCCCCAAATGTCTGACTC	CTTGTGGCTCTGTCTGTAGGG
E3	CCAGGCCTTCAACCACTATC	ATGGTGGTGGCCATACAAAT
E4	TTCCATGAGGCAGGCTGAAC	CCTCCTTGCATAGTAAGCGTT

2.10 Semi-quantitative PCR

10 pmol forward primer was 5' end labeled using 10U T4 PNK (New England Biolabs, USA) and 10.2 pmol γ -³²P-ATP (60 μ Ci; Perkin Elmer, USA). Labeled primers were then purified by illustra Microspin G-25 column (GE Healthcare Life Sciences, USA) and added into primer mixture containing 100 pmol unlabeled reverse primer and 90 pmol unlabeled forward primer. cDNA was amplified by PCR using Go-Taq polymerase (Promega, USA) with the radiolabeled primer mixture for 22 cycles. The primer pairs used for amplification of cDNA have been shown in Table 2.15. The PCR products were separated by 8% (w/v) native polyacrylamide gel (Acrylamide/Bis 29:1; Biorad, USA) in 1x TBE buffer. Radioactive signals from the gel were subsequently exposed to storage phosphor imaging screen which were read by Typhoon Trio variable mode imager (GE Healthcare Life Sciences, USA). The analysis of band intensity was done with ImageQuant TL software (GE healthcare life sciences, USA). For image formation, X-ray films (Kodak, USA) were exposed with the gel for 12-48 h in X-ray cassette at -80 ° C and then processed by Kodak Model 2000 X-12 ray film processor.

Splicing patterns of pSXN minigenes were analyzed by measuring the percentage of exon inclusion band's intensity. The average exon inclusion percentage was derived from at least three experimental replicas.

Table 2.15 List of radioactive primers

pSXN vector	
pSXN_βglobinEx1_F	AGGTGAACGTGGATGAAGTTGGTGGTG
pSXN_βglobinEx2/cTNTEx18.R	CGTGCAGCCTTTGACCTACTAGTGTG

2.11 Protein extraction and western blotting

2.11.1 Sample preparation

K562 cells were washed with 1x phosphate-buffered saline (PBS) once and lysed in NP-40 lysis buffer (50mM Tris-HCl pH8, 150mM NaCl, 1% Triton X-100, 10% glycerol, 1mM EDTA) with protease inhibitors cocktail (Roche, Switzerland). Lysates were incubated on ice for 30 minutes followed by 30 minutes centrifugation at maximum speed, and the resulting supernatants were collected. Protein concentration was measured using the Bradford assay (Bio-Rad, USA). 20 µg proteins from each sample were boiled for 10 min with 5XSDS loading buffer before loading into sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE).

2.11.2 SDS-PAGE and transfer

20µg of proteins from each sample were separated by 10-12.5% polyacrylamide gel (Acrylamide/Bis 37.5:1; Biorad, USA) in 1x SDS running buffer [25 mM Tris, 192 mM glycine, 0.1% SDS]. Proteins were transferred to PVDF membrane (Biorad, USA) in transfer buffer [25 mM Tris, 192 mM glycine, 20% 4 methanol] at 100 V for 1 h.

2.11.3 Detection

Membrane was blocked in 5% milk in 1x TBST [144 mM NaCl, 20 mM Tris, 0.05% Tween 20, pH7.6] for 1 h and washed with 1x TBST thrice, each for 5 min. Next, the membrane was incubated with the primary antibody diluted in 1x TBST at 4°C overnight followed by three washes with 1x TBST and incubation with the secondary anti-mouse/rabbit/goat IgG antibodies (BioLegend, USA/Sigma-Aldrich, USA/Santa cruz biotechnology, USA) diluted in 1x TBST with or without 5% milk. After three washes, proteins were detected by Western Lightning Plus ECL (Perkin Elmer, USA).

Subsequently, the membrane was blocked for one h in 5% (w/v) non-fat dry milk (BioRad, USA) in 1x TBST [144 mM NaCl, 20 mM Tris, 0.05% Tween 20, pH7.6] for 1 h and washed with 1x TBST thrice, each for 5 min. Next, the membrane was incubated with the primary antibody diluted in 1x TBST at 4°C overnight followed by three washes with 1x

TBST and incubation with the secondary anti-mouse IgG antibodies (BioLegend, USA) diluted in 1x TBST (Table 2.16). After three washes in TBST, the protein bands on the membrane were visualized by Western Lightning Plus enhanced chemiluminescence substrate (Perkin Elmer, USA) and exposed to Xray films (Carestream Health, USA).

Table 2.16 List of antibodies

Primary Antibody		Secondary Antibody	
anti-SRSF1	1:1000	Anti-mouse IgG	1:10000
anti-SRSF3	1:1000	Anti-mouse IgG	1:10000
anti-PTBP1	1:1000	Anti-mouse IgG	1:10000
anti-hnRNP A1	1:1000	Anti-mouse IgG	1:10000
anti-hnRNP A2/B1	1:1000	Anti-mouse IgG	1:10000
β -actin	1:10000	Anti-mouse IgG	1:10000
Tubulin	1:1000	Anti-mouse IgG	1:10000

2.12 RNA pull-down

RNA pull-down was performed using Pierce Magnetic RNA-Protein Pull-down kit (USA). First, 50 pmol synthetic RNA (IDT, USA, Table 2.17) was ligated with desthiobiotinylated cytidine bisphosphate at the 3' end and then purified by ethanol precipitation. Subsequently, labeled RNA was bound to the streptavidin magnetic beads at room temperature followed by incubation with 120 μ g HeLa nuclear extract at 4 ° C for 2 h with agitation. After incubation, magnetic beads were washed three times with wash buffer [20 mM Tris (pH 7.5), 10 mM NaCl, 0.1% Tween-20] and the bound proteins were eluted using biotin elution buffer. Bound proteins were analyzed by SDS-PAGE and silver staining.

Table 2.17 List of RNA oligonucleotides for RNA pull-down

BIM.Ex3ESE2.WT	UAGUC AAGAUACAGAACAACUCAACCACAA
BIM.Ex3ESE2.MUT	UAGUC AACAUAUUAAACA U U CAACCACAA
BIM.Ex3ESE3.WT	UCAACC ACAAGGAUUUCUCAUGAUA
BIM.Ex3ESE3.MUT	UCAACC CAUAAGCAUUUC CAUGAUA

Note: Core sequences of *cis*-acting elements are shown in bold. Mutations are labeled in red.

2.13 Silver staining

Protein gels were first incubated in 50% methanol twice, then in 5% methanol and in water with 60 mM DTT once for fixation. Each incubation was 15 min long with constant shaking. Subsequently, the gel was incubated in 0.1% silver nitrate solution (Sigma-Aldrich, USA) for 15 min followed by incubation with developing solution (3% sodium carbonate (Sigma-Aldrich, 5 USA) and 0.05% formaldehyde). The protein bands which were absent or less intense in the mutant control samples were analyzed by LC-MS/MS, a service provided by NTU mass spectrometry core facilities, and confirmed by western blot.

2.14 CRISPR guide RNA Vector construction and ssODN design

In order to study the effects of different *BIM* E3 SNPs on E3 splicing and TKI response, we intended to introduce minor alleles of these SNPs into the endogenous *BIM* genes in K562 by the Clustered Regularly Interspaced Short Palindromic Repeats/CRISPR associated protein 9 (CRISPR/Cas9) genome editing tool. To introduce SNP 3c and SNP 4a in K562 *BIM* E3, 20 nt guide sequences targeting human *BIM* E3 were cloned into the pSpCas9n(BB)-2A-GFP (PX461; Addgene) vector which expresses the mutated “nickase” version of the Cas9 enzyme. A pair of appropriately spaced guide RNAs (gRNAs) would target the Cas9 nickase (Cas9n) to simultaneously introduce single-stranded nicks on both strands of *BIM* E3. The double-stranded breaks (DSBs) from double nicking would then be repaired via homology directed repair (HDR) in presence of a repair template (Ran et

al., 2013b). To achieve HDR, single-stranded sense DNA oligonucleotides (ssODNs) purchased from IDT were used as repair template. ssODNs contained flanking sequences of 60 nt on each side of the SNP minor allele that are homologous to the target region (Table 2.19).

Since the CRISPR/Cas9 genome editing approach was not efficient in introducing minor alleles of SNPs into endogenous *BIM* gene, we also used another experimental strategy to study the effect of SNPs on splicing and TKI resistance. In this second approach, we firstly wanted to knockout all endogenous *BIM* genes in K562. This would be followed by transfecting K562 with a *BIM* expression plasmid with different allelic variants.

To generate *BIM* knockout in K562, firstly we designed a pair of 20 nt guide sequences each targeting two different regions of *BIM* EXON 2A (E2A). We cloned these 20 nt guide sequences individually into the pSpCas9 (BB)-2A-GFP (PX458; Addgene) vector which expresses the wild-type Cas9 enzyme. A pair of gRNAs expressed from these plasmids would induce Cas9 mediated double strand breaks (DSBs) at the two sites was expected to delete intervening DNA segment due to precise end joining (Zheng et al., 2014).

All guide sequences (shown in Table 2.18) were designed using the CRISPR design tool, CHOPCHOP (Montague et al., 2014).

PX458 and PX461 vectors were digested with BbsI at 37°C for 3 h. A pair of complementary DNA oligonucleotides containing guide sequence and corresponding restriction overhangs was first annealed by heating equal molar amounts of the two oligonucleotides at 37°C for 30 min followed by 95 °C for 5 min and gradually cooling to room temperature in 1 h. The annealed oligonucleotides were then 5' end phosphorylated by T4 polynucleotide kinase (T4PNK; New England Biolabs, USA) in 1x 10 ligation buffer [50 mM Tris-HCl, 10 mM MgCl₂, 1 mM ATP and 10 mM DTT] and ligated into PX 458 or PX 461 vector that was cut by BbsI.

Table 2.18 Design of CRISPR guide sequences

CRISPR guide sequence designs to introduce SNP 3c and 4a in <i>BIM</i> E3	
BIM.Ex3.gRNA 1.F	CACCgAAAGATCACCTATATCCTCT
BIM.Ex3.gRNA 1.R	AAACAGAGGATATAGGTGATCTTTc
BIM.Ex3.gRNA 2.F	CACCgTGCTTTGGATTTATATTTAC
BIM.Ex3.gRNA 2.R	AAACGTAAATATAAATCCAAAGCAc
BIM.Ex3.gRNA 3.F	CACCgTTTACTGGCTTAGATTTGTA
BIM.Ex3.gRNA 3.R	AAACTACAAATCTAAGCCAGTAAAc
BIM.Ex3.gRNA 1-modified .F	CACCgTCCAAAGCACAGTGAACCTCT
BIM.Ex3.gRNA 1-modified .R	AAACAGAGTTCAGTGTGCTTTGGAc
CRISPR guide sequence designs to generate <i>BIM</i> Knockout	
BIM.Ex2A.gRNA 1.F	CACCgAGTTCTGAGTGTGACCGAGA
BIM.Ex2A.gRNA 1.R	AAACTCTCGGTCACACTCAGAACTc
BIM.Ex2A.gRNA 2.F	CACCgTGTAGGGAGGTAGGGGCCCC
BIM.Ex2A.gRNA 2.R	AAACGGGGCCCCCTACCTCCCTACAc

Note: Guide sequence (20 nt) is in black font, BbsI overhang is in blue font. An extra g (shown in bold lower case) is added at the 5' of the sgRNA where the 20-nt guide sequence does not begin with g because the U6 RNA polymerase III promoter used to express the sgRNA prefers a guanine (g) nucleotide as the first base of its transcript. All DNA oligonucleotides are in 5' to 3' direction (Ran et al., 2013a).

Table 2.19 Design of ssODNs

ssODN SNP 3c	TTTTTTTTTTTTTAACAGTAGTCATCCTAGAGGATATAGGTGATCTT TCACTGTGCTTcGGATTTATATTTACTGGCTTAGATTTGTATGGCC ACCACCATAGTCAAGATACAGAACAAC
ssODN SNP 4a	TTTTTTTTTTTTTAACAGTAGTCATCCTAGAGGATATAGGTGATCTT TCACTGTGCTTTGaATTTATATTTACTGGCTTAGATTTGTATGGCC ACCACCATAGTCAAGATACAGAACAAC

Note: ssODNs contain flanking sequences of 60nt on each side of the SNP minor allele which is indicated in red color.

Cas9 gRNA and Cas9n plasmids were transformed into competent DH5 α *E. coli* cells via heat shock. Selection for successfully transformed *E. coli* cells was done using lysogeny broth agar plates containing 70 μ g/mL of ampicillin. Colonies were grown in 3 ml of LB

overnight at 37°C. The plasmids were then extracted using MiniPrep kit (Omega Bio-tek, USA). All purified plasmids were sequenced (1st Base, Singapore) to ensure insertion of correct guide sequences.

2.15 Nucleofection

For introducing minor alleles of SNPs in *BIM* E3 in K562, 2×10^6 cells were resuspended in 100 µl SF reagent (Lonza) with 2.5 µg of both Cas9n/gRNA CRISPR plasmids and 1 µl of 100 µM respective ssODNs. For introducing *BIM* knockout in K562, 2×10^6 cells were resuspended in 100 µl SF reagent (Lonza) with 2.5 µg of both Cas9/gRNA CRISPR plasmids. Nucleofection was performed according to manufacturer's instruction by using the FF-120 program with Amaxa 4D-Nucleofector (Lonza, Switzerland). After 24 hr, cells were sorted for GFP expression.

2.16 Fluorescence-activated cell sorting (FACS)

Post 24 h of nucleofection, cells were resuspended in culture media and sorted for expression of green fluorescent protein (GFP) using the BD FACS Aria cell sorter (Becton Dickinson, USA). FACS sorting was done for the top ~2% of GFP positive cells in order to enrich for cells that received high levels of the CRISPR/Cas9 plasmids. The sorted cells were plated on four 96-well plates for single-clone isolation. Cells were then allowed to expand for 3-4 weeks. When the cells were confluent, 50% of cells were extracted for DNA extraction and genotyping.

2.17 Genomic DNA extraction and genotyping CRISPR-edited *BIM*

K562 genomic DNA was extracted by adding lysis buffer [100 mM NaCl, 100 mM Tris-HCl (pH 7.5), 1 mM EDTA (pH 7.5)] to cells and incubating at 55°C for 1 h. DNA was then subjected to ethanol precipitation by adding 2x volume of 100% ethanol and 0.1x volume of 3 M sodium acetate (pH 5.2) and incubated at -20 °C overnight. DNA was centrifuged at 13500 rpm for 15 min and then washed with 70% ethanol twice before resuspending the pellets in 50 µl water.

Genotyping of SNP minor alleles introduced in K562 cell line was done by PCR amplification followed by restriction fragment length polymorphism (RFLP) analysis. PCR was done using GoTaq polymerase (Promega, USA) and PCR conditions (Table 2.20). PCR primers (forward primer- 5'-CTGTTCTCCATAGAGGCTGTGC-3' and reverse primer- 5'-CAACCCAAGGAGTAAGCTTCATGG-3') upstream and downstream of E3 were used outside the region spanned by the homology arms to avoid false detection of residual repair template. The correct band size was confirmed by agarose gel electrophoresis. Subsequently the PCR products were purified using QIAquick PCR purification kit (Qiagen, Germany). The presence of SNP 3c and SNP 4a in K562 was assessed using RFLP analysis of the purified PCR products followed by agarose gel electrophoresis. A correctly introduced SNP 3c or 4a would create a restriction site recognized by Hpy1881 or APOI, respectively. The successful introduction of a SNP minor allele was confirmed by 2% (w/v) ethidium bromide-stained agarose gel by UV light visualization of smaller DNA fragments generated by the restriction enzyme digestion. Positive single clones (showing restriction digestion into smaller DNA fragments) were selected and further sequenced (1st Base, Singapore) to ensure introduction of SNP minor alleles.

BIM CRISPR knockout clones were genotyped by PCR using GoTaq polymerase (Promega, USA) and PCR conditions (Table 2.20). A forward primer (5'-AGTTCCGCAAGCTGTTGAC-3') upstream and a reverse primer (5'-GAATCTGCTTGAAGGCTGTGT-3') downstream of the gRNA cleavage sites was used. Wild-type genomic fragment for K562 and truncated genomic fragments for knockout clones were resolved by 2% (w/v) ethidium bromide-stained agarose gel and visualized by UV light. Positive clones (showing truncated PCR products) were selected and sequenced (1st Base, Singapore) to confirm deletion of the intended genomic region.

Table 2.20 Conditions for PCR

Cycling step	Temperature (°C)	Time	
Initial denaturation	95	2 min	
Denaturation	95	30 sec	} 35 cycles
Annealing	56	50 sec	
Extension	72	50 sec	
Final extension	72	5 min	
Cooling	4	∞	

2.18 *BIM* expression plasmid construction

The *BIM* expression plasmid was generated to study the effects of *BIM* SNPs on splicing and TKI response. The idea was to knockout all endogenous *BIM* genes in K562, and then rescue it by transfecting cells with a *BIM* expression plasmid. After determining that the *BIM* plasmid recapitulated endogenous *BIM* splicing and reconstituted all BIM protein isoforms, the plasmid would then be mutated to harbour the minor alleles of the SNPs.

The *BIM* expression plasmid was constructed to include all *BIM* exons and shortened flanking introns (Figure 2.2). In order to generate the *BIM* plasmid, five fragments of *BIM* were amplified from Human Genomic DNA (Promega, USA) using PCR: (i) exon 1 and its downstream 200 nt intronic sequence using primers Ex1.F.NheI and In1.R, (ii) exon 2 and 200 nt of its upstream and downstream introns using primers In1.F and In2.R and (iii) E3 and 422 nt of its upstream and 200 nt of its downstream introns using primers In2.F and In3.R (iv) E4 and 200 nt of its upstream and downstream introns using primers In3.F and In4.R and (v) exon 5 (151 nt) and 200 nt of its upstream intron using primers In4.F and Ex5.R.EcoRI. Subsequently, PCR was done using fragment (i) as template with primers Ex1.F.NheI and In1.R.ClaI (present in intron 1) EcoRI. This generated a larger fragment named E1-I1 which contained exon 1 with NheI site at the 5' end and intron 1 with ClaI- EcoRI sites at the 3'end. Fragments (ii), (iii), (iv) and (v) were merged into one by overlap PCR using primers In1.F and Ex5.R.EcoRI primer. This generated a larger fragment named I1-E5 with EcoRI at the 3'end. The *BIM* expression plasmid was

constructed in a two-step process. In the first step, the fragment E1-I1 was sub-cloned into the *Nhe*I and *Eco*R1 sites in the pcDNA 3.1 plasmid to generate an intermediate vector. Next, fragment I1-E5 was cloned into the *Cla*I/*Eco*R1 sites of the vector (Figure 2.2). Primer sequences for generating *BIM* expression plasmid have been shown in Table 2.21. All restriction enzymes were purchased from New England Biolabs, USA.

BIM expression plasmids were transformed into competent DH5 α *E. coli* cells via heat shock. Selection for successfully transformed *E. coli* cells was done using lysogeny broth agar plates containing 70 μ g/mL of ampicillin. Colonies were grown in 3 ml of LB overnight at 37°C. The plasmids were then extracted using MiniPrep kit (Omega Bio-tek, USA). Plasmids were digested with restriction enzymes to confirm the insertion of specific *BIM* fragments before sequencing. Purified plasmid clones showing the correct insert size were sent for sequencing (1st Base, Singapore) to ensure insertion of correct *BIM* fragments.

Table 2.21 Primer sequences for generating *BIM* expression plasmid

Ex1.F.NheI	AATTTAGCTAGCACTTCGCTCCGCGC
In1.R	GTAATTTTTTTTTTGGTAGAGGGTACCCGACTTTTCTTTAA AGAGAACCTGCCC
In1.F	GGGCAGGTTCTCTTTAAAGAAAAGTCGGGGTACCCTCTACC AAAAAAAAAATTAC
In2.R	GAAATAAGCCTGTCATGGAAGAGCCAGCCTGGGCGACAG
In2.F	CTGTCGCCCAGGCTGGCTCTTCCATGACAGGCTTATTC
In3.R	GAGTCTTCTACATAACTCTGGTTAAAGAATAAGACCAAACA ACCCAAGGAGTAAGC
In3.F	GCTTACTCCTTGGGTTGTTTGGTCTTATTCTTTAACCAGAG TTATGTAGAAGACTC
In4.R	CCCTGCAAAGGGGAAGACATTATTTTATTTAAGTGCCATTA CTTAGTCATAAAAC
In4.F	GTTTTATGACTAAGTAATGGCACTTAAATAAAATAATGTCT TCCCCTTGCAGGG
Ex5.R.EcoRI	TGCTGGGTGAATTCGGTTTGAACAAGCAAAATGTCTGCATG G
In1.R.ClaI.EcoEI	CTCCCAGAATTCAAAAATCGATGAATAAACAAATTAGCTCT CACC

Figure 2.2 Generation of *BIM* expression plasmid by overlap PCR

Exons are represented by boxes, introns with solid lines. Arrows represent primers. Dotted lines represent deleted sequences. See text for details of plasmid construction.

3 RESULTS

3.1 *BIM* minigene plasmids

The *BIM* minigene plasmids, provided by collaborator Dr Ong Sin Tiong, are chimeric constructs containing adenovirus first exon (U) followed by *BIM* exon 3 (E3) and flanking intronic sequences, and *BIM* exon 4 (E4) and upstream intronic sequence (Figure 3.1) (Juan et al., 2014; Ng et al., 2012). E4 is fused with adenovirus exonic sequence D, which has a PAS. The full-length minigene contains the 2.9 kb intronic polymorphic sequence present in TKI sensitive patients, while the deletion minigene has the 2.9 kb intronic sequence removed (representative of TKI resistant patients). The $\Delta 10$ and $\Delta 11$ minigenes are essentially shorter versions of the full-length and deletion minigenes, respectively. The $\Delta 10$ minigene contains only the 322 nt fragment at the 3' end of the 2.9 kb polymorphic region, which is sufficient to suppress E3 splicing. $\Delta 11$ minigene has the 322 nt fragment removed, so it recapitulates the effects of the deletion minigene. The minigenes were expected to generate two transcripts, U-E3 and U-E4-D, as E3 and E4 are alternative exons (Figure 3.1) (Juan et al., 2014; Ng et al., 2012).

Figure 3.1 Schematic of the *BIM* minigenes

The *BIM* minigenes contained an adenovirus exonic sequence (U) followed by *BIM* E3 and E4 fused with the adenovirus exonic sequence, D. Boxes represent exons and lines represent introns. The 2.9 kb polymorphic fragment has been depicted in red line.

3.2 $\Delta 10$ and $\Delta 11$ mimic splicing patterns of full-length and deletion minigenes

In order to verify if $\Delta 10$ and $\Delta 11$ minigenes exhibited similar splicing patterns to the full-length and deletion minigenes, respectively, K562 cells were transfected with the minigene plasmids, followed by RNA extraction and reverse transcription. PCR as well as q-PCR was performed to analyze E3 versus E4 splicing. Both $\Delta 11$ and deletion minigene showed enhanced splicing of E3 over E4 in comparison to $\Delta 10$ and full length minigene respectively (Figure 3.2). Consistent with our collaborators findings, our results confirm that the splicing pattern of $\Delta 10$ is similar to full length, while that of $\Delta 11$ is similar to deletion, further verifying that the 322 nt fragment at the 3' end of the 2.9 kb deletion region is sufficient to suppress E3 splicing.

For our subsequent experiments, we used $\Delta 10$ as the non-deletion minigene and $\Delta 11$ as the deletion minigene, as PCR mutagenesis was more efficient using these smaller plasmids as templates.

Figure 3.2 E3 versus E4 splicing pattern in *BIM* minigenes

(A) PCR and (B) qPCR analysis of splicing patterns of *BIM* E3 versus E4 in different minigenes. Both deletion and $\Delta 11$ minigenes showed enhanced splicing of E3 over E4 in comparison to full length and $\Delta 10$ minigenes, respectively. The open arrows indicate the location of the primers used to analyze the U-E3 and U-E4-D transcripts. Experiments were performed in triplicate. qPCR data is shown as mean $\pm$ standard error of the mean. Statistical significance is indicated as follows **, $P < 0.01$ compared to cells transfected with the WT $\Delta 10$ and full length minigenes respectively using Student's t-test.

3.3 Multiple *cis*-acting elements regulate *BIM* E3 splicing

Cis-acting elements regulating *BIM* E3 splicing were identified by minigene deletion analysis in *BIM* E3 and the upstream intronic region after the 2.9 kb polymorphic sequence. The identified ESE sequences were further validated by point mutation and heterologous context assays.

3.3.1 *BIM* E3 is rich in ESEs and ESSs

As typical binding sites for splicing factors encompass 6-8 nucleotides (Cartegni et al., 2002), in order to identify the ESEs and ESSs in *BIM* E3, 23 serial deletions of 10 nucleotides each were introduced in *BIM* E3 in the context of $\Delta 10$ (deletion) and $\Delta 11$ (non-deletion) minigenes (Figure 3.3). To more precisely map the *cis*-acting elements, we introduced additional 23 staggered deletions of 10 nucleotides, for which each staggered deletion spans from middle to a deletion to middle of the next deletion (Figure 3.3). This approach of introducing deletions and staggered deletions to map the *cis*-acting splicing elements has been previously done in our lab (Tang et al., 2016). The combined analysis of the two sets of deletions that are staggered by 5 nucleotides would minimize the possibility that the splicing effects were due to the inadvertent creation of splicing elements within the new junction sequences. The first two nucleotides of the 5' end of E3, which are part of the 3'ss, were excluded from the deletions. Serial deletions were extended to the intronic region downstream of E3 (17 nt) (Figure 3.3), as this region contained polyadenylation regulatory elements which could influence E3 splicing. K562 cells were transfected with the wild-type ($\Delta 10$ and $\Delta 11$) or deletion mutant minigenes and the E3 to E4 ratios were then compared to those of the wild-type minigene by qRT-PCR 48-h post transfection.

Figure 3.3 Schematic map of serial deletions (D) and staggered deletions (SD) in *BIM* E3

D1 to SD23 indicate deletions in the exon while –D1 to –SD8 indicate deletions in the intron upstream E3. Stop codon is indicated with a traffic stop sign. The conserved PAS has been highlighted in grey. Exonic sequence is shown in upper case while intron is shown in lower case.

Numerous *cis*-acting splicing elements were identified within *BIM* E3 (Figure 3.4). As expected, transfection with the deletion mutants showed varied effects on E3 inclusion. Deletion of sequences D1, SD1, D4, D8, SD9, D10, SD10, SD11, D12 and SD17 caused a significant reduction in splicing of E3 over E4 (Figure 3.4). This strongly suggests that these sequences contain potential ESEs which recruit splicing activators to promote splicing, since their deletion caused a reduction in E3 splicing. In contrast, deletions of sequences D2, SD2, SD3, SD5, D6, SD6, D13, SD13, D14, SD14, D15, and SD16 showed enhanced splicing of E3 over E4 (Figure 3.4). This suggests that these sequences contain potential ESSs which recruit splicing repressors to suppress splicing, since their deletion caused an increase in E3 splicing. Deletion of other sequences had neutral effects on E3 splicing, indicating that these sequences had no *cis*-acting splicing elements (Figure 3.4). Moreover, apart from a few cases (e.g. SD10, D10 and SD11), the results of the first set of deletions and staggered deletions were consistent with each other.

Figure 3.4 Numerous *cis*-acting splicing elements are present in *BIM* E3

Schematic of serial deletions (D) and staggered deletions (SD) introduced in *BIM* E3 are shown above the graphs. Real time PCR analysis showing E3/E4 ratio and Log_2 (E3/E4) in the mutant plasmids, compared to the wild type (WT), in the context of both, $\Delta 10$ and $\Delta 11$ minigenes. ESEs are depicted in blue bars, while ESSs are depicted in red bars. WT and sequences having neutral effects on splicing are depicted in grey bars. Experiments were performed in triplicate and data is shown as mean $\pm$ standard error of the mean. Statistical significance is indicated as follows: *, $P < 0.05$; **, $P < 0.01$ compared to cells transfected with the WT minigene using Student's t-test.

3.3.2 E3 PolyA signal and flanking regions exhibit enhancer-like effects

The deletions of the PAS and flanking sequences (SD19 to SD23) caused significant decrease in inclusion of E3 over E4 in comparison to the WT minigene (Figure 3.5). This suggests that these sequences manifest enhancing effects on splicing due to interaction between the polyadenylation and splicing machineries. The results were consistent in the context of both $\Delta 10$ and $\Delta 11$ minigenes, further validating our analysis. Our findings are in concordance with previous studies, which have shown the mutual stimulation between polyadenylation and splicing of terminal exons (Kyburz et al., 2006; Millevoi et al., 2006; Niwa and Berget, 1991).

Figure 3.5 Polyadenylation signal and flanking region within *BIM* E3 promote E3 splicing

Schematic of serial deletions (D) and staggered deletions (SD) introduced in polyA regulatory sequences are shown above the graphs. D20 sequence containing the conserved PAS, AAUAAA is highlighted in grey. E3/E4 ratio of mutant plasmids compared to WT, in the context of both, Δ10 and Δ11 minigene by qRT-PCR. Experiments were performed in triplicate and data is shown as mean ± standard error of the mean. Statistical significance is indicated as follows: *, P<0.05; **, P<0.01 compared to cells transfected with the WT minigene using Student's t-test.

3.3.3 Intronic sequences upstream of *BIM* E3 3'ss contain splicing elements

In order to get a complete picture of the *cis*-acting elements regulating *BIM* E3 splicing, deletions were also introduced in *BIM* intronic region directly upstream of E3 and immediately downstream of the 2.9 kb polymorphic fragment. This work was performed by CN Yang undergraduate program student, Yue Wan Lin under my direct supervision, and I confirmed the effects of some mutants.

A total of 8 non-overlapping serial and staggered deletions of 10 nucleotides each were generated in *BIM* intron 2 (106 nt) in $\Delta 10$ and $\Delta 11$ minigenes (Figure 3.3). Deletions were not introduced in the last 17 nt of the intronic sequence as removal of this segment, which corresponds to the 3'ss of E3, would decrease splicing efficiency. From the experimental results, the deleted sequences corresponding to -D6, -SD6 and -SD8 are likely to contain ISSs while deleted sequences corresponding to -D1, -SD1, -D2, -SD2, -D3, -SD3, -D4 and -D5 are likely to contain ISEs (Figure 3.6). The deleted sequence corresponding to -SD1 contains the predicted BPS for intron 2 of the *BIM* gene, which is an essential *cis*-acting splicing element. Our experimental results are consistent with this prediction, as shown by the drastic reduction in E3/E4 ratio in both $\Delta 10$ and $\Delta 11$.

The identified ESEs/ESSs/ISEs/ISSs regulating *BIM* E3 were consistent in the context of both $\Delta 10$ and $\Delta 11$ as shown by the very high correlation coefficient ($R^2 = 0.9644$; Figure 3.7). A summary map of enhancers and silencers regulating *BIM* E3 has been generated based on the deletion scan analysis (Figure 3.8). Since enhancer sequences are useful potential targets for ASOs, we focused on three ESE regions: ESE1 (D1+SD1 sequences), ESE2 (SD9+SD10 sequences) and ESE3 (SD11+D12 sequences) for further analysis (Figure 3.8).

Figure 3.6 Intronic sequences upstream E3 contain ISEs and ISSs

Schematic of serial deletions (D) and staggered deletions (SD) introduced in intron 2 upstream of E3 are shown above the graphs. Real time PCR analysis showing the ratio of E3 to E4 transcripts in the mutant plasmids, compared to the wild type (WT), in the context of both, $\Delta 10$ and $\Delta 11$ minigene. ISEs are depicted in blue bars, while ISSs are depicted in red bars. WT and sequences having neutral effects on splicing are depicted in grey bars. Experiments were performed in triplicate and data is shown as mean $\pm$ standard error of the mean. Statistical significance is indicated as follows: *, $P < 0.05$; **, $P < 0.01$ compared to cells transfected with the WT minigene using Student's t-test.

Figure 3.7 E3/E4 correlation for deletion mutants across $\Delta 10$ and $\Delta 11$

The effects on E3 splicing caused by introduction of deletion and staggered deletion sequences across $\Delta 10$ and $\Delta 11$ show a positive correlation. X axis shows $\log_2$ (E3/E4) ratio of mutant plasmids compared to wild type (WT) in context of $\Delta 10$ minigene, while Y axis shows $\log_2$ (E3/E4) ratio of mutant plasmids compared to wild type (WT) in context of $\Delta 11$ minigene.

Figure 3.8 Map of enhancers and silencers for *BIM* E3 splicing by deletion analysis

For each deletion (D) or staggered deletion (SD), the E3/E4 ratio is considered as an average of the mean E3/E4 ratio from both $\Delta 10$ and $\Delta 11$ minigenes. To derive the final map, we took the average normalized E3/E4 ratio between each deletion and staggered deletion and applied it to the 5 nt regions where each deletion and staggered deletion overlapped, marking them as either enhancer (blue), silencer (red), or neutral (uncolored) based on the following criteria: an ESE/ISE is defined as such if there is $\geq 30\%$ decrease in the mean E3/E4 ratio compared to WT after the deletion is made; an ESS/ISS is defined as such if there is $\geq 30\%$ increase in the mean E3/E4 ratio compared to WT after the deletion is made. For strong ESE/ISE and ESS/ISS, the cutoff is a $\geq 50\%$ decrease or increase in the mean E3/E4 ratio compared to WT, respectively, after the deletion is made.

3.3.4 Point mutations within ESEs of *BIM* E3 reduced enhancer effects

Our linker scan analysis revealed many *cis*-acting splicing elements in *BIM* E3. *In silico* analysis using SFmap, Human Splicing Finder and SpliceAid (Akerman et al., 2009; Desmet et al., 2009; Giulietti et al., 2013; Paz et al., 2010) predicted that the ESEs and ESSs are bound by SR proteins like SRSF1 and hnRNPs like hnRNP A1 or hnRNP A2/B1 (Figure 3.9). However these predictions may not be accurate due to many possible false positives. Hence, in order to identify *trans*-acting factors binding to the *cis*-acting elements, we also used a systematic RNA pull-down approach.

Figure 3.9: *In silico* predictions of *trans*-acting factors binding to the E3 ESEs and ESSs

ESEs are indicated in blue while ESSs are indicated in red. The underlined sequences indicate the ESE regions (named ESE1-3) that we focused to identify potential *trans*-acting factors associating with them. The predicted *trans*-acting factors are mapped to their respective binding motifs.

For RNA pull-down assay, mutant control sequences are required to identify the specific proteins that interact with wild-type sequence but not with the mutant. Hence, we used the *in silico* tool, Human Splicing Finder (Desmet et al., 2009) to predict the outcome of introducing nucleotide substitutions in the three important exonic enhancer regions (ESE1: D1+SD1 sequences, ESE2: SD9+SD10 sequences and ESE 3: SD11+D12 sequences). Point mutations were introduced such that the enhancer sites were disrupted as much as possible without creating new enhancers/silencers (Desmet et al., 2009). For each enhancer region, 3 or 4 mutants were generated and each mutant contained one or more nucleotide substitutions (Figure 3.10). This point mutation assay would also aid in validating the effects of ESE1, ESE2 and ESE3 sequences derived from the deletion scan analysis.

Figure 3.10 List of *BIM* E3 WT ESE sequences and the corresponding mutants

The three ESE regions have been indicated in blue boxes within E3 (green box). The WT ESE sequences have been shown. The corresponding point mutations introduced within the ESE regions have been depicted in different colors.

Real time PCR analysis suggested that all the mutants caused a significant reduction in E3 splicing in different levels in the context of both $\Delta 10$ and $\Delta 11$ minigenes (Figure 3.11). And the effects were consistent with those caused by the respective deletion mutants. This further suggests that these sequences indeed harbor ESEs as mutating these sites caused a reduction in E3 splicing. The mutants (generated by introducing point mutations) that demonstrated the most similar splicing pattern to the respective deletion mutants were used as controls for subsequent RNA pull down assay.

Figure 3.11 Point mutations reduced enhancer effects

Real time PCR analysis showing the ratio of E3 to E4 transcripts in the mutant plasmids compared to the wild type, in the context of both $\Delta 10$ and $\Delta 11$ minigene. The corresponding deletion and staggered deletion mutants served as positive controls. (A) ESE1 region, (B) ESE2, (C) ESE3 region. Experiments were performed in triplicate and data is shown as mean $\pm$ standard error of the mean. Statistical significance is indicated as follows: *, $P < 0.05$; **, $P < 0.01$ compared to cells transfected with the WT minigene using Student's t-test.

3.3.5 Validation of ESE sequences in a heterologous context

Next, to further confirm splicing enhancer effects of the *BIM* E3 ESE sequences, we tested the ESEs in the context of the heterologous PSXN minigene. The PSXN minigene has been previously used to test splicing enhancer activity of RNA motifs. Briefly, the functional enhancer effect is verified by inserting test sequences in the weakly spliced alternative exon between β -globin exon 1 and exon 2 and observing if the presence of ESEs promotes inclusion of the weakly spliced alternative exon (Coulter et al 1997).

We analyzed the splicing effects of the wild-type and mutated ESE sequences by cloning the sequences in the PSXN minigene (Figure 3.12A). K562 cells were transiently transfected with the PSXN test constructs, and the splicing pattern was determined by semi quantitative-radioactive PCR using β -globin primers specific for this vector. All three ESE sequences caused varied levels of exon inclusion (Figure 3.12B). Of the three, ESE2 was found to be the strongest enhancer. Insertion of this sequence into the weakly spliced exon caused a very high level of exon inclusion (Figure 3.12B, lane 7). This effect was dramatically reduced by insertion of the respective mutated sequence in the alternative exon (Figure 3.12B, lane 8). ESE3 also showed strong enhancer effects with the respective mutation causing drastic exon skipping (Figure 3.12B, lanes 9-10). However ESE1 caused very mild inclusion of the alternative exon (Figure 3.12B, lane 3). Also enhancing the 5'ss sequence for more efficient base-pairing to U1 snRNA did not cause a large increase in exon inclusion (Figure 3.12B, lane 5). The respective mutated sequences caused reduction in exon inclusion, confirming that the mutations disrupted the mild splicing enhancer effects (Figure 3.12B, lanes 4 and 6). In summary, the ESE sequences identified by deletion and point mutation analysis were further validated in the heterologous context, confirming their enhancer effects.

Figure 3.12 Testing ESEs in heterologous context

(A) Schematic diagram of the pSXN plasmid. Test sequences were cloned into the gray exon using SalI and BamHI sites. Arrows depict RT-PCR primers for detecting minigene transcripts. (B) The pSXN 13 containing enhancer sequence and the pSXN-linker containing short neutral sequence were the positive and negative control. Exon inclusion suggests the enhancing effect of the inserted sequences (lanes 1, 3, 5, 7, 9), while almost complete absence of exon inclusion implies the disruption of the enhancer activity (lanes 2, 4, 6, 8, 10). Experiment was done at least in triplicate in K562.

3.4 *BIM* E3 splicing is regulated by length of PolyU tract at 3'ss

Polypyrimidine tracts are defined by a pyrimidine-rich region upstream of the 3'ss AG that normally have a preference of Us over Cs (Sheth et al., 2006). Binding of U2AF65 at the polypyrimidine tract facilitates recruitment of U2 snRNP at the branch point driving the subsequent splicing steps (Gozani et al., 1998). A previous study has reported that a range of 10-15 pyrimidines at the 3'ss of an exon ensures its normal splicing (Fu et al., 2011) while in another study, a stretch of 11 continuous uridines was found to be highly competent in ensuring optimal splicing (Coolidge et al., 1997). The polypyrimidine tracts from the 3'ss of the *BIM* E3 Δ 10 and Δ 11 minigenes encompass a stretch of 14 Uridines (Us) (Figure 3.13), while in humans this tract can range from 17 to 21Us (dbSNP rs748523945 and rs200489083). In order to determine the minimum number of pyrimidines required for maximal *BIM* E3 splicing, the length of the polyU tract at the 3'ss was varied in constant increments of 2 (6Us to 20Us) in the context of both Δ 10 and Δ 11 minigenes. Real time PCR analysis suggested that a minimum of 10 uridines (Us) in the pre-mRNA were required for E3 inclusion, as E3 inclusion was almost negligible when the polyU stretch was shorter than 10Us (Figure 3.13). Surprisingly, we observed that a stretch of 16Us was required for maximal E3 splicing. Increasing the number of Us beyond 16 did not cause further enhancement of E3 inclusion (Figure 3.13).

Figure 3.13 *BIM* E3 splicing is regulated by length of PolyU tract

(A) Schematic showing PolyU tract from *BIM* E3 minigenes, depicted in red color. The 3'ss conserved dinucleotide, AG has been underlined (B) Real time PCR analysis showing the ratio of E3 to E4 transcripts in the mutant plasmids normalized to $\Delta 10$ and $\Delta 11$ (containing 14 Us). Experiments were performed in triplicate and data is shown as mean $\pm$ standard error of the mean. Statistical significance is indicated as follows: *, $P < 0.05$; **, $P < 0.01$ compared to cells transfected with the WT minigene using Student's t-test.

3.5 Identification of *trans*-acting factors regulating *BIM* E3 splicing

Next, we wanted to identify the *trans*-acting factors binding to the ESE sequences verified by deletion, mutation and heterologous context assays. For that, we used the systematic RNA pull-down approach and candidate overexpression/knockdown approach.

3.5.1 RNA pull-down approach

RNA pull-down allows the enrichment of RNA binding proteins using RNA end-labeled with desthiobiotin and streptavidin magnetic beads. The focus of my RNA pull-down assay was on ESE2 and ESE3 as they showed stronger enhancer effects in the heterologous minigene context. I purchased synthetic RNA oligonucleotides consisting of the wild-type or mutant ESE sequences and additional 5 flanking nucleotides at each end

from IDT (Table 2.17). These RNA oligonucleotides were first labeled with desthiobiotin and attached to streptavidin magnetic beads and then incubated with HeLa nuclear extract. Proteins bound to RNA were eluted and resolved by SDS-PAGE and detected by silver staining or western blot. Numerous proteins with varied molecular weights were pulled down together with wild-type RNA. With the assumption that binding of real splicing factors is disrupted by sequence mutations, the protein bands showing reduced intensity in the mutant as compared to the wild-type sample were selected for mass spectrometry at NTU-SBS core facility. From mass spectrometry results, proteins with high scores (high possibility of the identity to be true) were considered for further study.

In ESE2 RNA pull-down, one band showed reduced intensity in mutant sample compared to WT ESE2. From mass spectrometry analysis of this band, the only splicing factor identified was hnRNP A2/B1 (Figure 3.14A). The common splicing activators i.e. SR proteins were not found to be binding to ESE2. hnRNP A2/B1 was further confirmed to bind to ESE2 sequence while its binding was reduced in the mutant ESE2 by western blotting analysis (Figure 3.15A). Western blotting was also performed to test if hnRNP A1 binds to ESE2 region, as hnRNP A1 and hnRNP A2/B1 are closely related proteins (He and Smith, 2009). Indeed hnRNP A1 was found to bind to ESE2 sequence while its binding was reduced in the mutant ESE2 (Figure 3.15A).

In ESE3 RNA pull-down, several bands showed reduced intensity in mutant sample compared to WT ESE3. However, mass spectrometry analysis revealed only hnRNPs such as PTBP1 and hnRNP A1 binding to the ESE3 sequence (Figure 3.14B). No SR proteins (the typical splicing activators) were found binding to ESE3 by mass spectrometry analysis. Binding of hnRNP A1 and PTBP1 to the enhancer sequence was further confirmed by western blotting analysis (Figure 3.15B).

Figure 3.14 RNA pull-down assay for ESE2 and ESE3

(A) The protein band that showed reduced binding to ESE2 mutant versus ESE2 WT RNA sequence was identified as hnRNP A2/B1 by mass spectrometry analysis. (B) Proteins such as PTBP1 and hnRNP A1 showed reduced binding affinity to the ESE3 mutant sequence when compared to ESE3 WT RNA sequence as identified by mass spectrometry analysis. All RNA pull-down assays were performed twice.

Figure 3.15 Western blot confirmed protein binding to ESE2 and ESE3

(A) hnRNP A2/B1 and hnRNP A1 were confirmed to be bound to the ESE2 and this interaction was absent/reduced in mutant ESE2. (B) PTBP1 and hnRNP A1 were bound to the ESE3 WT and the binding was absent/reduced in mutant ESE3 respectively. Note: (A) the blot for hnRNP A1 consists of two separate gels.

As shown in the pull down, hnRNPs were mainly found to be interacting with ESEs and the interactions were compromised by mutations. Also, according to computational predictions, there are putative binding sites for hnRNP A2/B1 and hnRNP A1 within the ESE2 and ESE3 regions. Although hnRNPs are known repressors, they can also activate exon inclusion in some cases (Huelga et al., 2012). Next, to validate the RNA-pull down results, knockdown and overexpression assays for hnRNP A1 and hnRNP A2/B1 were performed to observe changes in endogenous *BIM* E3 splicing (data not shown). The knockdown and overexpression efficiency of hnRNP A1 was low. Hence its effects on the endogenous E3-containing *BIM* transcripts were not conclusive. Knockdown of hnRNP A2/B1 did not cause any splicing changes in endogenous *BIM* E3 splicing suggesting that hnRNP A2/B1 may not be a regulator of *BIM* E3 inclusion. Although PTBP1 was found to bind ESE3 by RNA-pull down, a previous knockdown experiment revealed PTBP1 to be a repressor of *BIM* E3 inclusion (Juan et al., 2014). The binding of PTBP1 to ESE3 found by RNA-pull down could be a weak interaction, and in cells this interaction could be compromised by the presence of stronger competitor activators binding to the same ESE. Hence further knockdown/overexpression study on PTBP1 was not conducted as we were interested in finding splicing activators binding to enhancer sites.

3.5.2 Candidate approach

Next, candidate *trans*-acting factors that bind to these ESE regions were predicted using the following *in silico* tools: SFmap, Human Splicing Finder and SpliceAid (Akerman et al., 2009; Desmet et al., 2009; Giulietti et al., 2013; Paz et al., 2010). *Trans*-acting factors predicted by more than one program were selected for further functional analysis.

According to these *in silico* tools: SRSF1 binds to ESE1 and ESE3, while SRSF3 binds to ESE1 and ESE2 and SRSF5 binds to ESE3 (Figure 3.9). Hence the effect of SRSF1, SRSF3 and SRSF5 on endogenous *BIM* E3 splicing was tested by knockdown and/or overexpression assays. All knockdown and overexpression assays were performed by final year project student, Loh Yan Ping, under my direct supervision, and I confirmed some of these results.

3.5.2.1 SRSF1 promotes *BIM* E3 inclusion

For overexpression of SRSF1, K562 cells were transfected with T7 tagged SRSF1 cDNA plasmid. Upregulation of SRSF1 increased production of endogenous *BIM* E3 containing transcripts, suggesting that SRSF1 may act as an activator of *BIM* E3 splicing (Figure 3.16C). To determine if the increased E3 splicing is due to binding of SRSF1 to *cis*-acting elements within E3 and E4, overexpression studies were done with co-transfection of *BIM* splicing minigenes ($\Delta 10$ and $\Delta 11$). Upon upregulation of SRSF1, it was observed that *BIM* E3 containing transcripts from the minigenes increased (Figure 3.17), which suggests that SRSF1 may regulate *BIM* E3 splicing by binding to E3 or nearby flanking intronic sequences in the minigene. To further validate the involvement of SRSF1 as an activator of *BIM* E3 inclusion, knockdown-rescue studies were performed. Knockdown of SRSF1 transcript levels using SRSF1 siRNA, resulted in a decrease of endogenous *BIM* E3-containing transcripts (Figure 3.18B). With subsequent rescue by plasmids containing siRNA resistant SRSF1 cDNAs, the knockdown splicing phenotype was reversed in a dose-dependent manner (Figure 3.18B). Taken together, the overexpression and knockdown-rescue results confirm that SRSF1 promotes *BIM* E3 inclusion.

Figure 3.16 SRSF1 overexpression promotes endogenous *BIM* E3 inclusion

(A) Relative expression of SRSF1 transcript levels by real time PCR. (B) Expression of SRSF1 protein level by western blot. The red arrow indicates SRSF1 overexpression (C) Real time PCR analysis showing the ratio of endogenous E3 to E4 transcripts (D) Relative expression of endogenous exon 2A-containing *BIM* transcripts. This represents the total level of *BIM* transcripts, as exon 2A is constitutive in all *BIM* isoforms. All relative expressions are determined by normalizing to K562 cells that were untreated (“U”). Results are represented as an average of three replicates and data are shown as mean $\pm$ standard deviation. *, $P < 0.05$.

Figure 3.17 SRSF1 overexpression promotes *BIM* E3 inclusion in minigenes

(A) Relative expression of SRSF1 transcript levels. (B) Real time PCR showing ratio of E3 to E4 transcripts from *BIM* minigenes. All relative expressions are determined by normalizing to K562 cells that were transfected with *BIM* splicing minigenes (Δ10 and Δ11) (“control”). Results are represented as an average of three replicates and data are shown as mean $\pm$ standard deviation. *, $P < 0.05$.

Figure 3.18 SRSF1 knockdown promotes E3 exclusion and rescue partially reverses it

(A) Relative expression of SRSF1 transcript levels by real time PCR. (B) Real time PCR analysis showing the ratio of endogenous E3 to E4 transcripts (C) Relative expression of endogenous exon 2A-containing *BIM* transcripts. This represents the total level of *BIM* transcripts, as exon 2A is constitutive in all *BIM* isoforms. All relative expressions are determined by normalizing to K562 cells that were not subjected to knockdown and rescue. Experiments were performed in triplicate and data is shown as mean $\pm$ standard error of the mean. Statistical significance is indicated using Student's t-test as follows: *, $P < 0.05$.

3.5.2.2 SRSF3 is likely a repressor of *BIM* E3 inclusion

Knockdown of SRSF3 using siRNA caused a 6-fold increase in endogenous *BIM* E3-containing transcripts (Figure 3.19C). This suggests that SRSF3 may act as a repressor in regulating *BIM* E3 inclusion, contradicting with the general role of SRSF3 as a splicing activator. However when subsequent rescue by plasmids containing siRNA resistant SRSF3 cDNAs was carried out, the knockdown splicing phenotype was not reversed upon rescue with increasing levels of SRSF3 (Figure 3.20C). Overall, our initial results suggested that SRSF3 repress *BIM* E3 inclusion, but subsequent experiments did not support the finding (see discussion, section 4.5.3).

Figure 3.19 SRSF3 knockdown promotes endogenous *BIM* E3 inclusion

(A) Relative expression of SRSF3 transcript levels by real time PCR. (B) Expression of SRSF3 protein level by western blot. (C) Real time PCR analysis showing the ratio of endogenous *BIM* E3 to E4 transcripts (D) Relative expression of endogenous exon 2A-containing *BIM* transcripts by real-time PCR. This represents the total level of *BIM* transcripts, as exon 2A is constitutive in all *BIM* isoforms. All relative expressions are determined by normalizing to K562 cells that were untreated (“U”). Results are represented as an average of three replicates and data are shown as mean $\pm$ standard deviation. **, P-value ≤ 0.01 .

Figure 3.20 SRSF3 rescue does not reverse the knockdown effects of SRSF3

(A) Relative expression of SRSF3 transcript levels by real time PCR. (B) Expression of SRSF3 protein level by western blot. The red arrow indicates SRSF3 overexpression (C) Real time PCR analysis showing the ratio of endogenous E3 to E4 transcripts (D) Relative expression of endogenous exon 2A-containing *BIM* transcripts. This represents the total level of *BIM* transcripts, as exon 2A is constitutive in all *BIM* isoforms. All relative expressions are determined by normalizing to K562 cells that were untreated (“U”). Experiments were performed in triplicate and data is shown as mean $\pm$ standard error of the mean. Statistical significance is indicated using Student’s t-test as follows: *, $P < 0.05$.

3.5.2.3 SRSF5 mildly promotes endogenous *BIM* E3 inclusion

Although, overexpression of SRSF5 using cDNA plasmid caused a 2-fold increase in endogenous *BIM* E3-containing transcripts (Figure 3.21), this effect was not recapitulated in the context of the *BIM* splicing minigenes ($\Delta 10$ and $\Delta 11$) (Figure 3.22). With inefficient knockdown of SRSF5 (only 15% decrease in SRSF5 mRNA level), the results were inconclusive to support the role of SRSF5 as an activator of *BIM* E3 inclusion (data not shown).

Figure 3.21 SRSF5 overexpression mildly promotes endogenous *BIM* E3 inclusion.

(A) Relative expression of SRSF5 transcript levels by real time PCR. (B) Real time PCR analysis showing the ratio of endogenous E3 to E4 transcripts (C) Relative expression of endogenous exon 2A-containing *BIM* transcripts. This represents the total level of *BIM* transcripts, as exon 2A is constitutive in all *BIM* isoforms. All relative expressions are determined by normalizing to K562 cells that were untreated (“U”). Experiments were performed in triplicate and data is shown as mean $\pm$ standard error of the mean. Statistical significance is indicated using Student’s t-test as follows: *, P<0.01, **P<0.05.

Figure 3.22 SRSF5 overexpression does not change *BIM* E3 splicing in minigenes

(A) Relative expression of SRSF5 transcript levels. (B) Real time PCR showing ratio of E3 to E4 transcripts from *BIM* minigenes. All relative expressions are determined by normalizing to K562 cells that were transfected with *BIM* splicing minigenes (Δ10 and Δ11) (“control”). Experiments were performed in triplicate and data is shown as mean ± standard error of the mean. Statistical significance is indicated using Student’s t-test as follows: *, P<0.01, **P<0.05.

To summarize this section, amongst the *trans*-acting factors regulating *BIM* E3 splicing, SRSF1 was confirmed to promote inclusion of E3, while SRSF3 was found to repress E3 splicing by knockdown. Further studies are required to establish the roles of hnRNP A1, hnRNP A2/B1 and SRSF5 on E3 splicing.

3.6 SNPs affect *BIM* E3 splicing in context of minigenes

The *BIM* deletion polymorphism can contribute to TKI resistance in many patients, however it does not account for all cases of inferior TKI responses. The multitude of *BIM* E3 regulating splicing enhancers and silencers revealed by the deletion scan analysis suggested that small sequence variations could affect the splicing of E3. Hence we were interested in studying the effect of SNPs on *BIM* E3 splicing. We searched the dbSNP database for SNPs throughout E3 and the upstream intronic region immediately downstream of the 2.9 kb polymorphic region. In our search, we found seven SNPs in the region of interest (Figure 3.23A) in May 2014. I replaced the major alleles with the minor alleles one at a time in both the $\Delta 10$ and $\Delta 11$ minigenes to observe the effects of each mutant on E3 splicing. K562 cells were transfected with the wild type ($\Delta 10$ and $\Delta 11$) or mutant minigenes. Changes in E3 to E4 splicing of mutant minigenes was then compared to the wild type minigene by qRT-PCR 48 h post transfection. Surprisingly, the minor alleles for four SNPs (SNP 3c, 4a, 5u, 6g) out of seven individually caused significant change in E3 splicing. The other three (SNP 1g, 2g, 7u) of seven SNPs did not cause significant change in splicing in either $\Delta 10$ or $\Delta 11$ minigenes (Figure 3.23B). The minor alleles for two SNPs, SNP 3c and SNP 6g, enhanced E3 to E4 ratio indicating that these alleles could be potentially associated with TKI resistance. The minor alleles for two other SNPs, SNP 4a and 5u, decreased the E3 to E4 ratio indicating that these SNPs could give additional sensitivity to TKIs (Figure 3.23B).

Figure 3.23 SNPs affect *BIM* E3 splicing

(A) Schematic of the seven SNPs in *BIM* E3 and upstream intronic region from dbSNP. Major alleles are highlighted in E3 and upstream intronic sequence in upper case and the alternative or minor alleles have been depicted below the major alleles in lower case. (B) E3/E4 ratios by real-time RT-PCR for $\Delta 10$ and $\Delta 11$ with different SNPs. Red and blue bars respectively highlight significant increases or decreases in E3/E4 ratios. Experiments were performed in triplicate and data is shown as mean $\pm$ standard error of the mean. Statistical significance is indicated as follows: *, $P < 0.05$; **, $P < 0.01$ compared to cells transfected with the WT minigene using Student's t-test (C) Table showing all reference SNP ID numbers.

Since individual patients may harbor different combination of SNPs between the two *BIM* alleles, we then explored combinations of the various alleles to determine their summative effect on *BIM* E3 splicing. All the possible combinations of alleles were introduced in both $\Delta 10$ and $\Delta 11$ minigenes to obtain the 16 different haplotypes. The wildtype $\Delta 10$ and $\Delta 11$ minigenes contain the major allele for all four SNPs and are representative of the most common haplotype. From qRT-PCR analysis, it is evident that the E3 to E4 ratio correlates well with the distribution of alleles i.e. for each combination or haplotype, the E3 to E4 splicing ratio depends on the ratio of splicing enhancing (blue bars) or splicing silencing (red bars) alleles (Figure 3.24).

For example, in the haplotypes that comprise two enhancing alleles and two silencing alleles i.e. SNP 3c+4a+5u+6g, SNP 3c+4a and SNP 5u+6g, the E3 to E4 ratio is near 1, with one exception ($\Delta 10$ SNP 3c+4a+5u+6g) (Figure 3.24). When a haplotype contained all the four enhancing alleles (SNP 3c+6g), the E3 to E4 ratio further increased in comparison to SNP 3c or SNP 6g alone (Figure 3.24). Thus by just changing two nucleotides, the E3 to E4 ratio could be boosted further not only in $\Delta 10$, but also in $\Delta 11$, in which the E3 to E4 ratio is already high due to the lack of the essential polymorphic silencer region. When the same haplotype had a major allele replaced by a minor silencing allele (SNP 3c+4a+6g and SNP 3c+5u+6g) the E3/E4 ratio dropped down proportionately to the silencing effect of the individual minor alleles (SNP 4a and SNP 5u) (Figure 3.24). When a haplotype contained all the four splicing silencing alleles (SNP 4a+5u), the E3 to E4 ratio was further decreased in comparison to SNP 4a or SNP 5u alone in both $\Delta 10$ and $\Delta 11$. When the same haplotype had a major allele replaced by a minor splicing enhancing allele (SNP 3c+4a+5u and SNP 4a+5u+6g) the E3 to E4 ratio surged proportionately to the splicing enhancing effect of the individual minor alleles (SNP 3c and SNP 6g) (Figure 3.24). To summarize, all the alleles showed an additive effect on E3 splicing and the change in splicing effects of all the allele combinations is consistent across $\Delta 10$ and $\Delta 11$.

Figure 3.24 Combinations of major and minor alleles of the four SNPs show additive effect

E3/E4 ratios by real-time RT-PCR for $\Delta 10$ and $\Delta 11$ with different combinations/haplotypes. For each combination, the E3/E4 splicing ratio depends on the ratio of enhancing (blue bars and font) or silencing (red bars and font) alleles. 2:2 results in a near 1 ratio (with one exception i.e. $\Delta 10$ SNP 3c+4a+5u+6g), 3:1 or 4:0 results in higher E3/E4 ratio, and 1:3 or 0:4 in lower E3/E4 ratio. The first bar for each graph

includes the major allele for all four SNPs representing the WT $\Delta 10$ and $\Delta 11$ minigenes normalized to 1. The major alleles are indicated in uppercase and the minor alleles are indicated in lowercase. Experiments were performed in triplicate and data is shown as mean $\pm$ standard error of the mean. Statistical significance is indicated as follows: *, $P < 0.05$; **, $P < 0.01$ compared to cells transfected with the WT minigene using Student's t-test.

Next, using the same set of data, we normalized the E3 to E4 ratio of all the $\Delta 11$ SNP haplotypes to $\Delta 10$. One of the haplotypes i.e. $\Delta 11$ SNP 4a+5u showed almost similar E3/E4 ratio as $\Delta 10$ (Figure 3.25). This indicates that a patient having this haplotype would most likely be sensitive to TKI, despite bearing the deletion polymorphism. On the contrary, a patient having the $\Delta 11$ SNP 3c+6g would be extremely resistant to TKI as seen by the dramatic increase (~ 150 -fold) in E3/E4 ratio (Figure 3.25). We also observed the result of normalizing E3/E4 ratio of $\Delta 10$ SNP haplotypes to $\Delta 11$. It was found that one of the haplotypes i.e. $\Delta 10$ SNP 3c+6g showed increased E3 to E4 ratio in comparison to $\Delta 11$ (~ 3 -fold) (Figure 3.25). This indicates that a patient having this combination of alleles would be definitely resistant to TKI.

Altogether, this analysis shows that other single-nucleotide variations change the *BIM* splicing patterns dictated by the deletion polymorphism, and thus they might modify the TKI responses in CML and EGFR-mutated NSCLC patients.

Figure 3.25 Comparison of 16 SNP haplotypes to reciprocal deletion or nondeletion alleles

E3/E4 ratios for all $\Delta 11$ SNP haplotypes are normalized to $\Delta 10$. E3/E4 ratios for all $\Delta 10$ SNP haplotypes are normalized to $\Delta 11$. E3/E4 ratios have been calculated based on the same set of data used in Figure 3.21.

3.7 Testing the effects of *BIM* E3 SNPs in K562 cell line

The discovery of four novel *BIM* splice-switching annotated SNPs suggested the intriguing possibility that SNPs in the E3 region could result in important differences in the generation of functional BIM protein, and thereby contribute to response heterogeneity in patients treated with TKIs. Hence, next we wanted to test the ability of these SNPs to affect splicing and TKI responses in K562 using the approaches shown in Figure 3.26 and described in the subsections below.

Figure 3.26 Approaches for testing effect of *BIM* SNPS in K562 cell line

Left, CRISPR/Cas9 design for introducing different allelic variants (vertical line in E3) in the genome by using two guide RNAs (above and below E3). Right, alternative approach by which all endogenous *BIM* alleles are removed by CRISPR/Cas9, and *BIM* with different allelic variants is reconstituted by transfection with plasmids (circular construct).

3.7.1 Introduction of SNP minor alleles in endogenous *BIM* by CRISPR editing

Firstly, we attempted to introduce SNP 3c and SNP 4a individually in K562 cells by homology-directed repair (HDR) via the Clustered Regularly Interspaced Short Palindromic Repeats/CRISPR associated protein 9 (CRISPR/Cas9) system. The wild-type Cas9 nuclease can induce DSB at the genomic target when guided by a gRNA. In contrast, the Cas9n can simultaneously create single-stranded nicks on both strands of the target DNA when guided by a pair of appropriately spaced and oriented gRNAs. This results in

DSB resulting from the double nicking (Ran et al., 2013b). We used the D10A nickase mutant of Cas9 (Cas9n) enzyme along with a pair of guide RNAs (gRNAs) offset by 5 bp to target the human *BIM* E3 (Figure 3.23A). Single-stranded nicks introduced by Cas9n on both strands of the *BIM* E3 would result in double-stranded breaks (DSBs). Cas9n increases the specificity of genome editing because DSBs would only occur if both gRNAs are able to bind the target regions within a defined space and any single-stranded nicks created are repaired without small insertions or deletions (indels) formation (Ran et al., 2013b).

The individual guide sequences, one targeting the sense strand while the other targeting the antisense strand of *BIM* E3, were cloned into the pSpCas9 (BB)-2A-GFP vector which expresses the mutated “nickase” version of the Cas9 enzyme. The plasmids obtained after cloning the respective guide sequences are referred as Cas9n/gRNA1 and Cas9n/gRNA2 plasmids respectively. The DSBs from double nicking would then be repaired via homology directed repair (HDR) in presence of a repair template (Ran et al., 2013b). Single-stranded DNA oligonucleotides (ssODNs) were used as repair templates to mediated HDR. ssODNs for SNP 3c and 4a were designed with the respective SNP minor allele positioned in the middle flanked by 60 nt homology arms on each side (Figure 3.27A).

To introduce SNP 3c and SNP 4a individually, K562 cells were then nucleofected with the pair of Cas9n/gRNA1 and Cas9n/gRNA2 plasmids and the respective ssODNs. After nucleofection, cells were sorted by GFP-FACS and single cells were isolated and expanded till confluent. Genotyping of the single-cell clones was done by PCR of extracted genomic DNA (of confluent cells) followed by RFLP analysis. To avoid unwanted detection of residual repair template, the PCR primers used were designed to anneal outside the region spanned by the homology arms. After restriction enzyme digestion of the PCR product, it was observed that several clones had the intended SNP minor alleles introduced as seen by the cleaved PCR product in agarose gel electrophoresis (see representative images in Figure 3.27B: SNP 3c, upper panel and SNP 4a, lower panel). None of the clones showed complete digestion of PCR product which means that not all alleles of a single cell had the substitution. Since K562 is a triploid cell

line, the higher percentage of uncleaved PCR product indicates that only one out of the three alleles had the SNP minor alleles introduced. After sequencing of the PCR product, we observed that all the SNP edited clones indeed had successful monoallelic substitution. However these clones also received additional random mutations due to non-homologous end joining (NHEJ) at the 3'ss of *BIM* E3 (see representative electropherogram Figure 3.27C: SNP 3c, upper panel and SNP 4a, lower panel). Hence, these clones could not be used for further study.

One of the clones, clone 11 for SNP 3c editing, had a biallelic substitution (see electropherogram Figure 3.27C upper panel). By plasmid cloning and sequencing, we observed a 14 nt deletion in this particular clone at the 3'ss. In order to correct this deletion, a 20 nt guide sequence targeting the sequence flanking the deletion in *BIM* E3 was designed and cloned into pSpCas9 (BB)-2A-GFP vector to get a plasmid named, Cas9n/gRNA1-modified. Another 20 nt guide sequence was designed targeting E3 and cloned into pSpCas9 (BB)-2A-GFP vector to get a plasmid named, Cas9n/gRNA3. The clone 11 for SNP 3c editing was then nucleofected with the Cas9n/modified gRNA1 and Cas9n/gRNA3 plasmids and the SNP 3c ssODN (see Figure 3.27D). Genotyping was then done after GFP single cell sorting and expansion. Using this strategy, it was possible to obtain clones (two of four screened clones) that showed successful elimination of the 14 nt deletion introduced in the clone and no further indels were observed in the gRNA target region (see electropherogram Figure 3.27E). Hence these clones had successful biallelic but not triallelic substitution for SNP 3c.

Given the inefficiency and difficulty encountered in introducing nucleotide substitutions by HDR mediated CRISPR genome editing, it would not be the best approach for testing the effect of the *BIM* SNPs in K562. Hence we switched to using another strategy to test the effect of SNPs in K562 (see section 3.7.2) which is currently in progress.

A **B** **C**

Figure 3.27 Introduction of SNP 3c and SNP 4a in endogenous K562 *BIM* by CRISPR

(A) Schematic of gRNAs targeting *BIM* E3 and ssODNs used as repair template to introduce SNP 3c and SNP 4a. Blue letters indicate the 20 nt guide sequence and orange letters highlight the PAM sequences in *BIM* E3. SNP 3 and SNP 4 major alleles have been indicated in red in *BIM* E3 and corresponding minor alleles have been indicated in red in the ssODNs. (B) Agarose electrophoresis of PCR products after RFLP analysis using restriction enzymes Hpy1881 and APO1 for SNP 3c (upper panel) and SNP 4a (lower panel) respectively. K562 and $\Delta 10$ are used as negative controls for restriction enzyme digestion. $\Delta 10$ SNP 3c and $\Delta 10$ SNP 4a are used as positive controls for restriction enzyme digestion for SNP 3c and SNP 4a editing respectively. (C) Representative electropherograms of the direct sequencing results of the PCR amplified *BIM* E3 region for a single cell clone for SNP 3c (upper panel) and SNP 4a (lower panel). The red arrow indicates the overlapping peaks due to the substitution. Additional mutations at 3'ss gRNA1 target site can be seen as continuous overlapping peaks. (D) Schematic of gRNAs targeting *BIM* E3 to eliminate the 14 nt deletion and SNP 3c ssODN. Blue letters indicate the 20 nt guide sequence and orange letters highlight the PAM sequences in *BIM* E3. (E) Electropherograms of the direct sequencing results of the PCR amplified *BIM* E3 region for a single cell clone for SNP 3c (left panel) after repair with the modified gRNA1. The

red arrow indicates the overlapping peaks caused by the substitution. Additional mutations at 3'ss were eliminated.

3.7.2 *BIM* knockout by CRISPR/Cas9 and rescue by *BIM* expression plasmid

Since the CRISPR/Cas9 genome editing approach was not efficient in introducing minor alleles of the SNPs into endogenous *BIM* gene, we used another experimental strategy to study the effect of SNPs on splicing and TKI resistance. In this second approach, we firstly wanted to knockout all endogenous *BIM* genes in K562. This would then be followed by transfecting cells with a *BIM* expression plasmid that would harbor different allelic variants of the SNPs.

To knockout endogenous *BIM* genes in K562, a pair of *BIM* CRISPR gRNAs targeting two different regions in human *BIM* gene was introduced in K562 cells. This would induce Cas9 mediated DSBs at the two sites. This strategy has been previously used whereby the DSBs at the two locations either led to deletion of the intervening DNA segment by precise end joining (Zheng et al., 2014) or indels at the deletion junctions (Canver et al., 2014).

I used different pairs of gRNAs targeting different regions of *BIM* gene to introduce deletions. Firstly, a pair of gRNA targeting two different locations in *BIM* exon 2A (E2A) was tested. Genotyping of the K562 single cell clones by PCR showed deletion in many clones. Sequencing of the PCR product revealed that the deletion of the intervening sequence resulted in precise end joining and this resulted in a 90 bp deletion in *BIM* E2A. Thus there was no frameshift of the resultant mRNA and hence the functional BIM protein was still produced (data not shown).

A second pair of gRNA, one targeting *BIM* E3 and another targeting *BIM* E4 was tested. However none of the single cell clones showed the deletion of the intervening sequence (data not shown). This is possibly because the intended deletion size was large and previously it has been demonstrated that the frequency of deletion is inversely

proportional to the size of deletion (Canver et al., 2014). Hence screening a large number of clones might be required to find a successful knockout clone.

A third pair of gRNA targeting two different locations in *BIM* E2A was tested. Since the first pair of gRNA resulted in deletion of intervening sequence and precise end joining, we designed the gRNA such that the distance between the Cas9 cut sites was 52 bp. Genotyping of the K562 single cell clones revealed that many clones had the deletion (Figure 3.28). Sequencing of the PCR product revealed that the deletion of the intervening sequence resulted in precise end joining which caused a 52 bp deletion in *BIM* E2A. From western blot analysis, most clones showed reduction in BIM protein levels, however none of the clones showed complete absence of BIM protein (data not shown). This was surprising given that the deletion was predicted to cause a frameshift and thus possibly NMD of the mRNA due to presence of a premature stop codon. Currently new gRNAs have been designed to delete exon 4 (E4). Deleting E4 should knockout functional BIM protein as E4 encodes for the proapoptotic BH3 domain.

Figure 3.28 Agarose electrophoresis of PCRs after targeting E2A with two CRISPR gRNAs

Most single cell clones show deletion of the targeted *BIM* E2A region as seen by the truncated PCR fragment in comparison to K562 cells.

The *BIM* expression plasmid has been constructed as described in Section 2.18. In the near future, it will be determined if the *BIM* plasmid recapitulates endogenous *BIM* splicing and reconstitutes all BIM protein isoforms. The plasmid will then be mutated to harbor the minor alleles for the SNPs to test their effects on *BIM* E3 splicing and TKI resistance. However, the clinical relevance of these SNPs recently became questionable (see Discussion).

4 DISCUSSION

A 2.9 kb deletion polymorphism in *BIM* intron 2 (present in 12.3 % of East Asians) leads to the production of *BIM* isoforms lacking the proapoptotic BH3 domain due to the preferential splicing of E3 over E4 encoding BH3 domain. This causes intrinsic TKI resistance in CML and EGFR NSCLC patients (Ng et al., 2012). In this project, we have mapped the *cis*-acting splicing elements within *BIM* E3 and upstream intronic region by deletion analysis in *BIM* splicing minigenes. The ESEs have been validated by point mutation analysis and in the context of a heterologous exon. The identified ESEs could further serve as targets for splice-switching ASOs repressing E3 inclusion. One *trans*-acting splicing factor, SRSF1 was confirmed to promote *BIM* E3 inclusion, while the effects of other splicing factors require further verification. In addition, regulation of *BIM* E3 splicing by varying lengths of polypyrimidine tract at the 3'ss has been determined. Lastly, using *BIM* splicing minigenes, we discovered the ability of four annotated SNPs located in *BIM* E3 to effectively alter the E3/E4 ratios. Moreover the combination of these alleles had an additive effect on E3 splicing. Accordingly, like the 2.9 kb *BIM* deletion, these four SNPs are predicted to modify the TKI responses in patients with CML and EGFR NSCLC.

4.1 *BIM* E3 is regulated by multiple regulatory *cis*-acting splicing elements

Deletion analysis revealed several *cis*-acting splicing elements in *BIM* E3 and upstream intronic region (Figure 3.8). The ESEs or ESSs are not concentrated at a particular region of the exon, rather they are distributed throughout the exon (Figure 3.4). The ESEs probably execute their effects by recruiting splicing activators. The splicing activators most likely antagonize the binding of splicing repressors or cooperate with other positive regulatory factors to facilitate the recognition of 3'ss. On the other hand, the ESSs possibly suppress splicing by recruiting splicing repressors (Chen and Manley, 2009). These repressors perhaps sterically block binding of splicing activators to the enhancer sequences. Splicing inhibition could also be due to the multimerization of repressors along pre-mRNA which prevent the 3'ss recognition by the U2AF35/65. Alternatively, some repressors could engage in protein-protein interactions causing 'looping out' of a

region of E3. This loop formation most likely causes steric hindrance to the recruitment of U2AF35/65 at the 3'ss (Chen and Manley, 2009).

Indeed, the splicing pattern change caused by the deletions and the respective staggered deletions were consistent, except for a few cases. For instance, SD10 and SD11 sequences contain enhancers while D11 does not (Figure 3.4). This discrepancy could be attributed to the junction effect, i.e. removal of D11 sequence probably created a new splicing regulatory site that compensated for the deletion. Another possibility could be that the first half of the SD10 sequence and the second half of the SD11 sequence contained major binding sites for different activators, while D11 contained minor binding site for the corresponding activators. Hence deletion of SD10 and SD11 sequences disrupted binding of the activators and reduced E3 splicing. On the other hand, deletion of D11 sequence caused negligible disruption in the activator binding, thus preventing decrease in E3 splicing.

Our finding that splicing of *BIM* E3 is regulated by multiple ESEs and ESSs is not surprising. In general, alternative exons contain weaker splice sites than constitutive ones. Also alternative exons are shorter, and flanked by longer introns when compared to constitutive exons. These factors limit the recognition of these exons by the spliceosome, thus necessitating splicing regulation by numerous ESEs and ESSs (Kim et al., 2008). In the case of the alternative *BIM* E3, the 3'ss is very strong as predicted by *in silico* tools. This is unusual for alternative exons. Since the E3 containing transcripts are non-apoptotic, numerous ESSs might have evolved to reduce the levels of these transcripts. Yet, the ESEs in E3 probably evolved to modulate the levels of functional proapoptotic E4 containing transcripts. Several ISEs/ISSs were also identified, yet their effects on splicing were not as strong when compared to that of ESEs/ESSs (Figure 3.6). The ISSs are concentrated near the polymorphic region which contains the splicing silencer region. On the other hand, the ISEs are concentrated near the 3'ss and the BPS (-SD1 region) as a positive splicing element which, as expected, showed a very strong enhancer-like effects (Figure 3.6).

The ESEs found in this study could serve as targets for ASOs that can bind to the ESEs and switch splicing from E3 to E4 to increase the production of proapoptotic *BIM* isoforms. Indeed, based on the *cis*-acting elements revealed by deletion analysis, 18-mer ASOs have been tiled throughout the region downstream of the deletion to the end of E3 by Dr Liu Jun (postdoc). This work has been done in partnership with Ionis Pharmaceuticals (Carlsbad, CA, USA). Some of these ASOs have shown promising effects in restoring TKI sensitivity which have not been discussed in this thesis because of confidentiality clauses. These ASOs could be potentially used as adjuvant therapy to resensitize the polymorphism containing CML cells to TKIs.

The consistency in identification of *cis*-acting elements in the context of both deletion and non-deletion minigenes (Figure 3.7) suggests that the regulatory sequences exert their effects on splicing independently of the 2.9 kb intronic deletion region. This also implies that ASOs targeting the enhancers should not only aid in restoring sensitivity in TKI resistant patients (containing the *BIM* deletion polymorphism), but also could boost response to TKIs in TKI sensitive patients (lacking the *BIM* deletion polymorphism).

Given that BIM is an important proapoptotic protein, the *cis*-acting elements in *BIM* E3 elucidated by the deletion analysis could provide useful information in treatment of other diseases. For example, BIM upregulation has been shown to contribute to increased cardiomyocyte and neuronal cell death following ischemia (Ness et al., 2006; Qian et al., 2011). Thus ASOs targeting silencers in E3 could promote splicing of E3 thereby causing reduced expression of exon 4-containing BIM transcripts, preventing cell death.

4.2 Polyadenylation regulatory sequences favor splicing of *BIM* E3

Our finding that the PAS and flanking sequences showed splicing enhancer-like effects (Figure 3.5) are in concordance with previous findings, which have shown that splicing and polyadenylation mutually influence each other (Figure 4.1) (Kyburz et al., 2006; Millevoi et al., 2006; Niwa and Berget, 1991). For example, mutation of the highly conserved polyadenylation signal, AAUAAA, has been found to repress *in vitro* splicing of the terminal intron (Niwa and Berget, 1991). The synergistic stimulation of splicing of last exon and polyadenylation is believed to be achieved by the cooperation between the

splicing and the cleavage/polyadenylation machineries (Kyburz et al., 2006; Millevoi et al., 2006). Interactions between U2 snRNP and the cleavage and polyadenylation specificity factor (CPSF) subunits have been shown to contribute to the coupling of splicing and 3'-end processing (Kyburz et al., 2006). In another study, coordination of *in vitro* splicing and polyadenylation was attributed to the direct interaction between U2AF65 and cleavage factor Im (CFIm 59) (Millevoi et al., 2006). In *BIM* E3, the conserved PAS, AAUAAA is embedded in D20 sequence while the downstream flanking regions are part of other polyadenylation regulatory elements. Therefore decrease in *BIM* E3 splicing was observed when the polyadenylation regulatory elements (SD19 to SD23) were removed (Figure 3.5).

Figure 4.1 Schematic of mutual stimulation between splicing and polyadenylation.

Boxes represent exons; Lines represent introns. Abbreviations: U2AF, U 2 auxiliary factor; CPSF, cleavage and polyadenylation specificity factor ; CstF, cleavage stimulation factor ; PAS, polyadenylation signal; DSE, downstream element.

4.3 Validation of ESEs in a heterologous context

The three ESEs identified in deletion and mutational assays were also tested in a heterologous context (Figure 3.12A). These ESEs, by themselves, can enhance inclusion of the exon in which they were inserted. Consistent with the deletion and point mutation

analyses, ESE2 and ESE3 showed strong effects on exon inclusion while ESE1 weakly enhanced exon inclusion (Figure 3.12B). This suggests that these regions exert their enhancer effects individually without the need of nearby splicing regulatory sequences. Yet, it is possible that these ESEs may have a combinatorial effect on *BIM* E3 inclusion so further study is needed.

The weak enhancer effects of ESE1 region in the heterologous context could be attributed to the junction effect, i.e. insertion of the ESE1 sequence in the heterologous alternative exon probably created a novel silencer region or resulted in formation of secondary structure, thereby restricting splicing (Figure 4.2A and B). Another possibility is that insertion of the ESE1 in the heterologous exon created a steric block for binding of U1 snRNP to the 5'ss. In the context of *BIM* E3, this effect was not observed as *BIM* E3 is a terminal exon and is much longer (Figure 4.2A and B).

Since ESE1 region is located near the 3'ss in *BIM* E3, it could also be likely that it directly or indirectly supports binding of U2AF to the 3'ss, while in the heterologous exon it is not part of the 3'ss (Figure 4.2C and D). Another reason could be that binding of splicing factor (s) to the ESE1 in the context of *BIM* E3, instead of recruiting the spliceosome, abrogated the binding of a splicing repressor at a nearby silencer region, thereby improving exon inclusion. Thus, in this heterologous context (in the absence of those silencers), ESE1 had less impact on splicing (Figure 4.2E and F).

Figure 4.2 Hypotheses for ESE1 as a mild enhancer in heterologous context

See text for details.

4.4 Length of PolyU tract at 3'ss regulates *BIM* E3 splicing

The polypyrimidine tract of the 3'ss is one of the most essential splicing regulatory elements of a mammalian pre-mRNA (Coolidge et al., 1997). The recognition of the 3'ss in introns is achieved by the U2 snRNP auxiliary factor (U2AF) heterodimer, whereby the U2AF65 subunit binds to the polypyrimidine tract and the U2AF35 subunit recognizes the AG dinucleotide at the intron-exon boundary (Graveley et al., 2001). U2AF65 promotes U2 snRNP binding to the BPS of pre-mRNA by interacting with SAP155, a component of U2 snRNP. Therefore binding of U2AF65 at the 3'ss polypyrimidine tract is essential for the first trans-esterification reaction of splicing (Gozani et al., 1998). In introns containing long polypyrimidine tracts, U2AF35 is not required for the binding of U2 snRNP, known as the “AG-independent 3'ss” (Figure 4.3) (Pacheco et al., 2006). However, U2AF35 is essential for stabilizing U2AF65 binding in introns that contain short or degenerate polypyrimidine tracts. This is known as “AG-dependent 3'ss” (Figure 4.3) (Guth et al., 1999).

In a prior study, analyses of artificial mutations in polypyrimidine tract in minigenes suggested that a stretch of 10-15 pyrimidines was sufficient to ensure normal splicing when splicing is AG-independent (Fu et al., 2011). However, in our study, we observed that a stretch of at least 16 uridines was required for maximal *BIM* E3 splicing in the context of both the deletion and non-deletion minigenes (Figure 3.13). Our result is surprising because a previous study done in the context of a *cis*-competition splicing construct reported that in polypyrimidine tracts, uridines are preferred over cytidines, and a stretch of 11 continuous uridines were highly competent in ensuring splicing. Also restricting the number of uridines to six was sufficient to fully support splicing if located in proximity to the 3'ss AG (Coolidge et al., 1997). Surprisingly, in the context of *BIM*, less than 10Us resulted in almost undetectable E3 inclusion (Figure 3.13). Moreover, even a stretch of 12-14 continuous uridines in *BIM* E3 3'ss polypyrimidine tract did not ensure maximal splicing (Figure 3.13), despite the polypyrimidine tract being located close to 3'ss AG. This implies that *BIM* E3 splicing is AG independent, yet for some reason this exon needs an unusually long polyU tract.

A recent study demonstrated that hnRNP C can repress exonization of Alu elements and even exon inclusion by competing with U2AF65 for binding at the polyuridine tracts with a length of 9Us or more (Zarnack et al., 2013). The repression of Alu inclusion mediated by hnRNP C maintains the integrity of transcripts and so loss of hnRNP C could cause disease (Zarnack et al., 2013). Since *BIM* E3 originates from a long interspersed nuclear element (LINE) repeat and has been found to be repressed by hnRNP C (Juan et al., 2014), perhaps hnRNP C might be a competitor for U2AF65 binding at the polyU tract of E3. The increased length of polyU tract required for maximal E3 inclusion might be due to this interaction, granting enough space for U2AF65 to bind to the E3 polyU tract in spite of hnRNP C binding further upstream.

The polymorphic alleles of human *BIM* E3 PolyU tract are annotated as 17Us, 18Us, 19Us, 20Us or 21Us (dbSNP rs748523945 and rs200489083) which are within the optimal splicing range. Targeting ASOs at the polypyrimidine tract is not preferable due to unspecific binding to other polypyrimidine rich sequences in the genome. Yet, specificity could be increased by designing ASOs that bind a few (5-6) uridines of polypyrimidine tract and the flanking upstream or downstream region. Given the long stretch of uridines required for optimal E3 inclusion, these ASOs could potentially block U2AF65 binding and reduce E3 splicing.

Figure 4.3 Schematic showing AG-dependent 3'ss and AG independent 3'ss

Boxes represent exons; Lines represent introns. Abbreviations: U2AF, U 2 auxiliary factor; BPS, branch point sequence; ss; splice site; Y, Pyrimidine; R, Purine.

4.5 Identification of *trans*-acting factors regulating *BIM* E3 splicing

Our study on *trans*-acting factors binding to the ESEs could be of potential clinical relevance as small inhibitory molecules blocking the activity of *trans*-acting factors have been used (Douglas and Wood, 2011; Havens et al., 2013). Inhibiting an activator of *BIM* E3 splicing could promote the expression of E4-containing *BIM* isoforms that have the proapoptotic BH3 domain. Therefore, reduced expression of these activators could restore sensitivity to TKIs in patients with the deletion polymorphism.

4.5.1 Effects of hnRNP A1, hnRNP A2/B1 and PTBP1 on *BIM* E3 inclusion

RNA pull-down revealed binding of the common repressors, hnRNP A1, hnRNP A2/B1 and PTBP1 to *BIM* ESEs by mass-spectrometry and western blotting analysis (Figure 3.14, Figure 3.15). This is not surprising because although typically hnRNPs are repressors, in certain instances they have been found to activate exon inclusion (Huelga et al., 2012). Yet, knockdown assay suggested that hnRNP A2/B1 is likely not a regulator of *BIM* E3

splicing. It is possible that additional activator binding to the ESEs can compensate for the depletion of hnRNP A2/B1. Hence, knocking down hnRNP A2/B1 alone is not sufficient to see any change in E3 inclusion. The knockdown of hnRNP A1 was not effective to draw any conclusion. It is possible that hnRNP A1 is a true activator for *BIM* E3 inclusion. Hence, further studies are required to validate the role of hnRNP A1 on *BIM* E3 splicing by designing siRNAs to achieve efficient knockdowns.

Alternatively, there could be a possibility that these splicing repressors indeed bound to *BIM* ESEs *in vitro* and their interaction was disrupted by mutations. However in cells, these proteins preferentially interact with other RNA sequences which have higher binding affinity and so less amount of proteins are available to bind to *BIM* E3 and regulate splicing. Hence they may not be the major splicing regulators of E3 splicing.

Although PTBP1, another common splicing repressor was found to bind ESE3 WT by RNA-pull down assay, depletion of PTBP1 enhanced E3 inclusion in a prior study suggesting that PTBP1 is a repressor of *BIM* E3 inclusion (Juan et al., 2014). Thus *in vitro*, binding of PTBP1 to ESE3 could be a weak interaction, while in cells this interaction could be disrupted due to binding of other competitor activators to ESE3.

4.5.2 SRSF1 might protect cancer cells from apoptosis by promoting E3 inclusion

SRSF1 has been demonstrated to potentiate splicing of oncogenic mRNA isoforms. (Anczukow et al., 2012; Anczuków et al., 2015; Ghigna et al., 2005; Karni et al., 2007). In a previous study, SRSF1 overexpression led to the generation of non-apoptotic E3 containing BIM γ 1 and BIM γ 2 isoforms in breast cancers. This caused a reduction in the production of proapoptotic BIMEL, BIML and BIMS isoforms (Anczukow et al., 2012). In our study, SRSF1 overexpression in K562 cells increased endogenous *BIM* E3-containing transcripts (Figure 3.16C) which is in concordance with the previous findings (Anczukow et al., 2012). We also found SRSF1 to increase inclusion of *BIM* E3 in the context of the Δ 10 and Δ 11 minigenes (Figure 3.17C). The increase in E3 containing transcripts in Δ 11 was much lower in comparison to Δ 10 (Figure 3.17C). This is expected because Δ 11 lacking the polymorphic fragment already has a high level of E3 containing transcripts. Although overexpression assay suggested that SRSF1 enhanced *BIM* E3

inclusion, no binding was detected with ESE 2/3 in the RNA pull-down assay. One possibility is that SRSF1 interacts with ISE upstream of E3 to promote its inclusion. Alternatively, SRSF1 may enhance E3 inclusion by cross-regulating the expression of other master splicing regulators that control *BIM* E3 splicing.

SRSF1 was able to increase the E3/E4 ratio in the $\Delta 10$ minigene (containing the polymorphic silencer region) to a similar level as seen in $\Delta 11$ minigene (lacking the polymorphic region). Additionally, SRSF1 overexpression further enhanced the usage of E3 in the absence of polymorphic region ($\Delta 11$ minigene) (Figure 3.17C). These results imply that SRSF1 can neutralize the silencing effect of the polymorphic region and has a very strong enhancing effect on E3 inclusion. Therefore, inhibiting SRSF1 expression, which reduces E3-containing isoforms thereby increasing production of E4-containing transcripts, could be applied to sensitize the primary TKI resistant patients to the treatment. Concurrently, SRSF1 inhibition would also cause a reduction in production of oncogenic isoforms of *RON* (Ghigna et al., 2005), *MNK2* and *S6K1* (Karni et al., 2007). SRSF1, however is not only an essential splicing factor but is also involved in other cellular processes such as NMD, translation regulation and miRNA biogenesis (Das and Krainer, 2014), hence depleting SRSF1 might have detrimental effects. Alternatively, blocking the binding of SRSF1 to *BIM* using ASOs could be a better way to repress the E3 inclusion and redirect splicing to E4. And so, identifying the SRSF1 interacting region is needed and will be studied in the future.

4.5.3 SRSF3 may enhance apoptosis via *BIM* splicing changes

SRSF3, the smallest member of the SR family (Zahler et al., 1992) is a well-known proto-oncogene important for tumor initiation, proliferation and maintenance. Upregulation of SRSF3 is found in many cancers (Jia et al., 2010) such as epithelial ovarian cancer (Iborra et al., 2013). In our study, we found that knockdown of SRSF3 led to an increase in endogenous *BIM* E3-containing transcripts (Figure 3.19). This is surprising because in a previous study, it was demonstrated that substantial knockdown of SRSF3 promoted apoptosis of ovarian cancer cells (He et al., 2011). Therefore, our finding of SRSF3 as a likely repressor of *BIM* E3 inclusion, decreasing the non-apoptotic *BIM* isoforms was initially counter-intuitive. To add on, endogenous *BIM* E3 containing transcript levels

remained high despite rescue of SRSF3 in SRSF3-depleted K562 cells (Figure 3.20). These findings could be explained by two possibilities. Firstly, the SRSF3 dsRNA used in our study could have had unspecific effects thereby causing reduction in the levels of a different splicing repressor of *BIM* E3. Another possibility is that SRSF3 showed its repressor effects on E3 indirectly by regulating the expression of another splicing factor, as mutual cross-regulation of expression is a known phenomenon amongst several *trans*-acting factors (Jangi and Sharp, 2014; Spellman et al., 2007). In a recent study, SRSF3 was shown to regulate the expression of another splicing factor, polypyrimidine tract binding protein 1 (PTBP1). siRNA mediated depletion of SRSF3 suppressed the expression of PTBP1 in Human Embryonic Kidney 293T and an oral adenosquamous carcinoma cell line, CAL-27. However, overexpression of SRSF3 did not promote the expression of PTBP1 (Guo et al., 2015). PTBP1 is a well-known splicing repressor that was previously demonstrated to reduce the inclusion of *BIM* E3 (Juan et al., 2014). Thus, knockdown of SRSF3 in our study could have increased the inclusion of *BIM* E3-containing transcripts by inhibiting the expression of PTBP1. Yet, endogenous *BIM* E3-containing transcripts remained high after rescue with exogenous SRSF3 most likely due to inability of SRSF3 to enhance expression of PTBP1.

4.6 Effect of SNPs on *BIM* E3 splicing in minigene context

After the discovery of the *BIM* deletion polymorphism in 2012 (Ng et al., 2012), several other groups from Taiwan, China, and Japan found a correlation between the polymorphism and TKI resistance (Isobe et al., 2014; Lee et al., 2015; Ma et al., 2015; Nie et al., 2015; Zhao et al., 2014), except for two South Korean centres (Lee et al., 2013; Lee et al., 2015). This suggests that inter-ethnic differences between East Asians could be due to the existence of other splice-switching polymorphisms in *BIM*, conferring heterogeneity in TKI responses. This intrigued us to study the effect of SNPs on E3 splicing. SNPs have been linked to splicing alterations contributing to phenotypic heterogeneity in terms of disease susceptibility, disease severity, response to drug treatment or drug toxicity effects (Lu et al., 2012).

Within the 2.9 kb deletion polymorphic region, at least 13 SNPs were found using the dbSNP database. Multiple redundant *cis*-acting elements have been found to be located in the 2.9 kb silencer fragment. Thus, it is very unlikely that the SNPs within the 2.9Kbp region would change *BIM* E3 splicing (Juan et al., 2014). However, the existence of numerous *cis*-acting elements in E3 and upstream intron from our detailed deletion analysis suggested that minute deviations in these sequences could potentially affect splicing. Hence we studied the effect of 7 SNPs in E3 and upstream intronic region that were found in the dbSNP database, in May 2014. From our study, we found that the minor alleles for four out of seven SNPs individually and potentially altered splicing of E3 versus E4 in the context of the *BIM* splicing minigenes. Minor alleles of two SNPs (SNP 3c and SNP 6g) that increased E3 splicing could be associated with TKI resistance (Figure 3.23). On the other hand, minor alleles for two SNPs (SNP 4a and SNP 5u) that reduced E3 levels could augment sensitivity to TKIs (Figure 3.23). When the combination of all different splice switching alleles was introduced in both the $\Delta 10$ and $\Delta 11$ minigenes, it was observed that the splicing phenotype of each haplotype was the result of adding up the splicing effects of the individual alleles (Figure 3.24). In brief, all the alleles had a summative effect on E3 splicing. We found that, for some SNP combinations, the effect on the E3/E4 ratio was as influential as that of the *BIM* deletion polymorphism. For example, in the haplotype, SNP 3c+6g, splicing of E3 was drastically increased (Figure 3.24). Having this allele combination would undeniably confer TKI resistance in a patient as when E3/E4 ratio of $\Delta 10$ SNP 3c+6g was normalized to $\Delta 11$, this haplotype showed a 3 fold increase in E3/E4 ratio (Figure 3.25). In contrast, the haplotype SNP 4a+5u resulted in a drastic reduction in E3 splicing. Having this haplotype would certainly boost sensitivity to TKI and require a lower dose (Figure 3.24). A patient with the deletion polymorphism having SNP 4a+5u combination, most probably would be sensitive to TKI as seen by the almost similar E3/E4 ratio of this haplotype when compared to $\Delta 10$ (Figure 3.25).

Our findings from SNP analysis of *BIM* E3 indicate that small sequence variations in the genome could play a major role in changing the splicing patterns and subsequent phenotype. Indeed the effect of such nucleotide variations on splicing was recently studied by a group whereby the outcome of almost all single and double mutations in the

alternatively spliced exon 6 of Fas/CD95 death receptor was analyzed. It was observed that single mutations at >90% of exonic bases changed inclusion signifying that almost every nucleotide in that exon contains splicing regulatory information (Julien et al., 2016). Interestingly it was found that 74% of single nucleotide mutations that had neutral effects on splicing, caused an unexpected outcome on splicing when combined with another single nucleotide mutation (Julien et al., 2016). However in comparison to this study, we did not observe any such unexpected outcomes or epistasis in our study of combination alleles of the *BIM* E3 SNPs.

Most likely the *BIM* E3 SNPs showed their effects on splicing by changing the binding affinity of a splicing factor. The effect of SNP 6g on E3 splicing was consistent with that observed from deletion analysis. Introduction of SNP 6g most probably caused loss of binding for a splicing repressor thereby increasing E3 inclusion (Figure 4.4). According to the deletion scan analysis, SNP 3 and SNP 4 were located in regions with neutral effects on splicing (Figure 4.4). Introduction of SNP 3a most probably created a binding site for a splicing activator thereby causing increase in E3 inclusion. On the other hand, introduction of SNP 4a, most probably created a binding site for a splicing repressor thereby decreasing E3 inclusion. The effect of SNP 5u on E3 splicing was contrary to that observed from the deletion analysis (Figure 4.4). This could be due to two possibilities. The introduction of SNP 5u, probably instead of disrupting the binding of a splicing repressor, further increased binding affinity of the splicing repressor. Another possibility is that introduction of SNP 5u eliminated binding of a competitor splicing activator thereby reducing E3 splicing.

Figure 4.4 Schematic showing SNPs positions in map of enhancers and silencers

SNP 3 and SNP 4 are located in neutral regions while SNP 5 and SNP 6 are located in ESS region. Sequences in blue indicate enhancers, while sequences in red indicate silencers in *BIM* E3 and upstream intronic region. Stop codon is indicated with a traffic stop sign. The conserved PAS has been highlighted. Exonic sequence is shown in upper case while intron is shown in lower case.

4.7 Testing the effect of *BIM* E3 SNPs by CRISPR genome editing in K562 cells

Since each of the four SNPs (minor alleles) in *BIM* E3 independently caused significant change in E3 splicing, we hypothesized that CML or EGFR NSCLC patients having these SNPs could show varied responses to TKI treatment. If this hypothesis was correct, these *BIM* sequence variants could function as direct genetic modifiers of the TKI response. Consequently, they could be ideal biomarkers for predicting both upfront TKI resistance, and sensitivity to ASO-based combination therapies. Hence, as a next step we intended to test the effects of these SNPs on splicing and TKI resistance in K562 cell line and we are

at the very initial stage. We first attempted to introduce the SNP 3c and SNP 4a in K562 using CRISPR/Cas9 genome editing.

CRISPR-Cas is part of the bacterial adaptive immune system against viral or plasmid genetic elements (Bhaya et al., 2011). Three types (I-III) of CRISPR systems are known to be present across different bacterial hosts (Makarova et al., 2011). The type II CRISPR/Cas system which is derived from *Streptococcus pyogenes* is so far the simplest and the best adapted for genome editing. It uses the CRISPR-associated protein 9 which is an RNA-guided endonuclease (RGEN) that can catalyze site-specific cleavage of target DNA sequences (Garneau et al., 2010). Cas9 can be targeted to specific genomic sequences by a 20 nt guide sequence within an associated CRISPR RNA (crRNA). crRNA hybridizes with a trans-activating crRNA (tracrRNA) that recruits the crRNA into the Cas9 complex (Garneau et al., 2010; Sander and Joung, 2014). The 20 nt guide sequence binds to the target DNA via RNA-DNA complementary base-pairing. Cleavage of the target DNA site occurs if it is immediately upstream of a protospacer-adjacent motif (PAM), a three nucleotide NGG (Jinek et al., 2012; Sander and Joung, 2014).

The CRISPR/Cas9 system has been reconstituted for site-specific genome editing in diverse cell types and organisms whereby the crRNA and tracrRNA are fused together to form a single - guide RNA (sgRNA). This sgRNA complexes with Cas9 which can introduce DSB at a targeted genomic locus upon base pairing of guide sequence with target DNA sequence if it is 5' of the NGG sequence (Jiang et al., 2013; Jinek et al., 2012; Pattanayak et al., 2013). This DSB is then repaired through either (a) NHEJ which results in small insertion or deletion (indel) mutations, or (b) HDR process in the presence of a homologous repair donor to generate precise modifications (Hsu et al., 2014).

In our intensive attempt to introduce SNP 3c and SNP 4a minor alleles in *BIM E3* in K562 using HDR, although we succeeded in introducing the minor alleles, there were additional mutations observed in the other alleles. This observation is in agreement with previous studies that have shown CRISPR/Cas 9 editing by HDR to be less efficient and being compromised by additional insertions/deletions in the same gene (Inui et al., 2014; Kang et al., 2016). In one of the biallelic SNP 3c clones, it was also possible to eliminate the

unwanted deletion due to NHEJ. This was done by designing a 20 nt guide sequence targeting the sequence flanking the deletion in *BIM* E3 followed by HDR. Nevertheless, due to experienced unpredictability of this approach we decided to change our strategy. Indeed CRISPR/Cas9 has been successfully used for knocking out genes through NHEJ, yet its use for knocking-in specific mutations requires further optimization (Cong et al., 2013; Wang et al., 2013). Thus we sought to switch our approach to knocking out endogenous *BIM* gene by CRISPR, and transfecting K562 cells with *BIM* expression plasmid harboring different allelic variants of the SNPs.

Our study of the *BIM* SNPs was initially clinically relevant as four *BIM* E3 SNPs showed significant change in E3 splicing, and the 16 haplotype combinations showed additive effects on E3 splicing. However, the SNPs 1-6 have been dropped from the UCSC genome browser because they have no population data, probably indicating that these are no real SNPs but rare variants or sequencing errors. Also, SNP 3 and SNP 6 are absent in 1000 genomes while SNP 4 is common in Africans and SNP 5 is found in 1 Asian. Very recently, genotyping of > 400 chromosomes from patients and other individuals by our clinician collaborators did not detect the presence of any of the SNP minor alleles, except for the SNP 7. Hence, it seems that these are not actual SNPs except SNP 4 and SNP 7.

Nevertheless, the discovery of the four novel single nucleotide variations is still relevant as these nucleotide substitutions could occur as somatic variations leading to acquired resistance in *BIM* dependent TKI-mediated kinase driven cancers. Our study also opens the possibility for many other single-nucleotide substitutions within *BIM* E3 that may modify the TKI responses. Also, these may constitute rare variants or be present only in other ethnicities. Hence, the study of these initially annotated SNPs sheds light on how single nucleotide substitutions can lead to phenotypic variations via influencing splicing.

5 CONCLUSION AND FUTURE DIRECTIONS

The therapeutic limitations of TKIs associated with an aberrant *BIM* splicing event in CML and EGFR NSCLC patients, led us to study the regulation of splicing of *BIM* E3. In my PhD project, a map of ESEs, ESSs, ISEs and ISSs in *BIM* E3 and upstream intron has been elucidated by deletion scan analysis in *BIM* minigenes. The ESEs which are potential targets for splice-switching ASOs were further validated by point mutation and heterologous context analysis. Apart from being regulated by the auxiliary *cis*-acting elements, *BIM* E3 splicing was also affected by variations in lengths of polyU tract. Surprisingly, I found that an unusually long polypyrimidine tract of 16Us is required for maximal E3 inclusion, which is in contradiction to previous studies. Further, I found that the minor alleles for four annotated SNPs in *BIM* E3 individually and potentially affected E3 splicing in the context of *BIM* minigenes. And the combination of different alleles of each SNP showed additive effects on E3 splicing. Our findings reveal that these SNPs could affect splicing and TKI responses.

Despite the identification of multiple *cis*-acting elements regulating *BIM* E3 splicing, a more intensive characterization of the *trans*-acting factors binding to these elements would be of potential interest. Our study confirmed that SRSF1 indeed promotes inclusion of *BIM* E3. In future, the actual SRSF1 binding motifs needs to be determined which can be done by combination of deletion/point mutation minigenes and SRSF1 overexpression/knockdown. If there is no change in minigene E3 splicing, it could indicate that the deleted or mutated sequence contains the SRSF1 binding as E3 can no longer respond to the change in SRSF1 expression level. The actual binding site can be further validated by electrophoretic mobility shift assay (EMSA) using full length and truncated E3 with upstream intron. In addition, *in vitro* splicing and spliceosome assembly assay could be carried out to examine which stage of splicing is inhibited by SRSF1. Based on the preliminary results, the possibility of cross-regulation between SRSF3 and PTBP1 (a known repressor of *BIM* E3 inclusion) was hypothesized. To test this hypothesis, the effects of SRSF3 knockdown and overexpression on PTBP1 mRNA and protein levels can be determined. Also to ensure that the knockdown effects of the dsRNAs on SRSF3 is direct, multiple SRSF3 dsRNAs can be designed and tested to

prevent off-target effects bias. Ideally three different dsRNAs against our target of interest can confirm the results obtained if they produce the same results. The use of multiple dsRNAs can also increase the chance of getting a good knockdown of the splicing factors, hnRNP A1 and SRSF5.

Although four different annotated SNPs in *BIM* E3 have shown to change E3 splicing in the context of *BIM* splicing minigenes, their effect on E3 splicing and TKI response needs to be tested in cell lines. This will be done by knocking out all endogenous *BIM* genes in K562 by CRISPR/Cas9, and transfecting cells with *BIM* expression plasmid that harbour different allelic variants. Using RNA pulldowns, overexpression and knockdown experiments, the *trans*-acting factors that regulate splicing via the *cis*-acting splicing elements affected by the SNPs can be identified.

Overall, this thesis greatly expands our knowledge of *cis*-acting elements and *trans*-acting factors that regulate *BIM* E3 splicing, by identifying for the first time the *cis*-acting elements within and around this exon. We also learned that single-nucleotide variants can have a significant effect on *BIM* E3 splicing, and the combinations of alleles have an additive effect. Our findings may provide useful insights for potential therapies that manipulate *BIM* E3 splicing in order to attain beneficial outcomes in the context of tyrosine kinase-driven cancers and perhaps other pathologies.

6 REFERENCES

- Abdul-Manan, N., Williams, K.R., 1996. hnRNP A1 Binds Promiscuously to Oligoribonucleotides: Utilization of Random and Homo-Oligonucleotides to Discriminate Sequence from Base-Specific Binding. *Nucleic Acids Research* 24, 4063-4070.
- Adachi, M., Zhao, X., Imai, K., 2005. Nomenclature of dynein light chain-linked BH3-only protein Bim isoforms. *Cell Death and Differentiation* 12, 690-690.
- Adams, J., Cory, S., 2007. The Bcl-2 apoptotic switch in cancer development and therapy. *Oncogene* 26, 1324-1337.
- Adams, J.M., Cory, S., 1998. The Bcl-2 protein family: arbiters of cell survival. *Science* 281, 1322-1326.
- Aichberger, K.J., Mayerhofer, M., Krauth, M.T., Vales, A., Kondo, R., Derdak, S., Pickl, W.F., Selzer, E., Deininger, M., Druker, B.J., Sillaber, C., Esterbauer, H., Valent, P., 2005. Low-level expression of proapoptotic Bcl-2-interacting mediator in leukemic cells in patients with chronic myeloid leukemia: role of BCR/ABL, characterization of underlying signaling pathways, and reexpression by novel pharmacologic compounds. *Cancer Research* 65, 9436-9444.
- Akerman, M., David-Eden, H., Pinter, R.Y., Mandel-Gutfreund, Y., 2009. A computational approach for genome-wide mapping of splicing factor binding sites. *Genome Biology* 10, R30.
- Akiyama, T., Dass, C.R., Choong, P.F., 2009. Bim-targeted cancer therapy: a link between drug action and underlying molecular changes. *Molecular Cancer Therapeutics* 8, 3173-3180.
- An, X., Tiwari, A.K., Sun, Y., Ding, P.R., Ashby, C.R., Jr., Chen, Z.S., 2010. BCR-ABL tyrosine kinase inhibitors in the treatment of Philadelphia chromosome positive chronic myeloid leukemia: a review. *Leukemia Research* 34, 1255-1268.
- Anczukow, O., Rosenberg, A.Z., Akerman, M., Das, S., Zhan, L., Karni, R., Muthuswamy, S.K., Krainer, A.R., 2012. The splicing factor SRSF1 regulates apoptosis and proliferation to promote mammary epithelial cell transformation. *Nature Structural and Molecular Biology* 19, 220-228.
- Anczuków, O., Akerman, M., Cléry, A., Wu, J., Shen, C., Shirole, Nitin H., Raimer, A., Sun, S., Jensen, Mads A., Hua, Y., Allain, Frédéric H.T., Krainer, Adrian R., 2015. SRSF1-Regulated Alternative Splicing in Breast Cancer. *Molecular Cell* 60, 105-117.
- Anderson, S.L., Coli, R., Daly, I.W., Kichula, E.A., Rork, M.J., Volpi, S.A., Ekstein, J., Rubin, B.Y., 2001. Familial dysautonomia is caused by mutations of the IKAP gene. *The American Journal of Human Genetics* 68, 753-758.

- Andrés, A.M., Dennis, M.Y., Kretzschmar, W.W., Cannons, J.L., Lee-Lin, S.-Q., Hurle, B., Schwartzberg, P.L., Williamson, S.H., Bustamante, C.D., Nielsen, R., 2010. Balancing selection maintains a form of ERAP2 that undergoes nonsense-mediated decay and affects antigen presentation. *PLoS Genetics* 6, e1001157.
- Antonicelli, A., Cafarotti, S., Indini, A., Galli, A., Russo, A., Cesario, A., Lococo, F.M., Russo, P., Mainini, A.F., Bonifati, L.G., Nosotti, M., Santambrogio, L., Margaritora, S., Granone, P.M., Dutly, A.E., 2013. EGFR-targeted therapy for non-small cell lung cancer: focus on EGFR oncogenic mutation. *International Journal of Medical Sciences* 10, 320-330.
- Ast, G., 2004. How did alternative splicing evolve? *Nature Reviews Genetics* 5, 773-782.
- Bakhshi, A., Jensen, J.P., Goldman, P., Wright, J.J., McBride, O.W., Epstein, A.L., Korsmeyer, S.J., 1985. Cloning the chromosomal breakpoint of t(14;18) human lymphomas: clustering around JH on chromosome 14 and near a transcriptional unit on 18. *Cell* 41, 899-906.
- Bartram, C.R., Deklein, A., Hagemeijer, A., Vanagthoven, T., Vankessel, A.G., Bootsma, D., Grosveld, G., Fergusonsmith, M.A., Davies, T., Stone, M., Heisterkamp, N., Stephenson, J.R., Groffen, J., 1983. Translocation of C-Abl Oncogene Correlates with the Presence of a Philadelphia-Chromosome in Chronic Myelocytic-Leukemia. *Nature* 306, 277-280.
- Bauman, J.A., Li, S.D., Yang, A., Huang, L., Kole, R., 2010. Anti-tumor activity of splice-switching oligonucleotides. *Nucleic Acids Research* 38, 8348-8356.
- Bell, D.W., Lynch, T.J., Haserlat, S.M., Harris, P.L., Okimoto, R.A., Brannigan, B.W., Sgroi, D.C., Muir, B., Riemenschneider, M.J., Iacona, R.B., Krebs, A.D., Johnson, D.H., Giaccone, G., Herbst, R.S., Manegold, C., Fukuoka, M., Kris, M.G., Baselga, J., Ochs, J.S., Haber, D.A., 2005. Epidermal growth factor receptor mutations and gene amplification in non-small-cell lung cancer: molecular analysis of the IDEAL/INTACT gefitinib trials. *Journal of Clinical Oncology* 23, 8081-8092.
- Bennett, C.F., Swayze, E.E., 2010. RNA targeting therapeutics: molecular mechanisms of antisense oligonucleotides as a therapeutic platform. *Annual Review of Pharmacology and Toxicology* 50, 259-293.
- Berget, S.M., 1995. Exon recognition in vertebrate splicing. *The Journal of Biological Chemistry* 270, 2411-2414.
- Berget, S.M., Moore, C., Sharp, P.A., 1977. Spliced segments at the 5' terminus of adenovirus 2 late mRNA. *Proceedings of the National Academy of Sciences* 74, 3171-3175.
- Bhamidipati, P.K., Kantarjian, H., Cortes, J., Cornelison, A.M., Jabbour, E., 2013. Management of imatinib-resistant patients with chronic myeloid leukemia. *Therapeutic Advances in Hematology* 4, 103-117.

- Bhaya, D., Davison, M., Barrangou, R., 2011. CRISPR-Cas systems in bacteria and archaea: versatile small RNAs for adaptive defense and regulation. *Annual Reviews of Genetics* 45, 273-297.
- Black, D.L., 2003. Mechanisms of alternative pre-messenger RNA splicing. *Annual Review of Biochemistry* 72, 291-336.
- Bonnevie-Nielsen, V., Field, L.L., Lu, S., Zheng, D.-J., Li, M., Martensen, P.M., Nielsen, T.B., Beck-Nielsen, H., Lau, Y.-L., Pociot, F., 2005. Variation in antiviral 2' , 5' - oligoadenylate synthetase (2' 5' AS) enzyme activity is controlled by a single-nucleotide polymorphism at a splice-acceptor site in the OAS1 gene. *The American Journal of Human Genetics* 76, 623-633.
- Bouillet, P., Metcalf, D., Huang, D.C., Tarlinton, D.M., Kay, T.W., Kontgen, F., Adams, J.M., Strasser, A., 1999. Proapoptotic Bcl-2 relative Bim required for certain apoptotic responses, leukocyte homeostasis, and to preclude autoimmunity. *Science* 286, 1735-1738.
- Bouillet, P., Zhang, L.C., Huang, D.C., Webb, G.C., Bottema, C.D., Shore, P., Eyre, H.J., Sutherland, G.R., Adams, J.M., 2001. Gene structure alternative splicing, and chromosomal localization of pro-apoptotic Bcl-2 relative Bim. *Mammalian Genome* 12, 163-168.
- Bouillet, P., Purton, J.F., Godfrey, D.I., Zhang, L.C., Coultas, L., Puthalakath, H., Pellegrini, M., Cory, S., Adams, J.M., Strasser, A., 2002. BH3-only Bcl-2 family member Bim is required for apoptosis of autoreactive thymocytes. *Nature* 415, 922-926.
- Braunschweig, U., Gueroussov, S., Plocik, A.M., Graveley, B.R., Blencowe, B.J., 2013. Dynamic integration of splicing within gene regulatory pathways. *Cell* 152, 1252-1269.
- Canver, M.C., Bauer, D.E., Dass, A., Yien, Y.Y., Chung, J., Masuda, T., Maeda, T., Paw, B.H., Orkin, S.H., 2014. Characterization of genomic deletion efficiency mediated by clustered regularly interspaced palindromic repeats (CRISPR)/Cas9 nuclease system in mammalian cells. *The Journal of Biological Chemistry* 289, 21312-21324.
- Cartegni, L., Chew, S.L., Krainer, A.R., 2002. Listening to silence and understanding nonsense: exonic mutations that affect splicing. *Nature Reviews Genetics* 3, 285-298.
- Charlet, B.N., Savkur, R.S., Singh, G., Philips, A.V., Grice, E.A., Cooper, T.A., 2002. Loss of the muscle-specific chloride channel in type 1 myotonic dystrophy due to misregulated alternative splicing. *Molecular Cell* 10, 45-53.
- Chen, M., Manley, J.L., 2009. Mechanisms of alternative splicing regulation: insights from molecular and genomics approaches. *Nature Reviews. Molecular Cell Biology* 10, 741-754.
- Chen, Z., Fillmore, C.M., Hammerman, P.S., Kim, C.F., Wong, K.K., 2014. Non-small-cell lung cancers: a heterogeneous set of diseases. *Nature Reviews Cancer* 14, 535-546.

- Chipuk, J.E., Moldoveanu, T., Llambi, F., Parsons, M.J., Green, D.R., 2010. The BCL-2 family reunion. *Molecular Cell* 37, 299-310.
- Cho, D.H., Tapscott, S.J., 2007. Myotonic dystrophy: emerging mechanisms for DM1 and DM2. *Biochimica et Biophysica Acta* 1772, 195-204.
- Chow, L.T., Gelinas, R.E., Broker, T.R., Roberts, R.J., 1977. An amazing sequence arrangement at the 5' ends of adenovirus 2 messenger RNA. *Cell* 12, 1-8.
- Cilloni, D., Saglio, G., 2012. Molecular Pathways: BCR-ABL. *Clinical Cancer Research* 18, 930-937.
- Close, P., Hawkes, N., Cornez, I., Creppe, C., Lambert, C.A., Rogister, B., Siebenlist, U., Merville, M.-P., Slaugenhaupt, S.A., Bours, V., 2006. Transcription impairment and cell migration defects in elongator-depleted cells: implication for familial dysautonomia. *Molecular Cell* 22, 521-531.
- Cong, L., Ran, F.A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P.D., Wu, X., Jiang, W., Marraffini, L.A., Zhang, F., 2013. Multiplex genome engineering using CRISPR/Cas systems. *Science* 339, 819-823.
- Coolidge, C.J., Seely, R.J., Patton, J.G., 1997. Functional analysis of the polypyrimidine tract in pre-mRNA splicing. *Nucleic Acids Research* 25, 888-896.
- Costa, D.B., Halmos, B., Kumar, A., Schumer, S.T., Huberman, M.S., Boggon, T.J., Tenen, D.G., Kobayashi, S., 2007. BIM mediates EGFR tyrosine kinase inhibitor-induced apoptosis in lung cancers with oncogenic EGFR mutations. *PLoS Medicine* 4, 1669-1679; discussion 1680.
- Coulter, L.R., Landree, M.A., Cooper, T.A., 1997. Identification of a new class of exonic splicing enhancers by in vivo selection. *Molecular and Cellular Biology* 17, 2143-2150.
- Daley, G.Q., Van Etten, R.A., Baltimore, D., 1990. Induction of chronic myelogenous leukemia in mice by the P210bcr/abl gene of the Philadelphia chromosome. *Science* 247, 824-830.
- Das, S., Krainer, A.R., 2014. Emerging functions of SRSF1, splicing factor and oncoprotein, in RNA metabolism and cancer. *Molecular Cancer Research* 12, 1195-1204.
- De Conti, L., Baralle, M., Buratti, E., 2013. Exon and intron definition in pre-mRNA splicing. *Wiley Interdisciplinary Reviews. RNA* 4, 49-60.
- Deininger, M., 2009. 1126 International Randomized Study of Interferon Vs STI571 (IRIS) 8-Year Follow up: Sustained Survival and Low Risk for Progression or Events in Patients with Newly Diagnosed Chronic Myeloid Leukemia in Chronic Phase (CML-CP) Treated with Imatinib Clinically Relevant Abstract.

- Deininger, M.W., Druker, B.J., 2003. Specific targeted therapy of chronic myelogenous leukemia with imatinib. *Pharmacological Reviews* 55, 401-423.
- Deininger, M.W., Goldman, J.M., Melo, J.V., 2000. The molecular biology of chronic myeloid leukemia. *Blood* 96, 3343-3356.
- Desmet, F.O., Hamroun, D., Lalande, M., Collod-Beroud, G., Claustres, M., Beroud, C., 2009. Human Splicing Finder: an online bioinformatics tool to predict splicing signals. *Nucleic Acids Research* 37, e67.
- Dewson, G., Kluck, R.M., 2010. Bcl-2 family-regulated apoptosis in health and disease. *Cell Health and Cytoskeleton* 2, 9-22.
- Douglas, A.G., Wood, M.J., 2011. RNA splicing: disease and therapy. *Briefings in Functional Genomics* 10, 151-164.
- Douglas, A.G., Wood, M.J., 2013. Splicing therapy for neuromuscular disease. *Molecular and Cellular Neuroscience* 56, 169-185.
- Egle, A., Harris, A.W., Bouillet, P., Cory, S., 2004. Bim is a suppressor of Myc-induced mouse B cell leukemia. *Proceedings of the National Academy of Sciences* 101, 6164-6169.
- El Awady, M.K., Anany, M.A., Esmat, G., Zayed, N., Tabll, A.A., Helmy, A., Zayady, E., Rahman, A., Abdalla, M.S., Sharada, H.M., 2011. Single nucleotide polymorphism at exon 7 splice acceptor site of OAS1 gene determines response of hepatitis C virus patients to interferon therapy. *Journal of Gastroenterology and Hepatology* 26, 843-850.
- Elmore, S., 2007. Apoptosis: a review of programmed cell death. *Toxicologic Pathology* 35, 495-516.
- Enders, A., Bouillet, P., Puthalakath, H., Xu, Y., Tarlinton, D.M., Strasser, A., 2003. Loss of the pro-apoptotic BH3-only Bcl-2 family member Bim inhibits BCR stimulation-induced apoptosis and deletion of autoreactive B cells. *The Journal of Experimental Medicine* 198, 1119-1126.
- Epstein, F.H., Faderl, S., Talpaz, M., Estrov, Z., O'Brien, S., Kurzrock, R., Kantarjian, H.M., 1999. The biology of chronic myeloid leukemia. *New England Journal of Medicine* 341, 164-172.
- Essafi, A., Fernandez de Mattos, S., Hassen, Y.A., Soeiro, I., Mufti, G.J., Thomas, N.S., Medema, R.H., Lam, E.W., 2005. Direct transcriptional regulation of Bim by FoxO3a mediates STI571-induced apoptosis in Bcr-Abl-expressing cells. *Oncogene* 24, 2317-2329.
- Faber, A.C., Corcoran, R.B., Ebi, H., Sequist, L.V., Waltman, B.A., Chung, E., Incio, J., Digumarthy, S.R., Pollack, S.F., Song, Y., 2011. BIM expression in treatment-naive cancers predicts responsiveness to kinase inhibitors. *Cancer Discovery* 1, 352-365.

- Faustino, N.A., Cooper, T.A., 2003. Pre-mRNA splicing and human disease. *Genes and Development* 17, 419-437.
- Feero, W.G., Guttmacher, A.E., Wang, L., McLeod, H.L., Weinshilboum, R.M., 2011. Genomics and Drug Response. *New England Journal of Medicine* 364, 1144-1153.
- Fischer, S.F., Bouillet, P., O'Donnell, K., Light, A., Tarlinton, D.M., Strasser, A., 2007. Proapoptotic BH3-only protein Bim is essential for developmentally programmed death of germinal center-derived memory B cells and antibody-forming cells. *Blood* 110, 3978-3984.
- Fu, Y., Masuda, A., Ito, M., Shinmi, J., Ohno, K., 2011. AG-dependent 3' -splice sites are predisposed to aberrant splicing due to a mutation at the first nucleotide of an exon. *Nucleic Acids Research* 39, 4396-4404.
- Gagné, V., Rousseau, J., Labuda, M., Sharif-Askari, B., Brukner, I., Laverdière, C., Ceppi, F., Sallan, S.E., Silverman, L.B., Neuberg, D., 2013. Bim polymorphisms: influence on function and response to treatment in children with acute lymphoblastic leukemia. *Clinical Cancer Research* 19, 5240-5249.
- Garcia-Manero, G., Faderl, S., O'Brien, S., Cortes, J., Talpaz, M., Kantarjian, H.M., 2003. Chronic myelogenous leukemia: a review and update of therapeutic strategies. *Cancer* 98, 437-457.
- Garneau, J.E., Dupuis, M.E., Villion, M., Romero, D.A., Barrangou, R., Boyaval, P., Fremaux, C., Horvath, P., Magadan, A.H., Moineau, S., 2010. The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. *Nature* 468, 67-71.
- Gautschi, O., Ratschiller, D., Gugger, M., Betticher, D.C., Heighway, J., 2007. Cyclin D1 in non-small cell lung cancer: a key driver of malignant transformation. *Lung Cancer* 55, 1-14.
- Gazdar, A., 2009. Activating and resistance mutations of EGFR in non-small-cell lung cancer: role in clinical response to EGFR tyrosine kinase inhibitors. *Oncogene* 28, S24-S31.
- Ghigna, C., Giordano, S., Shen, H., Benvenuto, F., Castiglioni, F., Comoglio, P.M., Green, M.R., Riva, S., Biamonti, G., 2005. Cell motility is controlled by SF2/ASF through alternative splicing of the Ron protooncogene. *Molecular Cell* 20, 881-890.
- Giulietti, M., Piva, F., D'Antonio, M., D'Onorio De Meo, P., Paoletti, D., Castrignanò, T., D'Erchia, A.M., Picardi, E., Zambelli, F., Principato, G., Pavesi, G., Pesole, G., 2013. SpliceAid-F: a database of human splicing factors and their RNA-binding sites. *Nucleic Acids Research* 41, D125-D131.
- Gomez-Bougie, P., Wuillème-Toumi, S., Ménoret, E., Trichet, V., Robillard, N., Philippe, M., Bataille, R., Amiot, M., 2007. Noxa up-regulation and Mcl-1 cleavage are associated

to apoptosis induction by bortezomib in multiple myeloma. *Cancer Research* 67, 5418-5424.

Gozani, O., Potashkin, J., Reed, R., 1998. A potential role for U2AF-SAP 155 interactions in recruiting U2 snRNP to the branch site. *Molecular and Cellular Biology* 18, 4752-4760.

Graveley, B.R., Hertel, K.J., Maniatis, T., 2001. The role of U2AF35 and U2AF65 in enhancer-dependent splicing. *RNA* 7, 806-818.

Groffen, J., Stephenson, J.R., Heisterkamp, N., de Klein, A., Bartram, C.R., Grosveld, G., 1984. Philadelphia chromosomal breakpoints are clustered within a limited region, bcr, on chromosome 22. *Cell* 36, 93-99.

Gross, A., McDonnell, J.M., Korsmeyer, S.J., 1999. BCL-2 family members and the mitochondria in apoptosis. *Genes and Development* 13, 1899-1911.

Guo, J., Jia, J., Jia, R., 2015. PTBP1 and PTBP2 impaired autoregulation of SRSF3 in cancer cells. *Scientific Reports* 5, 14548.

Guth, S., Martínez, C., Gaur, R.K., Valcárcel, J., 1999. Evidence for substrate-specific requirement of the splicing factor U2AF35 and for its function after polypyrimidine tract recognition by U2AF65. *Molecular and Cellular Biology* 19, 8263-8271.

Hall, J. A., Georgel P.T., 2011. The Worlds of Splicing and Chromatin Collide, *RNA Processing*, Prof. Paula Grabowski (Ed.), InTech, DOI: 10.5772/19985.

Hammond, S.M., Wood, M.J., 2011. Genetic therapies for RNA mis-splicing diseases. *Trends in Genetics* 27, 196-205.

Havens, M.A., Duelli, D.M., Hastings, M.L., 2013. Targeting RNA splicing for disease therapy. *Wiley interdisciplinary reviews: RNA* 4, 247-266.

He, X., Arslan, A.D., Pool, M.D., Ho, T.T., Darcy, K.M., Coon, J.S., Beck, W.T., 2011. Knockdown of splicing factor SRp20 causes apoptosis in ovarian cancer cells and its expression is associated with malignancy of epithelial ovarian cancer. *Oncogene* 30, 356-365.

He, Y., Smith, R., 2009. Nuclear functions of heterogeneous nuclear ribonucleoproteins A/B. *Molecular Cell Life Science* 66, 1239-1256.

Hildeman, D.A., Zhu, Y., Mitchell, T.C., Bouillet, P., Strasser, A., Kappler, J., Marrack, P., 2002. Activated T cell death in vivo mediated by proapoptotic bcl-2 family member bim. *Immunity* 16, 759-767.

Hoffman, J.D., Hallam, S.E., Venne, V.L., Lyon, E., Ward, K., 1998. Implications of a novel cryptic splice site in the BRCA1 gene. *American Journal of Medical Genetics* 80, 140-144.

- House, A.E., Lynch, K.W., 2008. Regulation of alternative splicing: more than just the ABCs. *The Journal of Biological Chemistry* 283, 1217-1221.
- Hsu, P.D., Lander, E.S., Zhang, F., 2014. Development and applications of CRISPR-Cas9 for genome engineering. *Cell* 157, 1262-1278.
- Hsu, S.Y., Lin, P., Hsueh, A.J., 1998. BOD (Bcl-2-related ovarian death gene) is an ovarian BH3 domain-containing proapoptotic Bcl-2 protein capable of dimerization with diverse antiapoptotic Bcl-2 members. *Molecular Endocrinology* 12, 1432-1440.
- Hua, Y., Vickers, T.A., Baker, B.F., Bennett, C.F., Krainer, A.R., 2007. Enhancement of SMN2 exon 7 inclusion by antisense oligonucleotides targeting the exon. *PLoS Biology* 5, e73.
- Hua, Y., Vickers, T.A., Okunola, H.L., Bennett, C.F., Krainer, A.R., 2008. Antisense masking of an hnRNP A1/A2 intronic splicing silencer corrects SMN2 splicing in transgenic mice. *The American Journal of Human Genetics* 82, 834-848.
- Hua, Y., Sahashi, K., Hung, G., Rigo, F., Passini, M.A., Bennett, C.F., Krainer, A.R., 2010. Antisense correction of SMN2 splicing in the CNS rescues necrosis in a type III SMA mouse model. *Genes and Development* 24, 1634-1644.
- Hua, Y., Sahashi, K., Rigo, F., Hung, G., Horev, G., Bennett, C.F., Krainer, A.R., 2011. Peripheral SMN restoration is essential for long-term rescue of a severe spinal muscular atrophy mouse model. *Nature* 478, 123-126.
- Huelga, S.C., Vu, A.Q., Arnold, J.D., Liang, T.Y., Liu, P.P., Yan, B.Y., Donohue, J.P., Shiue, L., Hoon, S., Brenner, S., Ares, M., Jr., Yeo, G.W., 2012. Integrative genome-wide analysis reveals cooperative regulation of alternative splicing by hnRNP proteins. *Cell Reports* 1, 167-178.
- Iborra, S., Hirschfeld, M., Jaeger, M., Zur Hausen, A., Braicu, I., Sehouli, J., Gitsch, G., Stickeler, E., 2013. Alterations in expression pattern of splicing factors in epithelial ovarian cancer and its clinical impact. *International Journal of Gynecological Cancer* 23, 990-996.
- Inui, M., Miyado, M., Igarashi, M., Tamano, M., Kubo, A., Yamashita, S., Asahara, H., Fukami, M., Takada, S., 2014. Rapid generation of mouse models with defined point mutations by the CRISPR/Cas9 system. *Scientific Reports* 4, 5396.
- Isobe, K., Hata, Y., Tochigi, N., Kaburaki, K., Kobayashi, H., Makino, T., Otsuka, H., Sato, F., Ishida, F., Kikuchi, N., 2014. Clinical significance of BIM deletion polymorphism in non-small-cell lung cancer with epidermal growth factor receptor mutation. *Journal of Thoracic Oncology* 9, 483-487.
- Jabbour, E., Kantarjian, H., 2014. Chronic myeloid leukemia: 2014 update on diagnosis, monitoring, and management. *American Journal of Hematology* 89, 547-556.

- Jangi, M., Sharp, P.A., 2014. Building robust transcriptomes with master splicing factors. *Cell* 159, 487-498.
- Jia, R., Li, C., McCoy, J.P., Deng, C.X., Zheng, Z.M., 2010. SRp20 is a proto-oncogene critical for cell proliferation and tumor induction and maintenance. *International Journal of Biological Sciences* 6, 806-826.
- Jiang, W., Bikard, D., Cox, D., Zhang, F., Marraffini, L.A., 2013. RNA-guided editing of bacterial genomes using CRISPR-Cas systems. *Nature Biotechnology* 31, 233-239.
- Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J.A., Charpentier, E., 2012. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* 337, 816-821.
- Juan, W.C., Roca, X., Ong, S.T., 2014. Identification of cis-acting elements and splicing factors involved in the regulation of BIM pre-mRNA splicing. *PloS One* 9, e95210.
- Julien, P., Minana, B., Baeza-Centurion, P., Valcarcel, J., Lehner, B., 2016. The complete local genotype-phenotype landscape for the alternative splicing of a human exon. *Nature Communications* 7, 11558.
- Kang, M.H., Reynolds, C.P., 2009. Bcl-2 inhibitors: targeting mitochondrial apoptotic pathways in cancer therapy. *Clinical Cancer Research* 15, 1126-1132.
- Kang, X., He, W., Huang, Y., Yu, Q., Chen, Y., Gao, X., Sun, X., Fan, Y., 2016. Introducing precise genetic modifications into human 3PN embryos by CRISPR/Cas-mediated genome editing. *Journal of Assisted Reproduction and Genetics* 33, 581-588.
- Karni, R., de Stanchina, E., Lowe, S.W., Sinha, R., Mu, D., Krainer, A.R., 2007. The gene encoding the splicing factor SF2/ASF is a proto-oncogene. *Nature Structural and Molecular Biology* 14, 185-193.
- Keren, H., Lev-Maor, G., Ast, G., 2010. Alternative splicing and evolution: diversification, exon definition and function. *Nature Reviews Genetics* 11, 345-355.
- Kim, D.H., Behlke, M.A., Rose, S.D., Chang, M.S., Choi, S., Rossi, J.J., 2005. Synthetic dsRNA Dicer substrates enhance RNAi potency and efficacy. *Nature Biotechnology* 23, 222-226.
- Kim, E., Goren, A., Ast, G., 2008. Alternative splicing: current perspectives. *BioEssays : news and reviews in molecular, cellular and developmental biology* 30, 38-47.
- Klener, P., Jr., Andera, L., Klener, P., Necas, E., Zivny, J., 2006. Cell death signalling pathways in the pathogenesis and therapy of haematologic malignancies: overview of apoptotic pathways. *Folia Biologica* 52, 34-44.

- Kornblihtt, A.R., Schor, I.E., Allo, M., Dujardin, G., Petrillo, E., Munoz, M.J., 2013. Alternative splicing: a pivotal step between eukaryotic transcription and translation. *Nature Reviews Molecular Cell Biology* 14, 153-165.
- Krawczak, M., Reiss, J., Cooper, D.N., 1992. The mutational spectrum of single base-pair substitutions in mRNA splice junctions of human genes: causes and consequences. *Human Genetics* 90, 41-54.
- Kuribara, R., Honda, H., Matsui, H., Shinjyo, T., Inukai, T., Sugita, K., Nakazawa, S., Hirai, H., Ozawa, K., Inaba, T., 2004. Roles of Bim in apoptosis of normal and Bcr-Abl-expressing hematopoietic progenitors. *Molecular Cell Biology* 24, 6172-6183.
- Kuroda, J., Puthalakath, H., Cragg, M.S., Kelly, P.N., Bouillet, P., Huang, D.C., Kimura, S., Ottmann, O.G., Druker, B.J., Villunger, A., Roberts, A.W., Strasser, A., 2006. Bim and Bad mediate imatinib-induced killing of Bcr/Abl+ leukemic cells, and resistance due to their loss is overcome by a BH3 mimetic. *Proceedings of the National Academy of Sciences* 103, 14907-14912.
- Kurzrock, R., Kantarjian, H.M., Druker, B.J., Talpaz, M., 2003. Philadelphia chromosome-positive leukemias: from basic mechanisms to molecular therapeutics. *Annals of Internal Medicine* 138, 819-830.
- Kuyumcu-Martinez, N.M., Wang, G.S., Cooper, T.A., 2007. Increased steady-state levels of CUGBP1 in myotonic dystrophy 1 are due to PKC-mediated hyperphosphorylation. *Molecular Cell* 28, 68-78.
- Kyburz, A., Friedlein, A., Langen, H., Keller, W., 2006. Direct interactions between subunits of CPSF and the U2 snRNP contribute to the coupling of pre-mRNA 3' end processing and splicing. *Molecular Cell* 23, 195-205.
- Lee, J.K., Shin, J.Y., Kim, S., Lee, S., Park, C., Kim, J.Y., Koh, Y., Keam, B., Min, H.S., Kim, T.M., Jeon, Y.K., Kim, D.W., Chung, D.H., Heo, D.S., Lee, S.H., Kim, J.I., 2013. Primary resistance to epidermal growth factor receptor (EGFR) tyrosine kinase inhibitors (TKIs) in patients with non-small-cell lung cancer harboring TKI-sensitive EGFR mutations: an exploratory study. *Annals of Oncology* 24, 2080-2087.
- Lee, J.Y., Ku, B.M., Lim, S.H., Lee, M.Y., Kim, H., Kim, M., Kim, S., Jung, H.A., Sun, J.M., Ahn, J.S., Park, K., Ahn, M.J., 2015. The BIM Deletion Polymorphism and its Clinical Implication in Patients with EGFR-Mutant Non-Small-Cell Lung Cancer Treated with EGFR Tyrosine Kinase Inhibitors. *Journal of Thoracic Oncology* 10, 903-909.
- Lefebvre, S., Bürglen, L., Reboullet, S., Clermont, O., Burlet, P., Viollet, L., Benichou, B., Cruaud, C., Millasseau, P., Zeviani, M., 1995. Identification and characterization of a spinal muscular atrophy-determining gene. *Cell* 80, 155-165.

- Lewis, J.D., Izaurralde, E., Jarmolowski, A., McGuigan, C., Mattaj, I.W., 1996. A nuclear cap-binding complex facilitates association of U1 snRNP with the cap-proximal 5' splice site. *Genes and Development* 10, 1683-1698.
- Ley, R., Balmanno, K., Hadfield, K., Weston, C., Cook, S.J., 2003. Activation of the ERK1/2 signaling pathway promotes phosphorylation and proteasome-dependent degradation of the BH3-only protein, Bim. *The Journal of Biological Chemistry* 278, 18811-18816.
- Li, Y., Chen, Z.Y., Wang, W., Baker, C.C., Krug, R.M., 2001. The 3'-end-processing factor CPSF is required for the splicing of single-intron pre-mRNAs in vivo. *RNA* 7, 920-931.
- Lin, X., Miller, J.W., Mankodi, A., Kanadia, R.N., Yuan, Y., Moxley, R.T., Swanson, M.S., Thornton, C.A., 2006. Failure of MBNL1-dependent post-natal splicing transitions in myotonic dystrophy. *Human Molecular Genetics* 15, 2087-2097.
- Liu, H.-X., Cartegni, L., Zhang, M.Q., Krainer, A.R., 2001. A mechanism for exon skipping caused by nonsense or missense mutations in BRCA1 and other genes. *Nature Genetics* 27, 55-58.
- Livak, K.J., Schmittgen, T.D., 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods* 25, 402-408.
- López-Bigas, N., Audit, B., Ouzounis, C., Parra, G., Guigó, R., 2005. Are splicing mutations the most frequent cause of hereditary disease? *FEBS letters* 579, 1900-1903.
- Lorson, C.L., Hahnen, E., Androphy, E.J., Wirth, B., 1999. A single nucleotide in the SMN gene regulates splicing and is responsible for spinal muscular atrophy. *Proceedings of the National Academy of Sciences* 96, 6307-6311.
- Lu, Z.X., Jiang, P., Xing, Y., 2012. Genetic variation of pre-mRNA alternative splicing in human populations. *Wiley interdisciplinary reviews. RNA* 3, 581-592.
- Luciano, F., Jacquél, A., Colosetti, P., Herrant, M., Cagnol, S., Pages, G., Auberger, P., 2003. Phosphorylation of Bim-EL by Erk1/2 on serine 69 promotes its degradation via the proteasome pathway and regulates its proapoptotic function. *Oncogene* 22, 6785-6793.
- Lydon, N., 2009. Attacking cancer at its foundation. *Nature Medicine* 15, 1153-1157.
- Lynch, T.J., Bell, D.W., Sordella, R., Gurubhagavatula, S., Okimoto, R.A., Brannigan, B.W., Harris, P.L., Haserlat, S.M., Supko, J.G., Haluska, F.G., Louis, D.N., Christiani, D.C., Settleman, J., Haber, D.A., 2004. Activating mutations in the epidermal growth factor receptor underlying responsiveness of non-small-cell lung cancer to gefitinib. *The New England Journal of Medicine* 350, 2129-2139.

- Ma, J.Y., Yan, H.J., Gu, W., 2015. Association between BIM deletion polymorphism and clinical outcome of EGFR-mutated NSCLC patient with EGFR-TKI therapy: A meta-analysis. *Journal of Cancer Research and Therapeutics* 11, 397-402.
- Mahadevan, M., Tsilfidis, C., Sabourin, L., Shutler, G., Amemiya, C., Jansen, G., Neville, C., Narang, M., Barcelo, J., O'Hoy, K., et al., 1992. Myotonic dystrophy mutation: an unstable CTG repeat in the 3' untranslated region of the gene. *Science* 255, 1253-1255.
- Makarova, K.S., Haft, D.H., Barrangou, R., Brouns, S.J., Charpentier, E., Horvath, P., Moineau, S., Mojica, F.J., Wolf, Y.I., Yakunin, A.F., van der Oost, J., Koonin, E.V., 2011. Evolution and classification of the CRISPR-Cas systems. *Nature Reviews Microbiology* 9, 467-477.
- Marani, M., Tenev, T., Hancock, D., Downward, J., Lemoine, N.R., 2002. Identification of novel isoforms of the BH3 domain protein Bim which directly activate Bax to trigger apoptosis. *Molecular and Cellular Biology* 22, 3577-3589.
- Matera, A.G., Wang, Z., 2014. A day in the life of the spliceosome. *Nature Reviews Molecular Cell Biology* 15, 108-121.
- Melo, J.V., Chuah, C., 2008. Novel agents in CML therapy: tyrosine kinase inhibitors and beyond. *ASH Education Program Book 2008*, 427-435.
- Mendell, J.R., Rodino-Klapac, L.R., Sahenk, Z., Roush, K., Bird, L., Lowes, L.P., Alfano, L., Gomez, A.M., Lewis, S., Kota, J., Malik, V., Shontz, K., Walker, C.M., Flanigan, K.M., Corridore, M., Kean, J.R., Allen, H.D., Shilling, C., Melia, K.R., Sazani, P., Saoud, J.B., Kaye, E.M., 2013. Eteplirsen for the treatment of Duchenne muscular dystrophy. *Annals of Neurology* 74, 637-647.
- Mestre-Escorihuela, C., Rubio-Moscardo, F., Richter, J.A., Siebert, R., Climent, J., Fresquet, V., Beltran, E., Agirre, X., Marugan, I., Marín, M., 2007. Homozygous deletions localize novel tumor suppressor genes in B-cell lymphomas. *Blood* 109, 271-280.
- Midha, A., Dearden, S., McCormack, R., 2015. EGFR mutation incidence in non-small-cell lung cancer of adenocarcinoma histology: a systematic review and global map by ethnicity (mutMapII). *American Journal of Cancer Research* 5, 2892-2911.
- Millevoi, S., Loulergue, C., Dettwiler, S., Karaa, S.Z., Keller, W., Antoniou, M., Vagner, S., 2006. An interaction between U2AF 65 and CF Im links the splicing and 3' end processing machineries. *The EMBO Journal* 25, 4854-4864.
- Mitsudomi, T., Yatabe, Y., 2007. Mutations of the epidermal growth factor receptor gene and related genes as determinants of epidermal growth factor receptor tyrosine kinase inhibitors sensitivity in lung cancer. *Cancer Science* 98, 1817-1824.
- Monani, U.R., 2005. Spinal muscular atrophy: a deficiency in a ubiquitous protein; a motor neuron-specific disease. *Neuron* 48, 885-895.

Montague, T.G., Cruz, J.M., Gagnon, J.A., Church, G.M., Valen, E., 2014. CHOPCHOP: a CRISPR/Cas9 and TALEN web tool for genome editing. *Nucleic Acids Research* 42, W401-407.

Montes, M., Becerra, S., Sanchez-Alvarez, M., Sune, C., 2012. Functional coupling of transcription and splicing. *Gene* 501, 104-117.

Mordes, D., Luo, X., Kar, A., Kuo, D., Xu, L., Fushimi, K., Yu, G., Sternberg Jr, P., Wu, J.Y., 2006. Pre-mRNA splicing and retinitis pigmentosa. *Molecular Vision* 12, 1259.

Murthy, K.G., Manley, J.L., 1995. The 160-kD subunit of human cleavage-polyadenylation specificity factor coordinates pre-mRNA 3'-end formation. *Genes and Development* 9, 2672-2683.

Narla, G., Difeo, A., Reeves, H.L., Schaid, D.J., Hirshfeld, J., Hod, E., Katz, A., Isaacs, W.B., Hebbing, S., Komiya, A., McDonnell, S.K., Wiley, K.E., Jacobsen, S.J., Isaacs, S.D., Walsh, P.C., Zheng, S.L., Chang, B.L., Friedrichsen, D.M., Stanford, J.L., Ostrander, E.A., Chinnaiyan, A.M., Rubin, M.A., Xu, J., Thibodeau, S.N., Friedman, S.L., Martignetti, J.A., 2005. A germline DNA polymorphism enhances alternative splicing of the KLF6 tumor suppressor gene and is associated with increased prostate cancer risk. *Cancer Research* 65, 1213-1222.

Nembaware, V., Wolfe, K.H., Bettoni, F., Kelso, J., Seoighe, C., 2004. Allele-specific transcript isoforms in human. *FEBS Letters* 577, 233-238.

Ness, J.M., Harvey, C.A., Strasser, A., Bouillet, P., Klocke, B.J., Roth, K.A., 2006. Selective involvement of BH3-only Bcl-2 family members Bim and Bad in neonatal hypoxia-ischemia. *Brain Research* 1099, 150-159.

Ng, K.P., Hillmer, A.M., Chuah, C.T., Juan, W.C., Ko, T.K., Teo, A.S., Ariyaratne, P.N., Takahashi, N., Sawada, K., Fei, Y., Soh, S., Lee, W.H., Huang, J.W., Allen, J.C., Jr., Woo, X.Y., Nagarajan, N., Kumar, V., Thalamuthu, A., Poh, W.T., Ang, A.L., Mya, H.T., How, G.F., Yang, L.Y., Koh, L.P., Chowbay, B., Chang, C.T., Nadarajan, V.S., Chng, W.J., Than, H., Lim, L.C., Goh, Y.T., Zhang, S., Poh, D., Tan, P., Seet, J.E., Ang, M.K., Chau, N.M., Ng, Q.S., Tan, D.S., Soda, M., Isobe, K., Nothen, M.M., Wong, T.Y., Shahab, A., Ruan, X., Cacheux-Rataboul, V., Sung, W.K., Tan, E.H., Yatabe, Y., Mano, H., Soo, R.A., Chin, T.M., Lim, W.T., Ruan, Y., Ong, S.T., 2012. A common BIM deletion polymorphism mediates intrinsic resistance and inferior responses to tyrosine kinase inhibitors in cancer. *Nature Medicine* 18, 521-528.

Nie, W., Tao, X., Wei, H., Chen, W.S., Li, B., 2015. The BIM deletion polymorphism is a prognostic biomarker of EGFR-TKIs response in NSCLC: A systematic review and meta-analysis. *Oncotarget* 6, 25696-25700.

Nikrad, M., Johnson, T., Puthalalath, H., Coultas, L., Adams, J., Kraft, A.S., 2005. The proteasome inhibitor bortezomib sensitizes cells to killing by death receptor ligand TRAIL via BH3-only proteins Bik and Bim. *Molecular Cancer Therapeutics* 4, 443-449.

- Nilsen, T.W., Graveley, B.R., 2010. Expansion of the eukaryotic proteome by alternative splicing. *Nature* 463, 457-463.
- Niwa, M., Berget, S.M., 1991. Mutation of the AAUAAA polyadenylation signal depresses in vitro splicing of proximal but not distal introns. *Genes and Development* 5, 2086-2095.
- Noone, P.G., Knowles, M.R., 2001. 'CFTR-opathies': disease phenotypes associated with cystic fibrosis transmembrane regulator gene mutations. *Respiratory Research* 2, 328.
- Nowell, P.C., 1960. A minute chromosome in human granulocytic leukemia. *Science* 132, 1497-1501.
- O'Connor, L., Strasser, A., O'Reilly, L.A., Hausmann, G., Adams, J.M., Cory, S., Huang, D.C., 1998. Bim: a novel member of the Bcl-2 family that promotes apoptosis. *The EMBO Journal* 17, 384-395.
- O'Hare, T., Deininger, M.W., Eide, C.A., Clackson, T., Druker, B.J., 2011. Targeting the BCR-ABL signaling pathway in therapy-resistant Philadelphia chromosome-positive leukemia. *Clinical Cancer Research* 17, 212-221.
- O'Reilly, L.A., Cullen, L., Visvader, J., Lindeman, G.J., Print, C., Bath, M.L., Huang, D.C., Strasser, A., 2000. The proapoptotic BH3-only protein bim is expressed in hematopoietic, epithelial, neuronal, and germ cells. *The American Journal of Pathology* 157, 449-461.
- Pacheco, T.R., Coelho, M.B., Desterro, J.M., Mollet, I., Carmo-Fonseca, M., 2006. In vivo requirement of the small subunit of U2AF for recognition of a weak 3' splice site. *Molecular and Cellular Biology* 26, 8183-8190.
- Packham, G., Stevenson, F.K., 2005. Bodyguards and assassins: Bcl - 2 family proteins and apoptosis control in chronic lymphocytic leukaemia. *Immunology* 114, 441-449.
- Paez, J.G., Janne, P.A., Lee, J.C., Tracy, S., Greulich, H., Gabriel, S., Herman, P., Kaye, F.J., Lindeman, N., Boggon, T.J., Naoki, K., Sasaki, H., Fujii, Y., Eck, M.J., Sellers, W.R., Johnson, B.E., Meyerson, M., 2004. EGFR mutations in lung cancer: correlation with clinical response to gefitinib therapy. *Science* 304, 1497-1500.
- Pao, W., Chmielecki, J., 2010. Rational, biologically based treatment of EGFR-mutant non-small-cell lung cancer. *Nature Reviews Cancer* 10, 760-774.
- Pao, W., Miller, V.A., Politi, K.A., Riely, G.J., Somwar, R., Zakowski, M.F., Kris, M.G., Varmus, H., 2005. Acquired resistance of lung adenocarcinomas to gefitinib or erlotinib is associated with a second mutation in the EGFR kinase domain. *PLoS Medicine* 2, e73.
- Pao, W., Miller, V., Zakowski, M., Doherty, J., Politi, K., Sarkaria, I., Singh, B., Heelan, R., Rusch, V., Fulton, L., Mardis, E., Kupfer, D., Wilson, R., Kris, M., Varmus, H., 2004. EGF receptor gene mutations are common in lung cancers from "never smokers" and are

associated with sensitivity of tumors to gefitinib and erlotinib. *Proceedings of the National Academy of Sciences* 101, 13306-13311.

Passini, M.A., Bu, J., Richards, A.M., Kinnecom, C., Sardi, S.P., Stanek, L.M., Hua, Y., Rigo, F., Matson, J., Hung, G., 2011. Antisense oligonucleotides delivered to the mouse CNS ameliorate symptoms of severe spinal muscular atrophy. *Science Translational Medicine* 3, 72ra18-72ra18.

Pattanayak, V., Lin, S., Guilinger, J.P., Ma, E., Doudna, J.A., Liu, D.R., 2013. High-throughput profiling of off-target DNA cleavage reveals RNA-programmed Cas9 nuclease specificity. *Nature Biotechnology* 31, 839-843.

Paz, I., Akerman, M., Dror, I., Kost, I., Mandel-Gutfreund, Y., 2010. SFmap: a web server for motif analysis and prediction of splicing factor binding sites. *Nucleic Acids Research* 38, W281-285.

Pellegrini, M., Belz, G., Bouillet, P., Strasser, A., 2003. Shutdown of an acute T cell immune response to viral infection is mediated by the proapoptotic Bcl-2 homology 3-only protein Bim. *Proceedings of the National Academy of Sciences* 100, 14175-14180.

Perrone-Bizzozero, N., Bolognani, F., 2002. Role of HuD and other RNA-binding proteins in neural development and plasticity. *Journal of Neuroscience Research* 68, 121-126.

Perrotti, D., Jamieson, C., Goldman, J., Skorski, T., 2010. Chronic myeloid leukemia: mechanisms of blastic transformation. *Journal of Clinical Investigation* 120, 2254-2264.

Pinon, J.D., Labi, V., Egle, A., Villunger, A., 2008. Bim and Bmf in tissue homeostasis and malignant disease. *Oncogene* 27 Suppl 1, S41-52.

Porensky, P.N., Mitropant, C., McGovern, V.L., Bevan, A.K., Foust, K.D., Kaspar, B.K., Wilton, S.D., Burghes, A.H., 2011. A single administration of morpholino antisense oligomer rescues spinal muscular atrophy in mouse. *Human Molecular Genetics*, ddr600.

Puthalakath, H., Huang, D.C., O'Reilly, L.A., King, S.M., Strasser, A., 1999. The proapoptotic activity of the Bcl-2 family member Bim is regulated by interaction with the dynein motor complex. *Molecular Cell* 3, 287-296.

Qian, L., Van Laake, L.W., Huang, Y., Liu, S., Wendland, M.F., Srivastava, D., 2011. miR-24 inhibits apoptosis and represses Bim in mouse cardiomyocytes. *The Journal of Experimental Medicine* 208, 549-560.

Quintás-Cardama, A., Cortes, J., 2009. Molecular biology of BCR-ABL1-positive chronic myeloid leukemia. *Blood* 113, 1619-1630.

Ran, F.A., Hsu, P.D., Wright, J., Agarwala, V., Scott, D.A., Zhang, F., 2013a. Genome engineering using the CRISPR-Cas9 system. *Nature Protocols* 8, 2281-2308.

- Ran, F.A., Hsu, P.D., Lin, C.Y., Gootenberg, J.S., Konermann, S., Trevino, A.E., Scott, D.A., Inoue, A., Matoba, S., Zhang, Y., Zhang, F., 2013b. Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. *Cell* 154, 1380-1389.
- Rana, A., Shah, S.H., Rehman, N., Ali, S., Ali, G.M., Bhatti, S., Farooqi, A.A., 2011. Chronic myeloid leukemia: Attributes of break point cluster region-abelson (BCR-ABL). *Journal of Cancer Research and Experimental Oncology* 3, 62-66.
- Reed, J.C., 2003. Apoptosis-targeted therapies for cancer. *Cancer Cell* 3, 17-22.
- Ren, R., 2005. Mechanisms of BCR-ABL in the pathogenesis of chronic myelogenous leukaemia. *Nature Reviews Cancer* 5, 172-183.
- Richter-Larrea, J.A., Robles, E.F., Fresquet, V., Beltran, E., Rullan, A.J., Agirre, X., Calasanz, M.J., Panizo, C., Richter, J.A., Hernandez, J.M., 2010. Reversion of epigenetically mediated BIM silencing overcomes chemoresistance in Burkitt lymphoma. *Blood* 116, 2531-2542.
- Rosell, R., Moran, T., Queralt, C., Porta, R., Cardenal, F., Camps, C., Majem, M., Lopez-Vivanco, G., Isla, D., Provencio, M., Insa, A., Massuti, B., Gonzalez-Larriba, J.L., Paz-Ares, L., Bover, I., Garcia-Campelo, R., Moreno, M.A., Catot, S., Rolfo, C., Reguart, N., Palmero, R., Sanchez, J.M., Bastus, R., Mayo, C., Bertran-Alamillo, J., Molina, M.A., Sanchez, J.J., Taron, M., 2009. Screening for epidermal growth factor receptor mutations in lung cancer. *The New England Journal of Medicine* 361, 958-967.
- Rowley, J.D., 1973. A new consistent chromosomal abnormality in chronic myelogenous leukaemia identified by quinacrine fluorescence and Giemsa staining. *Nature* 243, 290-293.
- Rubin, B.Y., Anderson, S.L., 2008. The molecular basis of familial dysautonomia: overview, new discoveries and implications for directed therapies. *Neuromolecular Medicine* 10, 148-156.
- Sander, J.D., Joung, J.K., 2014. CRISPR-Cas systems for editing, regulating and targeting genomes. *Nature Biotechnology* 32, 347-355.
- Savage, D.G., Antman, K.H., 2002. Imatinib mesylate—a new oral targeted therapy. *New England Journal of Medicine* 346, 683-693.
- Savona, M., Talpaz, M., 2008. Getting to the stem of chronic myeloid leukaemia. *Nature Reviews Cancer* 8, 341-350.
- Sequist, L.V., Bell, D.W., Lynch, T.J., Haber, D.A., 2007. Molecular predictors of response to epidermal growth factor receptor antagonists in non-small-cell lung cancer. *Journal of Clinical Oncology* 25, 587-595.
- Sequist, L.V., Waltman, B.A., Dias-Santagata, D., Digumarthy, S., Turke, A.B., Fidias, P., Bergethon, K., Shaw, A.T., Gettinger, S., Cosper, A.K., Akhavanfard, S., Heist, R.S.,

- Temel, J., Christensen, J.G., Wain, J.C., Lynch, T.J., Vernovsky, K., Mark, E.J., Lanuti, M., Iafrate, A.J., Mino-Kenudson, M., Engelman, J.A., 2011. Genotypic and histological evolution of lung cancers acquiring resistance to EGFR inhibitors. *Science Translational Medicine* 3, 75ra26.
- Sharma, S.V., Bell, D.W., Settleman, J., Haber, D.A., 2007. Epidermal growth factor receptor mutations in lung cancer. *Nature Reviews Cancer* 7, 169-181.
- Sheth, N., Roca, X., Hastings, M.L., Roeder, T., Krainer, A.R., Sachidanandam, R., 2006. Comprehensive splice-site analysis using comparative genomics. *Nucleic Acids Research* 34, 3955-3967.
- Shigematsu, H., Lin, L., Takahashi, T., Nomura, M., Suzuki, M., Wistuba, II, Fong, K.M., Lee, H., Toyooka, S., Shimizu, N., Fujisawa, T., Feng, Z., Roth, J.A., Herz, J., Minna, J.D., Gazdar, A.F., 2005. Clinical and biological features associated with epidermal growth factor receptor gene mutations in lung cancers. *Journal of the National Cancer Institute* 97, 339-346.
- Singh, R.K., Cooper, T.A., 2012. Pre-mRNA splicing in disease and therapeutics. *Trends in Molecular Medicine* 18, 472-482.
- Slaugenhaupt, S.A., Blumenfeld, A., Gill, S.P., Leyne, M., Mull, J., Cuajungco, M.P., Liebert, C.B., Chadwick, B., Idelson, M., Reznik, L., 2001. Tissue-specific expression of a splicing mutation in the IKBKAP gene causes familial dysautonomia. *The American Journal of Human Genetics* 68, 598-605.
- Smith, C.W., Valcarcel, J., 2000. Alternative pre-mRNA splicing: the logic of combinatorial control. *Trends in Biochemical Sciences* 25, 381-388.
- Smith, D.L., Burthem, J., Whetton, A.D., 2003. Molecular pathogenesis of chronic myeloid leukaemia. *Expert Reviews in Molecular Medicine* 5, 1-27.
- Sordella, R., Bell, D.W., Haber, D.A., Settleman, J., 2004. Gefitinib-sensitizing EGFR mutations in lung cancer activate anti-apoptotic pathways. *Science* 305, 1163-1167.
- Soria, J.C., Mok, T.S., Cappuzzo, F., Janne, P.A., 2012. EGFR-mutated oncogene-addicted non-small cell lung cancer: current trends and future prospects. *Cancer Treatment Reviews* 38, 416-430.
- Spellman, R., Llorian, M., Smith, C.W., 2007. Crossregulation and functional redundancy between the splicing regulator PTB and its paralogs nPTB and ROD1. *Molecular Cell* 27, 420-434.
- Steiner, B., Truninger, K., Sanz, J., Schaller, A., Gallati, S., 2004. The role of common single - nucleotide polymorphisms on exon 9 and exon 12 skipping in nonmutated CFTR Alleles. *Human Mutation* 24, 120-129.

- Strasser, A., O'Connor, L., Dixit, V.M., 2000. Apoptosis signaling. *Annual Review of Biochemistry* 69, 217-245.
- Strasser, A., Harris, A.W., Huang, D.C., Krammer, P.H., Cory, S., 1995. Bcl-2 and Fas/APO-1 regulate distinct pathways to lymphocyte apoptosis. *The EMBO Journal* 14, 6136-6147.
- Tagawa, H., Karnan, S., Suzuki, R., Matsuo, K., Zhang, X., Ota, A., Morishima, Y., Nakamura, S., Seto, M., 2004. Genome-wide array-based CGH for mantle cell lymphoma: identification of homozygous deletions of the proapoptotic gene BIM. *Oncogene* 24, 1348-1358.
- Tan, T.-T., Degenhardt, K., Nelson, D.A., Beaudoin, B., Nieves-Neira, W., Bouillet, P., Villunger, A., Adams, J.M., White, E., Key roles of BIM-driven apoptosis in epithelial tumors and rational chemotherapy. *Cancer Cell* 7, 227-238.
- Tang, S.J., Luo, S., Ho, J.X.J., Ly, P.T., Goh, E., Roca, X., 2016. Characterization of the Regulation of CD46 Alternative Splicing. *Journal of Biological Chemistry*. M115. 710350.
- Tazi, J., Bakkour, N., Stamm, S., 2009. Alternative splicing and disease. *Biochimica et Biophysica Acta (BBA)-Molecular Basis of Disease* 1792, 14-26.
- Thielen, N., Ossenkoppele, G.J., Schuurhuis, G.J., Janssen, J.J., 2011. New insights into the pathogenesis of chronic myeloid leukaemia: towards a path to cure. *Netherlands Journal of Medicine* 69, 430-440.
- Tzifi, F., Economopoulou, C., Gourgiotis, D., Ardavanis, A., Papageorgiou, S., Scorilas, A., 2011. The role of BCL2 family of apoptosis regulator proteins in acute and chronic leukemias. *Advances in Hematology* 2012.
- Ule, J., Ule, A., Spencer, J., Williams, A., Hu, J.S., Cline, M., Wang, H., Clark, T., Fraser, C., Ruggiu, M., Zeeberg, B.R., Kane, D., Weinstein, J.N., Blume, J., Darnell, R.B., 2005. Nova regulates brain-specific splicing to shape the synapse. *Nature Genetics* 37, 844-852.
- Vigneri, P., Wang, J.Y., 2001. Induction of apoptosis in chronic myelogenous leukemia cells through nuclear entrapment of BCR–ABL tyrosine kinase. *Nature Medicine* 7, 228-234.
- Voit, T., Topaloglu, H., Straub, V., Muntoni, F., Deconinck, N., Campion, G., De Kimpe, S.J., Eagle, M., Guglieri, M., Hood, S., Liefwaard, L., Loubakos, A., Morgan, A., Nakielnny, J., Quarcoo, N., Ricotti, V., Rolfe, K., Servais, L., Wardell, C., Wilson, R., Wright, P., Kraus, J.E., 2014. Safety and efficacy of drisapersen for the treatment of Duchenne muscular dystrophy (DEMAND II): an exploratory, randomised, placebo-controlled phase 2 study. *The Lancet Neurology* 13, 987-996.

- von Bubnoff, N., Duyster, J., 2010. Chronic myelogenous leukemia: treatment and monitoring. *Deutsches Arzteblatt International* 107, 114-121.
- Wahl, M.C., Will, C.L., Luhrmann, R., 2009. The spliceosome: design principles of a dynamic RNP machine. *Cell* 136, 701-718.
- Wang, E.T., Sandberg, R., Luo, S., Khrebtkova, I., Zhang, L., Mayr, C., Kingsmore, S.F., Schroth, G.P., Burge, C.B., 2008. Alternative isoform regulation in human tissue transcriptomes. *Nature* 456, 470-476.
- Wang, G.S., Cooper, T.A., 2007. Splicing in disease: disruption of the splicing code and the decoding machinery. *Nature Reviews Genetics* 8, 749-761.
- Wang, H., Yang, H., Shivalila, C.S., Dawlaty, M.M., Cheng, A.W., Zhang, F., Jaenisch, R., 2013. One-step generation of mice carrying mutations in multiple genes by CRISPR/Cas-mediated genome engineering. *Cell* 153, 910-918.
- Ward, A.J., Cooper, T.A., 2010. The pathobiology of splicing. *The Journal of Pathology* 220, 152-163.
- Will, C.L., Luhrmann, R., 2011. Spliceosome structure and function. *Cold Spring Harbor perspectives in biology* 3.
- Willis, S.N., Adams, J.M., 2005. Life in the balance: how BH3-only proteins induce apoptosis. *Current Opinion in Cell Biology* 17, 617-625.
- Wirth, B., 2000. An update of the mutation spectrum of the survival motor neuron gene (SMN1) in autosomal recessive spinal muscular atrophy (SMA). *Human Mutation* 15, 228.
- Wojciechowski, S., Tripathi, P., Bourdeau, T., Acero, L., Grimes, H.L., Katz, J.D., Finkelman, F.D., Hildeman, D.A., 2007. Bim/Bcl-2 balance is critical for maintaining naive and memory T cell homeostasis. *The Journal of Experimental Medicine* 204, 1665-1675.
- Xargay-Torrent, S., López-Guerra, M., Saborit-Villarroya, I., Rosich, L., Campo, E., Roué, G., Colomer, D., 2011. Vorinostat-induced apoptosis in mantle cell lymphoma is mediated by acetylation of proapoptotic BH3-only gene promoters. *Clinical Cancer Research* 17, 3956-3968.
- Yano, S., Wang, W., Li, Q., Matsumoto, K., Sakurama, H., Nakamura, T., Ogino, H., Kakiuchi, S., Hanibuchi, M., Nishioka, Y., Uehara, H., Mitsudomi, T., Yatabe, Y., Nakamura, T., Sone, S., 2008. Hepatocyte growth factor induces gefitinib resistance of lung adenocarcinoma with epidermal growth factor receptor-activating mutations. *Cancer Research* 68, 9479-9487.

- Ying, H.-Q., Chen, J., He, B.-S., Pan, Y.-Q., Wang, F., Deng, Q.-W., Sun, H.-L., Liu, X., Wang, S.-K., 2015. The effect of BIM deletion polymorphism on intrinsic resistance and clinical outcome of cancer patient with kinase inhibitor therapy. *Scientific Reports* 5.
- Youle, R.J., Strasser, A., 2008. The BCL-2 protein family: opposing activities that mediate cell death. *Nature Reviews Molecular Cell Biology* 9, 47-59.
- Yuan, J., Lipinski, M., Degtarev, A., 2003. Diversity in the mechanisms of neuronal cell death. *Neuron* 40, 401-413.
- Yun, C.H., Boggon, T.J., Li, Y., Woo, M.S., Greulich, H., Meyerson, M., Eck, M.J., 2007. Structures of lung cancer-derived EGFR mutants and inhibitor complexes: mechanism of activation and insights into differential inhibitor sensitivity. *Cancer Cell* 11, 217-227.
- Zahler, A.M., Lane, W.S., Stolk, J.A., Roth, M.B., 1992. SR proteins: a conserved family of pre-mRNA splicing factors. *Genes and Development* 6, 837-847.
- Zahler, A.M., 2005. Alternative splicing in *C. elegans*. *WormBook : The Online Review of C. elegans Biology*, 1-13.
- Zammarchi, F., de Stanchina, E., Bournazou, E., Supakorndej, T., Martires, K., Riedel, E., Corben, A.D., Bromberg, J.F., Cartegni, L., 2011. Antitumorigenic potential of STAT3 alternative splicing modulation. *Proceedings of the National Academy of Sciences* 108, 17779-17784.
- Zarnack, K., Konig, J., Tajnik, M., Martincorena, I., Eustermann, S., Stevant, I., Reyes, A., Anders, S., Luscombe, N.M., Ule, J., 2013. Direct competition between hnRNP C and U2AF65 protects the transcriptome from the exonization of Alu elements. *Cell* 152, 453-466.
- Zhang, W.W., Cortes, J.E., Yao, H., Zhang, L., Reddy, N.G., Jabbour, E., Kantarjian, H.M., Jones, D., 2009. Predictors of primary imatinib resistance in chronic myelogenous leukemia are distinct from those in secondary imatinib resistance. *The American Society of Clinical Oncology* 27, 3642-3649.
- Zhao, M., Zhang, Y., Cai, W., Li, J., Zhou, F., Cheng, N., Ren, R., Zhao, C., Li, X., Ren, S., 2014. The Bim deletion polymorphism clinical profile and its relation with tyrosine kinase inhibitor resistance in Chinese patients with non-small cell lung cancer. *Cancer* 120, 2299-2307.
- Zheng, Q., Cai, X., Tan, M.H., Schaffert, S., Arnold, C.P., Gong, X., Chen, C.Z., Huang, S., 2014. Precise gene deletion and replacement using the CRISPR/Cas9 system in human cells. *Biotechniques* 57, 115-124.