# Function of Splicing factor Tra2-beta1 in Alternative Splicing and Translation

Den Naturwissenschaftlichen Fakultäten
der Friedrich-Alexander-Universität Erlangen-Nürnberg
zur
Erlangung des Doktorgrades

vorgelegt von
Yesheng Tang
aus P.R.China
2007

Als Dissertation genehmigt von den Naturwissenschaftlichen Fakultäten der Universität Erlangen-Nürnberg.

Tag der mündlichen Prüfung:

| | |
|---|---|
| Vorsitzender der Prüfungskomission: | Prof. Dr. Eberhard Bänsch |
| Erstberichterstatter: | Prof. Dr. Andreas Burkovski |
| Zweitberichterstatter: | Prof. Dr. Michael Wegner |

**To my family**

# ACKNOWLEDGMENTS

The work presented here was performed in the Institute of Biochemistry at Friedrich-Alexander University of Erlangen-Nürnberg.

I would like to give my sincerest thanks to Prof. Dr. Stefan Stamm for his motivated support and kind guidance as the supervisor during my PhD research. His insight opinions and valuable suggestions inspire me in both scientific and daily life. His encouragement stimulates my confidence of the ability to not only complete a degree but goal to excellence. Indirectly he shows me the possibility of balance in lives between the family and lab. As a result, slowly, I found myself grow to a deeper level of maturity and understanding of science and life in his influence.

I want to thank my colleagues in the Stamm's lab: Dr. Bettina Heinrich, Dr. Tatyana Novoyatleva, Dr. Zhaiyi Zhang, Dr. Ilona Rafalska, Dr. Jingyi Hui Natalya Benderska, Amit Khanna, and Dominique Olbert for their help and providing wonderful atmosphere. I specially thank for Dr. Shivendra Kishore, who is very kind and warmhearted to me whenever I had troubles. I enjoy and miss very much the open scientific discussion with him during the life in Germany. His rational advice and moral support will forever remain in my mind.

I am also thankful to all the members of the Institute of Biochemistry for freely sharing their resources with us, in particular, Prof. Wegner and Prof. Hannappel.

I appreciate my friends Rong Xie, Xia Zhao, Yu Zhu, Ulf Venne for their in time help during my life in Germany.

I would also like to thank all the teachers for my education and training.

Thank my family and parents for all their support.
(谢谢爸爸妈妈，妻子东昱和女儿甜甜)

# PUBLICATIONS

(Parts of this work are included in the following publications)

**Tang, Y.**, Kishore, S., Stamm, S. Shuttling Splicing factor Tra2beta1 binds with Ribosomal Protein L3 and regulates translation. Preparing

Novoyatleva T, Heinrich B, **Tang Y,** Benderska N, Ben-Dov C, Bracco L, Bollen M, Stamm S. Protein phosphatase 1 regulates the usage of tra2-beta1 dependent alternative exons. Human Molecular Genetics. 2007 Oct 2

Novoyatleva T, **Tang Y,** Rafalska I, Stamm S. Pre-mRNA Missplicing as a Cause of Human Disease. Prog Mol Subcell Biol. 2006;44:27-46.

Watermann DO, **Tang Y,** Zur Hausen A, Jager M, Stamm S, Stickeler E. (2006) Splicing factor Tra2-beta1 is specifically induced in breast cancer and regulates alternative splicing of the CD44 gene. Cancer Res. May 1;66(9):4774-80.

Stamm S, Riethoven JJ, Le Texier V, Gopalakrishnan C, Kumanduri V, **Tang Y,** Barbosa-Morais NL, Thanaraj TA. (2006) ASD: a bioinformatics resource on alternative splicing. Nucleic Acids Res. Jan 1;34(Database issue):D46-55.

Glatz DC, Rujescu D, **Tang Y,** Berendt FJ, Hartmann AM, Faltraco F, Rosenberg C, Hulette C, Jellinger K, Hampel H, Riederer P, Moller HJ, Andreadis A, Henkel K, Stamm S. (2005) The alternative splicing of tau exon 10 and its regulatory proteins CLK2 and TRA2-BETA1 changes in sporadic Alzheimer's disease. J Neurochem. Feb; 96(3):635-44. Epub 2005 Dec 20.

Stamm, S., Ben-Ari, S., Rafalska, I.,**Tang, Y.,** Zhang, Z., Toiber, D., Thanaraj, T.A., Soreq, H. Function of alternative splicing. Gene. 2005 Jan 3;344:1-20

**Tang, Y.,** Novoyatleva , T., Benderska, N., Kishore, S., Thanaraj, T.A. and Stamm, S. (2004) Analysis of alternative splicing in vivo using minigenes. In Westhof, Bindereif, Schön and Hartmann (eds.), Handbook of RNA Biochemistry. 2/ 755–782 Wiley-VCH.

# CONTENTS

# ZUSAMMENFASSUNG

# ABSTRACT

# ABBREVIATIONS

9G8 splicing factor, arginine/serine-rich 7
AD Alzheimer's Disease
AEdb alternative exon database
ASD alternative splicing database
ASePCR alternative splicing electronic RT-PCR
ASF/SF2 splicing factor arginine/serine-rich 1
ATP adenosine 5'-triphosphate
bp base pairs
BSA bovine serum albumin
CBP CREB binding protein
cDNA complementary DNA
CFTR cystic fibrosis transmembrane conductance regulator ATP-binding cassette subfamily C member 7
CLIP cross linking and immunoprecipitation
CLK CDC2-like kinase
CTD carboxyterminal domain (of RNA polymerase II)
$dH_2O$ distilled water
DMEM dulbeco's modified eagle medium
DMSO dimethyl sulfoxide
DNA deoxyribonucleic acid
dNTP deoxyribonucleotidtriphosphate
Dscam Down syndrome cell adhesion molecule
dsRBD double-stranded RNA-binding domain
DTT dithiothreitol
ECL enhanced chemiluminiscence
EDMD Emery-Dreifuss muscular dystrophy
EDTA ethylenediaminetetraacetic acid
EGFP enhanced green fluorescent protein
EMSA electrophoretic mobility shift assay
ESE exonic splicing enhancer
EST expressed sequence tag
FC fibrillar center
FCS fetal calf serum
FGFR fibroblast growth factor receptor
FMR-1 fragile X mental retardation syndrome 1 homolog
FTDP-17 frontotemporal dementia with Parkinsonism linked to chromosome 17
GC granular component
GH growth hormone
GnRH gonadotrophin releasing hormone
HEK human embryonic kidney
hnRNP heterogenous nuclear ribonucleoprotein
ICD interchromatin domain
ICH-1 interleukin-1® converting enzyme homologue 1

IGC interchromatin granule cluster
IL-4 interleukin-4
IP immunoprecipitation
IPTG isopropyl ®-D-1-thiogalactopyranoside
kDa kilodalton
KH domain hnRNP K homology domain
KLH keyhole limpet haemocyanin
mGluR1 metabotropic glutamate receptor
mRNA messenger RNA
ND10 nuclear domain 10
NMD nonsense-mediated decay
NOR nucleolar organizer region
Nova neuro-oncological ventral antigen
NPC nuclear pore complex
PBS phosphate buffered saline
PCR polymerase chain reaction
pht6 rat brain post natal 10 library human tra2-beta clone 6; YT521-B
PKC protein kinase C
PML promyelocytic leukemia
PMSF phenylmethanesulfonyl fluoride
PNC perinucleolar compartment
POD PML oncogenic domain
PP1 Protein Phosphatase 1
RNA ribonucleic acid
RNase ribonuclease
rpm revolutions per minute
RPL3 ribosomal protein L3
RRM RNA recognition motif
rRNA ribosomal RNA
RT-PCR reverse transcription followed by polymerase chain reaction
RUST regulated unproductive splicing and translation
SAF scaffold attachment factor (A or B)
SC35 splicing component, 35 kDa; splicing factor, arginine/serine-rich 2
SDS sodium dodecyl sulfate
SFRS splicing factor, arginine/serine-rich
snoRNP small nucleolar ribonucleoprotein
snRNP small nuclear ribonucleoprotein paricle
SR-protein serine-arginine- rich protein
SRp30c splicing factor, arginine/serine-rich 9
TBE tris-borate-EDTA buffer
TE tris-EDTA
TEMED N,N,N’,N’-tetramethylethylenediamine
Tra2 transformer 2
tRNA transfer RNA
TSH thyroid stimulating hormone
U1 70K U1 snRNP 70 kDa protein

U2AF U2 snRNP auxiliary factor (35 or 65 kDa)
UTR untranslated region

# Introduction:

### *mRNA Processing coupled with transcription and translation*

Eukaryotic pre-messenger RNA is subjected to a series of cellular events (Fig.1.1). These events include transcription (Kapranov, Willingham et al. 2007), capping of the 5' end, polyadenylation of the 3'end (Shatkin and Manley 2000), splicing (Blencowe, Nickerson et al. 1994), editing (Nishikura 2006), exporting and translation (Stewart 2007). Many evidence shows that all closed processes are coordinated and, in some cases, functionally coupled: Transcription and pre-mRNA splicing are extremely complex multi-molecular processes that involve protein-DNA, protein-RNA, and protein-protein interactions (Kornblihtt, de la Mata et al. 2004). Polyadenylation and splicing facilitate mRNA export to the cytoplasm while intact cap structure and poly(A) tail are required for targeting of the mRNA to the ribosome and the initiation of the translation. The cap and the poly(A) tail are also subject to regulatory mechanisms which control the mRNA stability and as consequence the protein expression. Genome-wide analyses revealed a preferential association of certain RNA-binding proteins with distinct functional classes of mRNAs, which suggests that biogenesis, export and translation of mRNA subpopulations may be coordinated differently. (Hieronymus and Silver 2003; Kim Guisbert, Duncan et al. 2005; Kohler and Hurt 2007)

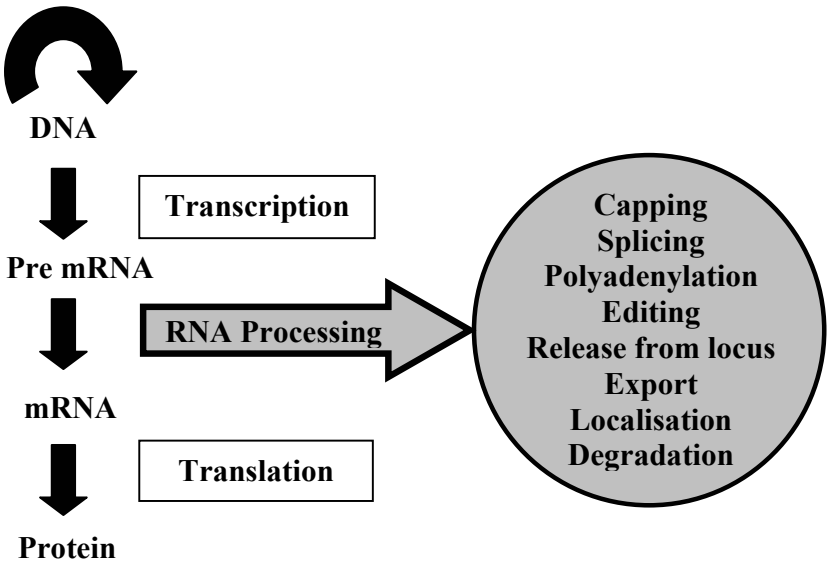



**Figure 1.1 Transcription, mRNA processing and translation**

### 1.1.1 Capping and Polyadenylation

The 5′ $m^7$GpppN cap and 3’ polyadenylation play essential roles in the life cycle of eukaryotic mRNA and are required for efficient pre-mRNA splicing, export, stability and translation.

The 5’triphosphate of the transcript is first shortened to diphosphate and then GMP is transferred generating an unusual triphosphate linkage: G(5’)ppp(5’)N. The guanosine base is subsequently methylated at N7 position. This cap can be further modified by methylating the 2’OH group of the ribose of the first and the second nucleotide (Fig. 1.2a) (Gu and Lima 2005).

Enzymatic synthesis and degradation of the mRNA cap is formed on nascent RNA by the sequential action of three enzymes (Fig. 1.2b): RNA triphosphatase, RNA guanylyltransferase and guanine-N7 RNA methyltransferase. RNA guanylyltransferase first forms a covalent lysyl-N–GMP adduct before transfer of the GMP to the 5′ diphosphate RNA end. Degradation of the RNA cap in the 5′-3′ decay pathway occurs through hydrolysis by the Dcp2–Dcp1 complex in a reaction that generates $m^7$GDP and 5′ monophosphate-terminated mRNA. Hydrolysis of the RNA cap in the 3′-5′ decay pathway is catalyzed by DcpS in a reaction that generates $m^7$GMP and diphosphate-terminated RNA. DcpS is also able to hydrolyze $m^7$GDP to release $m^7$GMP and phosphate (Fig. 1.2b) (Gu and Lima 2005).

3’ end formation starts with cleavage of the RNA guided by an AAUAAA sequence 10 to 30 nucleotides upstream of the cleavage site and a less conserved GU or U rich stretches downstream of the cleavage site. Some pre-mRNAs require also U-rich motifs located upstream of the cleavage site. After the cleavage a tail of 200-250 adenines is added to the 3’ end by poly(A) polymerase, with the exception of replication-dependent histone genes (in higher eukaryotes) (Fig.1.3) (Gu and Lima 2005).

The mammalian polyadenylation machinery consists of at least six multimeric protein factors (Fig.1.3). The site of cleavage in most pre-mRNAs lies between the highly conserved AAUAAA hexamer and a downstream sequence element (DSE), which is a U- or GU-rich motif. Cleavage itself occurs predominantly at a CA dinucleotide. In retroviruses it is also a common feature that U-rich upstream sequence elements (USE) located 5′ of the AAUAAA sequence are required for full efficiency of the poly(A) signal.

These *cis* elements are recognized by two multisubunit protein complexes. The 160 kDa subunit of the cleavage and polyadenylation specificity factor (CPSF) has been shown to interact with the AAUAAA. The DSE represents a platform for the interaction with the cleavage stimulatory factor (CstF) via its 64 kDa subunit. Two additional factors are essential to direct cleavage of the pre-mRNA: cleavage factor I (CF I) and cleavage factor II (CF II). CF I consists of three subunits that are able to directly interact with a pre-mRNA substrate. Poly(A) polymerase (PAP) itself is usually required for the cleavage reaction and together with CPSF directs poly(A) addition. Poly(A) binding protein PABP II binds the emerging poly(A) tail and in turn enhances the process of the poly(A) polymerase. (Fig.1.3) (Proudfoot, Furger et al. 2002).

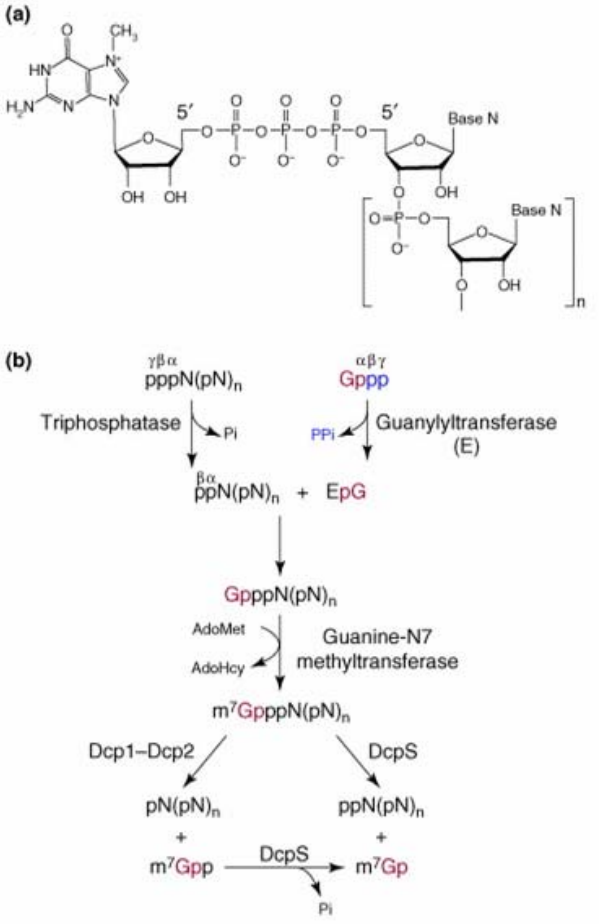



**Figure 1.2 mRNA cap structure and its metabolism. (a)** Chemical structure of the mRNA cap. (base N can be adenine, guanine, cytosine or uracil). **(b)** Enzymatic synthesis and degradation of the mRNA cap. Taken from (Gu and Lima 2005)

Both cap structure and poly(A) tail are essential for the mRNA transport to the cytoplasm, the initiation of the translation, and the protection of the mRNA from exonucleases. In addition, cap formation is required for transcription and pre-mRNA splicing. The cap needs to be in place by the time the first intron is spliced, because the cap binding complex, CBC, stimulates removal of this intron. Splicing of this intron is enhanced by interaction between U1 small nuclear ribonucleoprotein (snRNP) and CBC, which like other mRNA-binding proteins, associates with its target co-transcriptionally (Bentley 2002; Saguez, Olesen et al. 2005).

The fact that polyadenylation is strongly influenced by RNA splicing first became apparent through experiments that revealed the process of exon definition in RNA splicing (Niwa, MacDonald et al. 1992). It was found that splicing and polyadenylation on either side of the terminal exon were strongly enhanced by each other. Also in some cases, a domain in the U1A protein of U1 snRNP, which is homologous to the poly(A) polymerase (PAP) interacting domain of the 70 kDa subunit, can also be demonstrated to inhibit polyadenylation at a nearby site (Gunderson, Vagner et al. 1997). Examples of this type of regulation are well documented in retroviruses, where functional poly(A) signals may exist in the transcribed portion of the 5′ LTR sequence, as well as in DNA papillomaviruses, which again maximize their gene expression repertoire by employing internal poly(A) signals.

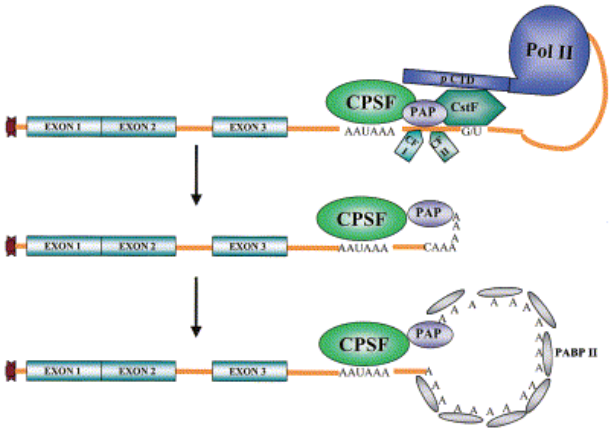



**Figure 1.3 Polyadenylation involves Separate RNA Cleavage and Poly(A) Synthesis.** Taken from (Proudfoot, Furger et al. 2002)

## 1.1.2 Splicing

### 1.1.2.1 Splicing sites

There are three major sequence elements (Table1.1) that mark the introns and are essential for their removal. The 5' splice site follows to the consensus YRG/GURRGU (the slash denoting the exon – intron border). In about 1% of all introns the first di-nucleotide is a GC instead of GU. The 3'splice site consensus is $Y_{12}$NYAG and is preceded by a polypyrimidine stretch. The third element is the branch point (YNCURAY) which is located 18 to 200 nucleotides upstream of the 3' splice site.

**Table 1.1: Sequence elements marking major class (GT-AG) intron.**

| Elements | Consensus Sequence* |
|---|---|
| 5' donor splice site | **YRG/GURAGU** |
| 3' acceptor splice site preceded by a polypyrimidine stretch | **$Y_{12}$NYAG /** |
| Branch point located 18-200nt upstream of the 3' splice site | **YNYURAY** |

*Symbols used: Y-Pyrimidine, R-Purine, N-any nucleotide, slash denotes the exon-intron border, Invariant nucleotides are underlined.

The major class introns (U2 type) have highly conserved di-nucleotides at the 5' and 3' termini (GT and AG respectively). In plants and metazoans another distinct minor intron class (U12 type) that have AT and AC termini also exist. Analysis of splice junction pairs from GenBank annotated mammalian genes showed that 99.24% conformed to canonical GT-AG, 0.69% to non-canonical GC-AG, 0.05% to AT-AC and 0.02% to other non-canonical splice termini (Burset, Seledtsov et al. 2001).

### 1.1.2.2 Spliceosome assembly

In major class introns, commitment of pre-mRNA to splicing pathway occurs upon formation of the E complex (Fig.1.4). Assembly of the E complex involves the recognition of 5' splice site, the polypyrimidine tract and 3' splice site by U1 snRNA, U2 auxiliary factor 65 (U2AF65) and U2 auxiliary factor 35 (U2AF35) respectively by base pairing. The branch point is recognized by the splicing factor1 (SF1). Several non-snRNP splicing factors such as serine/arginine rich (SR) proteins also associate to the pre-mRNA

at this step. In addition, U4/U6∗U5 tri-snRNP can associate with the first exon near the 5' splice site in the E-complex. This association is ATP dependent.

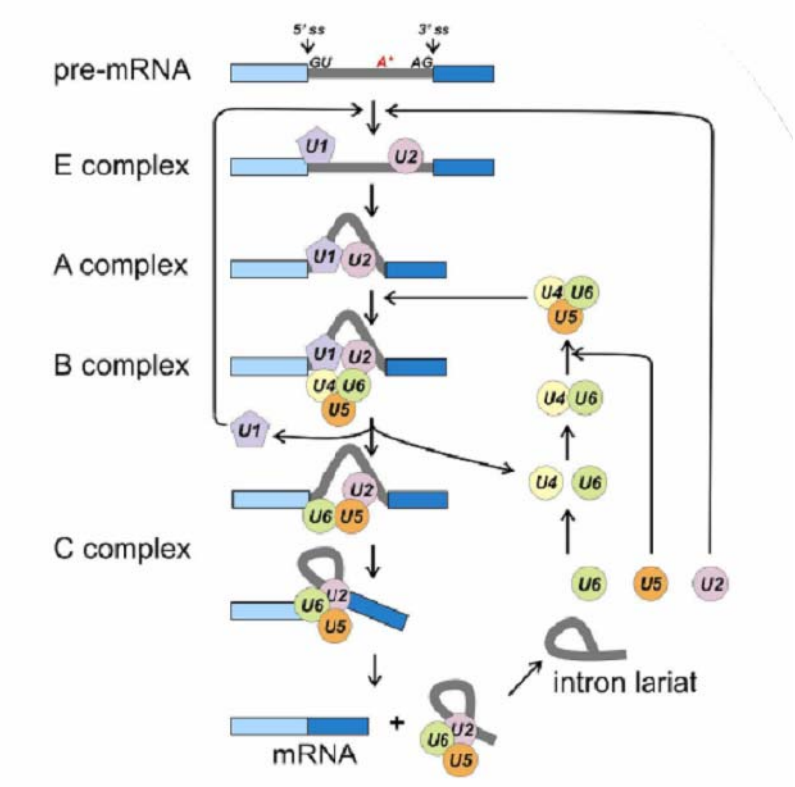



**Figure 1.4 The spliceosome cycle.** The U1, U2, U4, U5, and U6 snRNPs associate with the pre-mRNA and interact with each other in an ordered manner to form the spliceosome. Each formation of complex A,B,C and E are marked beside. The two transesterification reactions take place in the catalytic core of the spliceosome. After splicing, the spliceosome dissociates, and is re-assembled to take part in a new round of splicing cycle. Taken from Jingyi Hui's PhD thesis.

Next, U2 snRNP base pairs with the branch point during ATP dependent formation forms the A complex. Formation of next complex, the B complex, involves major rearrangements of the snRNP components associated with the pre-mRNA. The U6/U4 duplex is disrupted and a new duplex between U6 and the 5' splice site is formed, displacing the U1 snRNP. The 5' splice site is brought close to the branch point and the 3' splice site through U6/U2 snRNA base pairing and interaction of U5 snRNP with both exons near splice sites. At this point, U4 snRNP leaves the complex and the first catalytic

step of the splicing occurs, creating the intron lariat. Finally, U5 snRNP base pairs with both 5' and 3' exons, thus positioning the ends of the two exons for the second step of splicing. After the second step has been completed, the ligated exons and the lariat intron are released and the spliceosomal components dissociate and are recycled for further rounds of splicing (Fig.1.4).

The splicing reaction occurs in the spliceosome, which consists of five small nuclear ribonucleoproteins (snRNPs) and > 100 non-snRNP proteins (Zhou, Licklider et al. 2002) or around 200 splicing factors(Jurica and Moore 2003). Each snRNP is a tight complex composed of several proteins and a short RNA molecule. The RNA components of the five snRNPs are U1, U2, U4, U5, and U6 small nuclear RNAs (snRNAs). After U1, U2, U4, and U5 snRNAs are transcribed by RNA polymerase II, they are immediately exported to the cytoplasm, where a set of seven common proteins is assembled onto each of them. Since these seven common proteins are recognized by anti-Sm antibodies, they are named Sm proteins. Sm proteins bind to the highly conserved Sm site present in the snRNA, and are assembled in an ordered, stepwise manner to form the Sm core structure. Then the monomethylated guanosine cap (m7G) is converted to the 2,2,7-trimethylated guanosine (TMG) form, and the assembled snRNP cores are imported back to the nucleus, where the association of snRNP-specific proteins completes the biogenesis of snRNP complexes. The common(Sm or Sm like) and specific protein components of snRNPs are listed in Table.1.2(Will and Luhrmann 2001; Zhou, Licklider et al. 2002).

### 1.1.2.3 Splicing regulation

#### *1.1.2.3.1 Mechanism of splice site recognition*

A common feature of both alternative splicing and constitutive splicing is that they both require the spliceosome assembly, during which numerous trans-acting factors interact with each other and with cis-elements within the pre-mRNA to form active spliceosomes for the catalysis of pre-mRNA splicing (Fig.1.5). In alternative splicing, special regulatory factors are required to bind to splice sites or specific sequences within the pre-mRNA and subsequently activate or repress the utilization of splice sites. The mechanisms of splice site recognition are still not very well understood. Alternative splicing is often regulated in a tissue- or developmental stage-specific manner.

**Table 1.2 The common (Sm or LSm) and specific protein components of snRNPs**

| Designation | snRNPs | | | | |
|---|---|---|---|---|---|
| | U1 | U2 | U4/U6 | U5 | U4/U6·U5 tri-snRNP |
| Sm G | √ | √ | √ | √ | √ / √ |
| Sm F | √ | √ | √ | √ | √ / √ |
| Sm E | √ | √ | √ | √ | √ / √ |
| Sm D1 | √ | √ | √ | √ | √ / √ |
| Sm D2 | √ | √ | √ | √ | √ / √ |
| Sm D3 | √ | √ | √ | √ | √ / √ |
| Sm B/B' | √ | √ | √ | √ | √ / √ |
| LSm2 | | | ◆ | | ◆ |
| LSm3 | | | ◆ | | ◆ |
| LSm4 | | | ◆ | | ◆ |
| LSm5 | | | ◆ | | ◆ |
| LSm6 | | | ◆ | | ◆ |
| LSm7 | | | ◆ | | ◆ |
| LSm8 | | | ◆ | | ◆ |
| U1 C | ● | | | | |
| U1 A | ● | | | | |
| U1-70K | ● | | | | |
| U2 B'' | | ♥ | | | |
| U2 A | | ♥ | | | |
| SAP 61 | | ♥ | | | |
| SAP 62 | | ♥ | | | |
| SAP 114 | | ♥ | | | |
| SAP 49 | | ♥ | | | |
| SAP 130 | | ♥ | | | |
| SAP 145 | | ♥ | | | |
| SAP 155 | | ♥ | | | |
| p14* | | ♥ | | | |
| U4/U6-15.5K | | | ■ | | ■ |
| U4/U6-20K | | | ■ | | ■ |
| U4/U6-27K | | | ■ | | ■ |
| U4/U6-60K | | | ■ | | ■ |
| U4/U6-61K | | | ■ | | ■ |
| U4/U6-63K | | | ■ | | ■ |
| U4/U6-90K | | | ■ | | ■ |
| U5-15K | | | | ▲ | ▲ |
| U5-40K | | | | ▲ | ▲ |
| U5-52K | | | | ▲ | ▲ |
| U5-100K | | | | ▲ | ▲ |
| U5-102K | | | | ▲ | ▲ |
| U5-116K | | | | ▲ | ▲ |
| U5-200K | | | | ▲ | ▲ |
| U5-220K | | | | ▲ | ▲ |

* p14 is a recently identified U2 snRNP specific protein that interacts with the branch site adenosine (Will and Luhrmann 2001; Zhou, Licklider et al. 2002)

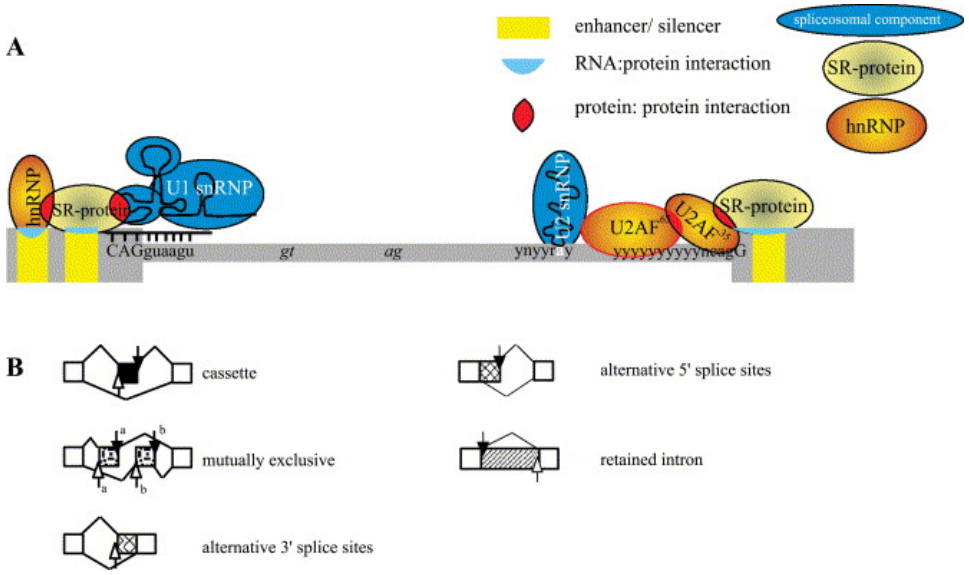



**Figure 1.5 Splice-site selection and types of alternative splicing events.** (A) Exons are indicated as boxes, the intron as a thick line. Splicing regulator elements (enhancers or silencers) are shown as yellow boxes in exons or as thin boxes in introns. The 5′ splice site (CAGguaagu) and 3′ splice site $(y)_{10}$ncagG, as well as the branch point (ynyyray), are indicated (*y*=c or u, *n*=a, g, c or u). Upper-case letters refer to nucleotides that remain in the mature mRNA. Two major groups of proteins, hnRNPs (orange) and SR or SR-related proteins (blue), bind to splicing regulator elements. Factors at the 3′ splice site include U2AF, which facilitates binding to U2 snRNP to the branchpoint sequence. In exons with weak polypyrimidine tracts, the binding of U2AF is facilitated by the SR proteins binding to exonic enhancers. (B) Types of alternative splicing events: Alternative exons are shown as boxes with different shading. Flanking constitutive exons are shown as white boxes. The open arrow indicates the position of the alternative 3′ splice site analyzed; a closed arrow indicates the position of the 5′ splice sites analyzed. Taken from (Stamm, Ben-Ari et al. 2005)

#### *1.1.2.3.2 Cis Splicing regulatory elements*

Introns are marked by short loosely conserved sequences located near their 5' and 3' termini. It is remarkable that the splicing assembly can with such precision locate relatively small exons in a pre-mRNA, excise huge intervening introns and splice exons to generate mature mRNA. However the sequences of the splice sites and the branch points are clearly insufficient for the intron recognition and removal. Auxiliary splicing signals play an important role in splice site recognition. These auxiliary signals are classified according to their location (exonic or intronic) and their functional effects on

splicing (activation or repression) as exonic splicing enhancer (ESE), exonic splicing silencer (ESS), intronic splicing enhancer (ISE), and intronic splicing silencer (ISS).

#### *1.1.2.3.3 Alternative splicing and regulation*

The released draft of the human genome sequence revealed a surprisingly low number of genes, with more recent estimates of under 25 000 genes. To generate the estimated proteome of at least 250 000 proteins, diverse posttranscriptional mechanisms are used. One major mechanism is alternative pre-mRNA splicing (Lander ES 2001). ~ 50%–75% of multi-exon genes undergo alternative splicing (AS), generating multiple mRNA isoforms and greatly increasing human proteomic diversity (Lander ES 2001; Modrek, Resch et al. 2001). Alternative splicing can generate an astonishing diversity of proteins. The Drosophila Dscam gene can generate more than 38,000 different mRNAs by alternatively splicing exons at four positions (Celotto and Graveley 2001).

The mechanism of splicing has been determined in great detail ((Jurica and Moore 2003; Timothy W. Nilsen 2003). In contrast, it is not yet fully understood how splice sites are selected. The major problem is the degeneracy of splicing regulatory sequences, such as the 5′, 3′ splice sites, branch points and exonic/intronic sequence elements. These can only be described as consensus sequences that are loosely followed (Black 2003). As a result, it is not possible to accurately predict splicing patterns from genomic sequence. The accurate recognition of splice sites in vivo is the result of a combinatorial regulatory mechanism (Smith and Valcarcel 2000).

Many effect are associated with final splicing patterns. These include ligand affinity and signaling capabilities of receptors, intracellular localization of proteins, ion channel properties and DNA binding affinity and activity of the transcription factors. A substantial part of the alternatively spliced exons show that tissue or cell type specific patterns of expression are regulated during development or in response to external stimuli. For example, serum deprivation alters usage of the serine/arginine-rich protein 20 (SRp20) exon 4 (Jumaa, Guenet et al. 1997). Neuronal activity changes the alternative splicing pattern of clathrin light chain B, the NMDAR1 receptor, and c-fos. Finally, programmed cell death is concomitant with a change in the alternative splicing patterns of several cell death regulatory proteins (reviewed in Jiang and Wu, 1999).

#### *1.1.2.3.4 Splicing factors*

The splicing factor could bind to several proteins and either promote the splice site usage (enhancer) or suppress it (silencer). Proteins binding to regulatory sequence elements can be classified into two groups: serine/arginine rich (SR) proteins or heterogeneous nuclear ribonucleoproteins (hnRNPs). RNA recognition by these proteins occurs through one or more RNA recognition motifs (RRMs). It is well known that SR or SR-like proteins can promote the formation of complexes containing U1 snRNP bound to the 5' splice site and U2 snRNP bound to the pre-mRNA branch site. They can also facilitate the recruitment of U4/U6 and U5 snRNPs. In addition, SR and SR-like proteins can bridge the introns by interacting with themselves and the core spliceosomal components. In short, serine/arginine rich domains (RS domains) serve as protein-protein interaction modules that recruit other components of the splicing machinery (Bourgeois, Lejeune et al. 2004) or could also be involved in direct RNA contacts to promote spliceosome formation (Shen and Green 2004).

#### *1.1.2.3.5 SR and SR-related proteins*

SR proteins constitute a family of non-snRNP proteins required for pre-mRNA splicing. They are highly conserved throughout the metazoan. All SR proteins have a characteristic structural organization, which consists of one or two N-terminal RNA recognition motifs [RRMs or so called RNA binding domains (RBDs)], that function in sequence-specific RNA binding, and a variable-length C-terminal arginine/serine-rich (RS) domain required for protein-protein interaction with other RS domains. Ten SR proteins have been identified so far in the human SR protein family. A number of additional splicing factors containing RS domains are structurally and functionally related to SR proteins and are collectively referred to as SR related (SR like) proteins or SRrps (Table.1.4). However the definition of a SR related protein is not fixed.

Although SR proteins display distinct RNA binding specificities, the consensus sequences that they recognize are rather degenerate. In several cases, sequences identified as binding sites for one SR protein can also be recognized by other SR proteins. The overlapping and promiscuous RNA binding specificities of SR proteins may partially account for their apparent redundancy in their function (Table.1.3).

**Table 1.3 binding sequences of SR related proteins**

| Protein | Binding site | Method | Reference |
|---|---|---|---|
| ASF/SF2 | RGAAGAAC | SELEX | Tacke R, Manley JL, EMBO J. 1995 17;14(14):3540-51. |
| | AGGACRRAGC | SELEX | above |
| | SRSASGA | Functional | Liu H-X, Zhang M, Krainer AR. Genes & Dev 12:1998–2012 |
| SRp30c | AGSAS | SELEX | Paradis C, Cloutier P, etal, RNA. 2007 Aug;13(8):1287-300 |
| Tra2-beta1 | (GAA)n | SELEX | Tacke R, Manley JL, EMBO J. 1995 17;14(14):3540-51. |
| PTB | CUCU | Functional | Oberstrass FC, Auweter SD, Science. 2005 Sep 23;309(5743):2054-7. |
| RBMY | MUCAA | Functional | Skrisovska L, Bourgeois CF, EMBO Rep. 2007 Apr;8(4):372-9. |
| SC35 | AGSAGAGUA | SELEX | Tacke R, Manley JL, EMBO J. 1995 17;14(14):3540-51. |
| | GUUCGAGUA | SELEX | above |
| | UGUUCSAGWU | SELEX | above |
| | GWUWCCUGCUA | SELEX | above |
| | GGGUAUGCUG | SELEX | Cavaloc Y, Bourgeois CF,et,al. 1999. RNA 5:468–483 |
| | GAGCAGUAGKS | SELEX | above |
| | AGGAGAU | SELEX | above |
| | GRYYCSYR | Functional | Liu HX, Chew SL. et.al 2000, Mol Cell Biol 20:1063–1071 |
| 9G8 | AGACKACGAY | SELEX | Cavaloc Y, Bourgeois CF,et,al. 1999. RNA 5:468–483 |
| | ACGAGAGAY | SELEX | above |
| SRp40 | UGGGAGCRGUYRGCUCGY | SELEX | Tacke R, Manley JL, EMBO J. 1995 17;14(14):3540-51. |
| | ACDGS | Functional | Liu H-X, Zhang M, Krainer AR. 1998. Genes & Dev 12:1998–2012 |
| SRp20 | CAUC | Functional | Hargous Y, Hautbergue GM, etal, EMBO J. 2006 Nov 1;25(21):5126-37 |
| SRp55 | USCGKM | SELEX | Liu H-X, Zhang M, Krainer AR. 1998. Genes & Dev 12:1998–2012 |
| FOX1 | UGCAUGU | Functional | Auweter SD, Fasan R, et.al, EMBO J. 2006 Jan 11;25(1):163-73 |
| DX16 | CCGUNUNKNW | SELEX | Yuan L, Zhou J, etal, Mol Cell Biochem. 2007 Aug;302(1-2):119-24. |
| B52 | GRUCAACCDNGGCGAACNG | SELEX | Shi H, Hoffman BE. et,al, 1997, Mol Cell Biol 17:2649–2657 |

N: any nucleotide; R: purine; Y: pyrimidine; S: G or C; K: U or G; W: A or U; D: A, G, or U; M: A or C

#### *1.1.2.3.6 Heterogeneous nuclear ribonucleoproteins (hnRNPs)*

Primary transcripts synthesized by RNA polymerase II are termed heterogeneous nuclear RNAs (hnRNAs) because of their diverse composition. The nascent hnRNAs are immediately bound by a family of proteins, termed hnRNPs, resulting in the formation of hnRNP complexes. The hnRNP proteins are among the most abundant proteins in the nucleus. They all share a common structure containing RNA binding domains and auxiliary domains, which are composed of clusters of certain amino acids, and might mediate protein-protein interaction or facilitate protein localization.

**Table 1.4 SR, SR related proteins and hnRNPs**
Data from (Krecic and Swanson 1999; Graveley 2000) and our collection

| *SR proteins* | *SR-related proteins* | ***hnRNPs*** |
|---|---|---|
| SRp20 | *snRNP components* | hnRNP A0 (CUG-BP) |
| SC35 | U1-70K | hnRNP A1 |
| SRp46 | U5-100K | hnRNP A2/B1 |
| SRp54 | U4/U6-27K | hnRNP A3 |
| SRp30c | hLuc7p | hnRNP C1 |
| ASF/SF2 | *U2 Auxiliary factor* | hnRNP C2 |
| SRp40 | U2AF65 | hnRNP D |
| SRp55 | U2AF35 | hnRNP E1 (PCBP1) |
| SRp75 | *Splicing regulators* | hnRNP E2 (PCBP2) |
| 9G8 | hTra2α | hnRNP F |
| hTra2β | P54 | hnRNP G |
| | FUSIP1 | hnRNP G-T |
| | SR-A1 | hnRNP H1 |
| | SRrp53 | hnRNP H2 |
| | Clasp | hnRNP H3 |
| | SRrp86 | hnRNP I (PTB) |
| | SRrp129 | hnRNP K |
| | *Splicing coactivators* | hnRNP L |
| | SRm160 | hnRNP M |
| | SRm300 | hnRNP P2 (TLS/FUS) |
| | *RNA helicases* | hnRNP R |
| | hPrp16 | hnRNP U |
| | HRH1 | hnRNP Q |
| | *Protein kinases* | HAP/SAF-B |
| | Clk/Sty 1-4 | |

#### *1.1.2.3.7 Other splicing factors*

In addition to SR proteins and hnRNPs, other splicing factors also play an important role in regulating alternative splicing. The neuro-oncological ventral antigen-1

(NOVA-1) can bind to specific intronic sequences within glycine α 2 receptor (GlyR α 2) and gamma aminobutyric acid (A) receptor (GABAA) pre-mRNAs, and stimulate the inclusion of neuronspecific exons (Jensen, Dredge et al. 2000). Another intriguing example is the apoptosis-promoting protein TIA-1, which can activate splicing of human growth factor receptor 2 (FGFR-2) and Drosophila male-specific-lethal 2 (msl-2) pre-mRNAs by binding to the intronic U-rich sequences immediately downstream of the weak 5' splice site and facilitating the recruitment of the U1 snRNP to the 5' splice site. (Del Gatto-Konczak, Bourgeois et al. 2000) Considering that factor could cross talk among different cell process, not many "pure" splicing factors are identified.

#### *1.1.2.3.8 Alternative splicing database*

Data on alternative splicing fall into two categories.

(i) experimentally determined and characterized splice events from specific genes, as reported in bibliography databases such as MEDLINE, or in curated nucleotide and protein sequence databases such as EMBL and SWISS-PROT. Efforts to create data sets based on these data include:

Alternative Exon Database: http://www.ebi.ac.uk/asd/aedb/

ASDB: http://cbcg.nersc.gov/asdb

AsMamDB: http://166.111.30.65/ASMAMDB.html

(ii) computationally determined splice events observed through examination of alignments of EST/cDNA sequences with one another or with genomic DNA sequences—these include:

AltExtron: http://rhodos.bioinf.mdc-berlin.de/asforms

ASAP: http://www.bioinformatics.ucla.edu/ASAP

Hollywood: http://hollywood.mit.edu/Login.php

HDBAS: http://jbirc.jbic.or.jp/h-dbas/

Prosplicer: http://prosplicer.mbc.nctu.edu.tw/

MASSE: http://maase.genomics.purdue.edu/

DEGEST: http://genome.ewha.ac.kr/DEGEST/

EDAS: http://www.ig-msk.ru:8005/EDAS/

PALSDB: http://binfo.ym.edu.tw/passdb/index.html

Arabidopsis: http://www.tigr.org/tdb/e2k1/ath1/altsplicing/splicing_variations.shtml

ARG: http://statgen.ncsu.edu/asg/

ASHESdb: http://sege.ntu.edu.sg/wester/ashes/

EASED: http://eased.bioinf.mdc-berlin.de/

ECgene: http://genome.ewha.ac.kr/ECgene/

Intronerator (*C. elegans*): http://hgwdev-hiram.cse.ucsc.edu/IntronWS120/index.html

SpliceDB:

http://www.softberry.com/berry.phtml?topic=splicedb&group=data&subgroup=spldb

spliceNest: http://splicenest.molgen.mpg.de/

ASDB: http://hazelton.lbl.gov/~teplitski/alt/

ASTRA: http://alterna.cbrc.jp/

TassDBl: http://helios.informatik.uni-freiburg.de/TassDB/

FastDB: http://www.fast-db.com/fastdb2/frame.html

ASIP (plant): http://www.plantgdb.org/ASIP/

### 1.1.3 Export

mRNA splicing is intimately connected to the export of mature transcripts from the nucleus. Recent study of three yeast RBP export factors — Nab2, Npl3 (also known as Nop3) and Nab4 (also known as Hrp1), which function as hnRNP shuttling proteins — identified three discrete subpopulations of mRNAs, encoding proteins that are involved in transcription, ribosomal biogenesis and intermediary metabolism, respectively(Kim Guisbert, Duncan et al. 2005; Keene 2007).

Injection experiments in *Xenopus* oocytes have shown that spliced mRNAs were more efficiently exported than their synthetic intron-less counterparts, pointing to a functional coupling of splicing and export (Luo and Reed 1999). Simplified model in metazoan has been shown in Fig.1.6 for mRNP quality control steps associated with the functional couplings of transcription, splicing and export(Kohler and Hurt 2007).

The transport of RNA molecules from the nucleus to the cytoplasm is fundamental for gene expression. The different RNA species that are produced in the nucleus are exported through the nuclear pore complexes via mobile export receptors. Small RNAs (such as tRNAs and microRNAs) follow relatively simple export routes by

binding directly to export receptors. Large RNAs (such as ribosomal RNAs and mRNAs) assemble into complicated ribonucleoprotein (RNP) particles and recruit their exporters via class-specific adaptor proteins. Export of mRNAs is unique as it is extensively coupled to transcription (in yeast) and splicing (in metazoa), Orthologous proteins or complexes between yeast and metazoan are shown in Fig.1.6 which taken from(Kohler and Hurt 2007). Several SR proteins (SRp20, 9G8) are recruited in a hyperphosphorylated form to play function in the splicing machinery. They remain bound to the spliced transcript and are exported to the cytoplasm, where they dissociate from the transcript and/or play role in translation, then are re-imported to nuclear(Caceres, Screaton et al. 1998). However the mechanism of SR protein in export is not clear.

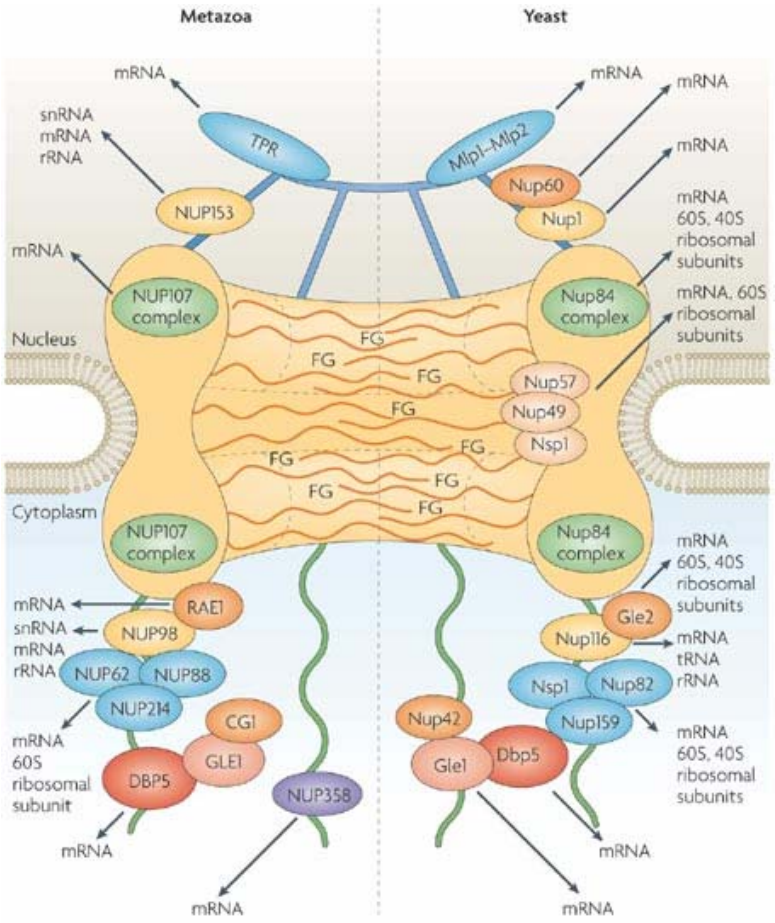



**Figure 1.6 Nuclear pore complexes (NPCs)** The NPC is formed by ~ 30 different nuclear pore proteins (nucleoporins). Orthologous proteins

or complexes between yeast and metazoa are shown in the same colour. Taken from (Kohler and Hurt 2007).

### 1.1.4 Coupling of pre-mRNA processing to translation

A comprehensive proteomic analysis of the human spliceosome ((Zhou, Licklider et al. 2002; Jurica and Moore 2003) reveals that at least 30 out of the 145 spliceosomal proteins are either known or candidate participants in the coupling between splicing and other gene expression steps (Kornblihtt, de la Mata et al. 2004). For example, the Wilms tumour 1 (WT1) gene is alternatively spliced to produce two protein isoforms: one regulates transcription and the other binds CTE-containing mRNAs in the cytoplasm and increases their translation during development.(Keene 2007)

A typical sample comes from SR(serine/arginine-rich) proteins. SR proteins were originally identified as essential splicing factors. A Model for the multiple roles of SR proteins in mRNA metabolism is shown in Fig.1.7.

SR proteins can potentially regulate the fate of alternatively spliced mRNAs and may provide an important mechanism for regulating mRNA surveillance (Zhang and Krainer 2004). Recent work has implicated these proteins in numerous additional steps of mRNA metabolism, including nuclear export, RNA stability, mRNA quality control, and translation (Huang and Steitz 2005).

Shuttling SR protein SF2/ASF can associate with translating ribosomes and enhance translation of reporter mRNAs both in vivo and in vitro. In addition, in endogenous, cytoplasmic SF2/ASF associated with the translation machinery is hypophosphorylated, suggesting that the phosphorylation state of the Arg-Ser-rich (RS) domain may influence the role of SF2/ASF in cytoplasmic RNA processing (Sanford, Gray et al. 2004; Sanford, Ellis et al. 2005).

Another nucleo-cytoplasmic SR protein, SRp20, functions in internal ribosome entry site (IRES)-mediated translation of a viral RNA. It interacts with the cellular RNA-binding protein, PCBP2, a protein that binds to IRES sequences within the genomic RNAs of certain picornaviruses and is required for viral translation (Bedard, Daijogo et al. 2007).

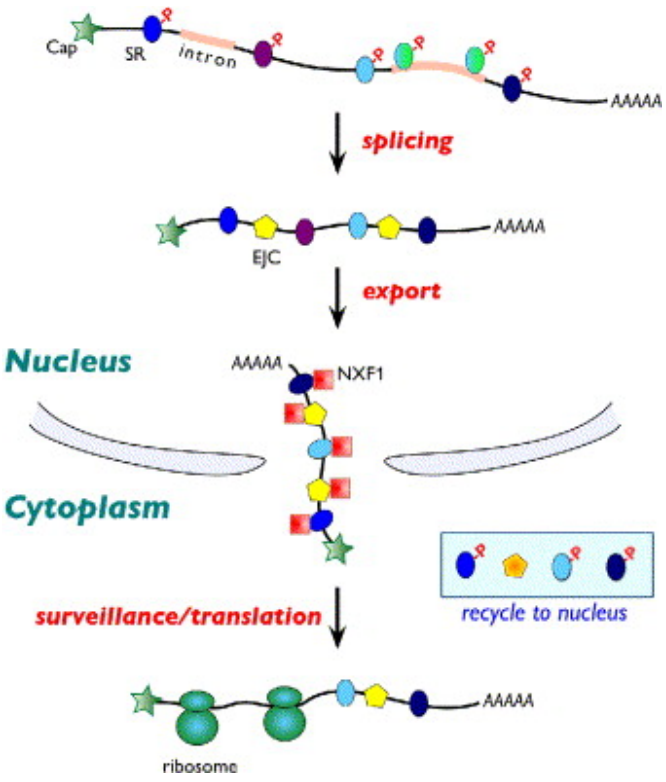



**Figure 1.7 A Model for the Multiple Roles of SR Proteins in mRNA Metabolism**

Hyperphosphoryalted SR proteins are enriched on exonic splicing enhancers (ESE) and participate in splicing. Then SR protein are dephosphorylated to hypophosphorylated forms and in the deposition of exon junction complexes (EJC) upstream of exon-exon boundaries in the spliced RNA. Some shuttling SR proteins then help to recruit the export adapter NXF1 to export mRNA with mRNP through the nuclear pore. The mRNAs that contain inappropriate nonsense codons are subjected to non sense mediated decay (NMD). Translation is followed by RNP remodeling, which lead to the release and re-import of re-phosphorylated SR proteins into the nucleus. Taken from (Huang and Steitz 2005)

### 1.1.5 RNA Degradation

The cell has evolved multiple surveillance mechanisms to assure that only perfect mRNAs are ultimately translated into proteins. This multitude of overlapping quality control pathways guarantees that the cell has mechanisms in place to pinpoint the many types of errors that could arise in the transcript as it is polymerized, processed, exported and translated. The process includes nonstop decay, nonsense-mediated decay (NMD), Staufen-mediated decay (SMD) and mRNA degradation after translation (Fig. 1.8) (Fasken and Corbett 2005).

Both translation-dependent RNA surveillance mechanisms, NMD and nonstop decay, depend on some method of decoding the transcript to detect the presence of stop codons. NMD is best known for its ability to target transcripts that contain PTCs for decay. It is triggered by exon-junction-complex (EJC), that have been formed during pre-RNA processing, being downstream of the nonsense codon. Normally, these EJCs are removed during the first round of translation of the mRNA, but in the case of a premature stop codon, they are still present on the mRNA (Maquat 2005). In contrast, nonstop decay targets transcripts that lack a stop codon for destruction. The non-stop decay pathway releases ribosomes that have reach the far 3'end of a mRNA and guides the mRNA to the exosome complex for degradation. (Fasken and Corbett 2005).

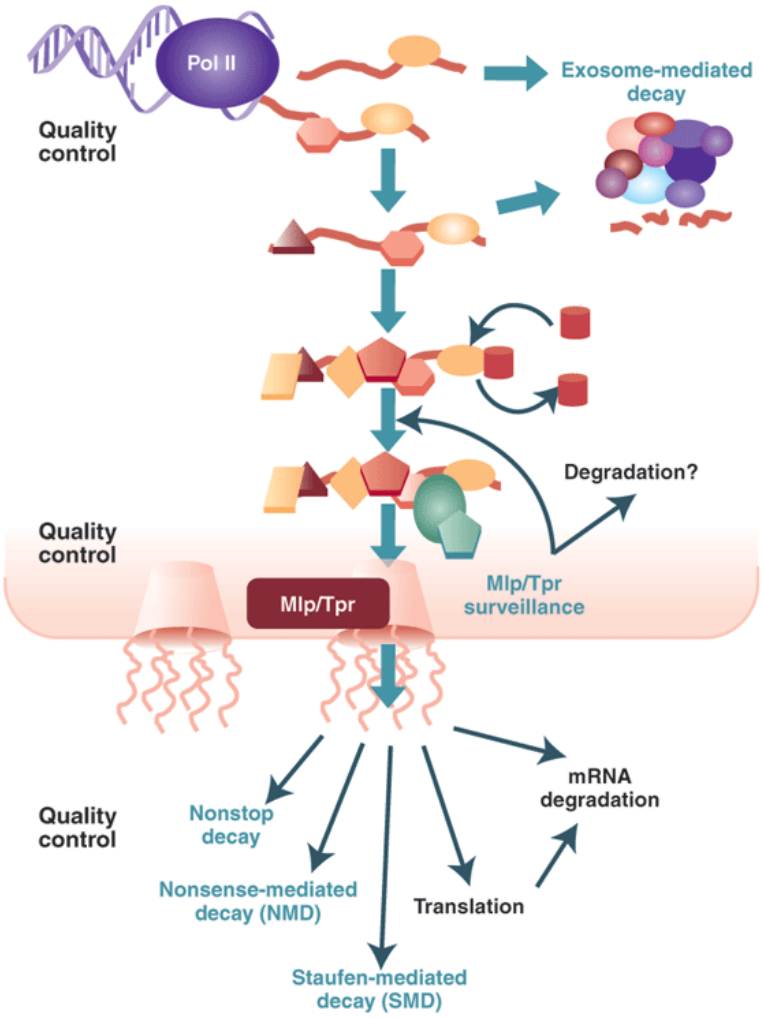



**Figure 1.8 Schematic of mRNA biogenesis and quality control.**

The known steps in mRNA processing where transcripts can be subject to quality control are shown in picture. Export factors bind prior to export, are indicated by the green shapes and each quality control step in the model is indicated by a green arrow. The different degration pathway in cytoplasm such as nonstop decay, nonsense-mediated dacay and staufen-mediated decay are listed. Taken from (Fasken and Corbett 2005)

# 2 Research overview

In this work, I investigated how the SR-like protein Tra2-beta1 regulated splicing and translation.

First, data on alternative splicing and splicing factors was systematically collected and annotated. It resulted in the improvement of a database of alternatively spliced exons collected from literature (AEdb) for ASD (Alternative Splicing Database)(Stamm, Riethoven et al. 2006). Based on this data, a custom splice array covering around 300 splicing factors was collected. This is a platform for a further research of alternative splicing and splicing factors.

An increasingly number of diseases are either associated with or caused by changes in alternative splicing. These diseases can be caused by mutation in regulatory sequences of the pre-mRNA or by changes in the concentration of trans-acting factors (Novoyatleva, Tang et al. 2006). Among them, we concentrated on the Alzheimer's Disease (AD) and breast cancer. We found that the regulation of *CD44* gene by *tra2-beta1* is associated with tumor progression and metastasis in breast cancer(Watermann, Tang et al. 2006). In sporadic AD patients, the amount of mRNAs of tau isoforms including exon 10, the htra2-beta1 isoform and an inactive form of clk2 are significantly increased. It suggest that a mis-regulation of alternative splicing seems to contribute to sporadic AD (Glatz, Rujescu et al. 2006).

Previous research shows that TRA2-BETA1 accumulates in the cytosol under cellular stress conditions. Yeast two hybrid studies showed that TRA2-BETA1 directly binds to RPL3, a protein of the large ribosomal subunit that plays a role in peptidyltransferase center formation. To identify the regulation of Translation by *tra2-beta1*, we confirmed the interaction between TRA2-BETA1 and RPL3 using in vitro pull down assays with recombinant proteins. Using sucrose gradient fractionation, we found that TRA2-BETA1 co-sediments with ribosomes and polysome fractions. Furthermore, CLIP (RNA Cross-Linking and ImmunoPrecipitation) of TRA2-BETA1 shows that most of the Tra2-beta1 targets from cytoslic RNA is ribosomal RNA. The CLIP targets were localized mainly on the large subunit of the ribosome, near the RPL3 binding sites in the 28S rRNA. TRA2-BETA1 with an inserted nuclear export signal strongly activates

luciferase reporter constructs that contain a TRA2-BETA1 binding motif, and this stimulation is regulated by the dephosphorylation in its PP1 (Protein Phosphatase 1) binding site.

# 3 Materials and methods

## *3.1 Materials*

### 3.1.1 Chemicals

| Product | Supplier | Product | Supplier |
|---|---|---|---|
| Acetone | Merck | Methanol | Carl Roth GmbH |
| Acetic acid | Carl Roth GmbH | [$^{35}$S]-Methionine | Amersham |
| 30% Acrylamide/Bis 37.5:1 | Sigma-Aldrich | Ni-NTA Agarose | Qiagen |
| | | Nonidet P-40 / Igepal CA-630 | Sigma-Aldrich |
| 40% Acrylamide/Bis 19:1 | Carl Roth GmbH | | |
| | | NTPs | Roche |
| Agar (Select Agar) | Sigma-Aldrich | dNTPs | Invitrogen, Sigma-Aldrich |
| Agarose UltraPure | Invitrogen | | |
| Ammoniumpersulfate | Sigma-Aldrich | Paraformaldehyde | Merck |
| Ampicillin | Sigma-Aldrich | PEG 3500 | Sigma-Aldrich |
| Aprotinin | Sigma-Aldrich | Pepstatin | Sigma-Aldrich |
| γ-[$^{32}$P]-ATP | Hartmann Analytics | Perhydrol 30% $H_2O_2$ | Merck |
| BSA | Merck | Phenol: Chloroform: Isoamyl alcohol | Sigma-Aldrich |
| Boric acid | Carl Roth GmbH | | |
| Bradford reagent (BioRad Protein Assay) | BioRad | PMSF | Sigma-Aldrich |
| | | Ponceau S solution | Sigma-Aldrich |
| Brilliant Blue R 250 | Sigma-Aldrich | Potassium acetate | Riedel de Haën |
| Bromophenol blue | Merck | Potassium chloride | Merck |
| Calciumchloride | Merck | Potassium dihydrogen phosphate | Merck |
| Cellfectin | Invitrogen | | |
| Chloramphenicol | Sigma-Aldrich | Protease Inhibitor Cocktail | Sigma-Aldrich |
| Chloroform: Isoamyl alcohol | Sigma-Aldrich | | |
| | | Protein A Sepharose | Amersham |
| Crystal violet | Merck | PTP 1B Inhibitor | Calbiochem |
| Dextrose | Sigma-Aldrich | RNase Inhibitor | Roche |
| DMSO | Sigma-Aldrich | SDS | Carl Roth GmbH |
| DTT | Merck | Sepharose CL-4B | Pharmacia |
| EDTA | Carl Roth GmbH | Silver nitrate | Merck |
| EGTA | Merck | Sodium acetate | Merck |
| Ethanol | Carl Roth GmbH | Sodium chloride | Carl Roth GmbH |
| Ethidium bromide | Sigma-Aldrich | Sodium deoxycholate | Sigma-Aldrich |
| Ficoll 400 | Fluka | Sodium dihydrogen phosphate | Merck |
| Formaldehyde | Merck | | |
| Forskolin | Calbiochem | Sodium fluoride | Sigma-Aldrich |
| Gelatin | Carl Roth GmbH | Sodium hydroxide | Merck |
| Gel/Mount | Biomeda | Sodium orthovanadate | Sigma-Aldrich |

| Product | Supplier | Product | Supplier |
|---|---|---|---|
| Gentamycin | Sigma-Aldrich | Sodium pyrophosphate | Merck |
| Glycerol | Sigma-Aldrich | di-Sodiumhydrogen phosphate | Merck |
| Glycerol 2-phosphate | Sigma-Aldrich | | |
| Glycine | Carl Roth GmbH | Sodium Thiosulphate | Merck |
| Guanidine hydrochloride | Fluka | Sucrose | Carl Roth GmbH |
| | | Superfect | Qiagen |
| Heparin | Sigma-Aldrich | Tautomycin | Calbiochem |
| HEPES | Sigma-Aldrich | TEMED | Sigma-Aldrich |
| HiPerfect | Qiagen | Trichloro acetic acid | Riedel de Haën |
| Hydrochloride | Merck | Tris base | Sigma-Aldrich |
| Imidazole | Carl Roth GmbH | TRIzol | Sigma-Aldrich |
| Isopropanol | Carl Roth GmbH | Triton X-100 | Carl Roth GmbH |
| p-Iodophenol | Sigma-Aldrich | Tryptone | Sigma-Aldrich |
| Kanamycin | Sigma-Aldrich | Tween 20 | Sigma-Aldrich |
| Leupeptin | Sigma-Aldrich | Urea | Merck |
| Luminol | Sigma-Aldrich | α-[$^{32}$P]-UTP | Hartmann Analytics |
| Magnesium chloride | Merck | Yeast Extract (Select Yeast extract) | Sigma-Aldrich |
| Magnesium sulfate | Merck | | |
| β-Mercaptoethanol | Merck | Xylene cyanole FF | Merck |

### 3.1.2 Commercially available Kits

| Product | Supplier | Product | Supplier |
|---|---|---|---|
| QIAGEN Plasmid Maxi Kit | Qiagen | QIAprep Spin M13 Kit | Qiagen |
| QIAEX II Gel Extraction Kit | Qiagen | Miniquick Spin RNA columns | Roche |
| QIAquick Gel extraction Kit | Qiagen | JetStar plasmid Maxi Kit | Genomed |
| RNeasy mini kit | Qiagen | TNT® T7 Coupled Reticulocyte Lysate System | Promega |
| TOPO TA cloning Kit | Invitrogen | | |

### 3.1.3 Enzymes, proteins and standards

| Product | Supplier | Product | Supplier |
|---|---|---|---|
| abl protein tyrosine kinase | New England Biolabs | Pwo Polymerase | PeqLab |
| | | Restriction endonucleases | New England Biolabs, Fermentas |
| Antarctic Phosphatase | New England Biolabs | | |
| | | RNase A | Roche |
| Benzonase | Sigma | Taq DNA polymerase | Invitrogen, PeqLab |
| Calf Intestinal alkaline phosphatase (CIP) | New England Biolabs | T4 DNA Ligase | New England Biolabs |
| DNase I | Roche | T4 Polynucleotide Kinase | New England Biolabs |
| MultiMark® Multi-colored Protein Standard | Invitrogen | | |
| | | T7 DNA Polymerase | New England Biolabs |
| peqGOLD Protein-Marker IV (Prestained) | Peqlab | | |
| | | T7 RNA Polymerase | Roche |

| | | | |
|---|---|---|---|
| Platinum Pfx polymerase | Invitrogen | TrackIt™ 100bp DNA ladder | Invitrogen |
| Precision Plus Protein Prestained Standards | BioRad | TrackIt™ 1kb DNA ladder | Invitrogen |
| Protein Phosphatase 1 | New England Biolabs | SuperScript II | Invitrogen |
| | | | |

### 3.1.4 Cell lines and media

| Cell Line | Description | ATCC number |
|---|---|---|
| Cos-7 | African green monkey kidney SV40 transformed | CRL-1651 |
| Hela | Homo sapiens cervical cancer cells | CCL-2 |
| HEK293 | Human embryonic kidney transformed with adenovirus 5 DNA | CRL-1573 |
| Neuro-2a | Neuroblastoma from mouse brain | CCL-131 |
| SF9 insect cells | Spodoptera frugiperda ( fall armyworm ) | CRL-1711 |

The eukaryotic cell lines were cultured in DMEM supplemented with 10% fetal calf serum (both from Invitrogen). For subculturing, 1 x Trypsin/EDTA (Invitrogen) was used. Insect cells were cultured in TNM-FH (BD Biosciences).

### 3.1.5 Bacterial strains and media

| Strain | Genotype | Reference |
|---|---|---|
| *E.coli* XL1-Blue MRF' | Δ(*mcr*A)183 Δ(*mcr*CB-*hsd*SMR-*mrr*) 173 *end*A1 *sup*E44 *thi*-1 *rec*A1 *gyr*A96 *rel*A1 *lac* [F´ *pro*AB *lac*IqZΔM15 Tn10 (Tetr)] | (Bullock, WO, Fernandez, JM and Short, JM , 1987) |
| *E.coli* CJ 236 | F´ cat (pCJ105 = pOX38::cat= FΔ(HindIII)::cat [$Tra^+$ $Pil^+$ $Cam^R$]/ ung-1 relA1 dut-1 thi-1 spoT1 | (Kunkel, Roberts et al. 1987) |
| One Shot® Top10 Chemically competent *E.coli* | $F^-$*mcr*A Δ(*mrr-hsd*RMS-*mcr*BC) Φ80*lac*ZΔM15 Δ*lac*X74 *rec*A1 *ara*D139 Δ(*ara-leu*)7697 *gal*U *gal*K *rps*L ($Str^R$) *end*A1 *nup*G | Invitrogen |
| DB3.1 | F- gyrA462 endA1 Δ(sr1-recA) mcrB mrr hsdS20(rB-, mB-) supE44 ara-14 galK2 lacY1 proA2 rpsL20(SmR) xyl-5 λ- leu mtl1 | Invitrogen |
| MAX Efficiency® DH10Bac™ | $F^-$*mcr*A Δ(*mrr-hsd*RMS-*mcr*BC) Φ80*lac*ZΔM15 Δ*lac*X74 *rec*A1 *end*A1 *ara*D139 Δ(*ara-leu*)7697 *gal*U *gal*K $\lambda^-$ *rps*L *nup*G /pMON14272 / pMON7124 | Invitrogen |

| LB Medium (1L) | | LB Agar (1L) | |
|---|---|---|---|
| 10 g | NaCL | 10 g | NaCL |
| 10 g | Tryptone | 10 g | Tryptone |
| 5 g | Yeast extract | 5 g | Yeast extract |
| | | 20 g | agar |

### 3.1.6 Antibiotics

| Antibiotic | Stock concentration | Working concentration | |
|---|---|---|---|
| | | Liquid culture | Agar plates |
| Ampicillin | 50 mg/ml | 100 µg/ml | 100 µg/ml |
| Chloramphenicol | 30 mg/ml | 15 µg/ml | 30 µg/ml |
| Gentamycin | 10 mg/ml | 7 µg/ml | 10 µg/ml |
| Kanamycin | 20 mg/ml | 50 µg/ml | 50 µg/ml |
| Tetracyclin | 5 mg/ml | 50 µg/ml | 50 µg/ml |

### 3.1.7 Antibodies

**Primary antibodies**

| Antibody | Organism | Dilution for Western Blot | Supplier |
|---|---|---|---|
| anti-beta-actin | Rabbit | 1:2000 | abcam |
| anti-Clk2 | Rabbit | 1:500 | abcam |
| anti-CCNDBP1 | Goat | 1:1000 | abcam |
| anti-eIF4B | Rabbit | 1:1000 | Cell signaling |
| anti-EPN2 | Rabbit | 1/3000 | Gife from Camilli's lab (1) |
| anti-ERK5 | Rabbit | 1:1000 | abcam |
| anti FLAG M2 | Mouse | 1:1000 | Sigma |
| anti GAPDH | Mouse | 1.2000 | abcam |
| anti GFP | Mouse | 1:4000 | Roche |
| anti GST | Mouse | 1:1000 | abcam |
| anti mAB 104 | Mouse | 1:100 | Gift from Neugebauer's lab |
| anti hnRNP G | Rabbit | 1:2000 | custom made(2) |
| anti-Phospho-hnG | rabbit | 1:1000 | custom made(3) |
| anti PP1γ1 (C-19) | Goat | 1:200 | Santa Cruz |
| anti- PHC2 | Mouse | 1:50 | Gift from Koseki's lab (4) |

| Antibody | Organism | Dilution for Western Blot | Supplier |
|---|---|---|---|
| anti RPL3 (Italy) | Rabbit | 1:100 | Gift from Italy Tina's lab |
| anti RPL3 | Rabbit | 1:1000 | ProteinTech |
| anti SF2/ASF | Mouse | 1:200 | Zymed Laboratories |
| anti S6 ribosomal | Rabbit | 1:100 | Cell signaliing |
| anti tra2-beta+alpha (ps568) | Rabbit | 1:2000 | custom made[5] |
| anti- UEV | Rabbit | 1/500 | abcam |
| anti- WHSC1/NSD2 | Rabbit | 1/2000 | abcam |
| anti-YT521B (PK2) | Rabbit | 1:3000 | custom made[6] |

Custom made antibodies were raised against the following peptides:

(1) Peptide: NFNGTVKDDFSEFDNLRTSKKPAES (Rosenthal, Chen et al. 1999)

(2) hnRNP G peptide: RDDGYSTKD

(3) Phospho hnRNP G peptide: RDDGYPSTKD

(4) (Isono, Fujimura et al. 2005)

(5) ps568/Tra peptide: GC(StBu)SITKRPHTPTPGIYMGRPTY (Stoilov, Daoud et al. 2004)

(6) YT521-B: RSARSVILIFSVRESGKFQCG and KDGELNVLDDILTEVPEQDDECG (Rafalska, Zhang et al. 2004)

**Secondary antibodies**

| Antibody | Organism | Dilution for Western Blot | Supplier |
|---|---|---|---|
| anti-mouse IgG-HRP | Sheep | 1:10000 | Amersham |
| anti-rabbit IgG-HRP | Donkey | 1:10000 | Amersham |
| anti-goat IgG-HRP | Donkey | 1:10000 | Santa Cruz |
| CY3-conjugated anti-rabbit | | 1:500 | Dianova |
| CY3-conjugated anti-mouse | | 1:500 | Dianova |

### 3.1.8 Brain Tissues

Human Brain Tissues conforming to 5 Alzheimer's Disease patients and 5 normal controls were obtained from the Maryland Brain and Tissue bank for Developmental Disorders, University of Maryland. Different brain sections were kindly dissected by Prof. Ingmar Blümke, Erlangen.

Brain samples were obtained from the Kathleen Price Bryan Brain Bank, Durham, USA (temporal and occipital cortex; AD group: Braak III to Braak V, n = 15; control group: no tau deposites, Braak I or II, n = 10) (Result Table.4.6, Table.4.7), and from the Wuerzburg-Hirnbank, Würzburg, Germany (supplementary motocortex; 8 control patients; 7 AD patients, Braak III to VI) (Table 2, result part). Samples from the Kathleen Price Bryan Brain Bank were matched in age (t=0,421; df=22; p=0,464), sex (Chi2=0,046; df=1; p=0,831) and post mortem interval (PM) (t=-1,081; df=21; p=0,292). The brains of the Würzburg Hirnbank showed differences concerning age (t=-2,342; df=13; p=0,036), but not concerning sex (Chi2=1,727; df=1, p=0,189) and PM (t=0,395; df=13; p=0,699). All tissues were obtained in accordance with the local ethics committee procedures.

### 3.1.9 Plasmids

**Minigenes**

| Name | Backbone | Description | Reference |
|---|---|---|---|
| SV9/10L/11 | Exontrap | Tau minigene | (Gao, Memmott et al. 2000) |
| pCR3.1 MGtra | pCR3.1TA | Tra2-beta minigene | (Stoilov, Daoud et al. 2004) |
| pXB (X16) | pCRneo | SRp20 minigene | (Jumaa, Guenet et al. 1997) |
| MG CD44 | Exontrap | CD44 v4v5 minigene | (Watermann, Tang et al. 2006) |
| pSMN2 | pCI | SMN2 minigene | (Lorson, Hahnen et al. 1999) |
| MG CLK2 | Exontrap | Clk2 minigene | (Glatz, Rujescu et al. 2006) |
| MG BACE1 | Exontrap | BACE1 minigene | None |
| MG RPL3 | Exontrap | RPL3 minigene | None |

**Clones from the lab collection or outside**

| Name | Backbone | Description | Reference |
|---|---|---|---|
| pEGFP-C2 | pEGFP-C2 | CMV-promoter, Kana$^r$/Neo$^r$, f1 ori | Clontech |
| pRK5-abl | pRK5 | c-Abl Kinase | (Stoss, Novoyatleva et al. 2004) |
| c-src wt | pcDNA3.1 | c-Src Kinase | (Wong, Besser et al. 1999) |
| Sik-YF | pcDNA3 | Constitutively active sik kinase | (Derry, Richard et al. 2000) |
| pRK5-fyn | pRK5 | Fyn kinase | (Stoss, Novoyatleva et al. 2004) |
| pSVL-Syk | pSVL | Syk kinase | (Zhang, Berenstein et al. 1996) |
| CSK | pcDNA3 | CSK kinase | (Nayler, Schnorrer et al. 1998) |
| AUG1 (pcDNA3-Rlk) | pcDNA3 | Rlk kinase | (Debnath, Chamorro et al. 1999) |
| pUHG10-3 (FER) | pUHD10-3 | FerH kinase | (Hao, Ferris et al. 1991) |
| pEGFP-DYRK-1A | pEGFP-C2 | DYRK1A kinase | (Sitz, Tigges et al. 2004) |
| HA-ACK2 | pCDNA3 | HA tagged ACK2 kinase | (Yang, Lo et al. 2001) |
| Bcr-abl | pCDNA3 | bcr-abl kinase | (Warmuth, Bergmann et al. 1997) |
| RAK | pCIneo | RAK kinase | (Craven, Cance et al. 1995) |
| lyn | pCMV-SPORT6 | lyn kinase | RZPD, Berlin |
| LAR | pCMV-SPORT6 | phosphotyrosine kinase LAR | RZPD, Berlin IMAGp998P2312140Q3 |
| MEG1 | pCMV-SPORT6 | phosphotyrosine kinase MEG1 | RZPD, Berlin IMAGp998I199578Q3 |
| MEG2 | pCMV-SPORT6 | phosphotyrosine kinase MEG2 | RZPD, Berlin IMAGp998G249687Q3 |
| PTP1B | pCMV-SPORT6 | phosphotyrosine kinase PTP1B | RZPD, Berlin IMAGp998C079948Q3 |
| PTPIA2β | pCMV-SPORT6 | phosphotyrosine kinase PTPIA2β | RZPD, Berlin IMAGp998B05 |

| Name | Backbone | Description | Reference |
|---|---|---|---|
| | | | 9626Q3 |
| PTPα | pCMV-SPORT6 | phosphotyrosine kinase PTPα | RZPD, Berlin IMAGp998O079750Q3 |
| SHP2 | pCMV-SPORT6 | phosphotyrosine kinase SHP2 | RZPD, Berlin IMAGp998K169625Q3 |
| STEP | pCMV-SPORT6 | phosphotyrosine kinase STEP | RZPD, Berlin IMAGp998A0212719Q3 |
| pEGFP-hnRNP G-C2 | pEGFP-C2 | EGFP tagged full length rat hnRNP G | Heinrich submitted |
| 9G8-pET28a | pET28a | Invitro translation of 9G8 | S. Kishore |
| Pcr-Flag-SF2Flag | pCR | Invitro translation of ASF/SF2 | Krainer AR |
| pEGFP-hTra2-beta | pEGFP-C2 | EGFP tagged human Tra2-beta1 | (Beil, Screaton et al. 1997; Nayler, Cap et al. 1998; Nayler, Schnorrer et al. 1998; Nayler, Stratling et al. 1998) |
| pEFGP-hTra2-beta1-RS1,2,3A | pEGFP-C2 | EGFP tagged human Tra2-beta1 with all serine residues in the first or/and second RS domain mutated to alanine | Y. Tang |
| pEFGP-hTra2-beta1-RS1,2,3E | pEGFP-C2 | EGFP tagged human Tra2-beta1 with all serine residues in the first or/and second RS domain mutated to glutamic acid | Y. Tang |
| hTra2-beta-HA | pFastBac-HTa | Human Tra2-beta-1 in *Drosophila* vector for generating bacmid | S. Kishore |

**Newly made clones**

| Name | Backbone | Description |
|---|---|---|
| TOPO-C313 | pCR4 | SMAR element |
| TOPO-hCLK2 | pCR4 | CLK2 minigene |
| Trap-hClk2 | pET01 Exontrap | CLK2 minigene |
| TOPO-MGclk2RT-EB | pCR4 | CLK2 minigene |
| TOPO-MGclk2RT-EB skipped | pCR4 | CLK2 minigene |
| TOPO-MGclk2RT-EB2 | pCR4 | CLK2 minigene |
| Topo-MusCD44-8 | pCR4 | CD44 isoforms from breast cancer |
| Topo-Mus-CD44-9 | pCR4 | CD44 isoforms from breast cancer |

| Name | Backbone | Description |
|---|---|---|
| Topo-Mus-CD44-11 | pCR4 | CD44 isoforms from breast cancer |
| Topo-Mus-CD44-17 | pCR4 | CD44 isoforms from breast cancer |
| Topo-Mus-CD44-18 | pCR4 | CD44 isoforms from breast cancer |
| Topo-MusCD44-19 | pCR4 | CD44 isoforms from breast cancer |
| opo-MusCD44-20 | pCR4 | CD44 isoforms from breast cancer |
| Topo-MusCD44-23 | pCR4 | CD44 isoforms from breast cancer |
| Topo-MusTra2a | pCR4 | Tra2 alpha isoforms from breast cancer |
| TOPO-hCD44-EFG | pCR4 | CD44 minigene |
| TOPO-CD44-40 | pCR4 | CD44 minigene |
| pcr3.1MGTra del SMAR | pCR 3.1 TA | SMAR element |
| pcr3.1MGTra pstI largest fragment | pCR 3.1 TA | SMAR element |
| SARE-MGTra del SMAR | pCR 3.1 TA | SMAR element |
| SARE-MGTra pstI largest | pCR 3.1 TA | SMAR element |
| topo-2flp-sfiI | pCR4 | SMAR element |
| Trap-clk2 | pET01 Exontrap | CLK2 minigene |
| TOPO-NES | pCR4 | Nuclear Export Signal |
| tra2-NES-fs | pEGFP-C2 | Nuclear Export Signa frame shift |
| MGTra-SMAR-TRANS | pCR 3.1 TA | SMAR element |
| topo-hTra ex1.4 | pCR4 | Tra2 isoforms |
| topo-hTra ex1.3.4 | pCR4 TA | Tra2 isoforms |
| Tra2NEStrans-1nt m | pEGFP-C2 | Nuclear Export Signal 1nt mutant |
| Tra2-beta1-C2-NES NO | pEGFP-C2 | Tra2-beta1 with Nuclear Export Signal |
| MGTra SM1del SM2trans | pCR 3.1 TA | SMAR element |
| MGTra SM1SM2 del | pCR 3.1 TA | SMAR element |
| MGTra linker 3 | pCR4 | SMAR element |
| Tra2-beta1-17-kozak | pCR4 | Tra2-beta1 to shulttling vector |
| Tra2-beta3-23-kozak | pCR4 TA | Tra2-beta3 to shulttling vector |
| Tra2-beta1-c2-NES A | pEGFP-C2 | Tra2-beta1 with Nuclear Export Signal |
| Tra2-beta1-c2-NES B | pEGFP-C2 | Tra2-beta1 with Nuclear Export Signal |
| topo-IL4-skip | pCR4 | IL4 isoforms |
| YFP-Tra2-beta3-A | YFP Shutling vector | Tra2-beta3 to shulttling vector |
| YFP-Tra2-beta3-B | YFP Shutling vector | Tra2-beta3 to shulttling vector |

| Name | Backbone | Description |
|---|---|---|
| topo-BACE1 | pCR4 | BACE1 minigene |
| YFP-Tra2-beta1 | YFP Shutling vector | Tra2-beta1 to shulttling vector |
| S280A Tra Mut | pEGFP-C2 | Tra2-beta1 mutant |
| ExonTrap hBACE1 | pET01 Exontrap | BACE1 minigene |
| Tra2-beta1-RATA-NES-C2 | pEGFP-C2 | Tra2-beta1 with Nuclear Export Signal, PP1 binding site mutated |
| pLCS 2BoxB | pBPLUGA | BoxB in ludiferase vector |
| pLCS 5BoxB | pBPLUGA | BoxB in ludiferase vector |
| Topo 2BoxB SalI BamHI | pCR4 | BoxB in ludiferase vector |
| Topo 5BoxB SalI BamHI | pCR4 | BoxB in ludiferase vector |
| pLCS Tra ESE2 | pBPLUGA | Tra2-beta1 motif to luciferase vector |
| pLCS Tra ESE1 | pBPLUGA | Tra2-beta1 motif to luciferase vector |
| pLCS Tra ESE2 b | pBPLUGA | Tra2-beta1 motif to luciferase vector |
| pLCS Tra ESE3 | pBPLUGA | Tra2-beta1 motif to luciferase vector |
| pLCS Tra ESE4 | pBPLUGA | Tra2-beta1 motif to luciferase vector |
| TOPO NRS | pCR4 | Nuclear Retaintion Signal from SC35 |
| Tra2 NRS trans C2 | pEGFP-C2 | Nuclear Retaintion Signal from SC35 |
| Tra RS2A C2 | pEGFP-C2 | Tra2-beta1 mutant |
| Tra2 NRS C2 | pEGFP-C2 | Tra2-beta1 with Nuclear Retaintion Signal from SC35 |
| Tra2 RS1E C2 | pEGFP-C2 | Tra2-beta1 mutant |
| Tra RS2E C2 | pEGFP-C2 | Tra2-beta1 mutant |
| TOPO TraYN155 | pCR4 | Tra2-beta1 to shulttling vector |
| TOPO TraYC155 | pCR4 | Tra2-beta1 to shulttling vector |
| Tra2 RS1A c2 | pEGFP-C2 | Tra2-beta1 mutant |
| TOPO Tra RS1A RS2A | pCR4 | Tra2-beta1 mutant |
| TOPO Tra RS1E RS2E | pCR4 | Tra2-beta1 mutant |
| Tra YN155 | pCMV | Tra2-beta1 to shulttling vector |
| Tra YC155 | pCMV | Tra2-beta1 to shulttling vector |
| TOPO BACE1 1# | pCR4 | BACE1 minigene |
| TOPO BACE1 5# | pCR4 | BACE1 minigene |
| TOPO BACE1 11# | pCR4 | BACE1 minigene |
| Tra RS1A RS2A c2 | pEGFP-C2 | Tra2-beta1 mutant |
| Tra RS1E RS2E c2 | pEGFP-C2 | Tra2-beta1 mutant |
| GATEWAY Entry-hRPL3 | pDONR221 | RPL3 in gateway entry clone |
| Gateway GFP hRPL3 | pcDNA-DEST53 | RPL3 in gateway pDEST53 |
| Gateway entry Tra2- | pDONR221 | Tra2-beta1 mutant in gateway entry clone |

| Name | Backbone | Description |
|---|---|---|
| beta1 RS3A | | |
| Gateway entry Tra2-beta1 RS3E | pDONR221 | Tra2-beta1 mutant in gateway entry clone |
| GATEWAY GFP TraRS3A | pcDNA-DEST53 | Tra2-beta1 mutant in gateway pDEST53 |
| GATEWAY GFP TraRS3E | pcDNA-DEST53 | Tra2-beta1 mutant in gateway pDEST clone |
| TOPO MG RPL3 | pCR4 | RPL3 minigene |
| Gateway entry Tra2-beta1 | pDONR221 | Tra2-beta1 in gateway entry clone |
| Tra EGFP 2nd RRR del | pEGFP-C2 | Tra2-beta1 mutant |
| Tra EGFP 2RRR del | pEGFP-C2 | Tra2-beta1 mutant |
| gateway entry tra2 del 207-222, 241-end | pDONR221 | Tra2-beta1 mutant in gateway entry clone |
| gateway GFP tra2-beta1 | pcDNA-DEST53 | Tra2-beta1 in gateway pDEST53 |
| EGFP Tra2 del 126-165 | pEGFP-C2 | Tra2-beta1 mutant |
| EGFP Tra2 del 223-240 | pEGFP-C2 | Tra2-beta1 mutant |
| gateway GFP tra2 del 207-222, 241-end | pcDNA-DEST53 | Tra2-beta1 mutant |
| Tra RS123A plus C2 | pEGFP-C2 | Tra2-beta1 mutant |
| Tra RS123E C2 | pEGFP-C2 | Tra2-beta1 mutant |
| Tra2 RS123A C2 | pEGFP-C2 | Tra2-beta1 mutant |
| gateway pDEST15 hRPL3 | pcDNA-DEST15 | RPL3 in gateway destination vector |
| TOPO MG RPL3 short | pCR4 | RPL3 minigene |
| MG RPL3 short | pET01 Exontrap | RPL3 minigene |
| pDEST27 RPL3 | pcDNA-DEST27 | RPL3 in gateway destination vector |
| TOPO pLCS ESE1 | pCR4 | Insert Tra2-beta1 binding motif to luciferase |
| TOPO pLCS ESE2 | pCR4 | Insert Tra2-beta1 binding motif to luciferase |
| TOPO pLCS ESE3 | pCR4 | Insert Tra2-beta1 binding motif to luciferase |
| TOPO pLCS ESE4 | pCR4 | Insert Tra2-beta1 binding motif to luciferase |
| TOPO pLCS EDA | pCR4 | Insert Tra2-beta1 binding motif to luciferase |
| TOPO pLCS MT | pCR4 | Insert Tra2-beta1 binding motif to luciferase |
| TOPO pLCS STOP | pCR4 | Insert Tra2-beta1 binding motif to luciferase |
| pDEST22 Tra2-beta1 | pDEST22 | Tra2-beta1 in gateway destination vector |
| pDEST22 RPL3 | pDEST22 | RPL3 in gateway destination vector |
| pDEST32 Tra2-beta1 | pDEST32 | Tra2-beta1 in gateway destination vector |
| pDEST32 RPL3 | pDEST32 | RPL3 in gateway destination vector |

| Name | Backbone | Description |
|---|---|---|
| Entry Tra2 Del RS2 (+RVDF) | pDONR221 | Tra2-beta1 mutant, for Y2H |
| Entry Tra2 Del RS2 (-RVDF) | pDONR221 | Tra2-beta1 mutant, for Y2H |
| Entry Tra2 RRM (+RVDF) | pDONR221 | Tra2-beta1 mutant, for Y2H |
| Entry Tra2 RRM (-RVDF) | pDONR221 | Tra2-beta1 mutant, for Y2H |
| Entry Tra2 Del RS1 | pDONR221 | Tra2-beta1 mutant, for Y2H |
| pDEST22 Tra2 Del RS2 (+RVDF) wrong | pDEST22 | Tra2-beta1 mutant, for Y2H |
| pDEST22 Tra Del RS2 (-RVDF) | pDEST22 | Tra2-beta1 mutant, for Y2H |
| pDEST22 Tra2 RRM (+RVDF) | pDEST22 | Tra2-beta1 mutant, for Y2H |
| pDEST22 Tra2 RRM (-RVDF) | pDEST22 | Tra2-beta1 mutant, for Y2H |
| pDEST22 Tra2 Del RS1 | pDEST22 | Tra2-beta1 mutant, for Y2H |
| Entry Tra2 Del 270-end | pDONR221 | Tra2-beta1 mutant, for Y2H |
| Entry Tra2 Del 241-end | pDONR221 | Tra2-beta1 mutant, for Y2H |
| pDEST22 Tra2 Del 270-end | pDEST22 | Tra2-beta1 mutant, for Y2H |
| pDEST22 Tra2 Del 241-end | pDEST22 | Tra2-beta1 mutant, for Y2H |
| pDEST22 9G8 | pDEST22 | for Y2H |
| pDEST22 ASF/SF2 | pDEST22 | for Y2H |
| pDEST22 SRp30c | pDEST22 | for Y2H |
| Entry 9G8 | pDONR221 | for Y2H |

### 3.1.10 Oligonucleotides

**Primers used for cloning and sequencing**

| Name | Orientation | Sequence 5' → 3' | Target |
|---|---|---|---|
| Tra2YC155 EcoRI | sense | GAATTCGGATGAGCGACAGCG GCGAGCA | Cloning human Tra2-beta1 to YC155 |
| Tra2YC155 XhoI | antisense | CTCGAGGATAGCGACGAGGTG AGTATGAT | Cloning human Tra2-beta1 to YC155 |
| YC155 seq | antisense | GGGGTGTTCTGCTGGTAGTG | Sequence for Cloning human Tra2-beta1 to YC155 |

| Name | Orientation | Sequence 5' → 3' | Target |
|---|---|---|---|
| Tra2YN155 EcoRV | antisense | GATATCCCATAGCGACGAGGT GAGTATGAT | Cloning human Tra2-beta1 to YN155 |
| YN155 seq | antisense | GAACTTCAGGGTCAGCTTGC | Sequence for Cloning human Tra2-beta1 to YN155 |
| TraEx2 rev | antisense | GCGTAGTGCTTTCTGATTCCA | Endogenous Tra2 RNA expression |
| TraEx3 for | sense | AATCCCGTTCTGCTTCCAG | Endogenous Tra2 RNA expression, also used for CLIP RTPCR |
| TraEx3 rev | antisense | CAGATCGGGACCTGGACTT | Endogenous Tra2 RNA expression |
| hHistoneH1 for | sense | CTCGCAGATCAAGTTGTCCA | Endogenous HistoneH1 RNA expression |
| hHistoneH1 rev | antisense | AAAAGGTGGTGGTGAGCATC | Endogenous HistoneH1 RNA expression |
| hMYC for | sense | GGAAGAAATTCGAGCTGCTG | Endogenous MYC RNA expression |
| hMYC rev | antisense | GCTGTCGTTGAGAGGGTAGG | Endogenous MYC RNA expression |
| rpL3 BamHI for | sense | ATAGGATCCCCTAAATGGGCAC TTTGCAT | For RPL3 minigene cloning |
| rpL3 NotI rev | antisense | ATAGCGGCCGCCTGAGCCTCA TCAACGAACA | For RPL3 minigene cloning |
| rpL3 XbaI rev | antisense | ATATCTAGACCCTCCAGGTTCC TTTCTGT | For RPL3 minigene cloning |
| hPP1 Ex3 for | sense | TGATTTGCTGCGACTTTTTG | Endogenous human PP1 RNA expression |
| hPP1 Ex5 rev | antisense | TTTTGCAACCACTTCTGCAC | Endogenous human PP1 RNA expression |
| hPP1 Ex5 for | sense | TTCAATCTATGGAGCAGATTCG | Endogenous human PP1 RNA expression |
| hPP1 Ex7 rev | antisense | AGTCCCGACTAGGCAGTGTC | Endogenous human PP1 RNA expression |
| hPolII for | sense | GAGTCCAGTTCGGAGTCCTG | Endogenous human PolII RNA expression |
| hPolII rev | antisense | ACCCTCAGGTTGTTCCACAC | Endogenous human PolII RNA expression |
| human actin for | sense | ACACTGTGCCCATCTACGAGG | Endogenous human actin RNA expression |
| Human Actin rev | antisense | AGGGGCCGGACTCGTCATACT | Endogenous human actin RNA expression |
| LambdaN SalI for | sense | GTCGACCCATGGACGCACAAA | Cloning LambdaN |
| LambdaN Bam rev | antisense | GGATCCCGGTGGGTTTGCAGC | Cloning LambdaN |

| Name | Orientation | Sequence 5' → 3' | Target |
|---|---|---|---|
| PS2 Ex4 for | sense | ACCCTGACCGCTATGTCTGT | Endogenous human PS2 RNA expression |
| PS2 Ex7 rev | antisense | ATGATGAGGTAGGCCTGCTG | Endogenous human PS2 RNA expression |
| | | | |
| Mus-GAPDH-for | sense | GCAGTGGCAAAGTGGAGATT | Endogenous mouse GAPDH RNA expression |
| Mus-GAPDH-rev | antisense | CATGAGCCCTTCCACAATG | Endogenous mouse GAPDH RNA expression |
| GAPDH 790bp rev | antisense | ACCTGGTGCTCAGTGTAGCC | Endogenous GAPDH RNA expression |
| pLCScodon for | sense | ATGGGGATTGGTGGCGACGAC | Cloning pLCS luciferase codon region |
| pLCScondon rev | antisense | TTACAATTTGGACTTTCCGCCC TTCTTGGC | Cloning pLCS luciferase codon region |
| Human U6 for | sense | CGCTTCGGCAGCACATATAC | Endogenous human U6 RNA expression |
| Human U6 rev | antisense | AAAATATGGAACGCTTCACGA | Endogenous human U6 RNA expression |
| EPN2 5UTR1 for | sense | TGGTGTGTGGGTGTCAAACT | Cloning EPN2 5'UTR |
| EPN2 5UTR2 for | sense | TCTACCGTGTTCCTCTGAAGC | Cloning EPN2 5'UTR |
| EPN2 5UTR rev | antisense | TGTTTTTCATCTGCCGTCTG | Cloning EPN2 5'UTR |
| PTK2 for | sense | TATTGGACCTGCGAGGGATT | Endogenous human PTK2 expression |
| PTK2 rev | antisense | TGACACCCTCGTTGTAGCTG | Endogenous human PTK2 expression |
| CCNDBP1 for | sense | AAGTTCTGTGAACAAGTCCATG C | Endogenous human CCNDBP1 expression |
| CCNDBP1 rev | antisense | TGCTCCATTTCTTCATGTGC | Endogenous human CCNDBP1 expression |
| CEP110 for | sense | CCAGAATCACCTTAACCATGTG | Endogenous human CEP110 expression |
| CEP110 rev | antisense | GGGCCAGAATTCTCTCCTTC | Endogenous human CEP110 expression |
| PHC2b for | sense | GCATGACCTGTTCCATTCAGC GG | EST cloning of PHC2 |
| PHC2b rev | antisense | CACTGCTGCTGTTGTTGCAG | EST cloning of PHC2 |
| PHC2a for | sense | AGGTGGAGGAGGAGCGCGGA | EST cloning of PHC2 |

| Name | Orientation | Sequence 5' → 3' | Target |
|---|---|---|---|
| PHC2a rev hLipin Ex6 for | sense | CCGCTGAATGGAACAGGTCATGC TTCCTAATGATATACCTCCATTCCA | EST cloning of PHC2 |
| hLipin Ex7 for | sense | AAAAGGACTGCCCCTCATCT | Endogenous human Lipin expression |
| hLipin Ex8 rev | antisense | CTCTCCCCACAGCCAAAG | Endogenous human Lipin expression |
| Lipin minigene f | sense | TTTGGCAGTTTCCTGCTTTT | Lipin minigene construction |
| Lipin minigene r | antisense | TTGTCAAACAACCCGAGACA | Lipin minigene construction |
| Phc2 gen for | sense | GCCACTAGGCCCGGTTAC | EST cloning of PHC2 |
| Phc2 gen rev | antisense | GGGGACGGCTGGATGTTAG | EST cloning of PHC2 |
| EPN2 gen for | sense | CCATCCCAAAACAATGGAAC | EST cloning of EPN2 |
| EPN2 gen rev | antisense | CACAGCCATGGACTCCACTC | EST cloning of EPN2 |
| CCNDBP1 gen for | sense | GCTGCAGCTCTTTTGATGCT | EST cloning of CCNDBP1 |
| CCNDBP1 gen rev | antisense | ACAATGTCATCCAGCTGTGC | EST cloning of CCNDBP1 |
| MusTra2B-for | sense | GTTAGACCGGTGCGGAGGT | Endogenous mouse Tra2-beta expression |
| MusTra2B-rev | antisense | CAACATGACGCCTTCGAGTA | Endogenous mouse Tra2-beta expression |
| T3 | sense | ATTAACCCTCACTAAAGGGA | sequencing in TOPO vector |
| T7 | antisense | TAATACGACTCACTATAGGG | sequencing in TOPO vector |
| M13rev | antisense | CAGGAAACAGCTATGAC | sequencing in TOPO vector |
| LSM1 exon2 | sense | AAGCACTTGGTTCTGCTTCG | Endogenous RNA expression |
| LSM1 exon2 | antisense | GTGATCAAATGCGTGAGGTG | Endogenous RNA expression |
| Human U1 for | sense | ATACTTACCTGGCAGGGGAG | Endogenous RNA expression |
| Human U1 rev | antisense | CAGGGGAAAGCGCGAACGCA | Endogenous RNA expression |
| Human U2 for | sense | ATCGCTTCTCGGCCTTTTGG | Endogenous RNA expression |
| Human U2 rev | antisense | TGGTGCACCGTTCCTGGAGG | Endogenous RNA expression |

| Name | Orientation | Sequence 5' → 3' | Target |
|---|---|---|---|
| HumanCD44V4 | sense | CATTCAAATCCGGAAGTGCT | Endogenous human CD44 RNA expression |
| HumanCD44V4 | antisense | GGTTGTGTTTGCTCCACCTT | Endogenous human CD44 RNA expression |
| HumanCD44V5 | sense | GGCACCACTGCTTATGAAGG | Endogenous human CD44 RNA expression |
| HumanCD44V5 | antisense | ACTGCAATGCAAACTGCAAG | Endogenous human CD44 RNA expression |
| Human CD44 EX2 for | sense | AATATAACCTGCCGCTTTGC | Endogenous human CD44 exon2 RNA expression |
| Human CD44 EX2 rev | antisense | CAGGTCTCAAATCCGATGCT | Endogenous human CD44 exon2 RNA expression |
| Human CD44 lastEX f | sense | GTGATCAACAGTGGCAATGG | Endogenous human CD44 last Exon RNA expression |
| Human CD44 lastEX r | antisense | CCACATTCTGCAGGTTCCTT | Endogenous human CD44 last Exon RNA expression |
| Human cd44v3 for | sense | CTGGGAGCCAAATGAAGAAA | Endogenous human CD44 RNA expression |
| Human cd44v8 rev | antisense | GAGGTCCTGTCCTGTCCAAA | Endogenous human CD44 RNA expression |
| HumanCD44exon4 | sense | CCTGAAGAAGATTGTACATCAGTCA | Endogenous RNA expression |
| HumanCD44exon4 | antisense | TGTGGGGTCTCTTCTTCCTC | Endogenous RNA expression |
| HumanCD44exon5 | sense | CCGCTATGTCCAGAAAGGAG | Endogenous RNA expression |
| HumanCD44exon5 | antisense | TCATCCTTGTGGTTGTCTGAA | Endogenous RNA expression |
| MGClk2exon3-for | sense | GACCGGAGGGTGTATGACC | Endogenous RNA expression |
| MGClk2exon5-rev | antisense | CGAAGGTCCCCTCTCCTAAG | Endogenous RNA expression |
| RSV-LTR-SalI-for | sense | ATAGTCGACTTGAAGCTGTCCCTGATGGT | For Stable trasfected minigene CLK2 and IL4 construction |
| InsulinSacI-rev | antisense | ATAGAGCTCCAGCACTGATCCACGATGC | For Stable trasfected minigene CLK2 and IL4 construction |
| InsulinAflII-rev | antisense | ATACTTAAGCAGCACTGATCCACGATGC | For Stable trasfected minigene CLK2 and IL4 construction |

| Name | Orientation | Sequence 5' → 3' | Target |
|---|---|---|---|
| Tra2Aexon3-for | sense | CACTCGATCCAGATCCCACT | Endogenous mouse Tra2alpha expression |
| Tra2Aexon5-rev | antisense | CCACCCGAATTCTTCTACCA | Endogenous mouse Tra2alpha expression |
| InsulinRsrII-rev | antisense | ATACGGACCGCAGCACTGATCCACGATGC | For Stable trasfected minigene CLK2 and IL4 construction |
| InsulinBclI-rev | antisense | ATATGATCACAGCACTGATCCACGATGC | For Stable trasfected minigene CLK2 and IL4 construction |
| ExonTrapInsulin-for | sense | CTGCCCAGGCTTTTGTCA | For Stable trasfected minigene CLK2 and IL4 construction |
| ESG-CLK2-StuI-For | sense | ATAAGGCCTAGGCTTTTGCAAAAAGCTCCCTGGGGGTGTCTACGGTGA | For Stable trasfected minigene CLK2 and IL4 construction |
| F3ES-IL4R-WF-Rev | antisense | TTTACTAGTCCGACCCCACTCATGTTC | For Stable trasfected minigene CLK2 and IL4 construction |
| ClkShtMluI-For | sense | CGCACGCGTAGTGCTCCACCTGCCTTG | For Stable trasfected minigene CLK2 and IL4 construction |
| ClkShtRsrII-Rev | antisense | TATCGGACCGAAGCCCCATATAACCCCAAC | For Stable trasfected minigene CLK2 and IL4 construction |
| ClkShtSacI-Rev | antisense | TATGAGCTCAAGCCCCATATAACCCCAAC | For Stable trasfected minigene CLK2 and IL4 construction |
| MusTra2A-for | sense | ACGCACTGGCCGTTGTAG | Endogenous mouse Tra2alpha expression |
| MusTra2A-rev | antisense | CACGAAGATCTCTCTCTGTTGTG | Endogenous mouse Tra2alpha expression |

**Primers used for Luciferase construct**

| Name | Orientation | Sequence 5' → 3' | Target |
|---|---|---|---|
| TraESE1 for | sense | TCGAGAAGGAAAAG | Cloning Tra2 ESE1 binding motif to Luciferase vector |
| TraESE1 rev | antisense | GATCCTTTTCCTTC | Cloning Tra2 ESE1 binding motif to Luciferase vector |
| TraESE2 for | sense | TCGAGGAAGAATGG | Cloning Tra2 ESE2 binding motif to Luciferase vector |
| TraESE2 rev | antisense | GATCCCATTCTTCC | Cloning Tra2 ESE2 binding motif to Luciferase vector |
| TraESE3 for | sense | TCGAGAAAGAAGTG | Cloning Tra2 ESE3 |

| Name | Orientation | Sequence 5' → 3' | Target |
|---|---|---|---|
| | | | binding motif to Luciferase vector |
| TraESE3 rev | antisense | GATCCACTTCTTTC | Cloning Tra2 ESE3 binding motif to Luciferase vector |
| TraESE4 for | sense | TCGAGAAAGAATGG | Cloning Tra2 ESE4 binding motif to Luciferase vector |
| TraESE4 rev | antisense | GATCCCATTCTTTC | Cloning Tra2 ESE4 binding motif to Luciferase vector |
| 2EDA for | sense | TCGAGAAGAAGACGAAGAAGACG | Cloning 2 EDA binding motif to Luciferase vector |
| 2EDA rev | antisense | GATCCGTCTTCTTCGTCTTCTTC | Cloning 2 EDA binding motif to Luciferase vector |
| 2TraESE1 for | sense | TCGAGAAGGAAAAGAAGGAAAAG | Cloning 2 Tra2 ESE1 binding motif to Luciferase vector |
| 2TraESE1 rev | antisense | GATCCTTTTCCTTCTTTTCCTTC | Cloning 2 Tra2 ESE1 binding motif to Luciferase vector |
| 2TraESE4 for | sense | TCGAGAAAGAATGGAAAGAATGG | Cloning 2 Tra2 ESE4 binding motif to Luciferase vector |
| 2TraESE4 rev | antisense | GATCCCATTCTTTCCATTCTTTC | Cloning 2 Tra2 ESE4 binding motif to Luciferase vector |
| 3TraESE1 for | sense | TCGAGAAGGAAAAGAAGGAAAAGAAGGAAAAG | Cloning 3 Tra2 ESE1 binding motif to Luciferase vector |
| 3TraESE1 rev | antisense | GATCCTTTTCCTTCTTTTCCTTCTTTTCCTTC | Cloning 3 Tra2 ESE1 binding motif to Luciferase vector |
| 3TraESE4 for | sense | TCGAGAAAGAATGGAAAGAATGGAAAGAATGG | Cloning 3 Tra2 ESE4 binding motif to Luciferase vector |
| 3TraESE4 rev | antisense | GATCCCATTCTTTCCATTCTTTCCATTCTTTC | Cloning 3 Tra2 ESE4 binding motif to Luciferase vector |

**Primers used for gateway cloning**

| Name | Orientation | Sequence 5' → 3' | Target |
|---|---|---|---|
| attB Tra del270 | antisense | GGGGACCACTTTGTACAAGAAAGCTGGGTTATAGTAAGGAGAAGGTGACC | Gateway cloning for Tra2 deletion from 270 |
| attB Tra del241 | antisense | GGGGACCACTTTGTACAAGAAAGCTGGGTTGTATGATCTGCTATAGTAGT | Gateway cloning for Tra2 deletion from 241 |

| Name | Orientation | Sequence 5' → 3' | Target |
|---|---|---|---|
| | | | |
| MutRS1A for | sense | AAGGCCAGGGCCCGAGCTGAAGC TAGGGCTAGAGCCAGAAGAAGCGC CCGAAGGCATT | Gateway cloning for Tra2 Mutant in first RS domain |
| MutRS1A rev | antisense | AGCCCTAGCTTCAGCTCGGGCCCT GGCCTTTGCTCTGGAACGCCTGGC ATCTTCCTTG | Gateway cloning for Tra2 Mutant in first RS domain |
| MutRS1E for | sense | AAGGAAAGGGAACGAGAAGAAGAA AGGGAAAGAGAAAGAAGAAGCGAA CGAAGGCATT | Gateway cloning for Tra2 Mutant in first RS domain |
| MutRS1E rev | antisense | TTCCCTTTCTTCTTCTCGTTCCCTTT CCTTTTCTCTGGAACGCCTTTCATC TTCCTTG | Gateway cloning for Tra2 Mutant in first RS domain |
| MutRS2A for | sense | CGCCCGCGCCCATAGACGAGCAC GTGCCAGGGCTTACGCTCGAGATT ATC | Gateway cloning for Tra2 Mutant in first RS domain |
| MutRS2A rev | antisense | AAGCCCTGGCACGTGCTCGTCTAT GGGCGCGGGCGCGAGCCCGTGAC CGG | Gateway cloning for Tra2 Mutant in first RS domain |
| MutRS2E for | sense | CGAACGCGAACATAGACGAGAACG TGAAAGGGAATACGAACGAGATTAT C | Gateway cloning for Tra2 Mutant in first RS domain |
| MutRS2E rev | antisense | ATTCCCTTTCACGTTCTCGTCTATG TTCGCGTTCGCGTTCCCGTGACCG G | Gateway cloning for Tra2 Mutant in first RS domain |
| MutRS3A a rev | sense | TCTGTATCCTCCACGAGCATAGTAA GGAGCAGGTGCCCGCCTTCTATAA A | Gateway cloning for Tra2 Mutant in second RS domain |
| MutRS3E a rev | antisense | TCTGTATCCTCCACGTTCATAGTAA GGTTCAGGTTCCCGCCTTCTATAAA | Gateway cloning for Tra2 Mutant in second RS domain |
| MutRS3A b rev | antisense | ATAGCGACGAGGTGCGTATGCTCG AGCTCTGGCACGTGCTCTGTATCC TCCACG | Gateway cloning for Tra2 Mutant in second RS domain |
| MutRS3E b rev | antisense | ATAGCGACGAGGTTCGTATTCTCG TTCTCTTTCACGTTCTCTGTATCCT CCACG | Gateway cloning for Tra2 Mutant in second RS domain |
| attB Tra2short rev | antisense | GGG GAC CAC TTT GTA CAA GAA AGC TGG GTC ATAGCGACGAGG | Gateway cloning for Tra2 Mutant |
| MutRS3A b for | sense | CGTGGAGGATACAGAGCACGTGCC AGAGCTCGAGCATACGCACCTCGT CGCTAT | Gateway cloning for Tra2 Mutant in second RS domain |

| Name | Orientation | Sequence 5' → 3' | Target |
|---|---|---|---|
| | | | |
| MutRS3E b for | sense | CGTGGAGGATACAGAGAACGTGAAAGAGAACGAGAATACGAACCTCGTCGCTAT | Gateway cloning for Tra2 Mutant in second RS domain |
| Tra del 2RRR for | sense | CACAGCCACAGCCATTCTCCCATGTCTACTCATGTTGGGAATCG | Gateway cloning for Tra2 deletion in RRR rich region |
| Tra del 2RRR rev | antisense | ATGGCTGTGGCTGTGATAATCTCGACTGTA | Gateway cloning for Tra2 deletion in RRR rich region |
| Del 207-222 FOR | sense | CCAACACCAGGAATTGACAGAGGATATGAT | Gateway cloning for Tra2 deletion |
| Del 207-222 REV | antisense | ATCATATCCTCTGTCAATTCCTGGTGTTGG | Gateway cloning for Tra2 deletion |
| Del 223-240 FOR | sense | CGCCGTCGGGATTACAGAGGAGGAGGTGGA | Gateway cloning for Tra2 deletion |
| Del 223-240 rev | antisense | TCCACCTCCTCCTCTGTAATCCCGACGGCG | Gateway cloning for Tra2 deletion |
| Del 126-165 FOR | sense | GTCTTGGAGTATTTGGGCTGTTTGAAAATGTAGATGATGC | Gateway cloning for Tra2 deletion |
| Del 126-165 REV | antisense | GCATCATCTACATTTTCAAACAGCCCAAATACTCCAAGAC | Gateway cloning for Tra2 deletion |
| Bait F | sense | AACCGAAGTGCGCCAAGTGTCTG | Checking gateway Y2H clones |
| Prey F | sense | TATAACGCGTTTGGAATCACT | Checking gateway Y2H clones |
| Bait and Prey R | antisense | AGCCGACAACCTTGATTGGAGAC | Checking gateway Y2H clones |
| attB NES for | sense | GGGGACAAGTTTGTACAAAAAAGCAGGCTTCATCGATCGGAACAGCAATG | Gateway cloning for NES started clone |
| attB RPL3 for | sense | GGGGACAAGTTTGTACAAAAAAGCAGGCTTTATGTCTCACAGAAAGTTC | Cloning for hRPL3 |
| attB RPL3 rev | antisense | GGGGACCACTTTGTACAAGAAAGCTGGGTCAGCTCCTTCTTCCTTTGC | Cloning for hRPL3 |
| Tra2-beta1 EcoRI for | sense | GAATTCATGAGCGACAGCGGCGAGCAGAACTAC | Cloning for Tra2-beta1 |
| Tra2-beta1 BamHI rev | antisense | GGATCCTTAATAGCGACGAGGTGAGTATGAT | Cloning for Tra2-beta1 |

**Primers used for minigene analysis**

| Name | Orientation | Sequence 5' → 3' | minigene |
|---|---|---|---|
| Exontrap-pcr-r | antisense | in Exontrap insulin exons (MoBiTec) | Clk2 |
| Exontrap-pcr-f | sense | in Exontrap insulin exons (MoBiTec) | |
| Exontrap-pcr-r | antisense | in Exontrap insulin exons (MoBiTec) | BACE1 |
| Exontrap-pcr-f | sense | in Exontrap insulin exons (MoBiTec) | |
| Exontrap-pcr-r | antisense | in Exontrap insulin exons (MoBiTec) | RPL3 |
| Exontrap-pcr-f | sense | in Exontrap insulin exons (MoBiTec) | |
| Globin-rev | antisense | AGACACCATGCATGGTGCACC | CD44v4v5 |
| Globin-for | sense | CCTGATCAGCGAGCTCTAG | |
| pCR3.1 RT revers | antisense | GCCCTCTAGACTCGAGCTCGA | Tra2-beta1 |
| MG Tra-Bam | sense | GGGCCAGTTGGGCGACCGGCGCGTCGTGCG | |
| MG Tra-Xho | antisense | GGGCTCGAGTACCCGATTCCCAACATGACG | |
| N5 INS | sense | GAGGGATCCGCTTCCTGCCCC | CD44v5 |
| N3 INS | antisense | CTCCCGGGCCACCTCCAGTGCC | |
| T7 | sense | TAATACGACTCACTATAGGG | SRp20 |
| X16R | antisense | CCTGGTCGACACTCTAGATTTCCTTTCATTTGACC | |
| INS1 | sense | CAGCTACAGTCGGAAACCATCAGCAAGCAG | Tau |
| INS3 | antisense | CACCTCCAGTGCCAAGGTCTGAAGGTCACC | |
| pCl for | sense | GGTGTCCACTCCCAGTTCAA | SMN2 |
| SMNex8 rev | antisense | GCCTCACCACCGTGCTGG | |

**Primers used for ExonHit splicing Microarray with AD samples**

| Name | Orientation | Sequence 5' → 3' | Target |
|---|---|---|---|
| SFRS14 for | sense | TGACCAAAATAGTTCTGCTTTCA | SFRS14 |
| SFRS14 rev | antisense | CTTTTGGCTCCTGGATGAGA | SFRS14 |
| CBP80 exon3 for | sense | TAGAAGGCTTGGCTGGTGTT | CBP80 |
| CBP80 exon6 rev | antisense | TTCAGTGTTGGCAAAGATGC | CBP80 |
| ZNF207 exon7 for | sense | CCTCCAATGACTCAAGCACA | ZNF207 |
| ZNF207 exon10 rev | antisense | GTCCAACTGGTGGATTACCG | ZNF207 |
| ARL6IP4 exon3 for | sense | CTTCTAGCTCCTCTTCTTCCTCCT | ARL6IP4 |
| ARL6IP4 exon4 rev | antisense | GTCTCAGGGTCCACCACCT | ARL6IP4 |

| Name | Orientation | Sequence 5' → 3' | Target |
|---|---|---|---|
| HnRNPH3 exon2 for | sense | AAATCAAACGGTATTGAGATGGA | HnRNPH3 |
| HnRNPH3 exon4 rev | antisense | ATCACCTCCTCGTCGCATT | HnRNPH3 |
| NSAP1 for | sense | ATACCACCAACCGGATGACA | NSAP1 |
| NSAP1 rev | antisense | TTTTTGCTGCTTGCCTCTG | NSAP1 |
| LSM7 for | sense | CTGCGAAGAGCCACACG | LSM7 |
| LSM7 rev | antisense | AAGTCCGCGGGAAACC | LSM7 |
| HnRNPA2B1 for | sense | TGCAGAAATACCATACCATCAA | HnRNPA2B |
| HnRNPA2B1 rev | antisense | TGCTACCACCAAAGTTTCCA | HnRNPA2B |
| CDK9 for | sense | CATGAAGGCTGCTAATGTGC | CDK9 |
| CDK9 rev | antisense | GTGATCTGGCTGCCCTTC | CDK9 |

## Oligoes used for CLIP

| Name | Orientation | SEQUENCE 5' → 3' | Target |
|---|---|---|---|
| X6aR for | sense | CAGACGGGGCACAAATA | For PCR in CLIP |
| TraEx3 for | sense | AATCCCGTTCTGCTTCCAG | Endogenous Tra2 RNA expression, also used for CLIP RT and PCR |
| TraEx3 RNA | RNA | 5'PHO-CUGGAAGCAGAACGGGAUU-3'BIOT3 | 5' phosphorylated and 3' with Biotin TEG, RNA oligo linkers in CLIP |
| 2OM RNA X6aR | RNA | CAGACGGGGCACAAAUA | 5' Biotin TEG - O-Me-RNA, for RNA oligo linkers in CLIP |

## Oligos for siRNA knockdown

| Name | Target | Supplier |
|---|---|---|
| hnRNP G siRNA | human hnRNP G | Santa Cruz |
| SFRS10 siRNA | human Tra2-beta1 | Santa Cruz |

## *3.2 Methods*

### 3.2.1 Plasmid DNA isolation

Large amounts of pure plasmid DNA (100-500 ug from 100ml) was isolated using QIAGEN Plasmid Maxi kit according to the manufacturer's protocol.

Smaller amounts of plasmid DNA were isolated using the alkaline lysis method first described by Birnboim and Doly (Birnboim and Doly, 1979). In brief, bacterial cells carrying the desired plasmid were cultured overnight at 37°C in 5ml LB medium containing the appropriate antibiotics. The cells were harvested by centrifugation for 5 minutes at 12,000 rpm. The pellet was resuspended in 250µl buffer P1. Equal volume of lysis buffer P2 was then added and the solution mixed gently by invertion. The cells were allowed to lyse for 5 minutes, followed by addition of the neutralization buffer P3. The tube was mixed gently by inversion and the solution was maintained on ice for 15 minutes. After centrifugation for 10 minutes at 12,000 rpm, the resulting supernatant was precipitated by adding 1 volume of isopropanol. Plasmid DNA was pelleted by centrifugation at 12,000 rpm for 10 minutes, washed with 70 % ethanol, air-dried and dissolved in 30µl of TE buffer. All the steps were carried out at room temperature in a conventional tabletop microfuge (Eppendorf Centrifuge 5415C).

| LB MEDIUM: | BUFFER P1: | BUFFER P2: |
|---|---|---|
| 10g Tryptone | 50 mM Tris-HCl, pH 8.0 | 200 mM NaOH |
| 10g NaCl | 10 mM EDTA | 1% SDS |
| 5g yeast extract | 100 µg/ml RNase A | |

| BUFFER P3: | BUFFER TE: |
|---|---|
| 3M Potassium acetate, pH 5.5 | 10 mM Tris-HCl, pH 8.0 |
| | 1 mM EDTA |

### 3.2.2 Maxi prep

For large-scale purification of plasmid DNA commercially available Kits from Qiagen or Genomed were used. The procedure was carried out according to the manufacturer's protocol. The dry pellet was diluted in TE or $dH_2O$.

### 3.2.3 Electrophoresis of DNA

DNA was resolved on 0.7-2% agarose gels prepared in 1 x TBE buffer. The electrophoresis was run for 80 min at 100 V. The gels were stained for 30 min in 0.5 mg/ml ethidium bromide and visualized under UV light, $\lambda$ = 260 nm.

1X TBE:

90 mM Tris-borate
20 mM EDTA

6 X GEL–LOADING BUFFER:

0.25% bromophenol blue
0.25% xylene cyanol FF
15% Ficoll 400 in $dH_2O$

### 3.2.4 Elution of DNA from agarose gels

DNA was purified from agarose gels where crystal violet was added to a final concentration of 2 µg per ml to detect DNA under visible light. Individual bands were excised and DNA was extracted using the Qiagen QIAEX II gel extraction kit according to the manufacturer's protocol.

6 X CRYSTAL VIOLET GEL–LOADING BUFFER:

0.25% crystal violet
15% Ficoll 400 in $dH_2O$

### 3.2.5 Determination of DNA concentration

Concentrations of nucleic acids in buffered solution were determined using a spectrophotometer (Eppendorf BioPhotometer 6131). Absorbance was measured at 260 nM in plastic cuvettes. The concentration was calculated using following formulas:

1 $A_{260}$ = 50 µg/ml double stranded DNA

1 $A_{260}$ = 37 µg/ml single stranded DNA

1 $A_{260}$ = 40 µg/ml RNA

The purity was determined with help of the A260/A280 ratio, which should be in the range of 1.8-2.0 for pure nucleic acids

### 3.2.6 PCR amplification of DNA

A standard PCR reaction to amplify DNA from a plasmid template contained 1-10 ng of plasmid DNA, forward and reverse primers (0.5 µM each), dNTPs (200 µM), 1 x Taq polymerase buffer, 1.5 mM $MgCl_2$ and 1 U Taq polymerase in total volume of 25 µl. When the amplification was made for cloning purposes, a high-fidelity polymerase, i.e. Platinum Pfx polymerase was used instead of Taq polymerase. The amplification was carried out in a Perkin Elmer GeneAmp PCR System 9700 thermocycler under the following conditions: initial denaturation for 2-4 min at 94°C; 25-35 cycles of 15-30 sec at 94°C, annealing at the Tm of the primers pair, extension of 1 min per 1 kb at 72°C (or 68°C for Pfx polymerase). After the last cycle the reaction was held for 5-10 min at the extension temperature to complete the amplification of all products.

### 3.2.7 DNA Ligation

When the vector ends were blunt or compatible with each other, the vector was dephosphorylated prior to ligation to prevent self-ligation. To remove 5' phosphates from the vector, 2 U of Calf intestinal phosphatase (CIP, Boehringer) was added to 5 µg of linearised vector in 1 x CIP buffer in 20 µl. The reaction was incubated for 1 hour at 37°C. CIP was subsequently inactivated by heating the reaction to 68°C for 20 min. A typical ligation reaction contained vector and insert at a ratio of about 1:3 (500-1000 ng total DNA), 1 x ligase buffer, 1 mM ATP and 200-400 U T4 DNA Ligase (New England Biolabs) in 15 µl. The incubation was carried out at 16°C for 12-16 h. After that, one third to one half of the ligation mixture was then transformed in *E coli* cells.

### 3.2.8 Preparation of competent E.coli cells

5 ml of LB medium were inoculated with a single bacterial colony and grown overnight at 37°C with vigorous shaking. 4 ml of this culture were transferred to 250 ml LB and grown to early logarithmic phase (OD600 = 0.3-0.6). The culture was centrifuged for 10 min at 2500 rpm at 4°C. The bacterial pellet was resuspended in 1/10 volume of cold TSB buffer and incubated on ice for 10 min. Cells were aliquoted into cold Eppendorf tubes and frozen in liquid nitrogen. Competent bacterial cells could then be stored at -80°C for several months.

TSB BUFFER:

10% PEG 3500
5% DMSO
10 mM $MgCl_2$
10 mM $MgSO_4$
in LB medium, pH 6.1

### 3.2.9 Transformation of E.coli cells

1-10 ng of plasmid DNA or a ligation reaction were added to 20 µl of 5 x KCM buffer and then the volume was equaled to 100 µl. Equal volume of competent cells was added. The reaction mixture was incubated on ice for 20 min followed by incubation at RT for 10 min. Then 1 ml of LB medium was added and the bacteria were incubated for 1 h at 37°C with vigorous shaking. Finally cells were plated on LB Agar plates containing appropriate antibiotic. Plates were incubated at 37°C until colonies were visible.

5 X KCM BUFFER:

500 mM KCl
150 mM $CaCl_2$
250 mM $MgCl_2$

### 3.2.10 Construction of minigenes:

Minigenes were constructed as described previously in Tang et al. 2004. In brief, most minigenes were amplified from genomic DNA and contained the alternatively

spliced exon and its flanking constitutive exons. In majority of the cases, these parts could be amplified by long-range PCR. If the introns were too long, several kilo-bases flanking the exons were amplified and ligated together to the multiple cloning sites of exon trap vectors (Mobitec, Göttingen, Germany). The alternative exons was located between two constitutive rat insulin exons. This chimeric gene was then analyzed similar to a genomic construct.

### 3.2.11 Site directed mutagenesis by overlap extension:

Four primers were designed to introduce mutations by this method which was first described by Higuchi et al 1989. One set of forward F and reverse R primer was complimentary to the extreme ends of the DNA template (Fig.3.1). The other set of forward MF and reverse MR primer carrying the desired mutation, were complimentary to each other and target the site where the mutation was desired. Mutant primers had 8-10 bases on either side of the mutation cassette to allow precise annealing. The first PCR was carried out with a proof reading polymerase to avoid any A-overhang. Individual PCRs were carried out to amplify fragments with F1 and R2, and with F2 and R1 respectively.

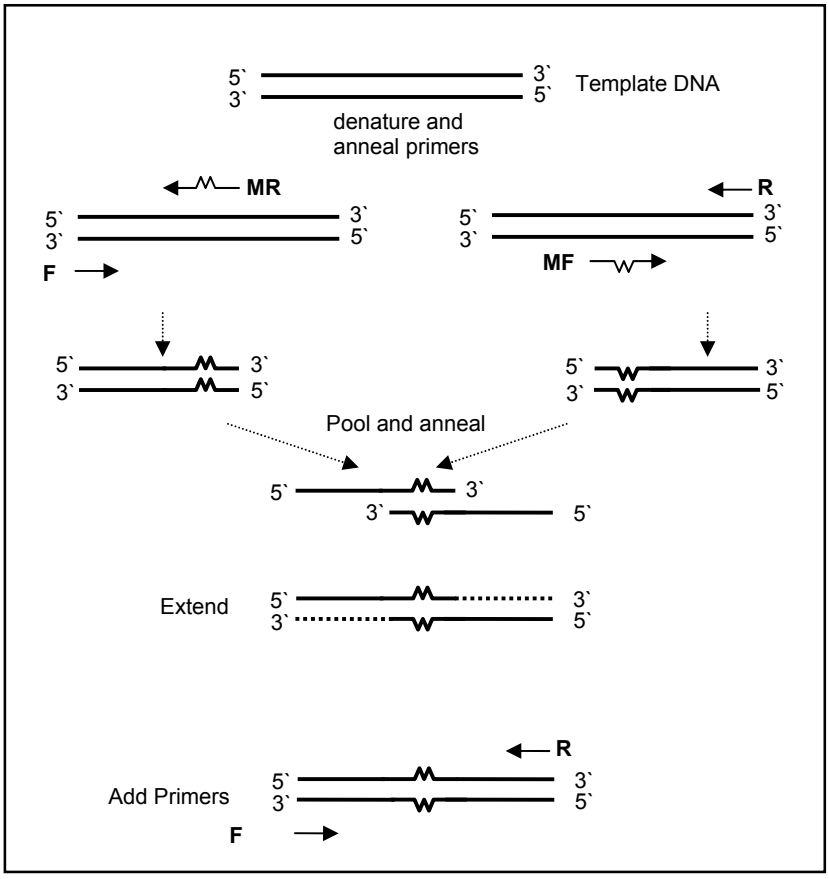



**Figure 3.1 Four primers strategy for mutations**

The amplified fragments were gel eluted to free them from any contaminating DNA template. 200 ng of the individual purified fragments were pooled together and allowed to anneal and extend without any addition of primer with dNTPs (200 µM), 1 x Taq polymerase buffer, 1.5 mM $MgCl_2$ and 1 U Taq polymerase in total volume of 25 µl. The amplification was carried out in a Perkin Elmer GeneAmp PCR System 9700 thermocycler under the following conditions: initial denaturation for 5 min at 94°C; 10 cycles of 30 sec at 94°C, annealing at 50°C, extension of 1 min per 1 kb at 72°C. After the last cycle the reaction was held for 5 min at the extension temperature to complete the amplification of all products. External primers were then added and the reaction was again supplemented with 1 U of Taq polymerase. Final PCR was performed with the following conditions: initial denaturation for 5 min at 94°C; 30 cycles of 30 sec at 94°C, annealing at 60°C and extension of 1 min per 1 kb at 72°C. The last cycle was followed by another 5 min of extension at 72°C. A part of the amplified fragment was run on the Agarose gel and the other subcloned into pCR4 for sequencing.

Another strategy based on 2 primers is: primers are complementary and the mutant site is designed inside the overlapping of the MR and MF primers. The PCR template should be circled plasmid and in a very low concentration (10ng per PCR reaction). 20-30 cycles PCR product is then digested by DpnI, 37C, 1hours which in principle digested most original template plasmids (with CpG). Left are the annealed PCR product with site mutation.

### 3.2.12 Radioactive labeling of 5' DNA ends

About 500 ng of purified DNA fragment (or 5-10 pmol oligonucleotide) were mixed with 1 x T4 polynucleotide kinase buffer, 50 µCi of γ-P 32 -ATP (3000 Ci/mmol) and 10 U T4 polynucleotide kinase (New England Biolabs) in 10 µl. The reaction was incubated for 1 hour at 37°C after which the kinase was inactivated at 68°C for 20 min.

### 3.2.13 Southern Blotting and hybridisation of DNA

DNA was separated on an 1% agarose gel. After electophoresis the gel was placed in denaturation buffer (1.5 M NaCl/0.5 N NaOH) for 30 min. The DNA was transferred on a nylon membrane in a denaturation buffer for 16 h. The DNA was crosslinked to the membrane on a UV transilluminator. Prehybridisation was carried out for 30 minutes at 68°C in a hybridisation buffer. Hybridisation was done in 10-20 ml of hybridisation buffer for 16 hours at 68°C. The membrane was washed two times in 2 x SSC/0.1% SDS for 10 minutes at 68°C. It was than washed (two times) in 0.5 x SSC/0.1% SDS for 10 minutes at 68°C. Finally, the membrane was washed in 0.1 x SSC/0.1% SDS for 10 minutes at 68°C. The membrane was then exposed to X-ray film or the signal was revealed on a phosphorimager.

| Hybridisation buffer: | 20 x SSC: |
|---|---|
| 0.5M phosphate buffer, pH 7.2<br>7% SDS | 3 M NaCl<br>0.3 M Na citrate |

### 3.2.14 Freezing, thawing and subculturing of eukaryotic cells

To freeze, cells were grown to mid logarithmic phase (about 75% of confluence) in 10 cm Petri dishes. They were collected by trypsinization with 1 x Trypsin/EDTA,

resuspended in 1 ml of the freezing medium (90% of the growth medium and 10% of DMSO). Vials were placed in Nalge Nunc Cooler giving a cooling rate of ~1°C/min while at -80°C. Cells were stored later in liquid nitrogen.

To thaw, cells were incubated at 37°C. The entire content of the tube was transferred to a 10 cm Petri dish and 10 ml of the growth medium were added. The dish was placed in the incubator at 37°C and 5% $CO_2$. When cells were attached to the plastic surface, the medium was removed and replaced with fresh one. The cells were maintained in the incubator until ready for the subculturing.

Cells were subcultured after reaching confluence. The monolayer was detached by adding 1 X Trypsin /EDTA and incubating at 37°C until single cell suspension was formed. 1/5 – 1/10 of this suspension was transferred to a new dish and mixed with the growth medium. Cells were maintained in the incubator at 37°C and 5% $CO_2$.

### 3.2.15 Transfection of eukaryotic cells

The procedure used for HEK293 cells was based on the one published by Chen and Okayama (Chen and Okayama, 1987). Exponentially growing cells were replated at a density of about 3 x $10^5$ cells / 8 $cm^2$. Growth medium was added and the cells were incubated at 37°C, 5% $CO_2$ for about 24 h, to reach 60-70% of confluence. For most applications cells were grown in 6-well plates, with 2 ml of growth medium per well. The transfection reaction for one well was made the following way. 1 to 5 µg of expression construct were mixed with 25 µl of 1 M $CaCl_2$ in final volume of 100 µl. Equal volume of 2 x HBS buffer was added drop by drop, with constant mixing. In order to form a precipitate, the solution was allowed to stay at RT for 20 min. After that, it was added to the growth medium. To express the transfected plasmid, cells were grown for additional 24 h at 37°C, 3 % $CO_2$.

2 X HBS:

280 mM NaCl
10 mM KCl
1.5 mM $Na_2HPO_4 \cdot 2H_2O$
12 mM Dextrose
50 mM Hepes
pH 6.95

### 3.2.16 *In vivo* splicing assay

To determine the influence of a protein on the splicing of selected minigenes, *in vivo* splicing assays were performed as described earlier (Stoss, Stoilov et al. 1999; Tang, Novoyatleva et al. 2004).

Eukaryotic cells were plated in a 6 well format. 24 hours after plating 1-2 µg of plasmid carrying the desired minigene were cotransfected with a protein expression construct per well. The plasmid for expressing the protein was usually transfected in increasing amounts from 0-3 µg to a concentration dependent effect. To ensure a constant amount of transfected DNA the parental vector of the expression plasmid missing only the gene to be expressed was added in decreasing amounts. After cells were incubated for 14-18 hours at 37 °C and 3% $CO_2$ total RNA was isolated and RT-PCR was performed.

If the effect of siRNA knockdown on different minigenes was studied, siRNA was transfected shortly after plating cells in a 24 well format. After 24-28 hours incubation at 37 °C and 5% $CO_2$, 250 ng minigene were transfected per well using the calcium phosphate method. After incubation for 16 hours at 37 °C and 5% $CO_2$ total RNA was isolated and RT-PCR performed.

The reverse primer used for RT was specific for the vector which carried the minigene, to avoid reverse transcription of endogenous RNA.

PCR was carried out using minigene specific primers, which amplify alternatively spliced minigene products. A control reaction using RNA instead of DNA was included.

Optimized PCR conditions were used for each minigene:

**Tau minigene**

| | | |
|---|---|---|
| 94 °C | 2 min | |
| 94 °C | 1 min | 30 |
| 60 °C | 1 min | cycles |
| 72 °C | 48 sec | |
| 72 °C | 10 min | |

**SMN2 minigene**

| | | |
|---|---|---|
| 94 °C | 4 min | |
| 94 °C | 20 sec | 25 cycles |
| 62 °C | 20 sec | |
| 72 °C | 20 sec | |
| 72 °C | 5 min | |

**Tra minigene**

**CD44v5 minigene**

| | | |
|---|---|---|
| 94 °C | 2 min | |
| 94 °C | 20 sec | 33 |
| 65 °C | 20 sec | cycles |
| 72 °C | 40 sec | |
| 72 °C | 2 min | |

| | | |
|---|---|---|
| 94 °C | 5 min | |
| 94 °C | 20 sec | 30 cycles |
| 72 °C | 50 sec | |
| 72 °C | 7 min | |

**pXB (X16) minigene**

| | | |
|---|---|---|
| 94 °C | 5 min | |
| 94 °C | 50 sec | 20 |
| 55 °C | 50 sec | cycles |
| 72 °C | 1 min | |
| 72 °C | 7 min | |

PCR reactions were resolved on a 2% agarose or 5% polyacrylamide gel. The image was analysed with Image J software (http://rsb.info.nih.gov/ij/download.html).

### 3.2.17 Isolation of total RNA

Total RNA was isolated from eukaryotic cells grown in 6-well plates. Cells were washed with 1 x PBS and the RNeasy Mini kit (QIAGEN) was used according to the manufacturer's protocol. RNA was eluted from the column in 30 µl of RNase-free $dH_2O$. However this procedure was applied only when the RNA of interest was larger than 200 bases.

Alternatively, for RNA smaller than 200 bases or when in lower concentrations (as in case of RNA immunoprecipitation), RNA was isolated using TRIzol reagent according to the manufacturer's protocol. 15-20µg of Glycogen per 1 ml was added to the reaction prior to ethanol precipitation for better recovery. After ethanol precipitation, the RNA pellet was dissolved in 20 µl of RNase-free $dH_2O$.

### 3.2.18 Gel shift assay

*5% native RNA gel setup mixture (20 ml)*
40% (w/v) acrylamide/bisacrylamide (80:1) 2.5 ml
20× TBE 0.5 ml
DMPC-treated $H_2O$ 17 ml
10% (w/v) ammonium persulfate (APS) 200 µl
TEMED 20 µl

*6× native RNA gel loading buffer*
0.025% (w/v) bromophenol blue
30% (v/v) glycerol

[ α -32P] CTP-labeled RNAs were incubated at 30°C under standard splicing conditions in HeLa cell nuclear extract or with different amount of recombinant Tra2-beta1 proteins. 5 µl of aliquot was removed at different time points and transferred to a new tube containing 1 µl of heparin (5 mg/ml). After 5 min of heparin-treatment at room temperature, 1 µl of native RNA gel loading buffer was added and the samples were kept on ice until all samples were loaded into a 5% native RNA gel.

### 3.2.19 RT–PCR

400 ng of total RNA (200 ng / µl), 5 pmol of reverse primer, 40 U of SuperScript II reverse transcriptase, and optionally 4 U of DpnI restriction endonuclease were mixed in 5 µl of RT buffer. To reverse transcribe the RNA, the reaction was incubated at 42°C for 45 min.

1/8 of a typical reverse transcription reaction was used to amplify cDNA. The reaction was reformed in a volume of 25 µl and contained 10 pmol of specific forward and reverse primers, 200 mM dNTPs, 1 x Taq polymerase buffer and 1 U of Taq DNA polymerase. The conditions of the PCR cycles were dependent on the template to be amplified.

RT BUFFER:

300 µl 5 X First strand synthesis buffer (Invitrogen)

150 µl 0.1 M DTT (Invitrogen)

75 µl 10 mM dNTPs

475 µl $dH_2O$

### 3.2.20 In vitro transcription

The sequences were amplified from TOPO vector by PCR using T7pro and RT primers for their flanking regions. T7pro contains the T7 promoter. After PCR amplification the cDNA was purified from a 2% Agarose gel using QIAquick Gel extraction Kit (Qiagen) following the manufacturer's protocol. The purified cDNA was eluted with 35 µl $dH_2O$ and used as template in the following transcription reaction:

5-10 µl cDNA

| | |
|---|---|
| 2.5 µl | 10 x transcription buffer |
| 1.0 µl | 10mM rATP |
| 1.0 µl | 10 mM rCTP |
| 1.0 µl | 10 mM rGTP |
| 1.0 µl | 10 mM rUTP |
| 2-3 µl | α-[$^{32}$P]-UTP (400Ci/mmol) |
| 1 µl | RNase Inhibitor |
| 1 µl | T7 RNA polymerase |
| | up to 25 µl with $dH_2O$ |

The reaction was incubated for 1 hour at 37 °C. Afterwards 1 µl DNase was added and incubated for 30 min at 37 °C. Finally the reaction was purified with miniquick Spin RNA columns (Roche) following the manufacturer's protocol. The purified RNA product was checked on a poly acrylamide (SIGMA) gel and used for electrophoretic mobility shift assay.

### 3.2.21 Isolation of nuclear extract and RNA immunoprecipitation

To isolate the nuclear extract, cells were trypsinized 24-36 hours after the transfection (section 3.2.14.) and washed in 30 volumes of PBS. The pellet was then resuspended in one packed cell volume of buffer A and allowed to swell on ice for 15 minutes. Cells were lysed with a 23G hypodermic needle and nuclei were recovered by centrifugation for 20 sec at 12,000*g* at RT. The crude nuclear pellet was resuspended in two-thirds of one packed cell volume of buffer C and incubated for 30 min at 4°C with stirring. The nuclear debris was pelleted by 5 min centrifugation at 12,000*g*. Collected nuclei were then resuspended in 0.6 ml of NET-Triton, sonicated several times on ice and centrifuged. The supernatant contained the nuclear extract.

For immunoprecipitaion, one day before anti-GFP was allowed to bind to Protein A sepharose in NET-Triton overnight at 4°C. The next day, BSA was added to a final concentration of 1mg/ml and allowed to incubate at 4°C for another 2-3 hours to block the sepharose beads. In parallel, anti-IgG in NET-Triton was added to the sepharose beads and allowed to mix at 4°C for a few hours. The nuclear extract was precleared for 1 hour at 4°C with the anti-IgG bound Protein A sepharose. Immunoprecipitation of the desired GFP-Tagged protein from the precleared nuclear extract was performed with Protein A Sepharose bound anti-GFP antibody overnight at 4°C, followed by 5 washes with cold RIPA buffer. RNA was isolated using the TRIzol reagent. After ethanol

precipitation, the RNA pellet was dissolved in RNase-free water and DNAase treated as per manufacturer's protocol.

| BUFFER A: | BUFFER C: | NET-TRITON: |
|---|---|---|
| 10 mM HEPES, pH 8.0<br>1.5 mM $MgCl_2$<br>10 mM KCl<br>1 mM DTT | 20 mM HEPES, pH 8.0<br>25% (v/v) Glycerol<br>420 mM NaCl<br>0.2 mM EDTA, pH 8.0<br>1 mM DTT<br>0.5 mM PMSF | 150 mM NaCl<br>50 mM Tris-HCl, pH 7.4<br>0.1% Triton X-100<br>1 x complete mini protease Inhibitor Cocktail (Roche)<br>0.5U/ul RNAse inhibitors |

All the three buffers were supplemented with Protease and RNAse inhibitors in appropriate concentrations.

### 3.2.22 CLIP

UV-crosslinking and immunoprecipitation was done according to the protocol published by the R.Darnell's lab (Ule, Jensen et al. 2003; Ule, Jensen et al. 2005). In brief, 1,000,000 HEK293 cells for each sample was grown in 10cm dish for 2 days. Cells in dish with medium was put on ice and irradiated to UV-crosslinking at 400 mJ/$cm^2$ for 2 times (dish with all steps except for UV-crosslinking was used as a control). The lysis was performed in RIPA buffer (3.2.24) (400ul per dish) in absence of 2ul Benzonase. Cells were harvested for 30 minutes in ice and then treated with low concentrated (1ng) RNAseA in 37$^oC$ for 10 minutes. The lysates were immunoprecipitated with 15ul anti-Tra (568) serum, 900ul RIPA rescue including 0.2U/ul RNAse inhibitor (Roche) and 55ul Dynabeads Protein A (Dynal Biotech ASA, Norway) overnight. Dynabeads were prepared before by washing 3 times with 0.1M $Na_3PO_4$, pH8.1 and twice with RIPA rescue buffer (3.2.24). The next day, beads were washed 3 times each with bufferA, bufferB and bufferC (bufferA: 1× PBS, 0.1% SDS, 0.5% deoxycholate, 0.5% NP-40; bufferB: 5× PBS, 0.1% SDS, 0.5% deoxycholate, 0.5% NP-40; bufferC: 50 mM Tris–Cl, pH 7.4, 10 mM $MgCl_2$, 0.5% NP-40.). Then CIP treatment and 3' RNA linker ligation (TraX3 RNA, 3.1.10, oligoes used for CLIP) were performed for overnight at 16$^oC$. In the following day, the beads were washed 3 times with bufferC and labeled 30minutes in 37$^oC$ with γ-ATP by T4 polynucleotide kinase. After labeling, beads were washed 4

times with bufferC, the samples were mixed in 1x protein loading buffer without DTT and β-Me (50mM Tris-Cl, Ph6.8, 2% SDS, 0.1% bromphenol blue, 10% glycerol) and boiled for 10min at 70$^{o}$C. Samples were loaded into 12% SDS-PAGE gel and transferred to PROTRAN nitrocellulose (Whatman GmbH, Germany). The membrane was rinsed with 1x PBS and exposed to X-ray film over night. The band in film which around 10kD above the normal size of Tra2-beta1 were cut and treated with 200ul proteinase K solution (4mg/ml proteinase K(Roche), 100 mM Tris–Cl, pH 7.5, 50 mM NaCl, 10 mM EDTA) at 37 °C for 20 min with shaking. 7M Urea solution was added to  for another 20 minutes in 37 °C and shaking. Phenol-chloroform extraction was performed to extract RNA. The 5' RNA-linker (X6aR, 3.1.10, "oligoes used for CLIP") was then ligated to the purified RNA over night at 16$^{o}$C. After DNAse treatment and phenol-chloroform extraction, the RNA was amplified by RT-PCR using specific oligos priming the linker sequences (3.1.10, "oligoes used for CLIP"). pCR4 TOPO cloning was performed and DNA was sequenced (Agowa, Berlin) to find the RNA binding targets.

### 3.2.23 Immunostaining

Cells grown on cover slips were fixed in 4% paraformaldehyde in 1 x PBS, pH 7.4 for 20 min at 4 °C. After fixing the cells were washed three times in PBS with 0.1% Triton X-100 and blocked in PBS containing 0.1% Triton X-100 and 3% NGS for 2 hours at room temperature. Cells were then incubated with the desired antibody (diluted in PBS with 0.1% Triton X-100 and 3%BSA) overnight at 4 °C. After washing three times in PBS with 0.1% Triton X-100, cells were incubated with CY3 or CY5 coupled secondary antibodies, diluted 1:500 in PBS with 0.1% Triton X-100 for 2 hours at room temperature. After washing for another three times in PBS with 0.1% Triton X-100 the cover slips were mounted on microscope slides with Gel-Mount. Finally stained cells were examined by confocal laser scanning microscopy.

### 3.2.24 Immunoprecipitation of proteins

20-24 hours after transfection, cells were washed with 1 x PBS. Cell lysis was performed for 25 min at 4 °C on ice in 200 µl RIPA buffer per well. Lysates were collected in Eppendorf tubes and cleared by centrifugation for 1 min at 12,000 rpm. The

supernatant was diluted with 3 volumes of RIPA rescue buffer and antibody recognizing the expressed protein or an attached tag was added. After incubation on a rotating wheel for 2 hours at 4 °C 50 µl Protein A Sepharose / Sepharose CL-4B (1:1) was added and the incubation continued over night under the same conditions. The Sepharose beads were pelleted for 1 min at 1000 rpm in a microcentrifuge, followed by 3-5 washes with 400 µl of 1 x HNTG buffer. 20 µl 3 x SDS sample buffer were added to the pellet and boiled for 5 min at 95 °C to denature the proteins. After spinning down shortly the supernatant was resolved by SDS polyacrylamide gel electrophoresis and analyzed by Western blot using appropriate antibodies.

Protein A Sepharose / Sepharose CL-4B preparation:

Protein A Sepharose beads were twice washed in 15 ml $dH_2O$ and pelleted at 500 rpm for 2 min. Then equal amount of Sepharose CL-4B was added and the beads were washed two more times in RIPA rescue buffer and stored in equal amount of RIPA rescue buffer at 4 °C.

**RIPA**

| | |
|---|---|
| 1% | NP-40 |
| 1% | Na-deoxycholate |
| 0.1% | SDS |
| 150 mM | NaCl |
| 10 mM | Na-phosphate, pH 7.2 |
| 2 mM | EDTA |
| 5 mM | β-glycerolphosphate |

**+ Freshly added**

| | |
|---|---|
| 4 mM | $Na_3VO_4$ |
| 1 mM | DTT |
| 1 mM | PMSF |
| 20 µg/µl | aprotinin |
| 1µg/µl | pepstatin |
| 1 µg/µl | leupeptin |
| 1 x | PIC |
| 100 U/ml | benzonase |

**4 x HNTG**

| | |
|---|---|
| 200 mM | HEPES, pH 7.5 |
| 450 mM | NaCl |
| 4 mM | EDTA |

**RIPA rescue**

| | |
|---|---|
| 20 mM | NaCl |
| 10 mM | Na-phosphate, pH 7.2 |
| 1 mM | NaF |
| 5 mM | β-glycerolphosphate |

**Freshly added**

| | |
|---|---|
| 2 mM | $Na_3VO_4$ |
| 1 mM | DTT |
| 1 mM | PMSF |
| 20 µg/µl | aprotinin |
| 1 µg/ µl | pepstatin |
| 1 µg/µl | leupeptin |
| 1 x | PIC |

| | |
|---|---|
| 40% | glycerol |
| 0.4% | Triton-X-100 |

**+ Freshly added to 1 x HNTG**

| | |
|---|---|
| 2 mM | $Na_3VO_4$ |
| 100 mM | NaF |
| 1 mM | DTT |
| 1 mM | PMSF |
| 20 µg/µl | aprotinin |
| 1 µg/µl | pepstatin |
| 1 µg/µl | leupeptin |
| 1 x | PIC |

### 3.2.25 Coupled *in vitro* transcription and translation

cDNA of desired genes was cloned in a vector downstream of a T7 promoter. Plasmid DNA was used for a coupled *in vitro* transcription/translation reaction using the TNT® reticulozyte lysate system (Promega) according to the supplier's manual. [$^{35}$S]-Methionine (1000 µCi) was used for labelling the proteins. Translation products were verified by SDS-PAGE and autoradiography.

### 3.2.26 Expression of HIS-tagged protein in the Baculovirus system

For expression of HIS-tagged proteins the Bac-to-Bac® Baculovirus Expression System from Invitrogen was used. This method is based on site-specific transposition of an expression cassette into a baculovirus shuttle vector (bacmid) propagated in *E. coli* (Luckow, Lee et al. 1993; Ciccarone, Polayes et al. 1997).

pFastBac vector containing the gene desired for expression was transformed into *E. coli* DH10Bac™. Generating bacmid from these cells followed the Bac-to-Bac manual from Invitrogen. The isolated bacmid was tested by PCR and sequenced with the forward primers used for cloning into pFastBac-HTa (gene specific) and M13 reverse primers.

Transfection of bacmid to generate virus was performed in 6 well plates according to the Bac-to-Bac manual from Invitrogen, using unsupplemented Grace's Medium and Cellfectin (both Invitrogen).

When cells showed typical signs of infection (about 72 hours after transfection) the medium was collected from each well (about 2 ml) and transferred into sterile 15 ml falcons. The P1 viral stock was stored at 4 C, protected from light. For amplification of

P1 viral stock SF9 cells were infected at a multiplicity of infection according to the manual.

Cells were harvested about 48h after infection and expression of recombinant protein was analyzed by SDS-PAGE and Western Blot or protein staining

### 3.2.27 Purification of HIS-tagged protein in insect cells

48 hours after infection, SF9 cells were centrifuged at 500 g for 10 min. The pellet was resuspended in 1 ml of denaturing lysis buffer. The suspension was lysed with a 19 G hypodermic needle and centrifuged at 14,000 rpm for 25 min in a 5417R centrifuge (Eppendorf). The supernatant was incubated for 1-2 hours at 4 °C with Ni-NTA agarose resin (Qiagen), equilibrated once with $dH_2O$ and once with denaturing binding/washing buffer, pH 7.8. After incubation, the resin was washed twice with denaturing wash buffer, pH 7.8, twice with denaturing wash buffer pH 6.4 and once with native buffer. Protein elution from the resin was performed with native buffer containing 250 mM Imidazol. Fractions of each step were run on SDS-PAGE.

**denaturing lysis buffer**

| | |
|---|---|
| 6 M | Guanidine HCl |
| 20 mM | $NaPO_4$, pH 7.8 |
| 500 mM | NaCl |

**denaturing binding /washing buffer**

| | |
|---|---|
| 8 M | Urea |
| 20 mM | $NaPO_4$, pH 7.8 |
| 500 mM | NaCl |

**denaturing washing buffer**

| | |
|---|---|
| 8 M | Urea |
| 20 mM | $NaPO_4$, pH 7.8 / pH 6.4 |
| 500 mM | NaCl |
| 0.1% | Triton X-100 |
| 30 mM | Imidazol |

**native buffer (pH 8.0)**

| | |
|---|---|
| 50 mM | $NaH_2PO_4$ |
| 300 mM | NaCl |
| 30 mM | Imidazol |

### 3.2.28 Determination of protein concentration

Concentration of proteins was determined using BioRad Protein Assay Kit which is based on the Bradford method (Bradford 1976).

Protein in 800 µl dH2O was mixed with 200 µl of 1 x Dye reagent and incubated for 5 min at room temperature. BSA concentrations were used as standard. Absorbance

was measured in a spectrophotometer at λ = 595 nM. Concentration of protein was read from a standard curve where OD595 was plotted against concentration of BSA standards.

### 3.2.29 Electrophoresis of proteins

Proteins were resolved using denaturing SDS polyacrylamide electrophoresis (Laemmli 1970). For the separating gel 7.5-15% acrylamide was used depending on the molecular weight of the proteins and for the stacking gel 4% acrylamid was used. The proteins were mixed with sample loading buffer, boiled for 5 min and cooled on ice prior to loading. Electrophoresis was performed at 100-150V for 2-2.5 hours in SDS gel running buffer.

| separating gel (10 ml) | | 4% stacking gel (10 ml) | |
|---|---|---|---|
| 2.5 ml | 1.5 M Tris-HCl, pH 8.8 | 2.5 ml | 0.5 M Tris-HCl, pH 6.8 |
| 100 µl | 10% SDS | 100 µl | 10% SDS |
| 100 µl | 10% APS | 100 µl | 10% APS |
| 10 µl | TEMED | 10 µl | TEMED |
| according to desired conc. | 30% Acrylamide / Bis | 620 µl | 30%Acrylamide / Bis |
| up to 10 ml $dH_2O$ | | up to 10 ml $dH_2O$ | |

### 3.2.30 Western Blot

After gel electrophoresis SDS polyacrylamide gels were transferred to a nitrocellulose membrane (Protran, Schleicher & Schüll). Before placing the membrane on the gel, gel and membrane were equilibrated shortly in protein transfer buffer. The transfer was performed for 45 min at 120 V. Afterwards the membrane was blocked for 1 hour in 1 x NET gelatine at room temperature. Primary antibody diluted in 1 x NET gelatine was added and incubated overnight at 4 °C or 2-4 hours at room temperature. After three 15 min washes in 1 x NET gelatine, incubation with HRP-coupled secondary antibody, diluted in 1 x NET gelatine followed for 1-2 hours at room temperature. The membrane was washed another three times and detection of antibodies was carried out with the ECL system. The membrane was incubated for 5 min with equal amounts of ECL1 and ECL2 solutions, exposed to an X-ray film (Fuji Super RX) and developed in a Kodak X-omat 1000.

| **Transfer buffer** | | **1 x NET gelatine** | |
|---|---|---|---|
| 192 mM | Glycine | 150 mM | NaCl |
| 25 mM | Tris base | 5 mM | Tris-HCL, pH 7.5 |
| 20% | Methanol | 0.05% | Triton X-100 |
| | | 0.25% | Gelatine |
| **ECL1** | | **ECL2** | |
| 4.5 mM | Luminol | 0.003% | $H_2O_2$ |
| 4.3 mM | p-Iodophenol | 100 mM | Tris-HCL, pH 9.5 |
| 100 mM | Tris-HCL, pH 9.5 | | |

For re-blotting membranes with another antibody membranes were stripped with stripping buffer for 5-10 min before washing with NET gelatine and incubation in new first antibody.

**Stripping buffer**

| | |
|---|---|
| 30% | Trichloroacetic acid |
| 1% | Acetic acid |
| 1% | Ponceau S solution |

### 3.2.31 Coomassie Blue Staining of protein gels

To detect proteins in SDS polyacrylamide gels Coomassie brilliant blue was used. The gel was placed in staining solution for 2-3 hours at room temperature or overnight at 4 °C. Afterwards the gel was washed 2-3 times in 50% Methanol /10% acetic acid and 2-3 times in 20% Methanol / 10% acetic acid.

**Coomassie staining solution**

| | |
|---|---|
| 2.5% | Coomassie Brilliant Blue R250 |
| 45% | Methanol |
| 10% | Acetic acid |

### 3.2.32 Silver staining of protein gels

1. Sink gel at first into water
2. Incubate gel for 5 minutes in 6ml of buffer 1 containing 1,5ml TCA, 25ul formaldehyde and 50% aceton
3. Rince it in distilled water 2 -3 times and shake it in water for 5 minutes
4. Add 50% aceton (buffer 2) and keep for 5 minutes until the edges will be white
5. Put gel into 60 ml of buffer (buffer 3), containing 100ul $Na_2SO_4$, expose gel for not

more than 1 min

6. Wash the gel with distilled water and shake gel in water for 3 minutes
7. Remove water and add buffer (buffer 4) containing 60ml $dH_2O$, 60ul formaldehyde, 800ul silver nitrate. Keep the gel in buffer for 7 minutes
8. Wash gel in water not keeping it there too long
9. Put the gel into the buffer (buffer 5) containing 1g $Na_2CO_3$, 25ul of formaldehyde, 25 ul $Na_2SO_3$ in water. Develop untill all the bands are visible
10. Take the gel and put it in 10% acetic acid (stop buffer)

### 3.2.33 Nuclear and Cytoplasm extraction

PERFORM ALL STEPS ON ICE

1. Collect cells (normally from 10cm dish, scraping or trypsinizing).
2. Wash cells once with ice-cold PBS and repellet.
3. Resuspend cells in 1ml ice-cold PBS and transfer to an eppendorf tube.
4. Pellet cells at 200g for 5 minutes.
5. Resuspend cells in 200ul Sucrose buffer with NP-40 by gently pipetting with a 1000ml tip, and incubate on ice for 5 minutes to lyse.
6. Pellet nuclei by centrifugation at 1500g for 5 minutes and transfer the supernatant (cytoplasmic fraction) to a new tube. (NOTE: It's best to leave the last 50ul at the bottom of the tube out of the cytoplasmic fraction, this reduces the likelihood of contaminating the cytoplasmic fraction with nuclear protein.)
7. Gently re-suspend the nuclei in 1ml Sucrose buffer without NP-40.
8. Pellet the nuclei at 1500g for 5 minutes. Discard supernatant. This and the above step removes leftover cytoplasmic contaminants using a sucrose cushion.
9. Gently resuspend nuclei in 200ul LOW salt buffer (nuclei should be semi-granular, and intact).
10. Add 0.2X volume HIGH salt buffer and gently flick tube.
11. Continue adding 0.2X HIGH salt buffer and mix them gently flicking until 1X volume has been added OR the nuclei begin to shrink and viscosity increases (it generally takes me about 0.4X volume with HeLa cells and 1 X volume with 293 cells).

12. Incubate tubes on the rotary platform in the cold room for 20 minutes

13. Centrifuge at 13,000g for 15 minutes.

14. Retain supernatant (nuclear fraction).

QC Controls: GAPDH, Tubulin and PARP Western Blots on both nuclear and cytosolic samples, Tra2-beta1 mainly in nuclear.

NOTE: This method is low-salt, so it does not disrupt cytoskeletal interactions, which means it will pellet most if not all of any cytoskeletal proteins. This includes nuclear cytoskeletal proteins.

15. for RNA extraction, above supernatant could be extracted by Phenol-Chloroform twice and precipitate by ethanol with 2ul glycogen.

Columned associated RNA extraction based on PARIS kit (Ambion, cat.AM1921)

**Reagents:**

*Low Salt Buffer-For 10ml*

200ul of 1M HEPES pH 7.9, 20mM

2.5ml of glycerol, 25%

15ul of 1M $MgCl_2$, 1.5mM

200ul of 1M KCl, 20mM

8ul of 250mM EDTA, 0.2mM

100ul of 100mM DTT (Freshly added!), 1mM

50ul of 100mM PMSF (Freshly added!), 0.5mM

6.927ml of $dH_2O$

*High Salt Buffer-For 10ml*

200ul of 1M HEPES pH 7.9, 20mM

2.5ml of glycerol, 25%

15ul of 1M $MgCl_2$, 1.5mM

2.67ml of 3M KCl, 800mM

8ul of 250mM EDTA, 0.2mM

100ul of NP-40, 1%

100ul of 100mM DTT (Freshly added!), 1mM

50ul of 100mM PMSF (Freshly added!), 0.5mM

4.357ml of $dH_2O$

*Sucrose Buffer w/o NP-40-For 10ml*

3.2ml of 1M Sucrose, 320mM

300ul of 0.1M $CaCl_2$, 1mM

20ul of 1M MgAc, 2mM

4ul of 250mM EDTA, 0.1mM

100ul of 100mM DTT (Freshly added!), 1mM

50ul of 100mM PMSF (Freshly added!), 0.5mM

6.326ml of $dH_2O$

*Sucrose Buffer w/ NP-40-For 1ml*

1ml of Sucrose Buffer w/o NP-40

5ul of NP-40, 0.5%

### 3.2.34 *Cell fractionation and sucrose gradient centrifugation*

293 cells with or without transfection were removed from 20cm plate by trypsin and washed 2 times with 10 ml of ice-cold PBS. The cell pellet was then resuspended in ice-cold sucrose buffer with NP40 (as described in 3.2.33) and incubated on ice for 5 -10 minutes for lysis. Nuclei and insoluble material were then pelleted at 10,000 rpm for 10 min in a cold microfuge. The supernatant (cytoplasmic fraction) were loaded onto a 10%-60% sucrose gradient containing 20 mM Tris (pH 7.5), 5 mM $MgCl_2$, 100 mM KCl and centrifuged for 10 h at 27,000 rpm in a Sorval SW28 rotor. The gradients were fractionated using home made fraction collector and the absorbance of 260nm was detected each in the Bio Photometer (Eppendorf).

### 3.2.35 Yeast Two hybrid

1. Make ON culture - vortex a single colony in 10 ml YPAD. Grow at 30$^{o}$C, 250 rpm, 16 hours (After ~ 16 h the OD600 should be between 0.2 and 0.3).
2. Take 10 ml of the ON culture and transfer to 100 ml YPAD. Grow for ~ 4 h, 30$^{o}$C, shaker. Check OD600 - should be between 0.4 - 0.6 (Clontech; our lab protocol - 0.6 - 0.9)
3. Transfer cells to 50 ml Falcons and centrifuge at 2000g for 5 min at RT

4. Wash with 50 ml $dH_2O$. Resuspend by vortexing. Centrifuge 2000 rpm 5 min at RT
5. Resuspend the cells in freshly prepared 1 x LiAc /0.5x TE, 2ml. Incubate the cells at room temperature for 10 minutes.
6. Combine 1ug plasmid and 100 ug denatured sheared salmon sperm DNA and add to 100 ul yeast cells.
7. Add 700 ul 1 x LiAc/40%PEG-3350/1xTE. Vortex 10 sec. Incubate at 30$^{o}$C for 30 min, 250 rpm
8. Add 88 ul DMSO. Mix (no vortex)
9. Heat shock 15 min at 42$^{o}$C
10. Chill on ice 10 min
11. Centrifuge 2000 rpm, 1 min. Discard super, resuspend in 5 ml YPAD. Shaker, 30$^{o}$C, 1 hour.
12. Centrifuge, resuspend in TE.
13. Plate on appropriate medium (-Leu-Trp-His). Colonies appear after 2-3 days.
14. Second plate on 5mM or 10mM 3AT added plate. 30$^{o}$C for one week.

**YPAD Medium and plate:**

| | |
|---|---|
| YPD(Clontech, cat.630409) | 50g |
| Adenine sulphate (SIGMA, A3159) | 100mg |
| Autoclaved, distilled water to 1 liter | |

For agar plates, add 20g bacteriological grade agar per liter of non autoclaved YPAD medium. Adjust the pH to 6.0 with HCl. Autoclave at 121$^{o}$C for 25 minutes. Cool to 55$^{o}$C and dispense into sterile Petri dishes. Store plates when solidified upside down at 4$^{o}$C.

**SC Medium and Plates:**

Based on protocol from Proquest two-hybrid system (Invitrogen).

Premixed reagents are purchased from:

Minimal SD Base (Clontech, cat.630411)

3-Amino-1,2,4-triazole (3AT) (Fluka, lot.1211633)

amino acid

-Leu DO supplement (Clontech, cat.630414)

-Trp DO supplement (Clontech, cat.630413)

-Leu-Trp-His DO supplement (Clontech, cat.630419)

-Trp-Ura DO supplement (Clontech, cat.630427)

-Leu-Trp DO supplement (Clontech, cat.630417)

-Leu DO supplement (Clontech, cat.630414)

-Leu-Trp-Ura DO supplement (Clontech, cat.630426)

**1X LiAc/0.5X TE,**

**1 x LiAc /0.5x TE,**

**1 x LiAc/40%PEG-3350/1xTE**

Based on protocol from Proquest two-hybrid system (Invitrogen).

### 3.2.36 Statistical evaluation

The density of each band was measured (TotalLab, Phoretix) and the relative ratio between the isoforms determined. Statistics were performed using the SPSS 11.0 Software (Statistical Package for Social Sciences, SPSS Inc, Chicago, 2001).

Student T-tests were used to check for differences concerning the relative expression between the two groups in the respective brain regions. Analyses used two-tailed estimation of significance, normally a significance level of $p<0.05$ was defined to be statistically significant but not always.

## *3.3 computational tools*

| Database/ software | URL | Description | Reference |
|---|---|---|---|
| ASD | http://www.ebi.ac.uk/asd | The alternative splicing database | (Thanaraj, Stamm et al. 2004; Stamm, Riethoven et al. 2006) |

| | | | |
|---|---|---|---|
| ClustalW | http://www.ebi.ac.uk/clustalw/index.html | Multiple Sequence alignment program for DNA or proteins | (Thompson et al., 1994) |
| Human BLAT search | http://www.genome.ucsc.edu/cgi-bin/hgBlat | Sequence alignment tool similar to BLAST | (Kent, 2002; Kent et al., 2002) |
| NCBI BLAST | http://www3.ncbi.nlm.nih.gov/BLAST/ | Finds regions of sequence similarity | (Altschul et al., 1990; Altschul et al., 1997) |
| Melina | http://melina1.hgc.jp/ | Motif searching | |
| NEBcutter | http://tools.neb.com/NEBcutter2/index.php | Diggestion site | |
| PMW | http://bioinformatics.org/sms/prot_mw.html | Protein Molecular Weight | |
| T-test | http://home.clara.net/sisa/t-test.htm | P-value and T-test | |
| LOGO | http://weblogo.berkeley.edu/ | LOGO search | |
| BCM | http://searchlauncher.bcm.tmc.edu/seq-util/Options/revcomp.html | Reverse Complement of sequence | |

# 4 Results

## *4.1 Genome wide analysis of alternative splicing*

### **4.1.1** Minigene construction and *in vivo* splicing assays

4.1.1.1 Overview of the Method

A minigene contains a genomic fragment including the alternative exon(s) and the surrounding introns as well as the flanking constitutively spliced cloned in a eukaryotic expression vector. Thus, the transfected minigenes should contain all RNA-elements necessary to show the same alternative splicing pattern as the corresponding endogenous alternatively spliced gene when compared in a specific cellular environment.

In comparison to a biochemical analysis, the major advantages of analyzing splicing patterns with minigenes in vivo are: that the length of the analyzed minigene is not limiting, that a large number of cell types can be analyzed and that the analysis is based on the in vivo situation. In addition, indirect effects, such as phosphorylation or cellular differentiation can be addressed. Several parameters can be changed to analyze factors that affect alternative exon recognition. Firstly, the cell type used for transfection can be changed, e.g., tropomyosin minigenes have been transfected in muscle and nonmuscle cells and clathrin light chain B minigenes were transfected into primary neuronal cultures, as well as nonneuronal cells. In both cases, the splicing pattern of the minigenes reflected the exon usage observed for the endogenous genes in the appropriate cell system and allowed the analysis of regulatory factors.

Secondly, parts of the minigene can be changed by site-directed mutagenesis. Often, alternative exons are surrounded by weak splice sites and their improvement leads then to a constitutive exon usage. Another parameter that is often analyzed by mutagenesis of minigenes are splicing enhancers or silencers.

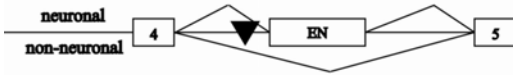
Finally, minigenes can be cotransfected with putative alternative splicing factors to identify possible trans-acting factors. This can be used to verify in vitro data collected in biochemical systems, to analyze genes that do not show splicing activity in vitro, or to analyze systems such as differentiated neurons where biochemical systems are difficult to apply.

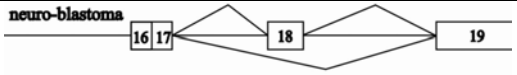
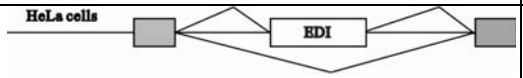
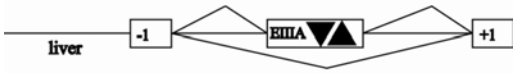
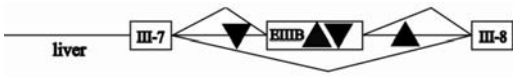
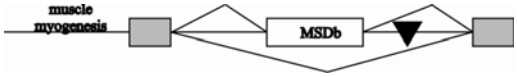
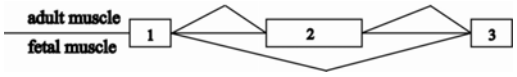
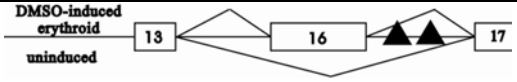
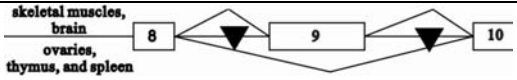
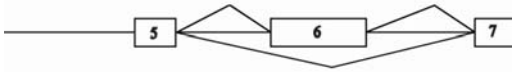
A minigene has to be constructed by either cloning genomic fragments or PCR amplified genomic DNA under the control of a suitable promoter. This construct is transfected into eukaryotic cells where the concentration of regulatory splicing factors are changed. They can be either elevated by co-transfection of an increasing amount of cDNAs expressing the factor or decreased with RNA interference. Furthermore factors can be modified by phosphorylation. In the next step, the splicing patterns of the minigene are analysed. This is done either by analyzing the RNA by RT-PCR or by indirect methods, such as the splicing dependent formation of a protein. The major advantage of the method is that almost every construct tested splices upon co-transfection. In contrast to in vitro splicing methods, there is no limit for the intron length. Furthermore, different cell types can be tested. The major disadvantage is that the method is prone to indirect effects, since intact cells are studied. (Tang, Novoyatleva et al. 2004)

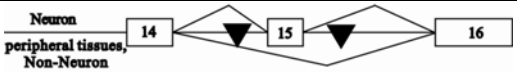
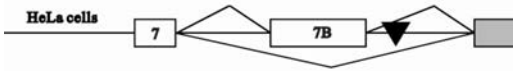
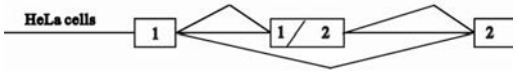
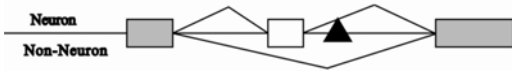
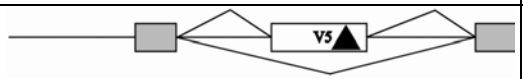
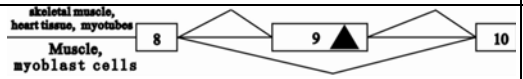
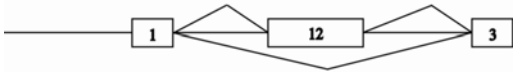
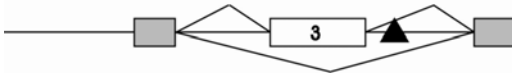
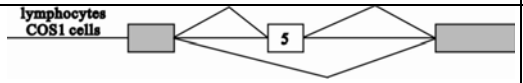
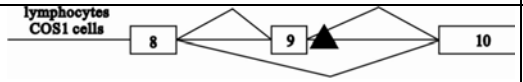
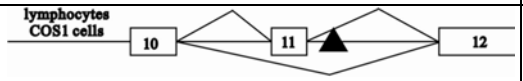
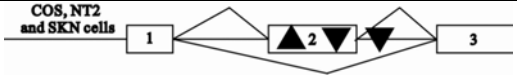
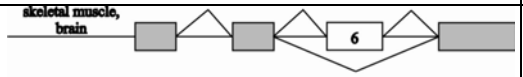
After minigenes were first cotransfected with splicing factors about 13 years ago (Caceres, Stamm et al. 1994) they have been wildly applied to study alternative splicing. Currently, basic constructs from at least 78 different genes have been reported in the literature (Table 4.1). The method has been proven to be extremely robust and reproducible. It was applied to test unknown factors for their involvement in alternative splicing, as well as to study regulatory elements, signal transduction pathways and basic splicing patterns of genes of interest.

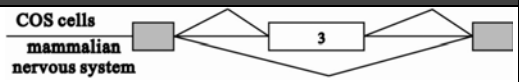
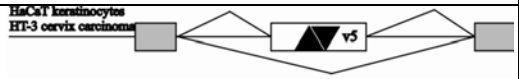
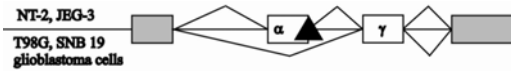
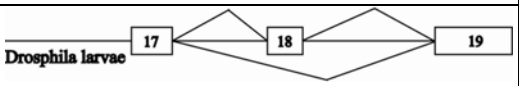
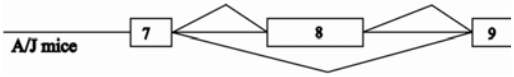
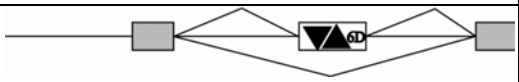
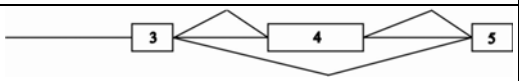
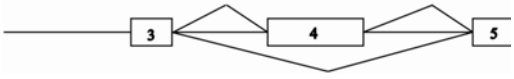
## *Table 4.1: Overview of existing minigenes*

**A:minigenes containing one cassette exon**

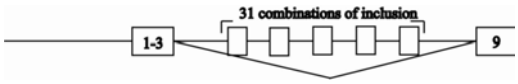
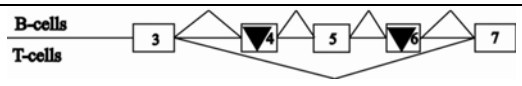
| name | species | tissue specificity minigene | Cell lines |
|---|---|---|---|
| **GABAA gamma2 24nt exon** | rat |  | HeLa |
| **Clathrin light chain B, exon EN** | rat |  | Primary rat neuronal cell, rat primary glia cells |
| **src, exon N1** | mouse |  | Neuronal Lans,neuroblastoma, HeLa,HEK293 |

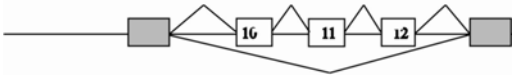
| name | species | tissue specificity | minigene | Cell lines |
|---|---|---|---|---|
| **NCAM, Exon 18** | mouse | neuro-blastoma |   | N2A,non muscle fibroblast,myoblast |
| **MHC-B, Exon N30** | human | neuronal |   | Neuronal retinoblastoma Y79 |
| **Fibronectin EDI(EDA)** | human | HeLa cells |   | HeLa |
| **Fibronectin EIIIA** | mouse | liver |   | NIH3T3, Hep3B,HepG2,HeLa,N-Mute Mouse liver |
| **Fibronectin EIIIB** | rat | liver |   | Human neuroblastoma platt, murine fibroblast 3T3,F11 |
| **Insulin receptor, Exon 11** | human | liver, muscle, kidney |   | HepG2, 3T3L1 adepocytes |
| **NCAM, Exon MSDb** | mouse | muscle myogenesis |   | Embryonic fibroblasts C3H10T1/2, bc3H1, HeLa,COS1 |
| **ccTNT Exon 5** | chicken | embryonic striated muscle adult striated muscle |   | Primary skeletal muscles from chicken embryo |
| **AMP Deaminase1 Exon 2** | rat | adult muscle fetal muscle |   | 3T3,Slo8 |
| **4.1R** | murine | DMSO-induced erythroid uninduced |   | MEL cells |
| **CASR Calcium-sensing receptor** | human | fetal kidney, HEK293 parathyroid, liver, thyroid, adult kidney |   | HEK293, Lymphoblastoid cell |
| **Caspase2** | human | skeletal muscles, brain ovaries, thymus, and spleen |   | HeLa, 293T |
| **CFTR** | human | FNMut promoter a-globin promoter |   | Hep3B |
| **CYP3A5** | human | |   | CaCo-2 |

| name | species | tissue specificity | minigene | Cell lines |
|---|---|---|---|---|
| **APP** | human<br>mouse | Neuron<br>peripheral tissues,<br>Non-Neuron |  | NIH3T3,P19,N2A,AtT20 |
| **hnRNPA1** | human | HeLa cells |  | HeLa |
| **DUP4-1 beta-globin** | human | HeLa cells |  | HeLa |
| **NMHC-B** | human | Neuron<br>Non-Neuron |  | Y79,HeLa |
| **CD44/Insulin** | mouse | |  | KLN205,LB172.3,HEK293 |
| **F1-gamma** | mouse | skeletal muscle,<br>heart tissue, myotubes<br>Muscle,<br>myoblast cells |  | C2C12,L929 |
| **MLH1** | human | |  | COS7 |
| **NF1 Fibromatosis Type1** | human | |  | Hep3B |
| **SMN** | Human<br>murine | |  | HEK293,NIH3T3,COS1, C2C12,U20S,H9,A9,HeLa |
| **Spastin Exon 5** | human | lymphocytes<br>COS1 cells |  | COS1 |
| **Spastin Exon 9** | human | lymphocytes<br>COS1 cells |  | COS1 |
| **Spastin Exon 11** | human | lymphocytes<br>COS1 cells |  | COS1 |
| **Tau Exon 10** | rat<br>human | Neuron COS1 cells,<br>PC12 cells<br>brain |  | Rat PC12(CRL-1721),Rat AR42J(CRL-1492), Monkey COS1(CRL-1650),N2A,HeLa RB |
| **Tau Exon 2** | human | COS, NT2<br>and SKN cells |  | SKN,COS,N-Tera2(NT2) |
| **Tau Exon 6** | human | skeletal muscle,<br>brain |  | Chinese Hamster |

| name | species | tissue specificity minigene | Cell lines |
|---|---|---|---|
| **Tau-Exon3/ Insulin** | human mouse |  | HN10,COS |
| **UL37** | HCMV Human cytomegalovirus |  | Human diploid Fribroblasts (HFP)cells |
| **CD44 Exon 5** | mouse |  | HaCa keratinocytes,HT-3,cervix carcinoma |
| **FGFR-1 Exon alpha** | human |  | NT-2,JEG-3,T98G, SNB19, glioblastoma cells |
| **Myosin heavy chain exon 18** | Droso-phila |  | Drosophila larvae |
| **FcgammaRIIA, Exon Tm** | human |  | HeLa, Dami cells Neutrophils |
| **Interleukin-3alpha Exon 8** | mouse |  | COS,3T3 |
| **DHFR,exon2A Reporter gene** | hamster |  | Chinese Hamster |
| **HIV-1, Exon 6D** | HIV-1 |  | CEM CD4&T-cell,HeLa-Tat |
| **SRp20 Exon 4** | human |  | Murine B, lymphoma K46 |
| **PPT Exon 4** | rat |  | NIH3T3,P19,N2A,AtT20 |

## B:minigenes containing multiple cassette exon

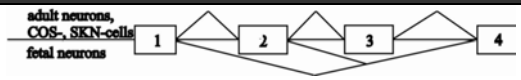
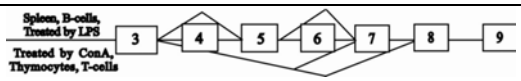
| name | species | tissue specificity minigene | Cell lines |
|---|---|---|---|
| **Fast sceletal TnT, Exon 4-7** | rat |  | COS,HeLa, Nonmuscle cells |
| **CD45 Exon4-6** | human mouse |  | B-cells, T-cells, hymoma cells(EL4,NIH3T3) |

| name | species | tissue specificity minigene | Cell lines |
|---|---|---|---|
| **spastin** | human |   | COS1 |

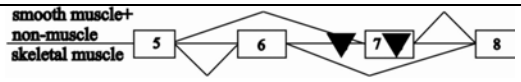
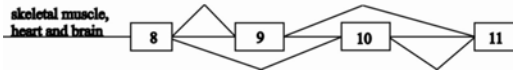
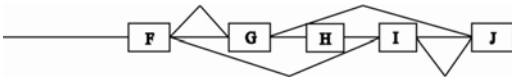
**C:minigenes containing a retained intron**

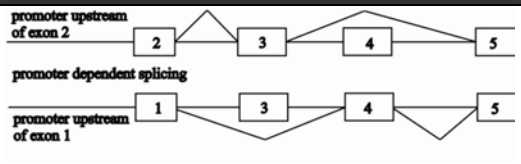
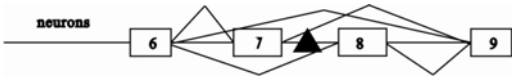
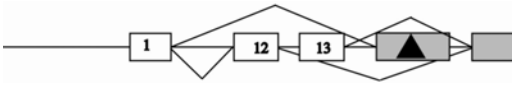
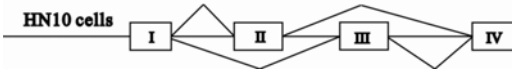
| name | species | tissue specificity minigene | Cell lines |
|---|---|---|---|
| **bGH intron D** | bovine |   | CHO |
| **Adenovirus/ Human tropomyosin** | human |  | COS,HEK293 |

**D:minigenes containing incremental combinatorial exons**

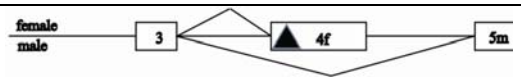
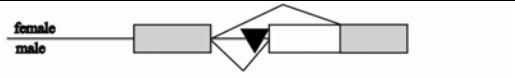
| name | species | tissue specificity minigene | Cell lines |
|---|---|---|---|
| **Tau Exon2,3** | human |   | COS,SKN, fetal Neurons |
| **CD45** | mouse |   | Spleen, B-cells, Thymocytes, T-cells |

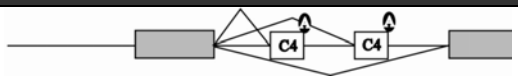
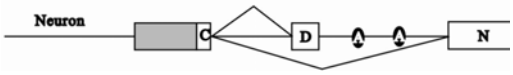
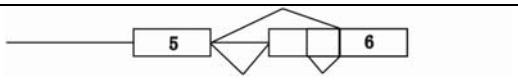
**E:minigenes containing mutually exclusive exons**

| name | species | tissue specificity minigene | Cell lines |
|---|---|---|---|
| **alpha-tropomyos** | rat |   | Muscle and nonmuscle cells |
| **Alpha-tropomyos** | human |   | COS1, myoblasts |
| **beta-tropomyos** | chicken |   | Mouse and quail muscle cells,HeLa |
| **beta-tropomyos** | rat |   | HeLa |
| **Pyruvatekina** | human |   | dRLh-84cells and hepatocytes |
| **Albumin Exon G,H** | rat |   | Nonhepatic cells,COS1,Hepatoma cells,HLE |

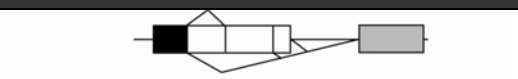
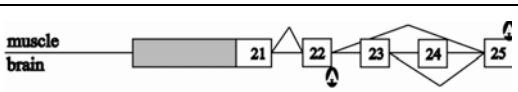
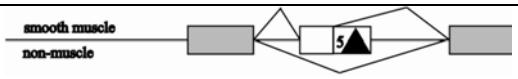
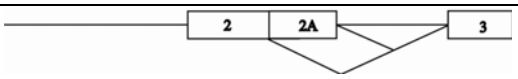
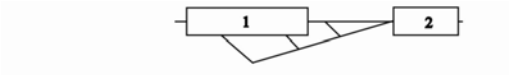
| name | species | tissue specificity minigene | Cell lines |
|---|---|---|---|
| **MLC Exon 1-4** | rat |  | HeLa |
| **FGFR K-SAM** | human |  | HeLa,293 |
| **APP Amyloid precursor protein** | rat human |  | Spleenocyt-es and thymocytes of mice |
| **BCR-ABL fusion gene** | human |  | 293T,CV1, CML,K562 EM3 |
| **FGF-R2** | rat |  | AT3,DT3 |
| **GlyRalpha 2** | human |  | 293 |
| **hTra2-beta** | human |  | HEK293 |

**F:minigenes containing alternative 3' splice site**

| name | species | tissue specificity minigene | Cell lines |
|---|---|---|---|
| **CT/CGRP** | human |  | HeLa,CHO, glioblastoma T98G,F9,teratocarcinoma cells |
| **dsx RO** | Droso-phila |  | Drosophila KC cells |
| **M-tra** | Droso-phila |  | Drosophila melanogaster |
| **BPV-1** | rat |  | HeLa |
| **CFTR** | human |  | COS7,IB3 |

| name | species | tissue specificity minigene | Cell lines |
|---|---|---|---|
| **Calcitonin/ DHFR** | hamster |   | 293 |
| **Calcitonin With PNP** | Human PNP:Ecoli |   | MTC,T98G, glioblastoma, HepG2 (hepatoma) |
| **CNV Nasi/pCV** | Droso-phia |   | Drosophila |
| **Thrombo-poietin** | human |   | Hep3B |

**G:minigenes containing alternative 5' splice site**

| name | species | tissue specificity minigene | Cell lines |
|---|---|---|---|
| **E1A** | Adenovirus |  | COS7 |
| **SERCA2 Exon 2a** | hamster |   | Mouse neuroblast-oma derived C1300 derived N2a cells |
| **Caldesmon** | human |   | COS M6,HeLa |
| **SWAP** | human |   | COS7 |
| **SV40 t-antigen** | human |   | HeLa,293 |
| **beta-globin beta-thalassemic allele** | human |   | HeLa,Neuron |
| **Adenovirus / human tropomyosin** | human |  | COS,293 |

**H:minigenes containing alternative 5' and 3' splice site**

| name | species | tissue specificity | minigene | Cell lines |
|---|---|---|---|---|
| **Fibronectin Exon IIICS** | human | liver, adult: skipping of CS1 and CS5 |   | COS7,HeLa |

### 4.1.1.2 Human Clk2 Minigene Construction and *in vivo* splicing assay

An alternative to cloning large pieces of DNA which we have to do is the usage of exon trap vectors**.** Here an alternative exon from human Clk2 is cloned between two constitutive exons, usually derived from insulin. This chimeric gene is then analyzed similar to a genomic construct. This approach has two advantages: (i) often exon trap constructs are easier and faster to clone and (ii) if these constructs behave like the endogenous gene, it is clear that the regulatory region is confined to the cloned exon.

The Clk2 pre-mRNA splicing in endogenous was studied using the in vivo splicing assay approach. Exon4 of Clk2 is an alternative spliced exon and the exon 4 skipping isoform generating a frameshift which results in an inactive form (Nayler, Schnorrer et al. 1998)(Fig.4.1.A). To study the factors and the cis-acting sequences that regulate the alternative splicing of the human Clk2 pre-mRNA, the human clk2 minigene containing exon3 to exon5 with introns was constructed by amplifying a fragment using primers hClk2-BAM-forward (5'-cgcggatccagtgctccacctgccttg-3') and hClk2-NOTI-reverse (5'-tatgcggccgcaagccccatataaccccaac-3'). The PCR product was cloned into a TA cloning vector (Invitrogen Life Technologies, Germany) and confirmed by sequencing. The insert was then cloned into the ExonTrap vector pET (MoBiTec, Goettingen, Germany)(Fig.4.1.A). The amplification of the resulting mRNA was performed using the two flanking insulin exons. The size of PCR products from hCLK2 minigene are: 490 and 578bp. The minigene splicing was studied by in vivo splicing assays. In a typical in vivo splicing assay, the minigene is cotransfected in HEK293 cells with increasing amounts of an expression vector carrying cDNA encoding splicing factors. Here we show the influence by hTra2-beta1 and CLK2 (Fig.4.1.B). The splicing pattern was analyzed by RT-PCR 16 hours after the transfection.

CLK2 is a SR protein kinase and it regulates a SR like protein tra2-beta1. Presence of the CLK2 kinase prevents the usage of exons 2 and 3, generating the htra2-beta3 mRNA (Stoilov, Daoud et al. 2004). Furthermore, tra2-beta1 utilizes a negative feedback loop to regulate the splicing of its own exon 2. The ratio between htra2-beta1 and htra2-beta3 is balanced in cells (Stoilov, Daoud et al. 2004). We are then interested in the potential effect for CLK2 by such tra2-beta balance: does tra2-beta1 extend it's regulation with feed back loop to own kinase CLK2?

As shown in Fig. 4.1.A, an increase of hTRA2-beta1 promotes skipping of clk2 exon 4, most likely through binding to an htra2-beta1 motif AAGAGCGA present in the 3' part of the clk2 exon 4. Similar to the situation in clk1 (Duncan, Stojdl et al. 1997), clk2 promotes skipping of its exon 4 generating a frameshift resulting in an inactive form. We then performed similar assays with the expression constructs encoding the inactive proteins hTRA2-BETA3 and CLK2-KR. CLK2-KR is an inactive CLK2 mutant because it cannot bind ATP. We tested this mutant to determine whether clk2tr protein might influence pre-mRNA splicing if the mRNA escapes nonsense-mediated decay under special conditions. We found that these inactive forms have no effect on clk2 pre-mRNA splicing (Fig. 4.1.B). Our previous research has revealed that CLK2 influences the splicing pattern of tra2-beta by promoting exon skipping and the formation of the inactive protein isoform TRA2-beta3(Stoilov, Daoud et al. 2004).This suggests that the amount of active TRA2-beta1 and CLK2 is controlled together through a feedback of alternative splicing decisions.

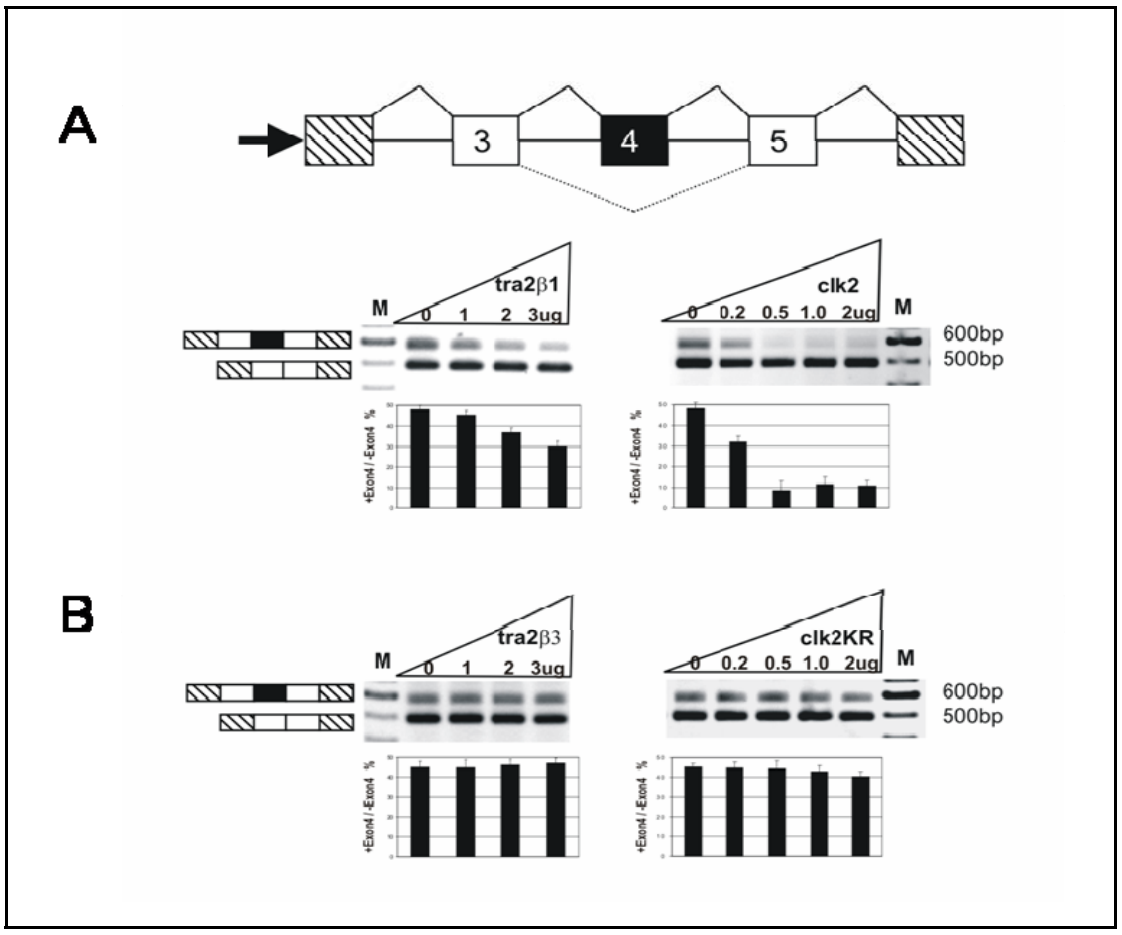



**Figure 4.1. Cotransfection of the CLK2 minigene with splicing factors.** A. hTRA2-beta1 and CLK2 promote skipping of clk2 exon 4. A minigene consisting of clk2 exon 3, 4 and 5 was cotransfected with increasing amounts of TRA2-beta1 and CLK2 expression cDNAs. The mRNA formed was analyzed by RT-PCR. Exons are shown as boxes, introns as lines. The alternative exon is black, insulin exons from the exon trap vector are striped. The structure of the PCR products is schematically indicated on the left. Numbers indicate µg of cDNA construct transfected, M: marker. B. hTRA2-beta3 and CLK2KR have no influence on clk2 exon 4 An experiment similar as in A was performed using constructs expressing htra2-beta3 and CLK2-KR. CLK2 KR is a catalytic inactive mutant of CLK2. In panels A and B, error bars indicate standard deviations of at least three independent experiments.

### 4.1.1.3 Human BACE1 Minigene Construction and *in vivo* splicing assay

Alzheimer's disease (AD) is the most common cause of age-related dementia. Pathologically, AD is characterized by the deposition in the brain of amyloid-beta peptides derived from proteolysis of amyloid precursor protein (APP) by beta-site APP cleaving enzyme 1 (BACE1) and gamma-secretase. To investigate a possible involvement of missplicing of the BACE1 gene in Alzheimer's Disease, we constructed a minigene and tested it with factors deregulated in Alzheimer's Disease.

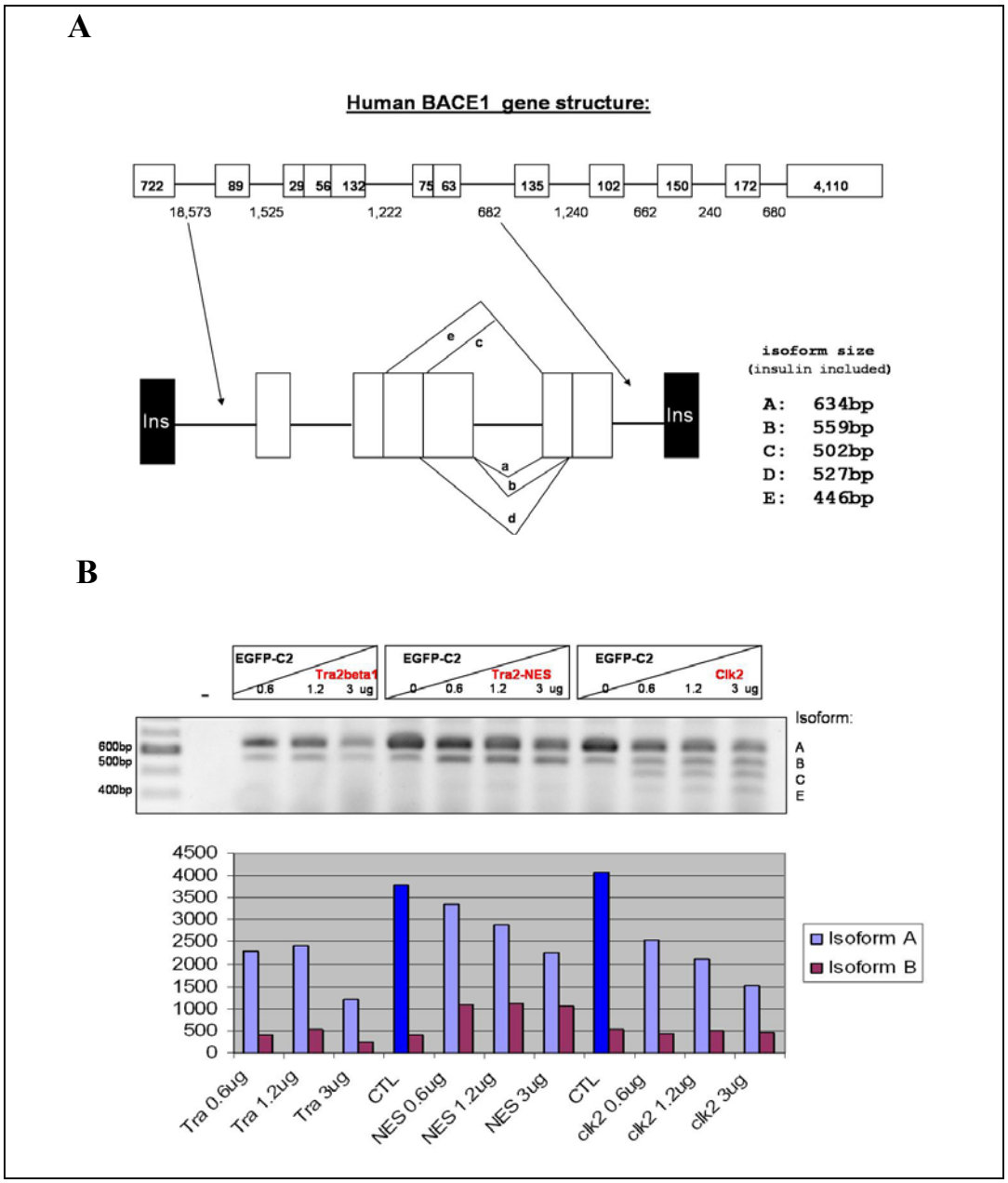



**Figure 4.2 BACE1 minigene.** A. human BACE1 gene structure and minigene construction. Five Isoforms A,B,C,D,E are found by RT-PCR. The size of isoforms basepairs are listed. B.Splicing factor Tra2-beta1 and Clk2 are co-transfection with BACE1 minigene. Ins: Insulin Exon from ExonTrap vector; NES: Nuclear Export Signal.

Over the last six years, numerous research groups in both academia and industry have synthesized inhibitors of BACE1 (isoform A) in the hope of developing a therapy to

halt or even reverse the progression of Alzheimer's disease. The BACE1 minigene was constructed in a similar way as the Clk2 minigene. The BACE open reading frame encodes a protein of 501 amino acids containing a 21-amino acid signal peptide followed by a 24-amino acid proprotein domain (Isoform A). The alternative splicing leads to a deletion of 132, 75 and 207 nucleotides in frame (isoform C, B, D) and these transcripts encoded 457, 476 and 432 prepro-BACE, designated as BACE-I-457, BACE-I-476 and BACE-I-432, respectively (Tanahashi and Tabira 2001). As shown in Fig.4.2.A, all isoforms differ in 3th and 4th Exons. To investigate RNA variants of the BACE1 gene in Alzheimer's Disease, exon 2,3 and 4 are inserted to Exontrap vector pET (MoBiTec, Goettingen, Germany). All the endogenous isoforms indicated variants are detected with minigene trasfected 293 cells. Cotransfection assay shows that Clk2 and Tra2NES strongly inhibit BACE1 variant A while variant B is continent which is predicted to have no coenzyme properties (Fig.4.2.B).

#### 4.1.1.4 Human RPL3 minigene

RPL3 regulates alternative splicing of its own pre-mRNA and forms a negative feedback loop (Cuccurese, Russo et al. 2005). Tra2-beta1 as a typical splcing factors binds to RPL3 which could indicate a new possible functions in translation (details in 4.2), RPL3 is alternative spliced and undergoes NMD by degrading the non-functional intron containing transcript (upstream of Exon4). Moreover, there are potential Tra2-beta RAAG rich binding site in the RPL3 alternative spliced sequence and Exon4. To know whether Tra2-beta1 regulates RPL3 mRNA splicing, an RPL3 minigene based the truncated variant was constructed (Fig.4.3.A).

Overexpression Tra2-beta1 slightly decrease RPL3 NMD degraded variant (Fig.4.3.B upper band) and seems stimulate RPL3 full length transcript (Fig.4.3.B).

Since both RPL3 and tra2-beta1 regulate *rpl3* pre-mRNA, we checked whether these two proteins' localization to special parcels in living cells. Interestingly, in HEK293 and Hela cells, we did not detect colocalization between Tra2-beta1 and RPL3 except during the mitosis. Tra2-beta1 localizes in nucleoplasm while RPL3 with EGFP tag stays in nucleolus (Fig.4.3.C). It indicates a potential difference of regulating pathways from two genes or in a different regulating process.

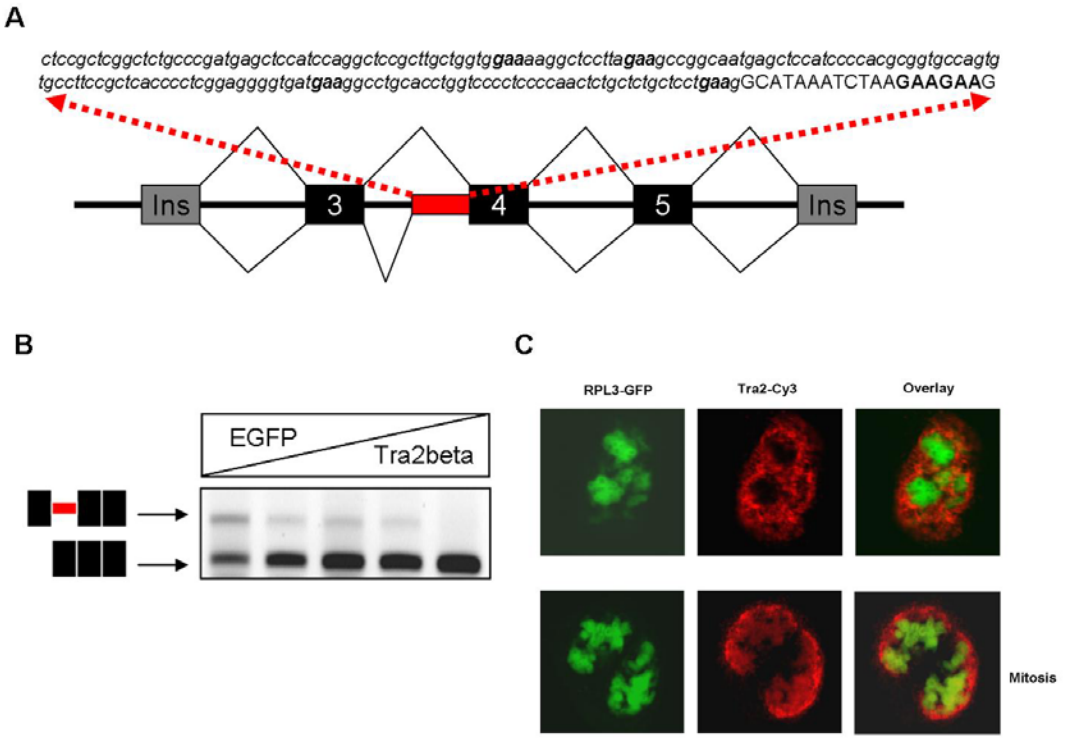



**Figure 4.3 RPL3 minigene construction and cotransfection with tra2-beta1.** A. human RPL3 gene structure. B. Splicing factor Tra2-beta1 is co-transfection with RPL3 minigene. C. Overlay of RPL3 and Tra2-beta1

### 4.1.1.5 A conserved region in human Tra2-beta minigene contributes to regulation of Exon III

#### *4.1.1.5.1 S/MAR like element in the first intron of Tra2-beta minigene*

We found that the first intron of Tra2-beta1 is conserved among human, rat and mouse, especially in the 2kb region upstream of Exon2 (Fig.4.4). Using bioinformatics tools we identified a S/MAR like element inside this region (Fig.4.5A). To the potential function of this sequence, around 1kb sequence span this S/MAR like element was deleted or inverted by PstI restriction in the MG-Tra minigene (MG-Tra-Del) (Fig.4.5B).

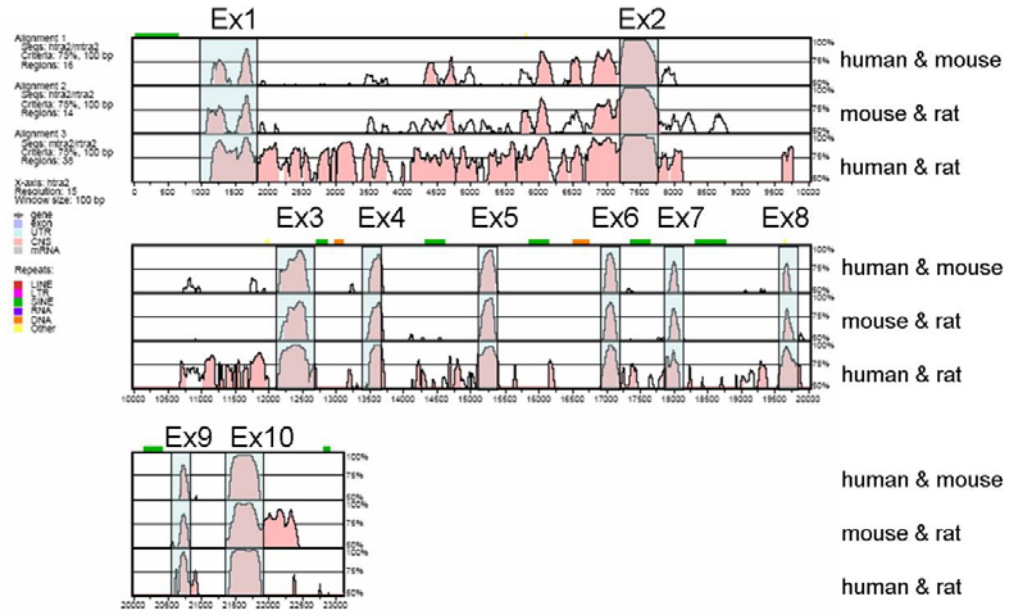



**Figure 4.4 conservation analysis of DNA sequence in Tra2-beta gene.** Tra2-beta1 DNA conservation was analysized among human, mouse and rat. The first line shows conservation between human and mouse, 2nd line shows conservation between mouse and rat, 3th line shows conservation between human and rat. Identity more than 60% are showed in pink peak. The localization of 10 exons is marked.
Online tools was used from: http://genome.lbl.gov/vista/index.shtml

Interestingly, the two minigenes gave a very different splicing pattern in vivo minigene assay as shown in Fig.4.5C lane1 and lane2: the most abundant product is Tra2-beta1(exon3 included) in MG-Tra, it changed to Tra2-beta3 (exon3 excluded) in a new constructed and 1kb conserved intron deleted MG-Tra-Del minigene. It shows that the conserved intron sequence contains a potential long distance regulation element to Tra2 Exon3 splicing.

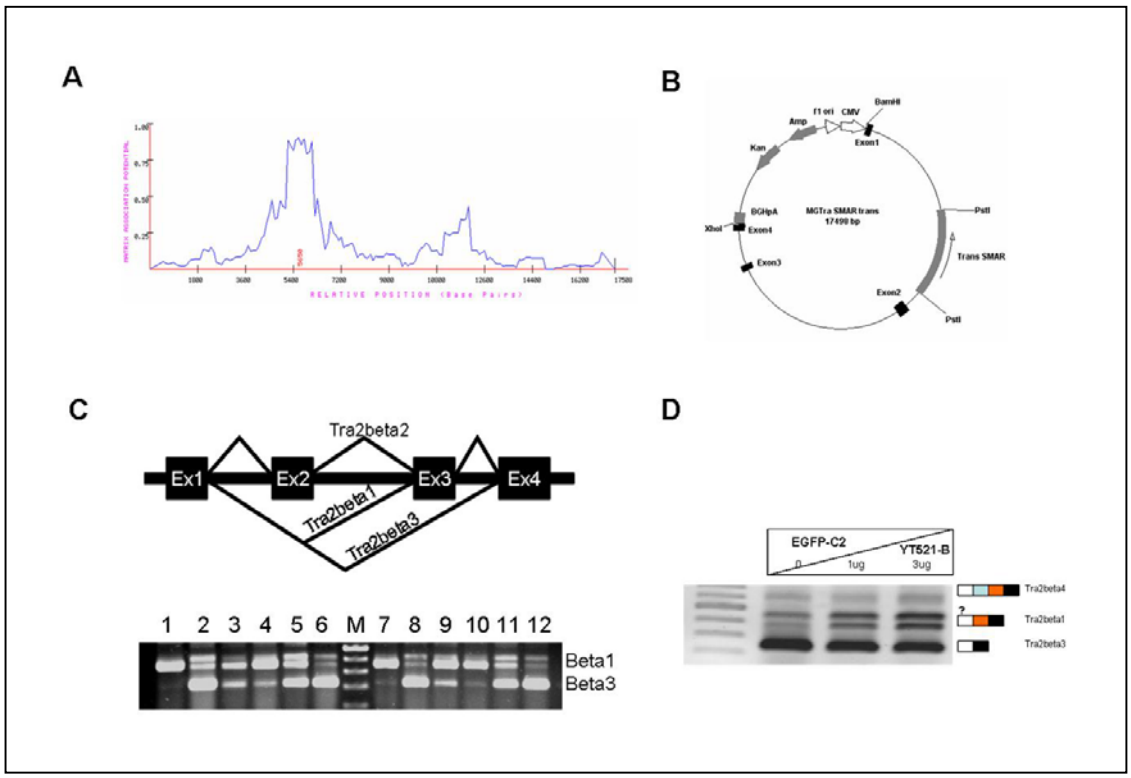



**Figure 4.5 S/MAR like element analysis in Tra2 minigene.** A. Tra2 minigene sequence analysis for potential S/MAR element Online tools from webpage: http://www.futuresoft.org/modules/MarFinder/, SMAR test: http://www.genomatix.de/. B. MG-Tra minigene construction and S/MAR like element localization. The intronic S/MAR like element was deleted by cutting out a fragment with two PstI sites. C. Co-transfection of splicing factors. With MG-Tra, MG-Tra -Del (MG-Tra with intronic S/MAR like element deletion), SarE-MG-Tra (Tra minigene with SarE), SarE-MG-Tra -Del (SarE-MG-Tra with intronic S/MAR like element deletion). Lane 1-12: MG-Tra with 1.EGFP, 3.CLK2, 4.Tra2-beta1; MG-Tra-Del with 2. EGFP, 5.SAF-B, 6.Tra2-beta1; SarE-MG-Tra with 7.EGFP, 9.CLK2, 10.Tra2-beta1; SarE-MG-Tra-Del with 8. EGFP, 11.SAF-B, 12 Tra2-beta1. D.Tra2-beta1 is stimulated by YT-521B in Tra2 Del minigene. 1 and 3 ug YT-521B was co-transfected with MG-Tra-Del minigene

Tra2-beta1 is the abundant variant in both endogenous and minigene tranfected cell as detected by RT-PCR. A new minigene was constructed since it is often have to detect changes in the tra2-beta1 isoform. With the conserved intron deletion, this minigene allows to detect a small change in tra2-beta1. The splicing factors are co-transfected with MG-Tra-Del to detect possible effect on the new minigene. Tra2-beta1 stimulates Tra2-beta3 by a feedback loop regulation which was shown in Fig.4.5C, lane6.

Clk2 and SAF-B increase Tra2-beta1 (Fig.4.5C, lane 3 and 5). Similar result was found in SarE inserted MG-Tra-Del (Fig.4.5C, lane7-12).

Previous research showed that YT-521B has no effect to Tra2-beta splicing, here we found that it slightly increased Tra2-beta1 when co-transfected with MG-Tra-Del (Fig.4.5D).

#### *4.1.1.5.2 Clk2 and IL4 with 2 SMAR element SARE and SARW in stable transfected 3T3 cell*

Scaffold Attachment Factor-B (SAF-B) is an interacting partner for tra2-beta1. It was also found in Yeast two hybrid screens which were designed to find novel binding proteins of splicing factor SRp30c, the CTD of the RNA polII and SR protein kinase CLK2. It serves as a molecular base to assemble a 'transcriptosome complex' in the vicinity of actively transcribed genes (Nayler, Stratling et al. 1998). SAF-B was independently isolated as a nuclear scaffold component binding S/MAR DNA (Renz and Fackelmayer 1996). Therefore we asked whether the S/MAR DNA is associated with general transcription and splicing.

Two S/MAR element SarE and SarW were selected to construct new minigenes. Vector contains those two S/MAR element and 2 FRT(or LoxP) recombinant sites were used for construction (from Prof. Juergen Bode's lab) and inserted in the human clk2 gene (Glatz, Rujescu et al. 2006) and mouse IL4 minigenes(Rafalska, Zhang et al. 2004). Plasmids which S/MAR replaced by similar size of Lamda DNA were used as control. Details of integrated minigene structures are shown in Fig.4.6A. The construct was then stably integrated into NIH/3T3 genome by homologous recombination integration. FISH result shows the integration site localized in chromosome telomere (Fig.4.6C).

As shown in Fig.4.6B, S/MARs significantly induces CLK2 exon4 inclusion. Interestingly, we found the transcription level was also increased when S/MARs exist. Other cell lines with different integration site were used for similar experiment, such as NIH/3T3 N1 and N15 (data not shown), all of them showed a marked influence by S/MARs on the alternative splicing pattern of CLK2. The effect is highly dependent on the integration site of the minigene.

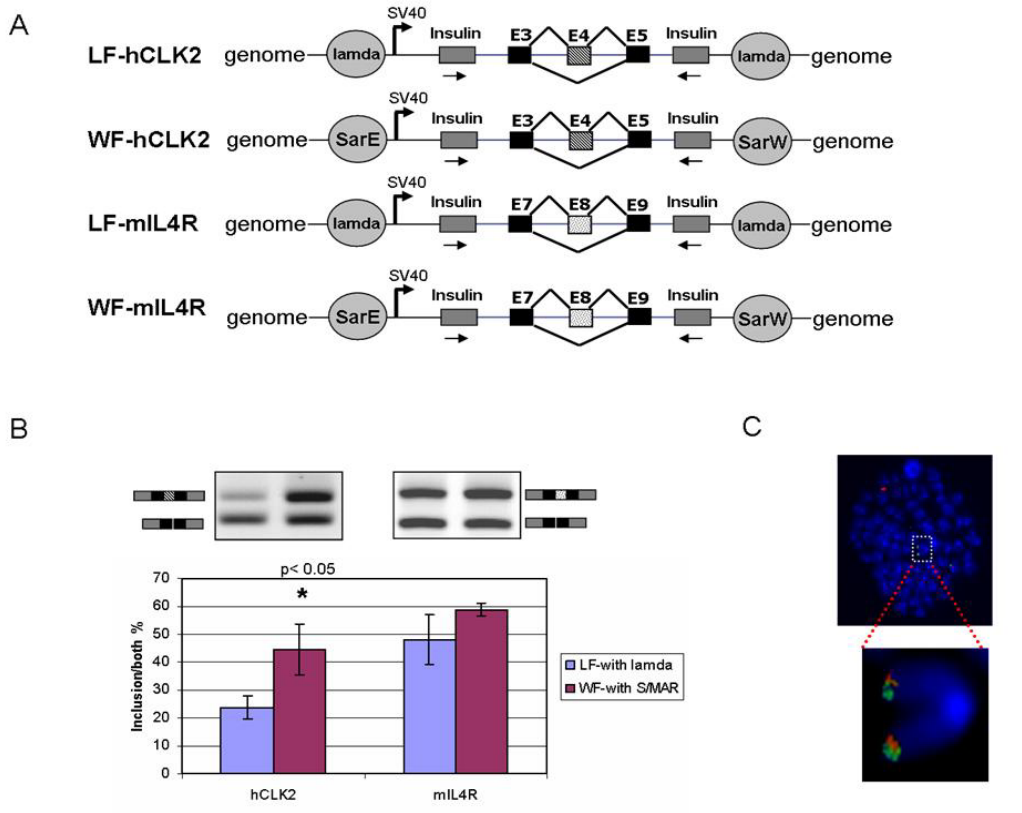



**Figure 4.6 stable transfected human CLK2 and mouse IL4R minigene.** A. integrated minigene structure in genome. S/MAR element (SarE and SarW) inserted minigenes were integrated to genome of NIH/3T3 N40 cell line (WF indicates minigene containing S/MAR element), S/MAR were changed to similar size lamda DNA as control (LF indicates minigene containing the lamda control). RT-PCR primers are marked by arrow. B. in vivo splicing assay for S/MAR coupled minigenes. Blue columns for RT-PCR from lamda DNA coupled control minigenes, red columns for RT-PCR from S/MAR coupled minigenes. Significant p value is marked above column. C. FISH for NIH/3T3 N40 cell's integration site. Chromosome are marked in blue, integration sites are marked in red and green

### 4.1.1.6 Troubleshooting of minigene assay

**• transfection efficiency**

The most crucial parameter for the success of an *in vivo* splicing experiment is the transfection efficiency, especially when cotransfections with putitative *trans*-acting factors are performed. We therefore usually employ EGFP tagged cDNA in cotransfection experiments that allow an easy monitoring of the transfection efficiency. Cell lines that allow high transfection efficiency should be used whenever possible. We

routinely use HEK293 cells and reach 90% transfection efficiency. Reasons for lower efficiencies are usually dense seeding of cells, a high passage number of cells or a deviation of the pH of the transfection solution caused by not transfecting in a 3% $CO_2$ atmosphere.

• **reproducibility**

*In vivo* splicing assays are generally well reproducible when several parameters are kept constant. For transfection cells should be always plated at the same density. It is also important to keep the time between seeding and transfection, as well as the actual transfection time constant.

• **Sporadic bands**

Sometimes, we observed sporadic bands whose appearance depended on the transfection time. When using new minigenes, the transfection time should therefore be optimized.

• **autoregulation**

Several splicing factors seem to autoregulate their expression levels e.g.tra2-beta1, hnRNA1, SF2/ASF. This can result in a substitution of the endogenous protein by the transfected cDNA, which means that the concentration of this splicing factor will not be dramatically changed. The autoregulation needs some time to occur and if observed, the time between transfection and cell harvesting can be shortened. Therefore, it is best to perform the analysis in transient transfection sytems. Western blots are needed to monitor effects on trans-acting factors.

• **contamination**

As with all PCR based methods, DNA contaminations are a major problem. It is therefore advisable to make aliquoted stocks of all solutions and if possible to separate the PCR setup form the DNA analysis.

• **heterodimers**

Often, the simultanous generation of two PCR products that differ only in the presence or absence of short exonic sequences results in the formation of a heteroduplex that

consists of two DNA strands differing by this exonic sequence. The heteroduplex usually migrates as a third PCR product. In our hands, heteroduplex formation increases when the reaction products are stored for longer time and if too many cycles in the PCR amplification are used. These parameters should therefore be minimized. If the problem badly persits, primers can be chosen that amplify each isoform individually.

### 4.1.2 Alternative splicing in Alzheimer's Disease

Alzheimer's disease (AD) is the most prevalent disease of dementia affecting more than 20 million people in worldwide (Blennow, de Leon et al.). This number highlights the need to study the mechanism of the disease. Several genes are identified to contribute Azheimer's disease's progress and are used as laboratory biomarkers, such as Tau(Ballatore, Lee et al. 2007), APP(Shen and Kelleher 2007), and presenilin(Cruts, Hendriks et al. 1996). All three genes are associated with alternative splicing regulation(Glatz, Rujescu et al. 2006) which raises our interest: whether alternative splicing plays important role in Alzheimer's Disease? More than 90% of Splicing factors themselves are alternative spliced by EST searching from BLAST (data not shown). So we firstly concentrate on the splicing factors' alternative splicing. To our hypothesis, a new oligo-junction array from manfully collected splicing factors was designed.

#### 4.1.2.1 Human Splicing factors collection and custom chip design

As shown in Table 4.2, in collaboration with Prof. Juan Valcarcel group, we collected around 300 published splicing and related factors and listed sequences from Ensemble and NCBI references. Based on every alternative spliced exon (1260 events) for each factor, 20mer and 60mer oligo junction array are designed and an array was produced by Exonhit company (www.exonhit.com).

**Table 4.2 Human splicing factors in chip and events percentage**

| Gene Name | Ensembl ref | NCBI ref |
|---|---|---|
| **15.5 tri-snRNP / NHP2L1** | ENSG00000003756 | NM_005066 |
| **9G8 / SFRS7** | ENSG00000005007 | AB007925 |
| **ABL1** | ENSG00000007392 | AL834470 |
| **ACINUS** | ENSG00000010244 | BQ434974 |
| **ARL6IP4** | ENSG00000010810 | BX537771 |

| Gene Name | Ensembl ref | NCBI ref |
|---|---|---|
| **ASF / SFRS1** | ENSG00000011304 | NM_000176 |
| **ASR2B** | ENSG00000013441 | NM_001212 |
| **BAT2** | ENSG00000013573 | NM_001240 |
| **BAZ1A** | ENSG00000015479 | NM_001241 |
| **BAZ2A** | ENSG00000033030 | NM_001261 |
| **BAZ2B** | ENSG00000048740 | NM_001280 |
| **BCAS2** | ENSG00000060138 | NM_001357 |
| **BUB3** | ENSG00000060688 | NM_001358 |
| **C21ORF66** | ENSG00000061936 | NM_001396 |
| **CBP20 / NCBP2** | ENSG00000063244 | NM_001402 |
| **CBP80 / NCBP1** | ENSG00000064607 | NM_001414 |
| **CCNT1** | ENSG00000065978 | NM_001415 |
| **CCNT2** | ENSG00000066427 | NM_001416 |
| **CDK9** | ENSG00000070756 | NM_001469 |
| **CF I-68kD / CPSF6** | ENSG00000071894 | NM_001533 |
| **CHERP** | ENSG00000075856 | NM_001967 |
| **CIRBP** | ENSG00000076053 | NM_002025 |
| **CLK1** | ENSG00000076108 | NM_002037 |
| **CLK2** | ENSG00000076650 | NM_002092 |
| **CLK3** | ENSG00000077312 | NM_002139 |
| **CLK4** | ENSG00000078269 | NM_002212 |
| **CPSF1** | ENSG00000079134 | NM_002370 |
| **CPSF2** | ENSG00000079785 | NM_002442 |
| **CPSF4** | ENSG00000080815 | NM_002486 |
| **CPSF5** | ENSG00000082258 | NM_002515 |
| **CRN / CRNKL1** | ENSG00000083896 | NM_002516 |
| **CUG-BP** | ENSG00000084072 | NM_002568 |
| **CUG-BP / BRUNOL4** | ENSG00000085872 | NM_002669 |
| **CUGBP2(BRUNO-L3)** | ENSG00000086589 | NM_002713 |
| **Cyp60 / PPIL2** | ENSG00000087087 | NM_002819 |
| **CypE / PPIE** | ENSG00000087365 | NM_002897 |
| **DDX1** | ENSG00000088247 | NM_002904 |
| **DDX10** | ENSG00000089280 | NM_002911 |
| **DDX11** | ENSG00000090621 | NM_002967 |

| Gene Name | Ensembl ref | NCBI ref |
|---|---|---|
| **DDX3** | ENSG00000092199 | NM_003016 |
| **DDX9** | ENSG00000092201 | NM_003017 |
| **DEK** | ENSG00000092208 | NM_003089 |
| **DGSI / DGCR14** | ENSG00000092277 | NM_003090 |
| **DNABINDING PROTEIN A** | ENSG00000096063 | NM_003093 |
| **DUSP11** | ENSG00000096746 | NM_003094 |
| **DYRK1A** | ENSG00000097007 | NM_003095 |
| **ECM2** | ENSG00000099622 | NM_003133 |
| **EEF1A1** | ENSG00000099783 | NM_003137 |
| **EIF2A** | ENSG00000099995 | NM_003142 |
| **EIF2B1** | ENSG00000100023 | NM_003146 |
| **EIF2B2** | ENSG00000100028 | NM_003252 |
| **EIF2B4** | ENSG00000100056 | NM_003431 |
| **EIF2-Beta / EIF2S2** | ENSG00000100109 | NM_003457 |
| **EIF2S3** | ENSG00000100138 | NM_003472 |
| **EIF3S10** | ENSG00000100410 | NM_003563 |
| **EIF3S4** | ENSG00000100419 | NM_003584 |
| **EIF4A1** | ENSG00000100603 | NM_003616 |
| **EIF4A2** | ENSG00000100650 | NM_003651 |
| **ERG** | ENSG00000100813 | NM_003675 |
| **EWS** | ENSG00000101213 | NM_003685 |
| **F23858** | ENSG00000101343 | NM_003750 |
| **FLJ10206** | ENSG00000101361 | NM_003755 |
| **FLJ20542** | ENSG00000101371 | NM_003769 |
| **FLJ90157** | ENSG00000101489 | NM_003787 |
| **FMR2** | ENSG00000101746 | NM_003819 |
| **FNBP1** | ENSG00000102241 | NM_003898 |
| **FNBP2** | ENSG00000102786 | NM_003902 |
| **FNBP3** | ENSG00000102978 | NM_003908 |
| **FRG1** | ENSG00000104824 | NM_003910 |
| **FUBP1** | ENSG00000104852 | NM_003913 |
| **FUBP3** | ENSG00000104859 | NM_003946 |
| **FUS** | ENSG00000104897 | NM_003992 |
| **FUSIP1** | ENSG00000104967 | NM_003993 |

| Gene Name | Ensembl ref | NCBI ref |
|---|---|---|
| **FXR1** | ENSG00000105618 | NM_004071 |
| **FXR2** | ENSG00000105705 | NM_004175 |
| **FYN** | ENSG00000106245 | NM_004247 |
| **G10** | ENSG00000107105 | NM_004396 |
| **GCIP-IP** | ENSG00000107164 | NM_004398 |
| **GRSF1** | ENSG00000107581 | NM_004432 |
| **Hcc1 (RNPC2)** | ENSG00000108561 | NM_004477 |
| **HDB/DICE1 / DDX26** | ENSG00000108654 | NM_004559 |
| **heln1 / ELAVL2** | ENSG00000108848 | NM_004593 |
| **hnRNP A0** | ENSG00000108883 | NM_004596 |
| **hnRNP A1 / HNRPA1** | ENSG00000109536 | NM_004597 |
| **hnRNP A2/B1** | ENSG00000109606 | NM_004630 |
| **hnRNP C / HNRPC** | ENSG00000109971 | NM_004638 |
| **hnRNP D / HNRPD** | ENSG00000110107 | NM_004697 |
| **hnRNP DL / HNRPDL** | ENSG00000110844 | NM_004698 |
| **hnRNP E1** | ENSG00000111325 | NM_004719 |
| **hnRNP E2 / PCBP2** | ENSG00000111361 | NM_004725 |
| **hnRNP F / HNRPF** | ENSG00000111406 | NM_004768 |
| **hnRNP G/ RBMXP1** | ENSG00000111605 | NM_004780 |
| **hnRNP H1 / HNRPH1** | ENSG00000111641 | NM_004792 |
| **hnRNP H3 / HNRPH3** | ENSG00000111786 | NM_004814 |
| **hnRNP I (PTB)** | ENSG00000111960 | NM_004818 |
| **hnRNP K** | ENSG00000111987 | NM_004860 |
| **hnRNP L** | ENSG00000112081 | NM_004939 |
| **hnRNP M** | ENSG00000112232 | NM_004960 |
| **hnRNP R / HNRPR** | ENSG00000112531 | NM_004966 |
| **hnRNP RALY** | ENSG00000112739 | NM_004993 |
| **HPRP3** | ENSG00000113240 | NM_005008 |
| **HPRP4 / PRPF4** | ENSG00000113580 | NM_005087 |
| **HsKin17/ KIN** | ENSG00000113649 | NM_005089 |
| **HSP71 / HSPA8** | ENSG00000114416 | NM_005105 |
| **HUD / ELAVL4** | ENSG00000114503 | NM_005131 |
| **HYPC** | ENSG00000115128 | NM_005157 |
| **ISGF3G** | ENSG00000115211 | NM_005243 |

| Gene Name | Ensembl ref | NCBI ref |
|---|---|---|
| **ISY1** | ENSG00000115241 | NM_005417 |
| **ITGB4BP** | ENSG00000115524 | NM_005520 |
| **KSRP / KHSRP** | ENSG00000115875 | NM_005626 |
| **Ku70 / G22P1** | ENSG00000115934 | NM_005778 |
| **LA-autoantigen/ SSB** | ENSG00000116001 | NM_005782 |
| **LSM1** | ENSG00000116350 | NM_005802 |
| **LSM2** | ENSG00000116560 | NM_005826 |
| **LSM3** | ENSG00000116752 | NM_005831 |
| **LSM4** | ENSG00000116754 | NM_005839 |
| **LSM7** | ENSG00000117360 | NM_005871 |
| **LSM8** | ENSG00000117569 | NM_005872 |
| **LUC7A** | ENSG00000117614 | NM_005877 |
| **LUCA15 (RBM5)** | ENSG00000117751 | NM_005968 |
| **MAGOH** | ENSG00000119718 | NM_005975 |
| **MAT3** | ENSG00000119953 | NM_006084 |
| **MGC13125** | ENSG00000121067 | NM_006112 |
| **MJD** | ENSG00000121774 | NM_006170 |
| **MOV10** | ENSG00000122566 | NM_006180 |
| **MSI1** | ENSG00000123144 | NM_006185 |
| **NDP52** | ENSG00000123596 | NM_006196 |
| **NF90 / ILF3** | ENSG00000123636 | NM_006275 |
| **NM_024038** | ENSG00000124193 | NM_006276 |
| **NOL1** | ENSG00000124214 | NM_006347 |
| **NOL3** | ENSG00000124256 | NM_006372 |
| **NOL4** | ENSG00000124487 | NM_006387 |
| **NOL5A** | ENSG00000124562 | NM_006392 |
| **NOL6** | ENSG00000124795 | NM_006425 |
| **NOSIP** | ENSG00000125676 | NM_006445 |
| **NOVA1** | ENSG00000125743 | NM_006558 |
| **NOVA2** | ENSG00000125835 | NM_006559 |
| **nPTB (PTBP2)** | ENSG00000125944 | NM_006560 |
| **NR3C1** | ENSG00000125970 | NM_006561 |
| **NSAP1** | ENSG00000125977 | NM_006590 |
| **NSEP1** | ENSG00000126005 | NM_006625 |

| Gene Name | Ensembl ref | NCBI ref |
|---|---|---|
| **NTRK2** | ENSG00000126698 | NM_006693 |
| **NUMA** | ENSG00000127054 | NM_006697 |
| **P14 (CGI110)** | ENSG00000128534 | NM_006701 |
| **P32(C1QBP)** | ENSG00000128739 | NM_006706 |
| **p54 / NONO** | ENSG00000129245 | NM_006758 |
| **p68 / DDX5** | ENSG00000129315 | NM_006802 |
| **PABP 4 / PABPC4** | ENSG00000129351 | NM_006805 |
| **PABP1** | ENSG00000130332 | NM_006924 |
| **PABPC3** | ENSG00000130520 | NM_006925 |
| **PABPC5** | ENSG00000130741 | NM_006929 |
| **POLR2C** | ENSG00000130811 | NM_006938 |
| **PPIG** | ENSG00000131051 | NM_007006 |
| **PPIL1** | ENSG00000131773 | NM_007007 |
| **PPIL3b / PPIL3** | ENSG00000131795 | NM_007056 |
| **PPM1G** | ENSG00000131876 | NM_007165 |
| **PPP1R8** | ENSG00000132463 | NM_007192 |
| **PRL1 / PLRG1** | ENSG00000132819 | NM_007279 |
| **PRP16 / DDX38** | ENSG00000133226 | NM_007318 |
| **PRP17 (CDC40)** | ENSG00000134453 | NM_007362 |
| **PRP19 (nmp200)** | ENSG00000135097 | NM_007363 |
| **PRP24 / SART3** | ENSG00000135250 | NM_012141 |
| **PRP31** | ENSG00000135316 | NM_012143 |
| **PRP4 Kinase PRPF4B** | ENSG00000135486 | NM_012207 |
| **PRP43 / DDX15** | ENSG00000135829 | NM_012218 |
| **PRP5** | ENSG00000136436 | NM_012245 |
| **PRPF18** | ENSG00000136450 | NM_012272 |
| **PRPF8** | ENSG00000136527 | NM_012311 |
| **PSF (SPFQ)** | ENSG00000136807 | NM_012321 |
| **PSIP1** | ENSG00000136875 | NM_012426 |
| **PTK6** | ENSG00000136937 | NM_012433 |
| **PUF60** | ENSG00000137168 | NM_013291 |
| **Q8N3B3** | ENSG00000137497 | NM_013293 |
| **QKI** | ENSG00000137656 | NM_013449 |
| **RAVER1** | ENSG00000138385 | NM_013450 |

| Gene Name | Ensembl ref | NCBI ref |
|---|---|---|
| **RBM8A (Y14)** | ENSG00000138398 | NM_014003 |
| **RBMS1** | ENSG00000138668 | NM_014239 |
| **RBP 7 / RBM7** | ENSG00000139218 | NM_014280 |
| **RDBP** | ENSG00000139343 | NM_014337 |
| **RENT1** | ENSG00000139910 | NM_014462 |
| **RNPC1** | ENSG00000140829 | NM_014463 |
| **RNPS1** | ENSG00000140939 | NM_014500 |
| **SAD1** | ENSG00000141759 | NM_014502 |
| **SAFB** | ENSG00000142546 | NM_014706 |
| **Sam 68 (KHDRBS1)** | ENSG00000143368 | NM_014829 |
| **SAP130 (SF3B3)** | ENSG00000143742 | NM_014884 |
| **SAP145 (SF3b145) / SF3B2** | ENSG00000144048 | NM_014977 |
| **SAP49 (SF3b49) / SF3B4** | ENSG00000144895 | NM_015033 |
| **SAP62 / SF3A2** | ENSG00000145833 | NM_015484 |
| **SF1** | ENSG00000146457 | NM_015542 |
| **SF3A1** | ENSG00000147140 | NM_015629 |
| **SF3A3** | ENSG00000147274 | NM_015891 |
| **SF3B1** | ENSG00000148053 | NM_015908 |
| **SF3b10** | ENSG00000149136 | NM_015953 |
| **SF3b14b** | ENSG00000151461 | NM_016047 |
| **SFRS12** | ENSG00000151657 | NM_016059 |
| **SFRS14** | ENSG00000151846 | NM_016090 |
| **SFRS2** | ENSG00000151923 | NM_016199 |
| **SFRS2IP** | ENSG00000152795 | NM_016200 |
| **SIP1** | ENSG00000153250 | NM_016333 |
| **SKIP** | ENSG00000153914 | NM_016424 |
| **SKIV2L** | ENSG00000154473 | NM_016638 |
| **SLM1** | ENSG00000155363 | NM_016652 |
| **SLM2 /KHDRBS3** | ENSG00000155966 | NM_016732 |
| **SLU7** | ENSG00000156508 | NM_017411 |
| **SmB/B' / SNRPB** | ENSG00000156976 | NM_017437 |
| **SmD1 / SNRPD1** | ENSG00000157540 | NM_017452 |
| **SmD2 / SNRPD2** | ENSG00000157554 | NM_017495 |

| Gene Name | Ensembl ref | NCBI ref |
|---|---|---|
| **SmD3 / SNRPD3** | ENSG00000159082 | NM_017612 |
| **SmE1 / SNRPE** | ENSG00000159086 | NM_017871 |
| **SMN1** | ENSG00000160201 | NM_018025 |
| **SMN2** | ENSG00000160633 | NM_018032 |
| **SNRPC** | ENSG00000160917 | NM_018047 |
| **SNRPF** | ENSG00000161265 | NM_018785 |
| **SNRPN** | ENSG00000161547 | NM_018834 |
| **SPF30** | ENSG00000161847 | NM_020180 |
| **SPF31 / DNAJC8** | ENSG00000161960 | NM_020666 |
| **SPF45** | ENSG00000162374 | NM_020701 |
| **SPOP** | ENSG00000162385 | NM_020963 |
| **SR140** | ENSG00000162613 | NM_021177 |
| **SR89(LUC7L)** | ENSG00000163486 | NM_021190 |
| **SRC** | ENSG00000163714 | NM_021952 |
| **SRm160 (SRRM1)** | ENSG00000164548 | NM_022173 |
| **SRm300 (SRRM2)** | ENSG00000164609 | NM_022719 |
| **SRp20 / SFRS3** | ENSG00000164985 | NM_022807 |
| **SRp30c (SFRS9)** | ENSG00000165119 | NM_022844 |
| **SRp40 / SFRS5** | ENSG00000165263 | NM_022875 |
| **SrP46** | ENSG00000165271 | NM_022917 |
| **SRp54(SFRS11)** | ENSG00000165630 | NM_024005 |
| **SRp55 / SFRS6** | ENSG00000165934 | NM_024038 |
| **SRp75 / SFRS4** | ENSG00000166291 | NM_024426 |
| **SRP9** | ENSG00000167005 | NM_024660 |
| **SRPK1** | ENSG00000167088 | NM_030653 |
| **SRPK2** | ENSG00000167971 | NM_030776 |
| **SSRP1** | ENSG00000167978 | NM_030979 |
| **STAU** | ENSG00000168000 | NM_031157 |
| **SUPT16H** | ENSG00000168066 | NM_031243 |
| **SWAP / SFRS8** | ENSG00000168233 | NM_031263 |
| **SWAP2** | ENSG00000168438 | NM_031287 |
| **SYNJ1** | ENSG00000168501 | NM_031314 |
| **SYNJ2** | ENSG00000168883 | NM_031370 |
| **TAF15** | ENSG00000169045 | NM_031372 |

| Gene Name | Ensembl ref | NCBI ref |
|---|---|---|
| **Tat-SF1** | ENSG00000169564 | NM_031989 |
| **TCEAL1** | ENSG00000169813 | NM_032025 |
| **TFE3(CA150)** | ENSG00000169976 | NM_032102 |
| **THOC1(HPR1)** | ENSG00000170860 | NM_032361 |
| **THOC2** | ENSG00000171566 | NM_032472 |
| **THOC3** | ENSG00000171960 | NM_032725 |
| **THOC3 / PSEN1** | ENSG00000172062 | NM_032758 |
| **THOC4** | ENSG00000172465 | NM_032905 |
| **TIA1** | ENSG00000172660 | NM_032940 |
| **TIA-R (TIAL)** | ENSG00000172780 | NM_033222 |
| **TIP39 / TFIP11** | ENSG00000173539 | NM_078480 |
| **Topo RS** | ENSG00000174044 | NM_080594 |
| **tra2 alpha** | ENSG00000174231 | NM_080598 |
| **tra2 beta / SFRS10** | ENSG00000174243 | NM_080832 |
| **U1-70KD** | ENSG00000174740 | NM_133370 |
| **U1-A / SNRPA** | ENSG00000175324 | NM_133452 |
| **U2-A' / SNRPA1** | ENSG00000176444 | NM_139168 |
| **U2AF1L2** | ENSG00000176773 | NM_139215 |
| **U2AF1RS3** | ENSG00000177733 | NM_145328 |
| **U2AF2** | ENSG00000178105 | NM_152235 |
| **U2AF35 / U2AF1** | ENSG00000179335 | NM_152688 |
| **U2AF65** | ENSG00000179950 | NM_153201 |
| **U5-100KD** | ENSG00000180855 | NM_172195 |
| **U5-116KD** | ENSG00000182338 | NM_172231 |
| **U5-15KD** | ENSG00000182944 | NM_177983 |
| **U5-40kD (hPRP8BP)** | ENSG00000183431 | NM_182648 |
| **UAP56 (BAT1)** | ENSG00000183684 | NM_182692 |
| **UPF2** | ENSG00000187239 | NM_182918 |
| **USA-CYP** | | NM_184244 |
| **WT1** | | NM_198216 |
| **YT521** | | NM_198291 |
| **ZBP1** | | NM_201997 |
| **ZNF124** | | NM_203446 |
| **ZNF207** | | NM_206853.1 |

| Gene Name | Ensembl ref | NCBI ref |
|---|---|---|
| | | U69127 |
| | | XM_047325.8 |

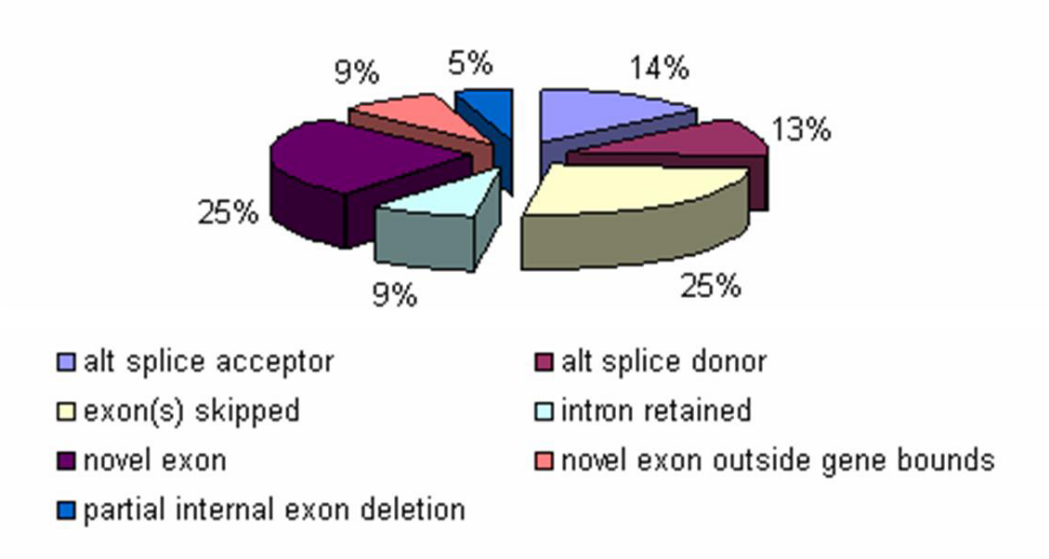



**Table 4.3 AD samples from Erlangen AD database**

| *Sample ID* | *Age* | *Gender* | *Diagnose* | *CERAD* | *Braak* | *PM* |
|---|---|---|---|---|---|---|
| **RZ92** | 72 | M | Control | 0 | 0-1 | 23 |
| **RZ99** | 56 | F | Control | 0 | 0 | 14 |
| **RZ145** | 86 | F | Control | 0 | 1 | 20 |
| **RZ85** | 63 | M | Control | 0 | 1 | 18 |
| **RZ104** | 60 | M | Control | 0 | 1-2 | 11 |
| | | | | | | |
| **RZ144** | 69 | F | AD | C | 6 | 18 |
| **RZ262** | 79 | F | AD | C | 6 | 24 |
| **RZ296** | 88 | M | AD | C | 5 | 4 |
| **RZ55** | 80 | M | AD | C | 5 | 12 |
| **RZ244** | 77 | F | AD | C | 5-6 | 24 |

4.1.2.2 Analysis of alternative splicing from Aizheimer's patients using splicing arrays

To analysis possible splicing changes in AD, we first identified up several samples from Erlangen Alzheimer's brain collections (supported by Dr. Kerstin Henkel). The custom filtered different age and gender as shown in Table 4.3 were selected.

For each sample, RT-PCR was performed in temporal Cortex (TC), frontal cortex (FC) and cereblllium (CE) from AD and control patients. We tested clk2 exon4, Tau exon10 and tra2 exon3. The result are shown in Fig 4.7. In temporal Cortex, we found a increase of Tau exon10 and Tra2 exon3. Clk2 exon4's usage is decreased (Fig.4.7, TC columns). This was confirmed by RT-PCR from another AD database (4.1.2.3). However the CE samples show difference as well (Fig.4.7, CE columns). Individual RT-PCR are shown in Figure 4.8.

Mixed temporal cortex of RZ244, RZ262 and RZ296 from AD samples with RZ145 control were chosen as probes for chip hybridization. Exons with change in regulation from chip result are listed in Table4.4. The chip hybridization shows around 20% splicing factors or 15% alternative splicing events are changed in AD samples.

Exons with red background were chosen for validation by RT-PCR. Junction oligos gave more details about alternative splicing from chips when was compared to exon based probes (Fig 4.9). However, often the signal is conflicting and sometimes oligos show opposite hybridization patterns. As shown in Table 4.5, if considering only the oligo B in alternative spliced exon, 57% of chip result were confirmed by RT-PCR when considering junction oligo C, D and E together, 67% RT-PCR were confirmed.

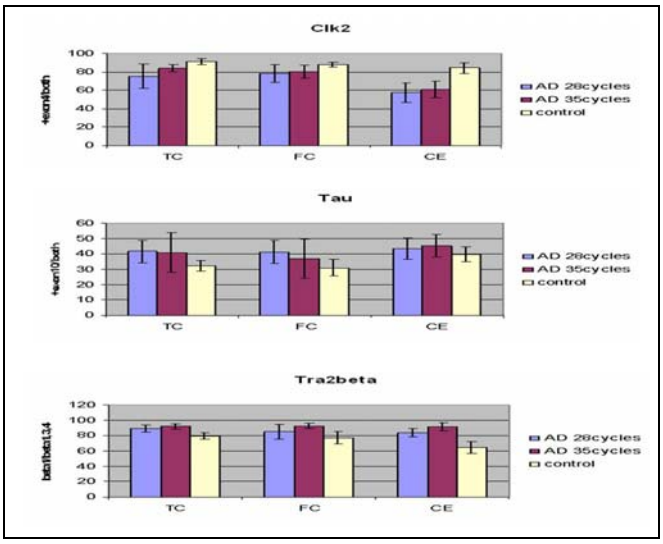



**Figure 4.7 summary of RTPCR based Erlangen AD samples.** Results from RT-PCR analysis performed in TC: temporal Cortex, FC: frontal cortex and CE: cereblllium from AD and control patients are shown on top of each panel. From each experiment, the ratio between exon 4 inclusion and exon 4 skipping from Clk2, exon 10 inclusion and exon 10 skipping from Tau , exon 3 inclusion and exon 3 skipping from Tra2beta was calculated. The mean of this ratio for control tissue was arbitrarily set to 100%. The mean and standard deviation of each set of experiments are included.

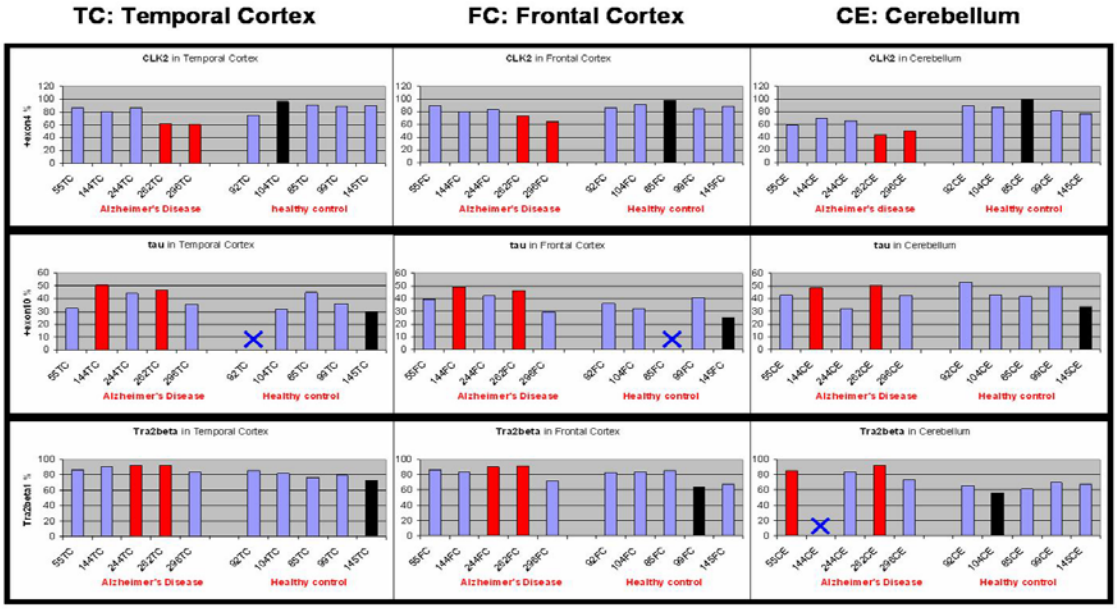



**Figure 4.8 RT-PCR of Tau Clk2 and Tra2 in AD brain tissues.** For each patient, 3 part region of brain are analysised: TC: Temporal Cortex; FC: Frontal Cortex; CE: Cerebllium. The ratio between two isoform are listed in the table : Clk2, +exon4/both. Tau, +exon10/both. Tra, +exon3/all. Significant AD samples are marked in red column while black column for control. Faint PCR signal was detected which is indicated by "X". Based on individual samples result, 3 from AD tissue and controls are selected to hybridize on custom splicing factors chip.

**Table 4.4 Result of custom splicing factor junction array**

| exon ID | status | gene name | splicing events | reference |
|---|---|---|---|---|
| 125.2.1 | **UP** | NSAP1/SYNCRIP | alt splice donor | NM_006372 |
| 189.4.1 | **pUP** | LSM7 | intron retained | NM_016199 |
| 227.3.2 | **pUP** | hnRNP A2/B1 | exon(s) skipped | NM_031243 |
| 015.111.1 | **UP** | CDK9 | exon(s) skipped | NM_001261 |
| 174.2.1 | **pUP** | LSM1 | exon(s) skipped | NM_014462 |
| 015.39.1 | **pUP** | CDK9 | novel exon | NM_001261 |
| 191.1.1 | **UP** | SRm300 (SRRM2) | novel exon | NM_016333 |
| 213.9.1 | **UP** | HUD / ELAVL4 | novel exon | NM_021952 |
| 232.6.3 | **UP** | PCBP2 | novel exon | NM_031989 |
| 163.1.1 | **pUP** | LSM4 | novel exon | NM_012321 |
| 102.7.1 | **UP** | SRC | novel exon | NM_005417 |
| 175.1.1 | **pUP** | Tat-SF1 | novel exon | NM_014500 |
| 175.2.2 | **pUP** | Tat-SF1 | alt splice acceptor | NM_014500 |
| 224.6.1 | **pUP** | DDX11 | exon(s) skipped | NM_030653 |
| 224.9.2 | **pUP** | DDX11 | exon(s) skipped | NM_030653 |
| 243.8.1 | **UP** | UAP56 (BAT1) | novel exon | NM_080598 |

| exon ID | status | gene name | splicing events | reference |
|---|---|---|---|---|
| 257.1.1 | **pUP** | ERG | novel exon | NM_182918 |
| 259.3.1 | **pUP** | SmB/B' / SNRPB | novel exon | NM_198216 |
| 263.1.3 | **UP** | QKI | novel exon | NM_206853.1 |
| 903.002.001 | **pUP** | BACE1 | alt splice acceptor | NM_012104 |
| 907.003.001 | **pUP** | FYN | novel exon | NM_002037 |
| 908.076.001 | **pUP** | HMGA1 | partial internal exon deletion | NM_002131 |
| 910.036.001 | **pUP** | ASR2 | alt splice donor | NM_182800 |
| **exon ID** | **status** | **gene name** | **splicing events** | **reference** |
| 27.1.1 | **DOWN** | CBP80 / NCBP1 | exon(s) skipped | NM_002486 |
| 50.3.1 | **DOWN** | ZNF207 | exon(s) skipped | NM_003457 |
| 193.1.3 | **DOWN** | ARL6IP4 | alt splice donor | NM_016638 |
| 158.10.2 | **DOWN** | hnRNP H3 / HNRPH3 | exon(s) skipped | NM_012207 |
| 179.3.3 | **DOWN** | SFRS14 | partial internal exon deletion | NM_014884 |
| 206.2.1 | **pDown** | MAT3 | exon(s) skipped | NM_018834 |
| 216.2.3 | **DOWN** | SNRPN | novel exon | NM_022807 |
| 227.4.1 | **DOWN** | hnRNP A2/B1 | novel exon | NM_031243 |
| 263.1.1 | **DOWN** | QKI | alt splice acceptor | NM_206853.1 |
| 09.11.1 | **DOWN** | EWS | exon(s) skipped | NM_005243 |
| 210.2.1 | **DOWN** | MOV10 | intron retained | NM_020963 |
| 016.357.1 | **DOWN** | DDX1 | exon(s) skipped | NM_004939 |
| 17.1.1 | **DOWN** | Ku70 / G22P1 | exon(s) skipped | NM_001469 |
| 45.2.1 | **DOWN** | SRP9 | exon(s) skipped | NM_003133 |
| 61.1.1 | **pDown** | NOL4 | exon(s) skipped | NM_003787 |
| 66.3.2 | **DOWN** | G10 | novel exon | NM_003910 |
| 113.1.1 | **DOWN** | SF3A1 | alt splice donor | NM_005877 |
| 117.8.1 | **pDown** | CypE / PPIE | exon(s) skipped | NM_006112 |
| 130.1.1 | **DOWN** | SLM2 /KHDRBS3 | exon(s) skipped | NM_006558 |
| 164.1.1 | **DOWN** | SAP130 (SF3B3) | novel exon | NM_012426 |
| 179.1.1 | **pDown** | SFRS14 | exon(s) skipped | NM_014884 |
| 184.1.1 | **pDown** | PRP31 | intron retained | NM_015629 |
| 209.5.1 | **DOWN** | ISY1 | novel exon | NM_020701 |
| 210.1.3 | **pDown** | MOV10 | alt splice acceptor | NM_020963 |
| 220.3.1 | **DOWN** | DDX3 | alt splice acceptor | NM_024005 |
| 234.2.1 | **DOWN** | SrP46 | partial internal exon deletion | NM_032102 |
| 235.1.1 | **DOWN** | THOC3 | exon(s) skipped | NM_032361 |
| 245.6.2 | **pDown** | YT521 | exon(s) skipped | NM_133370 |
| 257.3.1 | **pDown** | ERG | novel exon | NM_182918 |
| 259.5.1 | **DOWN** | SmB/B' / SNRPB | novel exon | NM_198216 |
| 259.5.1 | **pDown** | SmB/B' / SNRPB | alt splice acceptor | NM_198216 |

| exon ID | status | gene name | splicing events | reference |
|---|---|---|---|---|
| 901.002.002 | **pDown** | MAPT | exon skipped | NM_005910 |
| 901.009.001 | **pDown** | MAPT | exons skipped | NM_005910 |
| 905.009.001 | **pDown** | PSEN1 | alt splice donor | NM_000021 |
| 911.023.001 | **DOWN** | CLCN3 | exons skipped | NM_001829 |

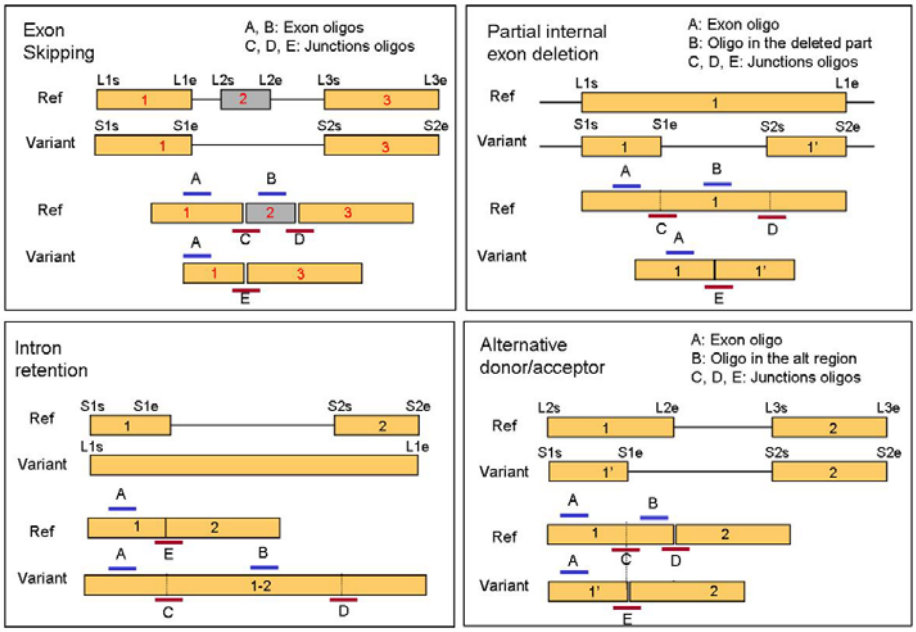



**Figure 4.9 junction oligo design in alternative splicing events.** Exon skipping, intron retention and different donor/acceptor site alternative splicing pattern are shown. Based on different splicing pattern, oligoB localized in alternative spliced exon and oligo C,D,E are junction sequence, average 4 different 20mer or 60mer oligos are designed in each type. The oligos sequence and variants are shown and analysised in BLAT customer track.

**Table 4.5 junction array result and RT-PCR analysis**

| | | | | | | | confirmed by RT-PCR | | |
|---|---|---|---|---|---|---|---|---|---|
| **Gene name** | **Event ID** | **oligoA** | **oligoB** | **oligoC** | **oligoD** | **oligoE** | **consider oligo B** | **consider oligo C ,D** | **consider oligo E** |
| **CBP80** | 27.1.1 | *-1.27* | ***-1.84*** | *-1.07* | *1.04* | *1.01* | | | |
| **ZNF207** | 50.3.1 | *-1.1* | ***-1.73*** | *-1.06* | *-1.59* | *-1.18* | Yes | Yes | No |
| **ARL6IP4** | 193.1.3 | *1.02* | ***-1.54*** | *-1.07* | *-1.12* | *1.03* | Yes | Yes | Yes |
| **hnRNP H3** | 158.10.2 | *-1.19* | ***-1.55*** | *-1.49* | | *-1.31* | No | No | Yes |
| **SFRS14** | 179.3.3 | *-1.13* | ***-1.5*** | *1.07* | *-1.11* | *1.1* | | | |
| **NSAP1** | 125.2.1 | *-1.32* | ***1.67*** | *-1* | *-1.85* | *1.05* | Yes | No | No |
| **LSM7** | 189.4.1 | *-1.1* | ***1.33*** | *1.19* | *1.12* | *-1.07* | Yes | Yes | Yes |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| **hnRNP A2/B1** | 227.3.2 | *-1.12* | ***1.26*** | *-1.39* | *-1.55* | *1.03* | No | Yes | Yes |
| **CDK9** | 015.111.1 | *1.08* | ***1.63*** | *-1.01* | *1.02* | *1.14* | | | |
| **LSM1** | 174.2.1 | *-1.11* | ***1.38*** | | | | No | | |
| | | | | | | | 4/7= **57%** | 4/6 = **67%** | 4/6= **67%** |

4.1.2.3 htra2-beta1 and clk2 feedback regulation and affect tau exon 10 usage

The TAU protein regulates the stability and assembly of microtubules by binding to microtubules through three or four microtubule-binding repeats (3R and 4R). The number of microtubule-binding repeats is determined by the inclusion or exclusion of the second microtubule-binding repeat encoded by exon 10 of the tau gene (Goedert, Ghetti et al. 2000).

We previously showed that tau exon 10 usage is increased by hTRA2-beta1 (Jiang et al. 2003) and decreased by CLK2 mediated phosphorylation (Hartmann et al. 2001). CLK2 binds and phosphorylates hTRA2-beta1 (Stoilov et al. 2004). Like almost all pre-mRNAs of splicing regulatory proteins, htra2-beta1 and clk2 pre-mRNAs undergo alternative splicing (Hanes et al. 1994; Nayler et al. 1998a). Skipping of exons 2 and 3 of htra2-beta pre-mRNA generates an inactive protein, hTRA2-beta3 (Stoilov et al. 2004), whereas skipping of exon 4 of clk2 pre-mRNA generates a frameshift resulting in the inactive variant, clk2tr (Duncan et al. 1997). clk2tr mRNA is subject to nonsense-mediated decay and, therefore, is not translated into any protein (Hillman et al. 2004). To test whether these proteins can influence each other's splice site selection at the pre-mRNA level, we employed a minigene construct consisting of the alternative exon flanked by the constitutive exons. As shown in Figure 1A, an increase of hTRA2-beta1 promotes skipping of clk2 exon 4, most likely through binding to a htra2-beta1 motif AAGAGCGA present in the 3' part of the clk2 exon 4. Similar to the situation in clk1 (Duncan, Stojdl et al. 1997), clk2 promotes skipping of its exon 4 generating a frameshift resulting in an inactive form (Figure 1A). We then performed similar assays with the expression constructs encoding the inactive proteins htra2-beta3 and Clk2-KR. Clk2-KR is an inactive CLK2 mutant because it cannot bind ATP. We tested this mutant to determine whether clk2tr protein might influence pre-mRNA splicing if the mRNA escapes nonsense-mediated decay under special conditions. We found that these inactive

forms have no effect on clk2 pre-mRNA splicing (Fig.4.1B). Interestingly, our previous research has revealed that CLK2 influences the splicing pattern of tra2-beta by promoting exon skipping and the formation of the inactive protein isoform TRA2-beta3(Stoilov, Daoud et al. 2004).This suggests that the amount of active TRA2-beta1 and CLK2 is controlled through a feedback of alternative splicing decisions.

Together, these results indicate that CLK2 promotes exon skipping in the tra2-beta, clk2 and tau genes resulting in the formation of the inactive htra2-beta3 isoform, the inactive clk2tr variant and the tau isoform lacking exon 10. The formation of CLK2 is influenced by hTRA2-beta1, which promotes the formation of the inactive clk2tr isoform, the mRNA of which is subjected to nonsense-mediated decay and is not transcribed into protein (Hillman, Green et al. 2004). Finally, TRA2-beta1 promotes the formation of the 4R tau isoform (Jiang, Tang et al. 2003; Wang, Wang et al. 2005). Given the intricate relationship between the three genes, we analyzed the distribution of the splicing variants in brain areas affected by AD.

#### *4.1.2.3.1 AD Samples selection for validation*

New brain samples are selected to perform validation. The brain samples were obtained from the Kathleen Price Bryan Brain Bank, Durham, USA (temporal and occipital cortex; AD group: Braak III to Braak V, n = 15; control group: no tau deposits, Braak I or II, n = 9) (Table 4.6), and from the Würzburg-Hirnbank, Würzburg, Germany (supplementary motocortex; eight control patients; seven AD patients, Braak III to VI) (Table 4.7). Samples from the Kathleen Price Bryan Brain Bank were matched in age ($t = 0421$; $df = 22$; $p = 0464$), sex ($\chi 2 = 0046$; $df = 1$; $p = 0831$) and postmortem interval ($t = -1081$; $df = 21$; $p = 0292$). The brains of the Würzburg Hirnbank showed differences concerning age ($t = -2342$; $df = 13$; $p = 0036$), but not concerning sex ($\chi 2 = 1727$; $df = 1$, $p = 0189$) and postmortem interval ($t = 0395$; $df = 13$; $p = 0699$). All tissues were obtained in accordance with the local ethics committee procedures.

**Table 4.6: Patient characteristics of the temporal and occipital cortex including case identification number, age, diagnostic category, gender, *post mortem* interval and neuropathological diagnostic category.** sample-number internal identification number, f

female, m male, PM *post mortem* interval (h), AD Alzheimer's disease, CON control subjects, ND no data, AR: age related

| Sample-number | Diagnostic category | Age | Gender | PM | *Neuropathological* diagnosis |
|---|---|---|---|---|---|
| 1 | AD | 91 | f | 6:00 | AD |
| 2 | AD | 90 | f | 2:20 | AD, Braak V |
| 3 | AD | 81 | f | 3:00 | AD |
| 4 | AD | 64 | f | 3:36 | AD, Braak V |
| 5 | AD | 100 | f | 5:50 | AD, Braak IV |
| 6 | AD | 77 | f | 4:00 | AD, Braak III |
| 7 | AD | 83 | m | 1:15 | AD, Braak IV |
| 8 | AD | 62 | m | 1:30 | AD, Braak V |
| 9 | AD | 77 | f | 4:05 | AD, Braak V |
| 10 | AD | 74 | m | 1:10 | AD, Braak IV |
| 11 | AD | 79 | f | 3:45 | AD, Braak V |
| 12 | AD | 72 | m | 3:30 | AD, Braak IV |
| 13 | AD | 85 | f | 2:45 | AD, Braak V |
| 14 | AD | 83 | m | 2:00 | AD, Braak V |
| 15 | AD | 79 | m | 2:00 | AD, Braak IV |
| 17 | CON | 85 | m | 2:00 | AR |
| 18 | CON | 73 | f | 0:30 | AR |
| 21 | CON | 81 | m | 10:30 | AR, Braak I - II |
| 22 | CON | 82 | m | 3:15 | AR |
| 23 | CON | 72 | f | 3:00 | AR |
| 24 | CON | 92 | f | 2:35 | AR |
| 25 | CON | 80 | f | 1:10 | AR |
| 27 | CON | 78 | f | 2:22 | AR |
| 29 | CON | 91 | m | 7:40 | AR, Braak I |

**Table 4.7: Patient characteristics of the supplementary motocortex including case identification number, age, diagnostic category, gender, *post mortem* interval and neuropathological diagnostic category.** Sample-number internal identification number, f female, m male, PM *post mortem* interval (h), AD Alzheimer's disease, CON control subjects, ND no data, AR: age related

| No. | Diagnostic category | Age | Gender | PM | *Neuropathological* diagnosis |
|---|---|---|---|---|---|
| 1 | CON | 68 | f | 13:00 | AR |
| 2 | CON | 76 | m | 18:05 | AR |
| 3 | CON | 64 | m | 13:55 | AR |
| 4 | CON | 84 | m | 16:30 | AR |
| 5 | CON | 63 | m | 10:30 | AR |
| 6 | CON | 88 | m | 15:00 | AR |
| 7 | CON | 71 | f | 11:00 | AR |
| 8 | CON | 74 | f | < 12:00 | AR |
| 9 | AD | 84 | f | < 20 | AD, Braak III - VI |
| 10 | AD | 83 | f | 19:30 | AD |
| 11 | AD | 89 | m | 02:00 | AD, Braak IV - V |
| 12 | AD | 90 | f | 08:00 | AD |
| 13 | AD | 81 | m | 13:40 | AD, Braak VI |
| 14 | AD | 76 | f | 09:00 | AD, Braak V |
| 15 | AD | 76 | f | 17:00 | AD, Braak IV - V |

#### *4.1.2.3.2 Increased usage of tau exon 10 in brain areas affected by AD*

First, we investigated whether the alternative splicing pattern of tau exon 10 is altered in human brain tissues from AD patients by semi-quantitative RT-PCR. The semi-quantitative nature of the assay was determined by RT-PCR using a defined quantity of tau cDNA. As shown in Figure 4.10, the assay was linear between 28 and 40 cycles and we, therefore, subsequently amplified cDNA using 37 cycles. Using different cDNAs (+/- exon 10), we found that the ratio between exon inclusion and exon skipping reflected the amount of cDNA in the starting reaction. To analyze the ratio of tau exon inclusion and exon skipping in tissue, mRNA was reverse transcribed using oligo-(dT) primers and cDNA was amplified using gene-specific primers located in tau exon 9 and 11. Using this assay, the ratio of exon 10 containing and lacking isoforms was determined in various brain regions. We analyzed the temporal cortex as this region is known to develop abundant neurofibrillary tangles in early stages of AD. As control regions from affected individuals, occipital cortex and supplementary motocortex were employed, which constantly demonstrate less NFT formation and functional impairment in AD.

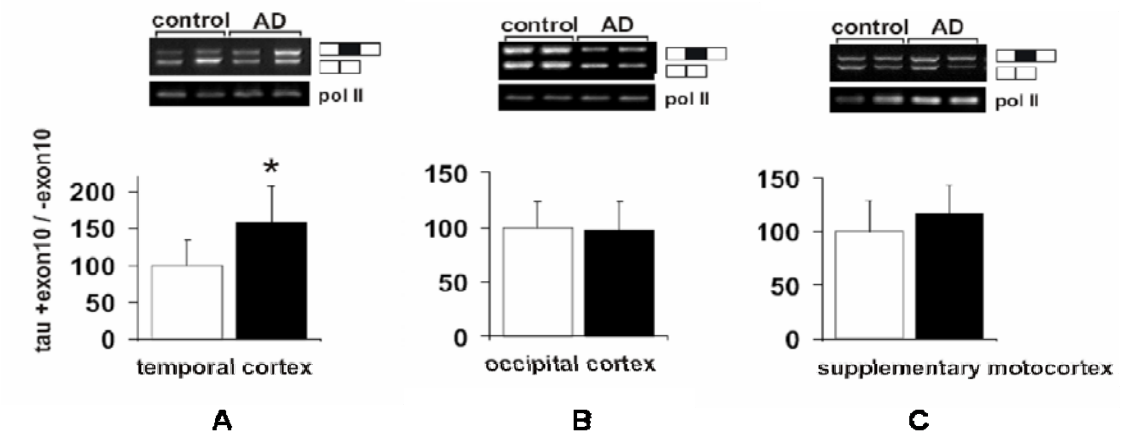



**Figure 4.10 Tau Exon 10 usage in brain regions affected with AD.** Results from RT-PCR analysis performed in temporal (A), occipital (B) and supplementary motocortex (C) from two representative AD and control patients are shown on top of each panel. The structure of the gene products is schematically indicated. Pol II: indicates mplification of polII as a loading control. From each experiment, the ratio between exon 10 inclusion and exon 10 skipping was calculated. The mean of this ratio for control tissue was arbitrarily set to 100%. The mean and standard deviation of each set of experiments are included.

These measurements were compared with tau isoform ratios in aged-matched individuals not suffering from AD. When compared with control patients, the mRNA from AD affected brains showed an increase of exon 10 in temporal cortex (t=-2,964; df=20; p=0,008; Figure 3). In contrast, there are no significant differences between those groups in control regions (occipital cortex: t=0,214; df=20; p=0,832; supplementary motocortex: t=-1,016; df=8; p=0,339). We conclude that there is a statistically significant increase of exon 10 containing tau mRNA in brain areas affected by AD.

Furthermore, to detect which factors influent Tau exon10 splicing, Tra2-beta1 /Tra2-beta1-NES, Tra2-beta3, Clk2, PP1, NIPP1, okadaic acid and vanadate were co-transfected with Tau minigene. As shown in Fig 4.11, Tra2-beta1 stimulates Tau exon10 inclusion while Tra2-beta1 with a NES tag (Nuclear Export Signal) enhances exon10 skipping (Fig.4.11, A). Tra2-beta3 has opposite effect with Tra2-beta1(Fig.4.11, C, left). Interestingly one of the Tra2-beta1 kinase CLK2 strongly enhances Tau exon10 skipping (Fig.4.11, C, right). High concentration of Okadaic acid shows different effect with vanadate to decrease exon10 inclusion (Fig.4.11,B), which is similar to the effect from low concentration of PP1 (protein phosphotase1, Fig.4.11, D, left) while high

concentration of PP1 inhibitor NIPP1 slightly enhances exon10 inclusion (Fig.4.11, D, right).

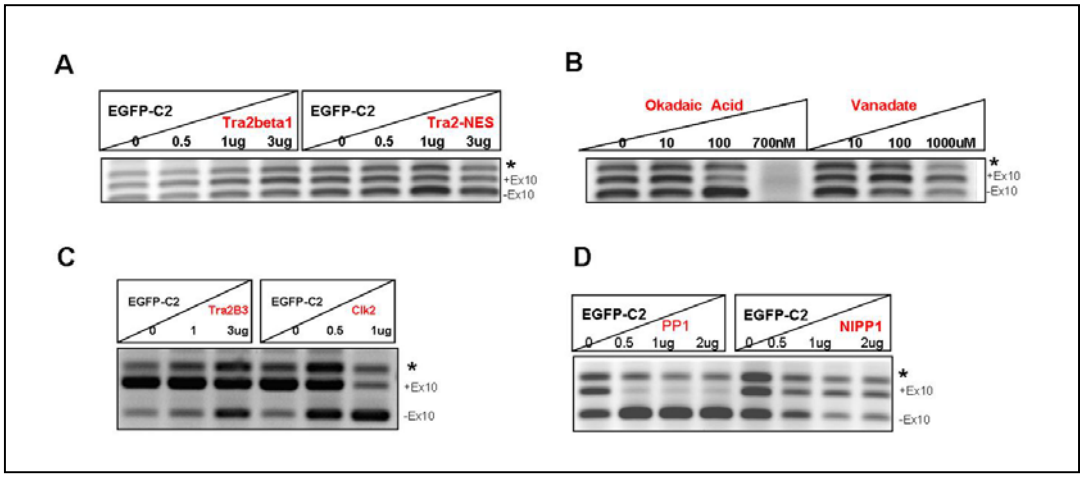



**Figure 4.11 Tau minigene co-transfected with different factors and chemical drug.** A. Tra21 and Tra2-NES are co-transfected with Tau minigene. Tra2-beta1 stimulate Tau Exon10 skipping. B. chemical drugs are co-transfected with Tau minigene. left: Okadaic Acid induces Tau exon10 skipping. right: vanadate slightly enhances exon10 inclusion variant in high concentration while the whole transcription are decreased. C. Tra2-beta3 and CLK2 are co-transfected with Tau minigene. Both Tra2-beta3 and CLK2 increase Tau exon10 skipping. D. PP1 and NIPP1 are co-transfected with Tau minigene. left: PP1 dramatically induces Tau exon10 skipping. right: PP1 inhibitor NIPP1 slightly enhances exon10 inclusion in high concentration while the whole transcription are decreased. The star indicates a hetrodimer.

### *4.1.2.3.3 The ratio between htra2-beta1 and htra2-beta3 is altered in brain areas affected by AD*

The inclusion of tau exon 10 is stimulated by hTRA2-beta1, a member of the SR-related protein family of splicing regulators (Jiang, Tang et al. 2003). Similar to most other alternative splicing regulators, the htra2-beta gene undergoes alternative splicing and forms two proteins: TRA2-beta1 that regulates exons containing the sequence GHVVGANR and TRA2-beta3, which is inactive.

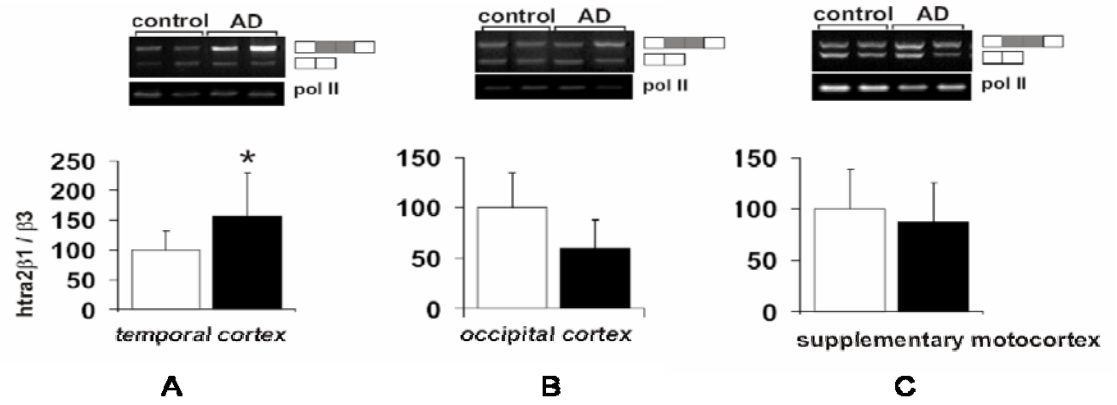



**Figure 4.12 Ratio of tra2-beta isoforms in brain regions from AD patients and controls.** Results from RT-PCR analysis performed in temporal (A), occipital (B) and supplementarymotocortex (C) from two representative AD and control patients are shown on top of eachpanel. The structure of the gene products is schematically indicated. Pol II: indicatesamplification of polII as a loading control. From each experiment, the ratio between exon 3 inclusion and skipping was calculated. The mean of this ratio for control tissue wasarbitrarily set to 100%. The mean and standard deviation of each set of experiments areincluded.

On the protein level, TRA2-beta1 and TRA2-beta3 differ in the presence of the first RS domain. On the RNA level both isoforms differ in the presence of exon 3. CLK2 promotes the formation of htra2-beta3 (Stoilov, Daoud et al. 2004). Since both tau and clk2 splicing are altered in brains of AD sufferers, we employed an RT-PCR assay to determine the ratio between tra2-beta1 and tra2-beta3 isoforms. The expression pattern of the isoforms hTra2-beta1 and htra2-ß3 was altered in the temporal cortex of AD patients when compared to controls (t=-2,060; df=15; p=0,057; Figure 4.12). Both control regions, comprising the occipital cortex (t=1,226; df=11; p=0,246) and supplementary motocortex (t=0,555; df=9; p=0,592), revealed no statistically significant differences between the AD and control group. We, therefore, conclude that the ratio between hTRA2-beta1 and hTRA2-beta3 is specifically altered in brain areas affected by AD.

#### *4.1.2.3.4 Expression of clk2 exon 4 is decreased in brain areas of AD patients*

Alternative splice site selection is frequently regulated by phosphorylation(Stamm 2002) and we have previously demonstrated that the SR-protein kinase CLK2 influences tau exon 10 splicing, where it strongly promotes exon 10 skipping (Hartmann, Rujescu et al. 2001). Similar to other clk kinases, CLK2 is regulated by alternative splicing. Skipping of exon 4 causes a frameshift resulting in a mRNA which becomes subjected to

nonsense-mediated decay and, therefore, is not translated into any protein (Hillman, Green et al. 2004).

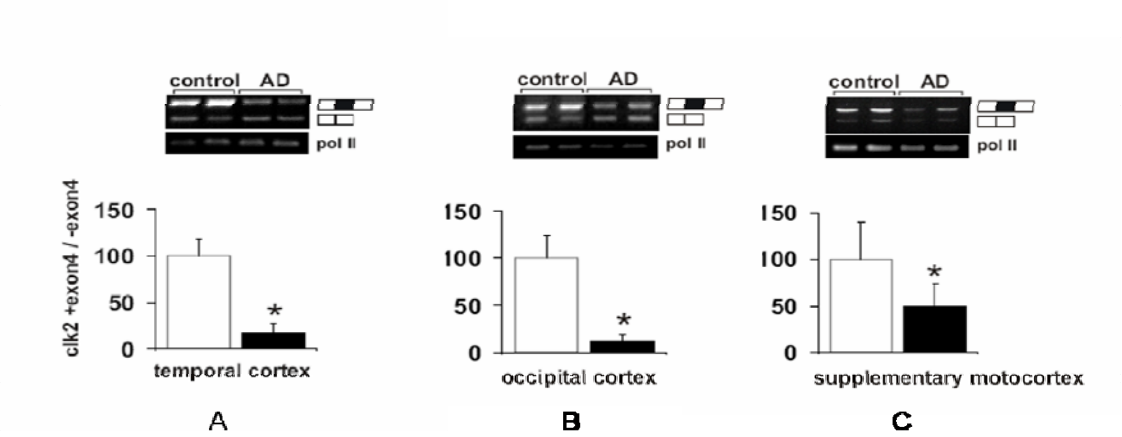



**Figure 4.13 Ratio of clk2 isoforms in brain regions from AD patients and controls.** Results from RT-PCR analysis performed in temporal (A), occipital (B) and supplementary motocortex (C) from two representative AD and control patients are shown on top of each panel. The structure of the gene products is schematically indicated. Pol II: indicates amplification of polII as a loading control. From each experiment, the ratio between exon 4 inclusion and exon 4 skipping was calculated. The mean of this ratio for control tissue was arbitrarily set to 100%. The mean and standard deviation of each set of experiments are included.

In order to determine whether clk2 is altered in brain regions affected by AD we examined clk2 exon 4 inclusion by RT-PCR (Figure 4.13). We found a strong decrease of exon 4 in brain areas affected by AD (temporal cortex: t=3,725; df=19; p=0,001). Unexpectedly, in the control regions (occipital cortex: t=12,549; df=17; p<0,001 and supplementary motocortex: t=2,545; df= 9; p=0,031) there was also a marked decrease of exon 4 when the AD controls were compared with age-matched controls. We conclude that there is a decrease of the mRNA encoding active CLK2 protein in the brain of individuals affected by AD, which extends beyond the areas specifically altered in AD.

Furthermore, more endogenous clk2 exon4 skipping variants are increased in high consentration of APP transcrected 293 cell lines (Fig. 4.14).

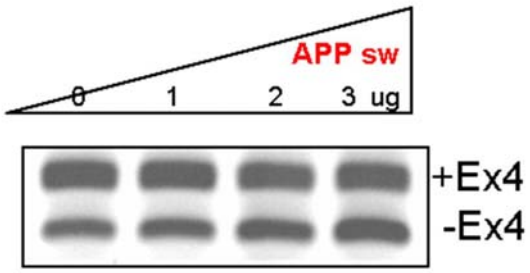



**Figure 4.14 endogenous Clk2 variants in APPsw transfected 293 cell line.** Plasmid encode APPsw was cotransfected with clk2 minigene to HEK293 cells, 48 hours after transfection, RNA were analysised by RT-PCR. APP sw: APP Swedish

### *4.1.2.3.5 Increased skipping of presenilin 2 exon 5 in Alzheimer's disease*

Recently, an alternatively spliced form of the presenilin 2 gene lacking exon 5 was found in human brains with sporadic Alzheimer's disease. It was shown that skipping of this exon was induced by the Imaizumi's group(Higashide, Morikawa et al. 2004). We thus determined PS2 exon 5 usage in our samples. PS2 exon 5 showed a stronger variation than other exons, since three classes of samples could be discriminated: one class of samples contained both exon 5 inclusion and skipping variants, one class contained only exon 5 inclusion and one class only exon 5 skipping (Figure 4.15 A-C). As can be seen in Figure 4.15 D, samples completely lacking exon 5 usage are found only in individuals suffering from Alzheimer's disease, being more frequent in temporal cortex than in occipital cortex. The relative frequency of the presenilin isoform skipping exon 5 was significantly higher in AD patients compared to controls in the temporal cortex (chi2=7.44; df=2; p=0.024; Figure 4.15). Both control regions, comprising the occipital cortex (chi2=2.37; df=2; p=0.306) and supplementary motocortex (chi2=0.00; df=2; p=1.0), revealed no statistically significant differences between the AD and control group. Thus, the usage of presenilin exon 5 seems to be specifically altered in brain areas affected by AD.

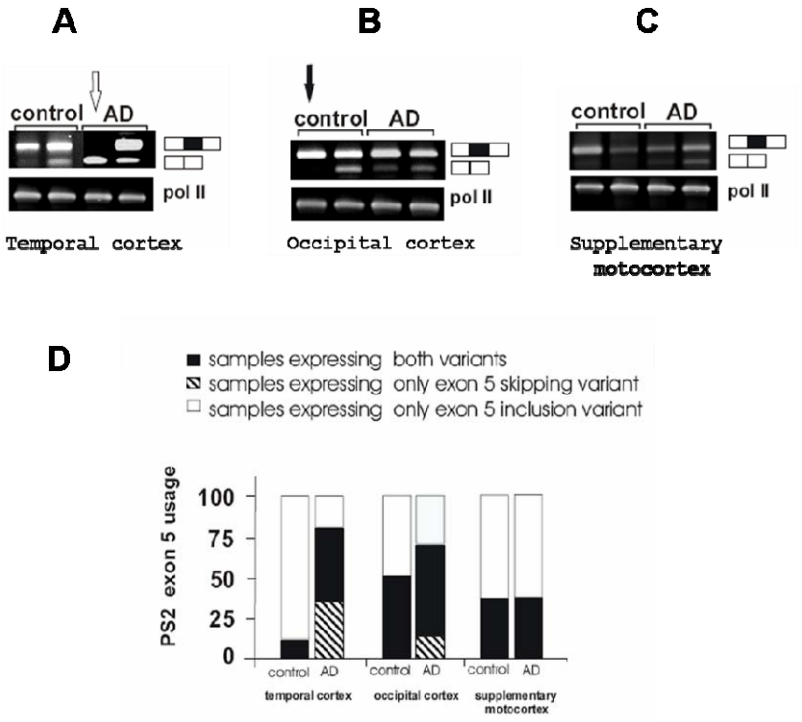



**Figure 4.15 Presenilin 2 isoforms in brain regions from AD patients and controls.** Results from RT-PCR analysis performed in temporal (A), occipital (B) and supplementary motocortex (C) from two representative AD and control patients are shown. The structure of the gene products is schematically indicated. Pol II: indicates amplification of polII as a loading control. Since a substantial number of samples completely lacked presenilin exon 5 inclusion (open arrow) or skipping (black arrow), no ratios of isoforms could be determined. We therefore classified samples as expressing both variants or only the variant either including or excluding exon 5. This evaluation is shown in the panel D.

## 4.1.3 Change of alternative splicing in Cancer

### 4.1.3.1 Tra2 protein and RNA level's stimulation in Breast Cancer

Cancer is presently responsible for about 25% of all deaths in the world (Jemal, Murray et al. 2005). It was recently shown that alternative splicing changed in cancer (Julian P. Venables 2006). Some research pointed out that the mutations in splicing regulatory elements resulted in changes in the splicing pattern of many cancer-related genes (Pajares, Ezponda et al. 2007).

To understand how alternative splicing contributes to tumorigenesis, several splicing factor including tra2-beta1 were analysised in cancer tissues. Expression analysis was done in histologically confirmed breast cancers and corresponding pairs of nonpathologic breast tissue. CD44 alternative splicing was determined by RT-PCR using

primers directed against the constitutive exons 5 and 16 of the CD44 gene (Fig.4.16 A). Tra2-beta1 expression in breast cancer development was monitored by Tra2-ß-specific RT-PCR (Fig.4.16 B and C) and Western blot (Fig.4.16D).

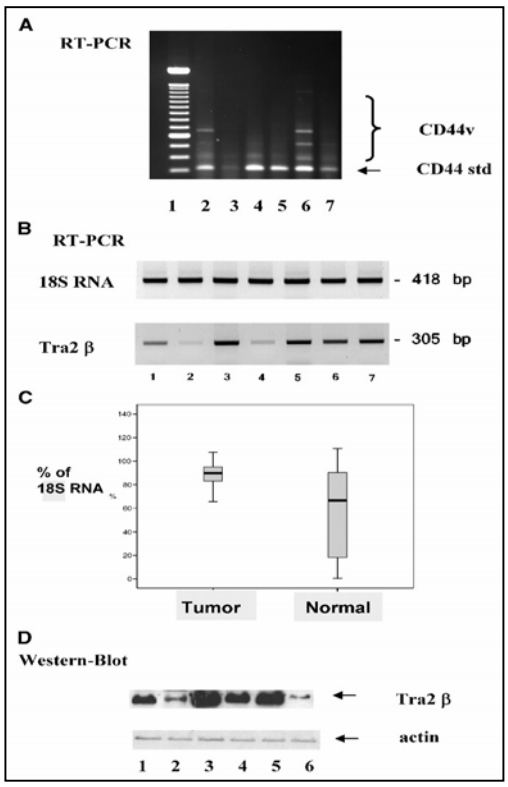



**Figure 4.16. Specific induction of alternative CD44 splicing is accompanied by induction of Tra2-beta1 in breast cancer.** CD44 splicing and expression of Tra2-beta1 in breast cancer and corresponding pairs of normal tissue was determined by low-cycle RT-PCR and Western blot. *A,* RT-PCR analysis of CD44 with equal amounts of total cell RNA (5µg). *Arrow,* constitutive CD44 splicing (*CD44std*); *dashed line,* alternative CD44 splicing. *Lane 1,* marker; *lane 2,* tumor 1; *lane 3,* normal tissue 1; *lane 4,* tumor 2; *lane 5,* normal tissue 2; *lane 6,* tumor 3; *lane 7,* normal tissue 3. *B,* Tra2-beta1 RT-PCR analysis with equal amounts of total cell RNA (5 µg). *Lane 1,* tumor 1; *lane 2,* normal tissue 1; *lane 3,* tumor 2; *lane 4,* normal tissue 2; *lane 5,* tumor 3; *lane 6,* normal tissue 3; *lane 7,* positive control (HeLa). *C,* statistical analysis of Tra2-beta1 RNA normalized to 18S RNA in 36 breast cancers (tumor) and corresponding normal breast tissue (normal). *Thick lines,* median (50% percentile); *gray boxes,* 25% to 75% percentile; *thin lines,* minimal and maximal value (without rouge results). *D,* Tra2-ß protein expression analysis by Western blot with human Tra2-ß-specific and actin-specific antibody as control (equal amounts of protein). *Lane 1,* tumor 1; *lane 2,* normal tissue 1; *lane 3,* tumor 2; *lane 4,* normal tissue 2; *lane 5,* tumor 3; *lane 6,* normal tissue 3. (Watermann, Tang et al. 2006)

The data show that breast cancer specimen indicated an induction of alternative CD44 splicing compared with nonpathologic tissue (Fig.4.16A). This induction was accompanied by an increase in Tra2-beta1 expression. We analyzed the amount of Tra2-beta1 RNA in 36 breast cancer samples compared with normal breast tissue of the same patient. All values were normalized to 18S RNA. RT-PCR revealed an over expression of Tra2-beta1 RNA in breast cancer samples. The mean of Tra2-beta1 RNA in breast cancer was 85.7 ± 23.4% of 18S RNA (mean ± SD) in contrast to 57.7 ± 35.6% in normal breast

tissue (Fig.4.16C). This difference was statistically highly significant ($P < 0.0001$). RT-PCR analysis of SRp40 as a control revealed constant levels in normal and cancerous tissues (data not shown). Western blot results confirmed the induction of Tra2-beta1 on the protein level (Fig.4.16D). (Watermann, Tang et al. 2006)

### 4.1.3.2 Analysis CD44 alternative splicing in Cancer tissue

To detect for cancer specific alternative splicing in the CD44 pre-mRNA, RT-PCR based on V4-V10 was performed. Since the analysis of CD44 alternative splcing generated too many bands, including several poorly detected variants, we designed specific primers: we made forward primers from v4 and v5 with reverse primer in constitutive exon G or H. As shown in Fig.4.17, v4-v10 are stimulated in lung cancer, mainly v4 is increased in liver cancer and v4 with v5 are stimulated in breast cancer. It clearly shows that distinguished CD44 alternative splicing pattern in different cancer types (Fig.4.17).

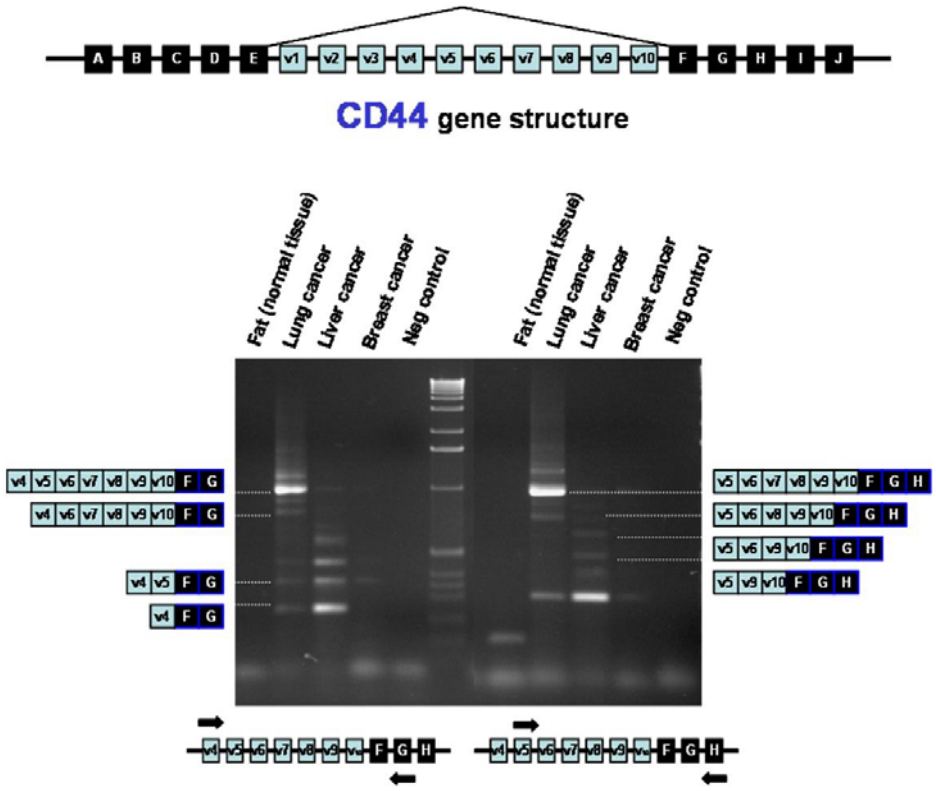



**Figure 4.17 CD44 alternative splicing in different cancer tissues.** The CD44 gene structure are indicated above, 10 alternative spliced exon v1-v10 are marked in light box while constitutive exon in black box. Variants are indicated beside RT-PCR bands. Primers designed for RT-PCR are shown arrow.

### 4.1.4 AEdb in Alternative Splicing database (ASD)

Humans have fewer genes than previously anticipated. Since humans contain not many more genes than for example fruit flies, they have to use other mechanisms to create their large transcriptome and proteome. Alternative splicing in the processing of pre-mRNA is such a mechanism. ASD aims to analyse this mechanism on a genome-wide scale by creating a database that contains all alternatively spliced exons from human, and other model species (Thanaraj, Stamm et al. 2004; Stamm, Riethoven et al. 2006).

AEdb forms the manually curated component of ASD database. It is a literature based data set containing sequence and properties of alternatively spliced exons, functional enumeration of observed splice events, characterization of observed splice regulatory elements, and a collection of experimentally clarified minigene constructs.

**AEdb-Sequence.** AEdb-Sequence is a literature based, manually curated database of alternative exons. We used "alternative splicing" as a keyword to search PubMed bibliography data and collected information on following features from the resultant research articles: organism, splicing mechanism, tissue-specificity, regulation during development stages, disease association, regulatory features of the exon, and the sequence of the alternatively spliced exon as well as its flanking constitutive exons. It is seen that more than half the number of AEdb-Sequence entries are from human (Table 4.8A). As is in the case of AltSplice data, cassette exon events outnumber other event types. A particularly interesting point is that the AEdb collection reports polyA-mediated alternative 3' exon events. The data set reports splice events that are specific to each of the three types of cell states, such as tissue type, development stage, and disease state. Roughly 10% of the entries report events that introduce premature stop codons and this data set can serve the studies on nonsense mediated decay of transcripts. Finally, 10% of the reported exons are from non-coding regions of the genes.

**AEdb-Function.** The function database is a literature based, manually curated database of known functions of the alternative exons. Functional differences between the protein isoforms generated by alternative splicing are enumerated from the literature and are organized into 11 well-defined categories, such as "Modulation of protein interaction"

or “Internal structural change” (Table4.8B). An analysis of the function of alternative exons based on this data set has previously been published (Stamm, Ben-Ari et al. 2005).

**AEdb-Motif.** Alternative splice site selection is partially regulated by weak binding of proteins to highly degenerate regulatory sequences. As a first attempt to understand the combinatorial control behind this regulation, we collected splice regulatory motifs described in literature and expanded upon the previous collections of intronic regulatory sequences (Ladd and Cooper 2002), exonic regulatory sequences (Bourgeois, Lejeune et al. 2004; Zheng 2004), and disease-causing mutations (Stoilov, Meshorer et al. 2002). The collection reports 153 enhancer sequences and 81 silencer sequences (Table4.8C). The entries are annotated with value-added information, such as the experimental technique used, the nucleotide sequence of the motif, mutations that are studied and the protein that binds at the motif.

**AEdb-Minigenes.** A minigene is a genomic fragment that includes the alternative exon and the surrounding introns as well as the flanking constitutively spliced exons. Constructs derived by cloning the insert in an eukaryotic expression vector are increasingly used to study alternative splicing (Stoss, Stoilov et al. 1999; Tang, Novoyatleva et al. 2005). We compiled all minigenes described in the literature. The splicing patterns and deduced regulatory sequences are represented in a graphic format. The minigene collection includes 82 entries for which a total of 97 regulatory sequences are ascribed. The reported minigene constructs representing cassette exon events outnumber those for other event types (Table4.8D). The minigene entries are linked to appropriate entries in AEdb-Sequence data set, which allows the user to quickly identify experimentally useful minigenes by searching the database.

**Table 4.8. Statistics on AEdb data.** AEdb sequence, function, motif and minigene databases are summarized and analysised. Based on distribution and function, entry numbers with species name are listed.

| **A: AEdb-Sequence data statistics** | |
|---|---|
| **Distribution** | **Number of entries** |
| Organism distribution | Human (1283); mouse (413); rat (232); |

| | |
|---|---|
| | drosophila (100); others (227). |
| Event type distribution | Cassette Exon (1281).<br>Alternative Acceptor or Donor (395).<br>Intron Retention (154).<br>Mutually exclusive exons (130).<br>Alternative 3' exon by polyA variant (71). |
| Regulation associated with disease | 295 |
| Regulation associated with development | 282 |
| Regulation associated with tissue type | 312 |
| Regulation causing frameshift | 151 |
| Regulation introducing stop codons | 260. |
| Alternative exon being noncoding exon | 222. |
| Entries associated with AltSplice | 1198 (human and mouse entries) |
| | |
| **B: AEdb-Function data statistics – 354 entries.** | |
| **Functional Role** | **Number of entries** |
| Modulation of protein interaction | 136 |
| Internal structural change | 119 |
| Novel carboxyl terminus | 87 |
| Novel amino terminus | 38 |
| Association with disease | 81 |
| Intracellular location | 76 |
| Enzymatic activity | 64 |
| Channel activity | 54 |
| Others | 37 |
| | |
| **C: AEdb-Motif data statistics – 255 entries** | |
| **Type of regulator sequence** | **Number of entries** |
| Exon Enhancer | 97 |
| Exon silencer | 44 |

| | |
|---|---|
| Intron Enhancer | 56 |
| Intron silencer | 37 |
| | |
| **D :AEdb-Minigene data statistics – 82 entries** | |
| **Distribution** | **Number of entries** |
| Organism distribution | Human (46); mouse (17); rat (15); others (9) |
| Splicing mechanism distribution | Cassette exon (single exon, 45; multiple casette exons (3); incremental combinatorial exons (2)); Alternative acceptor or donor sites (17); Mutually exclusive exons (13); Intron retention (2); |
| Reported tissue specificity | 55 |
| Known regulatory factors | 32 |
| Deduced Enhancer and silencer sequences | 97 |
| Hyperlinks to AEdb-Sequence database | 78 (to 105 AEdb-Sequence entries). |

***Interface for AEdb-Sequence.*** The data can be queried by gene names and synonyms, database cross-references, type of splice events, and type of regulatory roles (such as introducing premature termination codons or frameshift). Further, the data can be queried for disease association and developmental specificity (Fig. 4.18)

***Interface for AEdb-Function.*** The data can be queried by gene names, protein keywords and database cross-references. Further, queries based on the functional enumeration of the isoform peptide sequence can be raised by selecting from a predefined list of functional categories (see Table 1 for the list of functional categories).

***Interface for AEdb-Motif.*** The interface allows free-text search. The search items include gene names, sequence of the regulatory motifs, and type of regulatory sequence (enhancer or silencer).

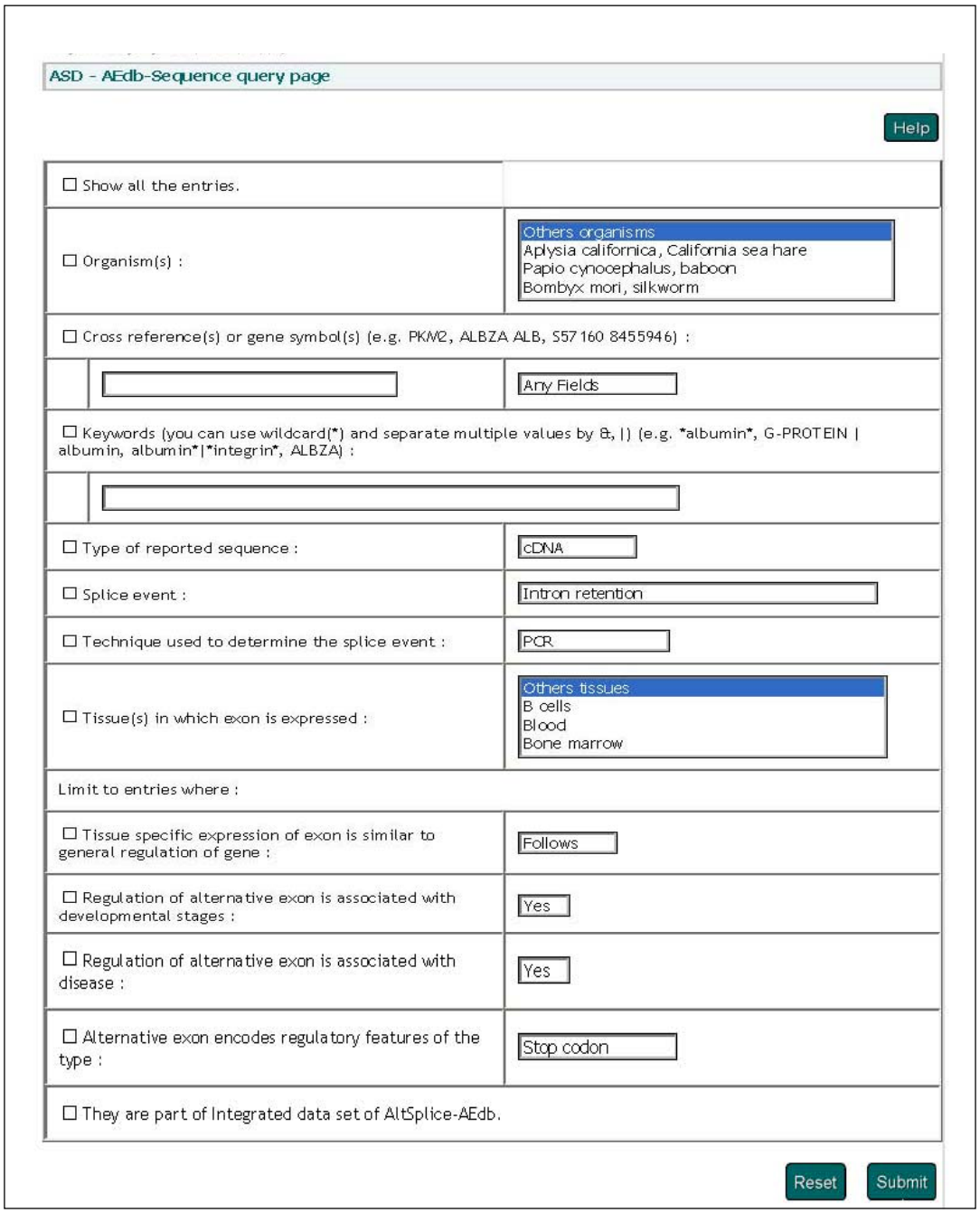



**Figure 4.18 Interface for AEdb-Sequence.** The data can be queried by gene names and synonyms, database cross-references, type of splice events, and type of regulatory roles

## *4.2 Tra2-beta1 as a splicing factor plays role in translation*

### 4.2.1 Identification of in vivo tra2-beta1 targets by CLIP

Using in vitro SELEX and mutagenesis approaches, slightly different binding sites for tra2-beta1 have been identified. The binding sites were generally purine rich, but differ in their degree of pyrimidine residues (Tacke, Tohyama et al. 1998; Stoilov, Daoud et al. 2004). To determine which RNA sequences bind to the protein in vivo, we performed cross-link and immunoprecipitation experiments. We used an antiserum that we previously developed against tra2-beta1 (Daoud, Da Penha Berzaghi et al. 1999) to immunoprecipitate endogenous RNA-tra2-beta1 complexes from HEK293 cells. The sequences of the CLIP targets are shown in supplemental Table 4.8 A. We used Melina CONSENSUS, http://melina1.hgc.jp/ to identify common sequences. This analysis of the RNA cross-linked to endogenous tra2-beta1 revealed a common purine-rich sequence that forms around a core RAAG sequence and is described by the weight matrix in Figure 1A. To our surprise, the majority of CLIP targets were from ribosomal RNA (Table 4.8 B).

**Table 4.8 A. CLIP gene targets**

| GENE | localization | Intron/Exon | PU value |
|---|---|---|---|
| **XIST** | CAACCCAAGGATGGAAGGCCCCTGTC ACAAAGCCTACCTAGATGGATAGAGG AC | E | 0.1 |
| **PTK2** | GACTCTCTCGAGGC | E | 0.3 |
| **RAC3** | GGGAGATTGGCTCTGTGAAATACCTG GAGTG | E | 0.3 |
| **EPN2** | GAAAGAGTTAGATGTGACCTC | E (5'UTR) | 0.2 |
| **PHC2** | GCATGACCTGTTCCATTCAGCGG | I (5'UTR) | 0.1 |
| **CCNDBP1** | GATCACCCTGAGAAAGCTGGTACGGG CCGCCACC | E (5'UTR) | 0.2 |
| **MYH9** | GCATCGCCCAGCTGGAGGAGGAG | E | 0.1 |
| **CEP110** | GACAAGAAGAGTTCAGGCAGGCCTGT GAGAGAGCC | E | 0.9 |
| **TIMM50** | GAAACAGGAGAGGATTTGAGATTAGG | I | 0.95 |
| **TIAM2** | CAACACAGTTCTGGTATTCGGGGTGC TATGGT | I | 0 |
| **ANLN** | GAAGATGACCGAGAGACCCTTGTCAG | E | 0 |

**Table 4.8 B. CLIP rRNA targets**

| No | location | subunit | Sequences (predicted motif marked in red) |
|---|---|---|---|
| 100 | | 5.8S rRNA | CGACTCTTAGCGGTGGATCACTCGGCTCGTGC GTCGATGAAGAACGCAGCTAG |
| 7 | | 5S rRNA | CCTGAACGCGCCCGATCTCGTCTGATCTCGGA AGCTAAGCAGGGTCGGGC |
| | | | |
| 97 | 2s | 18S rRNA | GCGTATATTAAAGTTGCTGCAGTTAAAAAGCTC GTAG |
| 5 | 3s | 18S rRNA | CAAAGTCTTTGGGTTC |
| 19 | 3s | 18S rRNA | CCATGACCCGCCGGGCAGCTTCCGGGAAACCA AAGTCTTTGGG |
| 106 | 2s | 18S rRNA (same as 52,53,120) | GCGTATATTAAAGTTGCTGCAGTTAAAAAGCTC GTAGTTGGATC |
| 108 | 1s | 18S rRNA | GAGTGTTCAAAGCAGGCCCGAGCCGCCTGGAT ACCGCAGCTAGGAATAATGGAATAGGACCGCG GTTCTATTTTGTTGGTT |
| 109 | 2s | 18S rRNA | GCGTATATTAAAGTTGCT |
| 112 | 3s | 18S rRNA | CCAAAGTCTTTGGGTTCCGG |
| 21 | 3s | 18S rRNA | CAAAGTCTTTGGGTTCCGGGGGGAG |
| 43 | 3s | 18S rRNA | GAAACCAAAGTCTTTGGGTTCCGGGG |
| 44 | 3s | 18S rRNA | CCATGACCCGCCGGGCAGCTTCCGGGAAACCA AAGTCTTTGGGT |
| 23 | 3s | 18S rRNA (part of No.21) | CAAAGTCTTTGGGTTCCGGG |
| 24 | 1s | 18S rRNA | CAAAGCAGGCCCGAGCCGCCTGGATACCGCAG |
| 203 | 3s | 18S Rrna | GGAAACCAAAGTCTTTGGGTTCCG |
| 101 | 2s | 18S rRNA or Unknown | GCGTATATTAAAGTTGCTGCAGTTAAAAAGCTC GTATTTGGATCTTGGGAG |
| 208 | 3s | 18S Rrna | CAAAGTCTTTGGGTTCCGGGGGGAGTATGGTT GCAAAG |
| 209 | 3s | 18S Rrna | ccaaagtCTTTGGGTTCCGGGGGGAGTATGGTTG C |
| 224 | 3s | 18s rrna | GACACGGAAAGTCTTTGGGTTCCGGGGGGAGT ATGG |
| 233 | 4s | 18s rrna | CCCTGCCCTTTGTACACACCGCC |
| 113 | F+4s | 28S and 18S Rrna | GGAACGTGAGCTGGGTTTAGACCGTCGTGAGAC AGGGGTCATAAGCTTGCGTTGATTAAGTCCCTGC CCTTTGTACACACCGCCCGT |
| | | | |
| 25 | F | 28S rRNA | CAGCGCCGTGGAGCCTCGGTTGGCCTCGGATAG CCGGTCCCCCGCCTGTCCC |
| 26 | B | 28S rRNA | CAGGGCGCCCTGGAATGGGTTCGCCCCGAGAG AGGGGCCCGTGCCTTGGAAAGCGTCGCGGTTC CGG |
| 28 | F | 28S rRNA (similar to No.25 but one nt different) | GCGGAGCCTCGGTTGGCCTCGGATAGCCGGTCC CCCGCCTGTCCCC |
| 32 | B | 28S rRNA | GGCGGGAGCCCCGGGGAGAGTTCTCTTTTCTT TGTGAAGGGCAGGGCGCCCTGGAATGGGTTCG CCCCGAGAGAGGGGCCCGTGCCTTGGAAAGC |

| | | | |
|---|---|---|---|
| | | | GTCGCGGTTCCGGCGGCGTCCGGTGAGCTCTCGC |
| 3 | D | 28S rRNA | GAGGGGCTCTCGCTTCTGG |
| 34 | E | HFM1 or 28S rRNA | CGGGGCCTCACGATCCTTCTGACCTTTTGGG |
| 36 | A | 28S rRNA | CGCGCCGGGGAGGTGGAGCACGAGCGCACGTGTTAGGACCCGAAAGATGGTGAAC |
| 37 | F | 28S rRNA | GCGCTAAACCATTCGTAGACGACCTGCTTCTGGGTC |
| 41 | C | 28S rRNA | GCACGGTGAAGAGACATGAGAGGTGTAGAATAAGTGGGAGGCCCCCGG |
| 47 | D | 28S rRNA | GGGCTCTCGCTTCTGGCGCCAAGCGCCC |
| 49 | F | 28S rRNA | GGAGCCTCGGTTGGCCTCGGATAGCCGGTCCCCCCGCCTGTCCCCG |
| 98 | B | 28S rRNA | GGGTTCAGATCCCCGAATCCGGAGTGGCGGAGATGGGCGCCGCGAGGCGTCCAGTGCGGTAACGCGACCGA |
| 22 | F | 28S rRNA | CAGGTTAGTTTTACCCTACTGATGATGTGTTGTTGCCATGGTAATC |
| 104 | F | 28S rRNA | CGAAGCTACCATCTGTGGGATTATGACTGAACGCCTCTAAGTCAGAATCCCGCCCAGGCGGAACGATACGGCAGCGCCGCGGAGCCTCGGTTGGCCTCGGATAGCCGGTCCCCCGCCTGTCC |
| 113 | F+4s | 28S and 18S Rrna | GGAACGTGAGCTGGGTTTAGACCGTCGTGAGACAGGGGTCATAAGCTTGCGTTGATTAAGTCCCTGCCCTTTGTACACACCGCCCGT |
| 116 | F | 28S rRNA | CCGCGCGCCGGGACCGGGG |
| 117 | A | 28S rRNA | GGAGGTGGAGCACGAGCGCACGTGTTAGGACCCGAAAGATGGTGAACTATGC |
| 119 | F | 28S rRNA | CTCGGTTGGCCTCGGATAGCCGGTCCCCCGCCTGTCCCCGCCGGCGGGCCGCCCCC |
| 206 | F | 28S Rrna | CCAGGATCTAAAAATAAAATCAGATCCAGGTTAGTTTTACCCTACTGATGATGTGTTGTTGCCATGGTAATCCTGCTCAGTACGAGAGGAACCGCAGGTTCAGACATTTGGTGTATGTGCTTGGCTGAGG |
| 207 | F | 28S Rrna | GAGCCAATGGGGCGAAGCTACCATCTGTGGGATTATGACTGAACGCCTCTAAGTCAGAATCCCGCCCAGGCGGAACGATACGGCAGCGCCGCGGAGCCTCGGTTGGCCTCGGATAGCCGGTCCCC |
| 202 | F | 28S Rrna | cgaagctaccaGTGGGATTATGACTGAACGCCTCTAAGTCAGAATCCCGCCCAGGCGGAACGATACGGCAGCGCCGCGGAGCCTCGGTTGGCCTCGGATAGCCGGTCCCCCGCCTGTCC |
| 221 | H | 28S Rrna | GGAAGAGCCCAGCGCCGAATCCCCGCCCCGCGGCGGGGCGCGGGACATGTGGCGTACGGAAGACCCGCCTGG |
| 114 | G | 28S rRNA | GATCAGACGTGGCGACCCGCT |
| 115 | H | 28S rRNA | CGGCGAGTGAACAGGGAAGAGCCCAGCGCCAATCCCCG |

To test whether these RNA sequences bind to tra2-beta1 in vitro, we performed gel shift analysis. We used bacculo-virus generated recombinant tra2-beta1 (Novoyatleva, Heinrich et al. 2007) and probes corresponding to the ribosomal CLIP targets. As shown

in Figure 4.19B, the probes bind to tra2-beta1 under in vitro conditions, demonstrating that tra2-beta1 can interact with them. We next mapped the CLIP targets to the known structure of rRNA. As shown in Figure 4.20 A, the majority of the rRNA CLIP targets originated from a region close to the sarcin/ricin loop of the large subunit, which is apparent when the CLIP targets are mapped to the three dimensional structure of the ribosomal RNA (Figure 4.20 C). In addition to hits on the large subunit, we identified several CLIP targets binding to the small subunit. Mapping them onto the three dimensional structure revealed a clustering in the XYZ region (Figure 4.20 C,D).

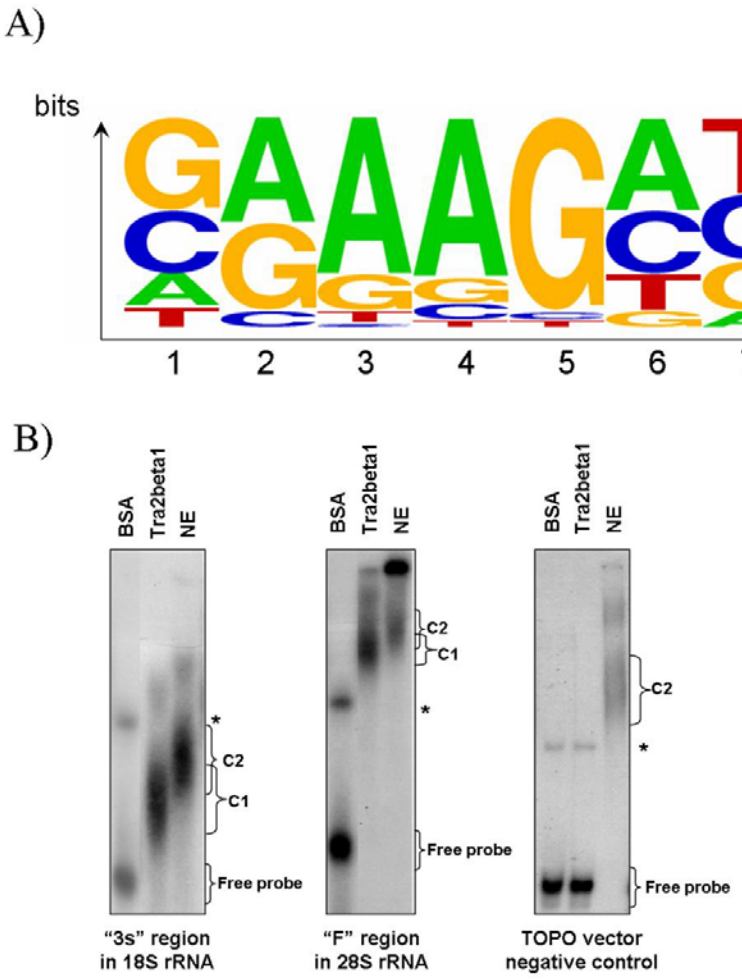
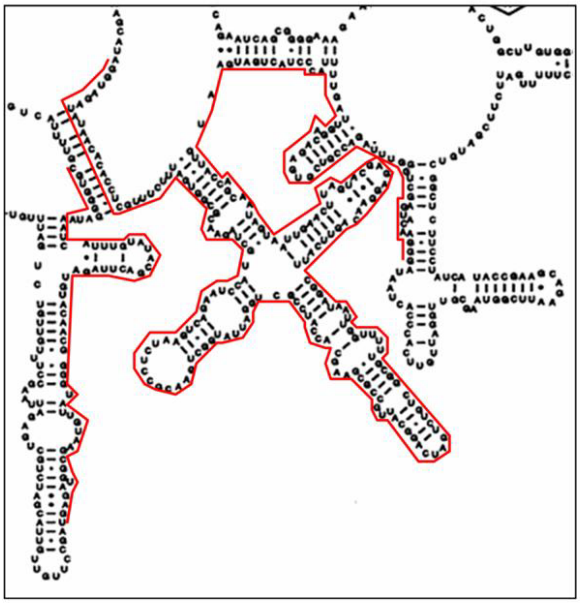



**Figure 4.19: CLIP analysis of tra2-beta1 binding sites**

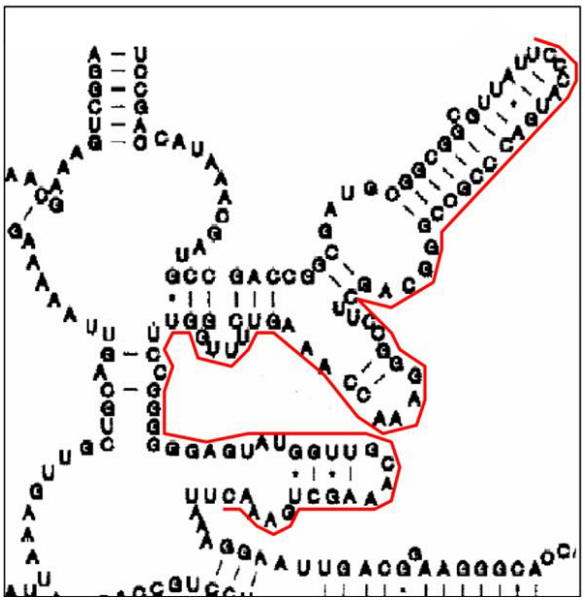
**A. CLIP motif matrix of tra2-beta1 binding.** The consensus sequence in CLIP signatures was determined by MIME and generated by WebLogo. The y axis shows the probability to find a base at the given position in bits. **B. Gel shift analysis of CLIP signatures.** CLIP

sequences corresponding to ribosomal RNA were analyzed by gel retardation assays using recombinant tra2-beta1 and nuclear extract (NE). C1: RNA complex with tra2-beta1, C2: RNA complex with nuclear extract, * indicates a band that is dependent on BSA and seen with some probes.

A



B



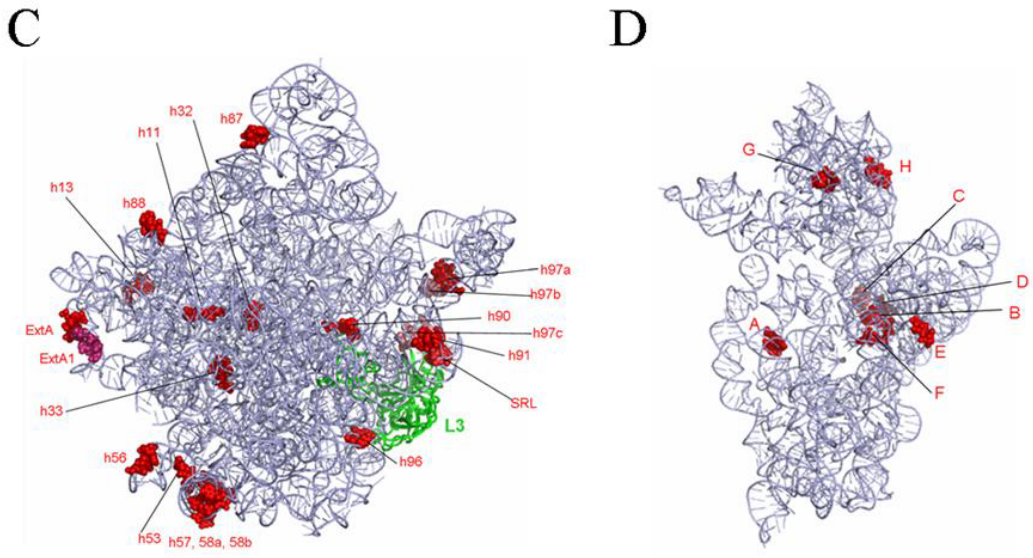



**Figure 4.20: Localization of CLIP signatures in the ribosomal model.** The localization of the CLIP sequences in the two dimensional structure of the ribosomes are shown. **A. Two dimensional structure of the large subunit that shows the most CLIP targets**. sequence reference: M11167. **B. Two dimensional structure of the small subunit.** Sequence reference: U13369. **C.Three dimensional structure of the large subunit.** The RNA residues corresponding to the CLIP targets are indicated in red. The ribosomal protein RPL3 is indicated in green. **D. 3D structure of the small subunit**. The RNA residues corresponding to the CLIP targets are indicated in red.

### 4.2.2 Tra2-beta1 binds to Rpl3

Tra2-beta1 was originally identified in two-hybrid screens using other SR-proteins as baits (Dauwalder, Amaya-Manzanares et al. 1996; Beil, Screaton et al. 1997). Subsequent two hybrid screens with tra2-beta1 revealed interactions with SR-proteins, hnRNP G (Elliott, Venables et al. 2000) and SAF-B (Nayler, Cap et al. 1998). In these screens we also repeatedly identified a ribosomal protein, RPL3 (Peltz, Hammell et al. 1999) as an interacting partner of tra2-beta1. Since we identified rRNA as an potential physiological target gene for tra2-beta1, we investigated the binding of tra2-beta1 to RPL3 in more detail.

First we verified the interaction between RPL3 and tra2-beta1 using a different method and employed GST-pull down of recombinant proteins. RPL3 was expressed in bacteria as a GST-tagged protein and coupled to glutathion-agarose. This affinity matrix was incubated with Flag-tagged-tra2-beta1 expressed in reticulate lysates. After washing,

the last supernatant and bound protein were analyzed by Western blot using an antiserum against the Flag-tag and tra2-beta1. As shown in Figure 4.21 A, left, tra2-beta1 is bound to RPL3. This experiment was repeated using His-tagged tra2-beta1 bound to Ni-agarose and GFP-RPL3 generated in reticualte lysates. GFP-RPL3 was applied to the tra2-beta1-agarose affinity matrix. As shown in Figure 4.21A right, RPL3 bound to the affinity matrix, but not to the free affinity matrix. This data confirm the binding of tra2-beta1 to RPL3 observed in yeast.

Since the proteins employed in the above experiments still contained other proteins derived from the reticulate lysates, we tested the binding between RPL3 and tra2-beta1 using recombinant purified proteins. We coupled purified bacterial derived GST-RPL3 to glutathione agarose beads and incubated it with bacculovirus derived His-tra2-beta1. As shown in Figure 4.21 B, lane 1, tra2-beta bound to the RPL3-loaded matrix. To test for unspecific binding, we used recombinant nuclear proteins YT521-B (Hartmann, Nayler et al. 1999) as a control and found that it is only present in the supernatant (Figure 4.21B, lane 3,4). The experiments were repeated with GST immobilized on agarose and in both cases, we did not observe binding to the affinity resin. These data show a direct binding between RPL3 and tra2-beta1.

The Tra2-beta1 protein contains two RS-domains flanking a central RNA recognition motif that also contains a protein phosphatase binding site in its fourth beta strand (Novoyatleva, Heinrich et al. 2007). We used yeast two hybrid deletion mutants to determine which part of tra2-beta1 binds to RPL3. The structure of the deletion proteins is shown in Figure4.21C. Growth on yeast indicator plates (Figure 4.21D) shows that deletion of the second RS-domain abolished interaction (construct 2), but deletion of the first RS domain had no effect, as shown by the growth of construct 4. Splitting the second RS domain for tra2-beta1 strongly reduced interaction with RPL3 (constructs 5 and 6). Together, these data show that tra2-beta1 binds with its second RS domain to RPL3.

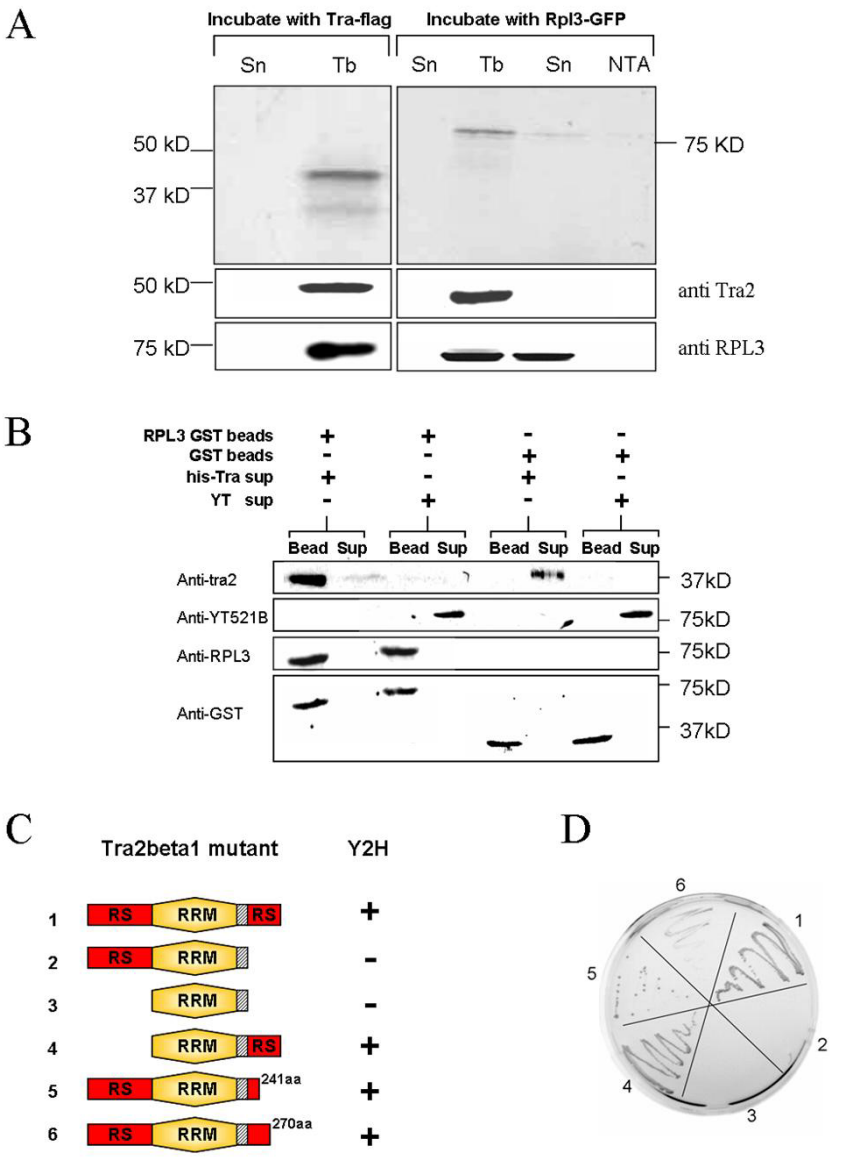



**Figure 4.21: tra2-beta1 binds to RPL3. A. Binding of in vitro translated tra2-beta1 and RPL3** (left) Recombinant Tra2-beta1 was expressed in reticulate lysates and loaded on a column that contained GST-RPL3. The affinity matrix was washed with PBS and the bound protein detected by autoradiography (right) In vitro translated RPL3 was incubated with bacculovirous derived recombinat tra2-beta1 coupled to a Ni-resin. After washing with PBS, the bound protein was detected by PAAG followed by autoradiography. Resin containing Ni was used as a negative control. **B) Binding of recombinant tra2-beta1 and RPL3**
His-tagged Tra2-beta1 and GST tagged RPL3 were generated in bacculovirus and bacteria, respectively. RPL3-GST was coupled to a GST-affinity matrix and incubated with his-tra2-beta1 or bacculo virus generated recombinant YT521-B. Protein bound to the affinity matrix (bound) and supernatant (sup) was detected by PAAG followed by western blot using the specific antisera **C. Schematic representation of tra2-beta1 mutants tested for RPL3 binding.** Both RS domains (RS1, RS2) and the RNA recognition motif (RRM) are indicated. The PP1 binding site is

indicated by a striped box. The interaction with RPL3 in yeast two hybrid system is indicated by + and -. **D. Binding of tra2-beta1 to RPL3 in yeast, A plate lacking Leu, Trp, His (LWH) indicates growth between RPL3 and the constructs numbered in panel A.**

### 4.2.3 Tra2-beta1 cosediments with ribosomal proteins.

We next determined whether RPL3 resides in the same cellular fractions as tra2-beta1 and separated RNP complexes using sucrose gradients sedimentation analysis. HEK293 cells were lysed and separated on a continuous 10 to 60% sucrose gradient. Fractions of the gradient were tested for the presence of rRNA by agarose gel electrophoreses and followed by staining the ribosomal RNA with ethidium bromide. 200 µl aliquots of the fractions representing approximately 40,000 cells were concentrated to 30µl and analyzed by gel electrophoresis using antisera against the endogenous proteins. As can be seen from Figure 4.22, tra2-beta1 is found in the RNP fraction in the beginning of the gradient. It also can be detected in the fractions that contains 60 and 80S ribosomes, but not the polyribosomal fraction. The distribution is similar to SF2/ASF an SR-protein previously shown to be involved in translation (Sanford, Gray et al. 2004). RPL3 can be detected predominantly in the polyribosomal and ribosomal fractions. As a negative control, we used YT521-B, a nuclear protein that does not bind to RPL3 (Figure 4.22B). As shown in Figure 4.22, YT521-B does not cosediment with ribosomal RNA or RPL3 in this assay. These data suggest that tra2-beta cosediments with ribosomal fraction from cellular lysates.

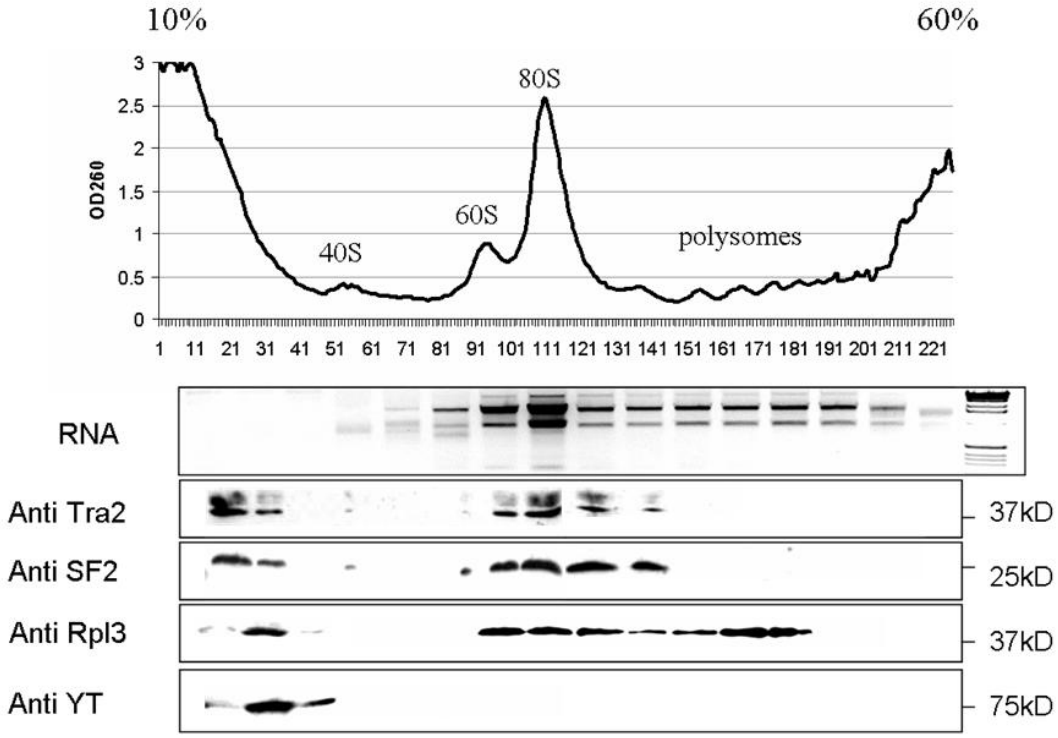



**Figure 4.22: Sucrose gradient analysis of tra2-beta1.** RNA derived from HEK293 cells were fractionated on a 10 to 60% sucrose gradient. A. OD260 from the gradient. The positions of the 40S, 60S and 80S ribosomal RNAs and the polysome fractions are indicated. B. Gradient fractions were directly separated on 1.5% agarose gels and stained with ethidiume bromide. C. Western blot detecting tra2-beta1. D. Western blot detecting RPL3 E. Western blot detecting SF2/ASF F. Western blot detecting YT521-B

### 4.2.4 Tra2-beta1 influences translation in reporter genes

Our data indicate that tra2-beta1 binds to ribosomal parts. We therefore next tested its functional involvement in translation. In several well-studied systems, tra2-beta1 regulates alternative splice site selection by binding to purine-rich enhancer sequences. For example, tra2-beta1 protein promotes inclusion of its own exon II into its pre-mRNA (Stoilov, Daoud et al. 2004) by binding to four enhancer sequences that are shown in Figure 4.23A. We tested the effect of these enhancer sequences on translation by cloning them into a luciferase reporter construct that was previously used to study the effect of SF2/ASF on translation (Sanford, Ellis et al. 2005). The reporter gene expresses luciferase controlled by an SV40 promoter and does not contain any introns. Each

enhancer sequence is within the luciferase reading frame (Figure 4.23B). These reporters were cotransfected with expression clones for tra2-beta1, tra2-beta1-NES, a mutant that contains a nuclear export signal and tra2-beta1-RATA-NES, a tra2-beta1 mutant with a nuclear export signal and a mutated protein phosphatase 1 binding site (Novoyatleva, Heinrich et al. 2007). The luciferase activity was controlled by cotransfected renilla plasmid. As can be seen in Figure 4.23C, introducing the tra2-beta1 enhancer stimulated translation of the reporter construct when compared to the EDAm control. Increasing the tra2-beta1 concentration by cotransfection did not statistical significantly increase the translation of the constructs. However, when we used a tra2-beta1-NES mutant, that increased tra2-beta1 concentration in the cytosol (Novoyatleva, Heinrich et al. 2007), we observed a strong increase of luciferase activity. This effect was less pronounced when the PP1 binding properties of tra2-beta1 were abolished in the tra2-beta1-NES-RATA mutant. Similar amounts of mutants were present in all transfections as determined by Western blot analysis against the transfected proteins (Figure 4.23C).

We next tested the action of tra2-beta1 on translation directly in a cell free system. We used reticulate lysates, where we expressed simultaneously renilla and firefly luciferase containing ESE4. To each reaction, either GST, GST-tra2-beta1 or Flag-tra2-beta1 generated by in vitro translation in reticulate lysates was added. As shown in Figure 4.23D, addition of tra2-beta1 significantly increased the production of the tra2-beta1 dependent firefly construct.

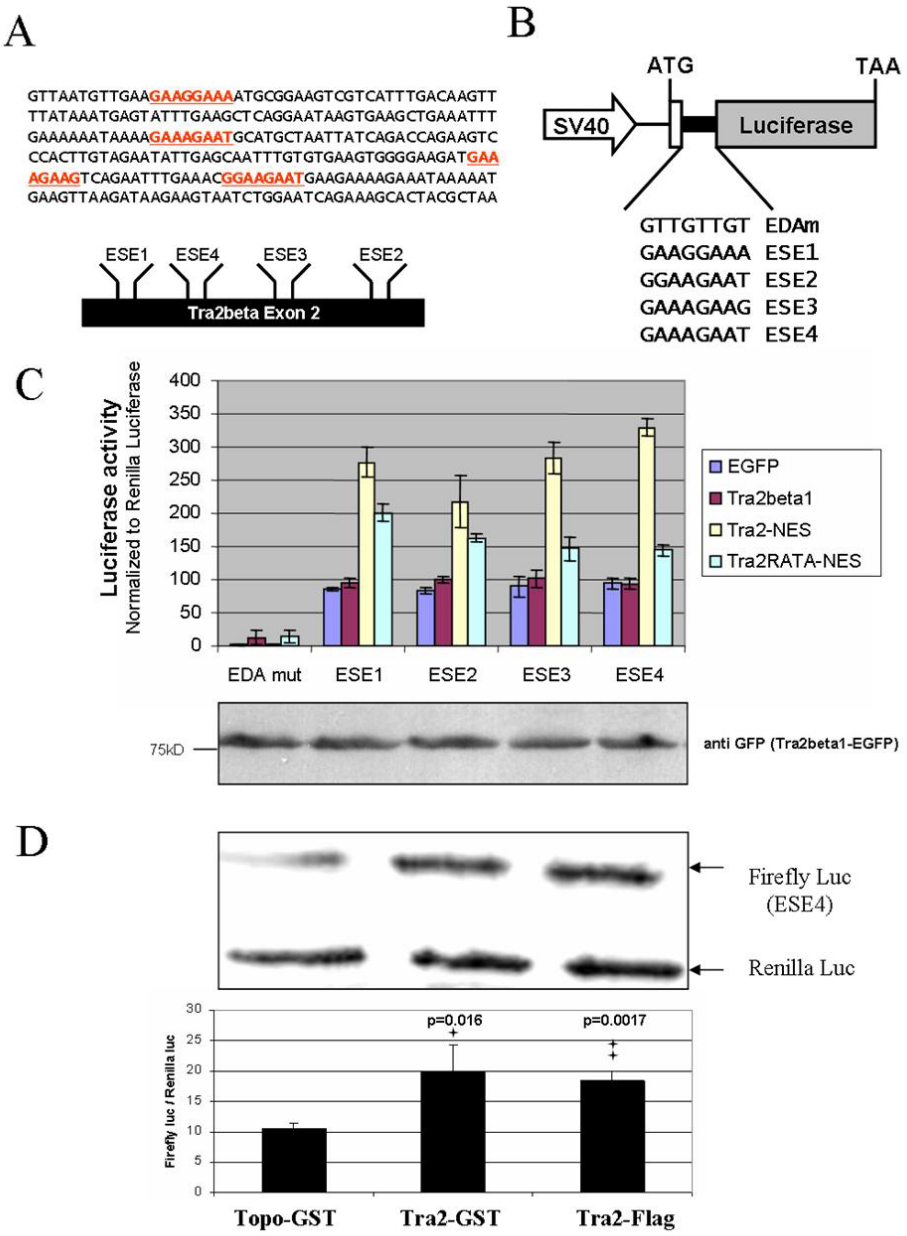



**Figure 4.23: tra2-beta1 promotes translation in reporter genes. A. Sequence of the tra2-beta1 dependent exon 2 of the tra2-beta1 pre-mRNA.** The four previously identified tra2-beta1 enhancers are underlined. **B. Structure of the luciferase construct used,** which is based on previously published construct (Sanford, Gray et al. 2004). The sequences cloned into the luciferase reading frame are indicated. EDA mutant is a mutant fibronectin exonic enhancer (Sanford, Gray et al. 2004) that serves as a negative control. **C. Luciferase assay with tra2-beta1 mutants and translation reporter.** 1 µg of the luciferase reporter genes described in panel B were transfected with 1µg of the expression constructs indicated. EGFP: expression clone for EGFP, tra2-beta1: expression clone for EGFP-tra2-beta1, tra2-NES: expression clone for tra2-beta1 containing a nuclear export sequence; tra2-RATA-NES: expression clone for tra2-beta1 containing an nuclear export site, but the protein phosphatase binding site of tra2-beta1 was destroyed by mutating it from RVDF to RATA. **D. In vitro**

**translation assay.** The ESE4 construct shown in panel D were in vitro translated using reticulate lysates. Renilla luciferase was used in the same reaction as a standard. The formed protein incorporated $^{35}S$ and was determined by autoradiography, which is shown in the PAGE gel. Below, a quantification of three experiments is shown. Error bars indicate the standard deviation; p-values from four independent experiment are indicated.

### 4.2.5 Tra2-beta1 regulates the translation of several genes identified by CLIP

We identified three tra2-beta1 CLIP targets where tra2-beta1 binds to regions in the 5' UTR and determined whether a reduction of tra2-beta1 by siRNA influences their translation. First, we analyzed cyclin D-type binding-protein 1 CCNDBP1, which is also named DIP1 or GCIP (Xia, Bao et al. 2000){Su, 2007 #6802}. The gene product can initiate from at least two alternative start codons, alpha and beta, giving rise to two isoforms of 27kD and 40kD, respectively. We identified a tra2-beta1 CLIP signature in the exon that contains the alpha start codon (Figure 4.24A). We analysed RNA from Hela cells where tra2-beta1 concentration was reduced by siRNA by Western blot and RT-PCR using the primers indicated in Figure 4.24A. RT-PCR analysis revealed that a reduction of tra2-beta1 concentration did not change the splicing pattern of CCNDBP1. However, Western blot analysis shows that the reduction of tra2-beta1 concentration resulted in a loss of the alpha isoform, whereas the beta isoform remained unchanged, suggesting that tra2-beta1 promotes translation of the alpha isoform, but does not influence RNA splicing in this region of the pre-mRNA.

The next CLIP signature was identified in epsin 2 (EPN2), a cytosolic protein that interacts with components of the clathrin coat. The CLIP signature was in an alternative exon six exons further upstream from the protein start site. Again, we observed no change in a RNA abundance of the expsin mRNA, but saw a strong dependency of the amount of protein detected by Western blot on the concentration of tra2-beta1.

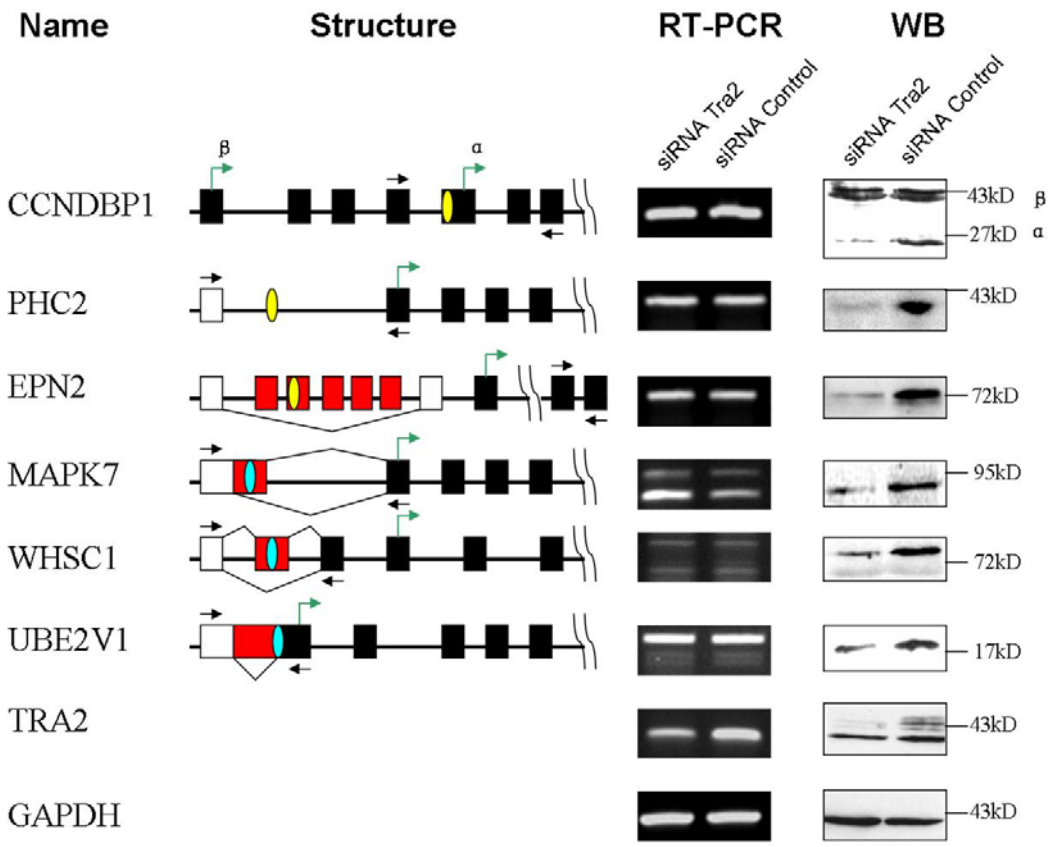



**Figure 4.24: Endogenous genes regulated by tra2-beta1 on the translational level. A. Gene structure of genes containing a CLIP target in the 5' UTR.** The CLIP target is indicated by a yellow oval. Arrows indicate translational start sites. Red boxes show alternative exons. The translational start sites are. Open boxes indicate non-translated exons. **B. RNA expression.** RT-PCR was performed with primers indicated in panel A. **C. Protein expression depends on tra2-beta1 concentration.** Tra2-beta1 was reduced in HeLa cells by siRNA treatment. Cell lysates were analyzed by western blot using antisera against the proteins depicted in panel A.

### 4.2.6 The tra2-beta1 binding site can be used to predict translational regulation by tra2-beta1

We next asked whether we can predict the influence of tra2-beta1 on translation using a bioinformatics approach. We scanned all human 5' UTRs for the existence of tra2-beta1 binding motifs located in alternative exons near start sites. We used the tra2-beta1 matrix shown in Figure 1A. As shown in Figure 6A, the gene for mitogen-activated protein kinase 7 (MAPK7), also known as ERK5, which contains an exon with an alternative 5' splice site downstream of the start codon. RT-PCR analysis shows that inclusion of this alternative exon is not dependent on the tra2-beta1 concentration.

However, reduction of tra2-beta1 by siRNA decreases the amount of translated MAPK7, indicating that tra2-beta1 influences translation and not splicing. Another example is the Wolf-Hirschhorn syndrome candidate 1, WHSC1. This gene encodes a nuclear protein and undergoes extensive alternative splicing, generating at least 22 isoforms. Here, the putative tra2-beta1 binding site is in an alternative cassette exons upstream of the start codon. Tra2-beta1 has no influence on the splice site selection, however its reduction by siRNA increases the production of the protein. The last example analysed was the ubiquitin-conjugating enzyme E2 variant 1 (UBE2V1) that contained a tra2-beta1 binding signature in a retained intron upstream of the start codon. In this case, tra2-beta1 reduction did not have an effect on processing of the RNA, but reduced the amount of the encoded protein.

Together, these data indicate that the presence of tra2-beta1 binding sites in the 5'UTR can indicate an effect of tra2-beta1 on translation.

## 4.3 Supplemental figure and table

### .3.1 Tra2/Tra2-NES analysis in small scale custom chip

Total RNA from Neuro2A transfected with Tra2-beta1-EGFP, Tra2-beta1-NES-EGFP was sent for DNA chip analysis. pEGFP-C2 transfected Neuro2A cells were taken as control.

| | | EGFP / Tra2-beta1 | | Tra2-beta1 / Tra2-beta1-NES | |
|---|---|---|---|---|---|
| function group | Transcript | median LR | IgnoreMe | median LR | IgnoreMe |
| apoptosis | apaf1_com | -0.368 | 0 | 0.185 | 6.777 |
| apoptosis | apaf-1L | 0 | 1 | 0 | 1 |
| apoptosis | Ayelet2 (luc7-homolog?) | -0.08 | 0 | 0.671 | 0 |
| apoptosis | bak1_altN | 0 | 1 | 0.094 | 8.088 |
| apoptosis | bak1_com | 0 | 1 | 0 | 1 |
| apoptosis | bax_alpha | -0.029 | 6.219 | 0.195 | 0 |
| apoptosis | bax_kappa | 0.216 | 0 | -0.038 | 0 |
| apoptosis | bcl2_alpha | 0.422 | 0 | 0 | 1 |
| apoptosis | bcl2_beta | -0.166 | 0 | -0.304 | 0 |
| apoptosis | Bcl2l13 | -0.014 | 0 | 0.058 | 0 |
| apoptosis | bcl-x_alt2 | -0.513 | 0 | 0 | 1 |
| apoptosis | bcl-X_com | -0.112 | 0 | -0.233 | 0 |
| apoptosis | bcl-x_gamma | 0 | 1 | 0 | 1 |
| apoptosis | bcl-x_long | -0.141 | 0 | 0 | 1 |
| apoptosis | BI1 | -0.102 | 0 | -0.258 | 0 |
| apoptosis | Bid | 0.404 | 0 | -0.467 | 0 |
| apoptosis | casp1 | -0.082 | 5.195 | 0 | 1 |
| apoptosis | casp2_alt | 0.255 | 0 | -0.078 | 21.732 |
| apoptosis | casp2_com | 0.025 | 0 | 0.474 | 0 |
| apoptosis | casp6 | -0.125 | 0 | 0.441 | 0 |
| apoptosis | Casp7 | -0.875 | 0 | 0.5 | 0 |
| apoptosis | casp8 | 0.077 | 0 | 0.459 | 0 |
| apoptosis | casp9L | 0.243 | 0 | -0.414 | 0 |
| apoptosis | casp9S | 0 | 1 | 0 | 1 |
| apoptosis | Ctnnbl1 (NAP) | -0.267 | 0 | 0.08 | 0 |
| apoptosis | Fas_beta | 0 | 1 | 0 | 1 |
| apoptosis | fas_long | 0.163 | 0 | -0.13 | 0 |
| apoptosis | FLIP-L | 0.082 | 0 | -0.159 | 0 |
| apoptosis | FLIP-S | 0.231 | 0 | 0 | 1 |
| apoptosis | ICAD_com | -0.335 | 0 | 0.174 | 0 |
| apoptosis | ICAD_L | -0.857 | 0 | 0 | 1 |
| apoptosis | LARD_altA | 0.034 | 113.31 | 0 | 1 |
| apoptosis | LARD_altB | 0.242 | 0 | -0.163 | 0 |
| apoptosis | madd_alt | 0.137 | 0 | 0 | 1 |
| apoptosis | madd_com | -0.317 | 0 | 0.082 | 0 |
| apoptosis | Mcl1 | 0.747 | 0 | -0.192 | 0 |

| | | | | | |
|---|---|---|---|---|---|
| apoptosis | traf2_com | 0.093 | 5.006 | 0.02 | 0 |
| apoptosis | TRAF2A | 0.021 | 0 | 0.073 | 0 |
| helicases | Ayelet1 (KIAA-homolog?) | 0.332 | 0 | -0.007 | 0 |
| helicases | Bat1a | -0.725 | 0 | 0.291 | 0 |
| helicases | Ddx15 | 0.165 | 0 | 0.231 | 0 |
| helicases | Ddx16 | 0.195 | 0 | -0.015 | 6.026 |
| helicases | Ddx27_alt1 | 0.283 | 0 | -0.078 | 0 |
| helicases | Ddx27_alt2 | 0.211 | 0 | -0.039 | 0 |
| helicases | Ddx27_com | -0.674 | 0 | 0.207 | 0 |
| helicases | Ddx41 | -0.221 | 0 | 0.135 | 0 |
| helicases | Ddx46 | 0.314 | 0 | -0.039 | 0 |
| helicases | ddx48 | 0.311 | 0 | -0.049 | 0 |
| helicases | Ddx5 (p68) | -1.314 | 0.322 | 0.867 | 0 |
| helicases | Ddx9 | 0.38 | 0 | 0.233 | 0 |
| hnRNPs | brPTB (ptb2) | 0.058 | 0 | 0.514 | 0 |
| hnRNPs | hnRNP A1 | -0.62 | 0 | 0.525 | 0 |
| hnRNPs | hnRNP A2/B1-old | 0.038 | 0 | 0.15 | 11.263 |
| hnRNPs | hnRNP AB_alt1 | 0 | 1 | 0.001 | 0 |
| hnRNPs | hnRNP AB_alt2 | -0.121 | 0 | 0.035 | 0 |
| hnRNPs | hnRNP C | -0.003 | 0 | -0.186 | 7.168 |
| hnRNPs | hnRNP G | -0.113 | 0 | 0.198 | 0 |
| hnRNPs | hnRNP H1 | -0.548 | 0 | 0.726 | 0 |
| hnRNPs | hnRNP K_alt1 | -0.849 | 0 | 0.543 | 0 |
| hnRNPs | hnRNP K_alt2 | -0.529 | 0 | 0.31 | 0 |
| hnRNPs | hnRNP K_alt3 | 0.27 | 0 | 0.614 | 0 |
| hnRNPs | hnRNP K_Pcom | -0.584 | 0 | 0.488 | 0 |
| hnRNPs | hnRNP L-old | -0.405 | 0 | 0.209 | 0 |
| hnRNPs | Pcbp2_alt | 0.229 | 0 | 0.245 | 0 |
| hnRNPs | Pcbp2_com | -1.572 | 0 | 0.785 | 0 |
| hnRNPs | Ptbp1_alt1 | 0.198 | 0 | -0.134 | 0 |
| hnRNPs | Ptbp1_alt2 | -0.303 | 0 | -0.077 | 0 |
| hnRNPs | Ptbp1_com | -1.261 | 0 | 0 | 1 |
| hnRNPs | Rbmxrt (hnRNP G) | 0.006 | 0 | 0.307 | 0 |
| hnRNPs | snRNP E | 0.225 | 0 | -0.231 | 0 |
| house-keeping genes | atubulin1 | -0.488 | 0 | 0.065 | 0 |
| house-keeping genes | atubulin4 | 0.757 | 0 | -0.595 | 0 |
| house-keeping genes | bActin | -0.038 | 0 | -0.042 | 0 |
| house-keeping genes | Gapd | 0.18 | 0 | -0.221 | 0 |
| house-keeping genes | Ndufc1 | 0.279 | 0 | -0.335 | 0 |
| house-keeping genes | Pgk1 | -0.157 | 0 | 0.095 | 0 |
| house-keeping genes | Rl13a | 0.227 | 0 | -0.106 | 0 |
| mRNA processing | CPSf1 | -0.049 | 0 | -0.163 | 0 |
| mRNA processing | Cpsf2 | 0.235 | 0 | 0.006 | 0 |
| mRNA processing | CPSf3 | 0.091 | 0 | -0.069 | 0 |
| mRNA processing | Cpsf4 | -0.859 | 0 | 0.226 | 0 |
| mRNA processing | Cpsf5 | -0.042 | 0 | 0.153 | 0 |
| mRNA processing | cstf3_alt | -0.21 | 0 | 0.298 | 0 |

| | | | | | |
|---|---|---|---|---|---|
| mRNA processing | cstf3_com | -1.125 | 0 | 0.617 | 0 |
| mRNA processing | Ncbp2 | -0.009 | 0 | 0.316 | 0 |
| mRNA processing | NLP4 | 0.178 | 0 | -0.338 | 0 |
| mRNA processing | Pabpc1 | -0.516 | 0 | 0.594 | 0 |
| mRNA processing | Pabpn1 | 0.175 | 0 | -0.644 | 0 |
| mRNA processing | Papolb | 0.467 | 0 | -0.226 | 0 |
| mRNA processing | Refbp1 | -0.408 | 0 | 0.574 | 0 |
| mRNA processing | Refbp2 | -0.411 | 0 | 0.505 | 0 |
| other genes | BChE | 0 | 1 | 0 | 1 |
| other genes | BCHE-old | 0.047 | 0 | -0.007 | 0 |
| other genes | Mapk8 | -0.181 | 0 | 0 | 1 |
| other genes | Mfap1_alt1 | 0.297 | 0 | 0.168 | 0 |
| other genes | Mfap1_alt2 | 0.461 | 0 | -0.236 | 0 |
| other genes | Pdyn_old | 0.117 | 0 | 0.022 | 0 |
| other genes | PenK1_old | -0.107 | 17.794 | -0.029 | 0 |
| other genes | Pon1 | 0 | 1 | 0 | 1 |
| other genes | Pon2 | 0 | 1 | 0 | 1 |
| other genes | Pon3 | 0.147 | 0 | 0.047 | 0 |
| other genes | Thyrosine hydorxylase | 0.545 | 0 | -0.604 | 0 |
| other spliceosomal component | Ayelet 4 | -0.063 | 0 | -0.038 | 8.364 |
| other spliceosomal component | Bcas2 | 0.142 | 0 | 0.147 | 0 |
| other spliceosomal component | Cd2bp2 | 0.081 | 0 | -0.229 | 0 |
| other spliceosomal component | Crnk1 | 0.239 | 0 | 0.359 | 0 |
| other spliceosomal component | Dnajc | 0.071 | 0 | -0.009 | 0 |
| other spliceosomal component | Fnbp3 (FB11) | 0.561 | 0 | -0.103 | 0 |
| other spliceosomal component | RBM17 | 0.069 | 9.222 | 0.038 | 0 |
| other spliceosomal component | Rbm8 | -0.656 | 0 | 0.573 | 0 |
| other spliceosomal component | Rnpc2 | 0.125 | 0 | 0.372 | 0 |
| other spliceosomal component | sam68(Khdrbs1) | -0.213 | 0 | 0.134 | 0 |
| other spliceosomal component | SLM1 | 0.267 | 0 | 0 | 1 |
| other spliceosomal component | SMNRP(sf30) | 0.207 | 0 | 0.141 | 0 |
| other spliceosomal component | Spop_alt | 0.447 | 0 | -0.136 | 0 |
| other spliceosomal component | Spop_com | 0.388 | 0 | -0.01 | 0 |
| other spliceosomal component | Thoc1 | 0.31 | 0 | 0.051 | 0 |
| other spliceosomal component | Wbp11 | 0.08 | 0 | -0.116 | 0 |
| other spliceosomal component | Wtap | 0 | 1 | 0 | 1 |
| snRNPs | Lsm2_com (smx5) | -0.817 | 0 | -0.177 | 8.118 |
| snRNPs | Lsm2_Pcom (smx5) | -0.219 | 0 | -0.15 | 0 |
| snRNPs | Lsm3 | -0.052 | 0 | -0.164 | 0 |

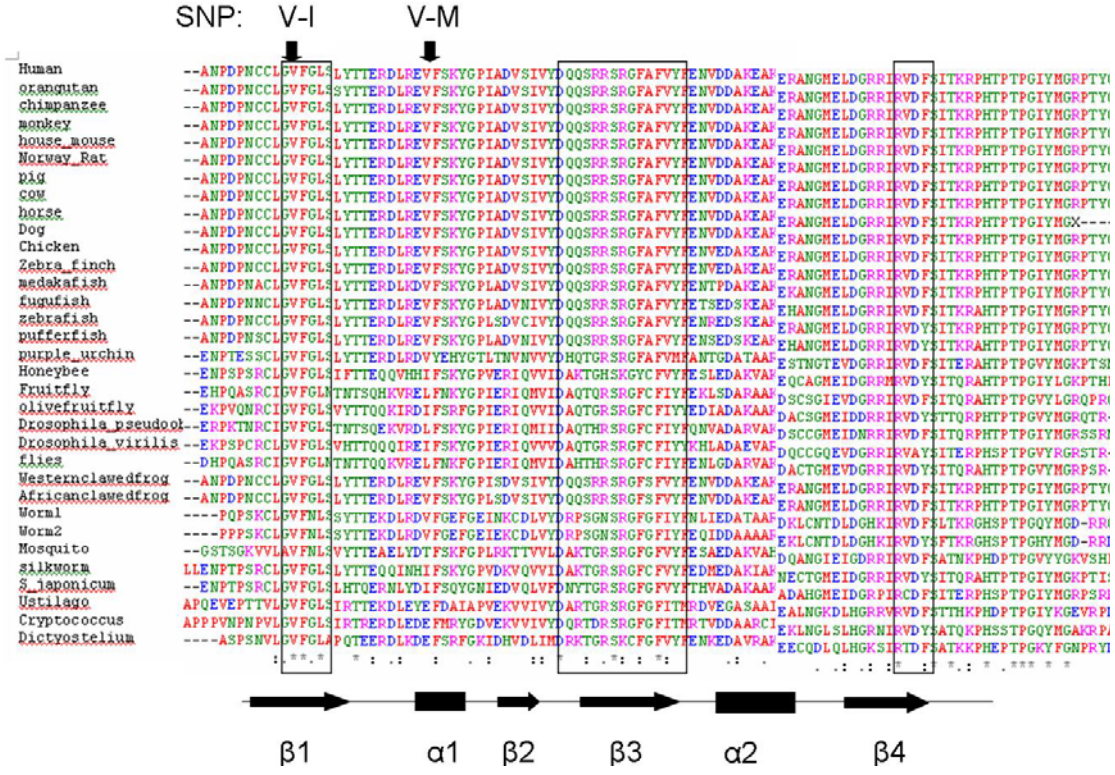
| | | | | | |
|---|---|---|---|---|---|
| snRNPs | Lsm4 | 0.119 | 0 | -0.501 | 0 |
| snRNPs | Lsm7 | -0.14 | 0 | -0.299 | 0 |
| snRNPs | p14-pending | 0.331 | 0 | 0.077 | 0 |
| snRNPs | Prpf3 (U4/U6-90kD) | 0.116 | 0 | -0.167 | 0 |
| snRNPs | Prpf4 (U4/U6-60kD) | -0.443 | 0 | 0 | 1 |
| snRNPs | Sf3a1 | 0.093 | 0 | 0.068 | 8.315 |
| snRNPs | sf3a2_alt1 | -0.381 | 0 | -0.135 | 0 |
| snRNPs | sf3a2_alt2 | 0 | 1 | 0 | 1 |
| snRNPs | Sf3a2_alt3 | -0.673 | 0 | -0.615 | 0 |
| snRNPs | sf3a2_com | -0.6 | 0 | -0.614 | 0 |
| snRNPs | Sf3a3 | -0.377 | 0 | 0.705 | 0 |
| snRNPs | sf3b1 (SAP155) | -0.37 | 0 | 0.405 | 0 |
| snRNPs | Sf3b4 | -0.187 | 0 | -0.306 | 0 |
| snRNPs | snRNP 1C | -0.134 | 0 | -0.272 | 0 |
| snRNPs | snRNP A | -0.093 | 0 | -0.633 | 0 |
| snRNPs | snRNP B2 | 0.139 | 0 | 0 | 1 |
| snRNPs | snRNP B-old | -0.093 | 0 | -0.502 | 0 |
| snRNPs | snRNP D1-old | 0.195 | 0 | 0.204 | 0 |
| snRNPs | snRNP d2 | 0.252 | 0 | -0.335 | 0 |
| snRNPs | snRNP d3 | 0.124 | 0 | -0.292 | 0 |
| snRNPs | snRNP G | -0.186 | 0 | -0.127 | 0 |
| snRNPs | snRNP N_alt1 | 0.535 | 0 | -0.451 | 0 |
| snRNPs | snRNP N_alt2 | -0.003 | 0 | -0.15 | 0 |
| snRNPs | snRNP N_alt3 | 0.126 | 0 | -0.313 | 0 |
| snRNPs | Tri-snRNP 27kD-pending (RY1?) | 0.091 | 0 | 0.083 | 0 |
| snRNPs | U1snRNP70_alt | 0.228 | 0 | -0.15 | 0 |
| snRNPs | U1snRNP70_com | 0.055 | 0 | -0.103 | 10.869 |
| snRNPs | U2A'-old | 0.133 | 0 | -0.014 | 0 |
| snRNPs | U4/U6-20kD-pending_alt1 | 0.094 | 0 | -0.092 | 0 |
| snRNPs | U4/U6-20kD-pending_alt2 | 0.44 | 0 | 0.067 | 0 |
| snRNPs | U4/U6-20kD-pending_alt3 | 0.116 | 0 | -0.045 | 0 |
| snRNPs | U4/U6-61kD | 0.07 | 0 | -0.169 | 0 |
| snRNPs | U5 116 kd-old | -0.15 | 0 | -0.08 | 0 |
| snRNPs | U5-102kd | 0.068 | 0 | -0.355 | 0 |
| snRNPs | U5-15_alt1 | 0.594 | 0 | -0.585 | 0 |
| snRNPs | U5-15_alt2 | -0.07 | 0 | -0.327 | 0 |
| snRNPs | U5-40 | 0.068 | 0 | 0.149 | 0 |
| spliceosome assembly | Cdc5l | 0.405 | 0 | -0.111 | 0 |
| spliceosome assembly | gemin2 (sip1) | 0.068 | 0 | 0.197 | 0 |
| spliceosome assembly | gemin3_alt | 0.322 | 0 | 0.086 | 9.739 |
| spliceosome assembly | gemin3_com | 0.279 | 0 | 0.116 | 0 |
| spliceosome assembly | gemin4 | 0.154 | 13.799 | 0.096 | 0 |
| spliceosome assembly | Gemin5 | 0.254 | 0 | -0.071 | 0 |
| spliceosome assembly | Gemin6 | 0.113 | 0 | 0.114 | 0 |

| | | | | | |
|---|---|---|---|---|---|
| spliceosome assembly | Gemin7 | 0.267 | 0 | -0.037 | 0 |
| spliceosome assembly | Plrg1 | -0.163 | 0 | 0.219 | 0 |
| spliceosome assembly | prp19 | -0.12 | 0 | -0.323 | 0 |
| spliceosome assembly | Prpf8 | -0.203 | 0 | 0.145 | 0 |
| spliceosome assembly | Sart1 | 0.328 | 0 | -0.245 | 0 |
| spliceosome assembly | SKIIP | 0.183 | 0 | -0.09 | 0 |
| spliceosome assembly | Slu7-pending | 0.274 | 0 | -0.085 | 14.873 |
| splicing factors phosphorylation | Abl1 | 0.075 | 0 | 0.476 | 0 |
| splicing factors phosphorylation | Cdc2a | -0.169 | 0 | 0.23 | 0 |
| splicing factors phosphorylation | Clk1 (sty) | 0.079 | 0 | 0.505 | 0 |
| splicing factors phosphorylation | Clk2 | -0.107 | 0 | -0.052 | 0 |
| splicing factors phosphorylation | Clk3_alt | -0.108 | 9.241 | 0.238 | 0 |
| splicing factors phosphorylation | Clk3_com | 0.327 | 0 | -0.169 | 0 |
| splicing factors phosphorylation | Clk4 | -0.374 | 0 | 0.416 | 0 |
| splicing factors phosphorylation | crk7 | 0.413 | 0 | 0 | 1 |
| splicing factors phosphorylation | Dusp11 | 0.299 | 0 | -0.014 | 0 |
| splicing factors phosphorylation | NIPP1 | -0.855 | 0 | -0.127 | 0 |
| splicing factors phosphorylation | pp2r2b | 0.296 | 0 | 0.416 | 0 |
| splicing factors phosphorylation | Ppm1g (pp2c) | -0.574 | 0 | -0.338 | 0 |
| splicing factors phosphorylation | Ppp2r5e | 0.129 | 0 | 0.748 | 0 |
| splicing factors phosphorylation | Prpf4b_alt1 | 0.563 | 0 | 0.25 | 0 |
| splicing factors phosphorylation | Prpf4b_alt2 | 0.556 | 0 | -0.117 | 6.151 |
| splicing factors phosphorylation | Prpf4b_com | -0.69 | 0 | 0 | 1 |
| splicing factors phosphorylation | SRpK1 | 0.017 | 0 | -0.171 | 0 |
| splicing factors phosphorylation | SRpk2 | -0.256 | 0 | 0.234 | 0 |
| splicing factors phosphorylation | topor | 0.115 | 0 | 0.186 | 5.464 |
| SR and SR-RELATED | 9G8 (sfrs7) | -0.16 | 0 | 0.316 | 0 |
| SR and SR-RELATED | ASF/SF2 (sfrs1) | -1.376 | 0 | 0.922 | 0 |
| SR and SR-RELATED | SC35_5 | -0.513 | 0 | 0.477 | 0 |
| SR and SR-RELATED | SC35_alt | 0.411 | 0 | -0.184 | 0 |
| SR and SR-RELATED | SC35_com | -0.481 | 0 | -0.003 | 0 |
| SR and SR-RELATED | SC35-old | 0 | 1 | 0.128 | 0 |
| SR and SR-RELATED | SRp20 (sfrs3) | 0.153 | 0 | 0.482 | 0 |
| SR and SR-RELATED | SRp25 (Arl6ip4) | 0.212 | 0 | -0.156 | 0 |
| SR and SR-RELATED | SRp30c (Sfrs9) | -0.046 | 0 | 0.074 | 0 |
| SR and SR-RELATED | SRp40 (sfrs5) | -1.44 | 0 | 0.861 | 0 |
| SR and SR-RELATED | SRp54 | -1.495 | 0 | 2.18 | 0 |
| SR and SR-RELATED | SRp55 (sfrs6)-pending | -0.189 | 0 | 0.183 | 0 |
| SR and SR-RELATED | SRp75_alt1 | -1.344 | 0 | 0 | 1 |

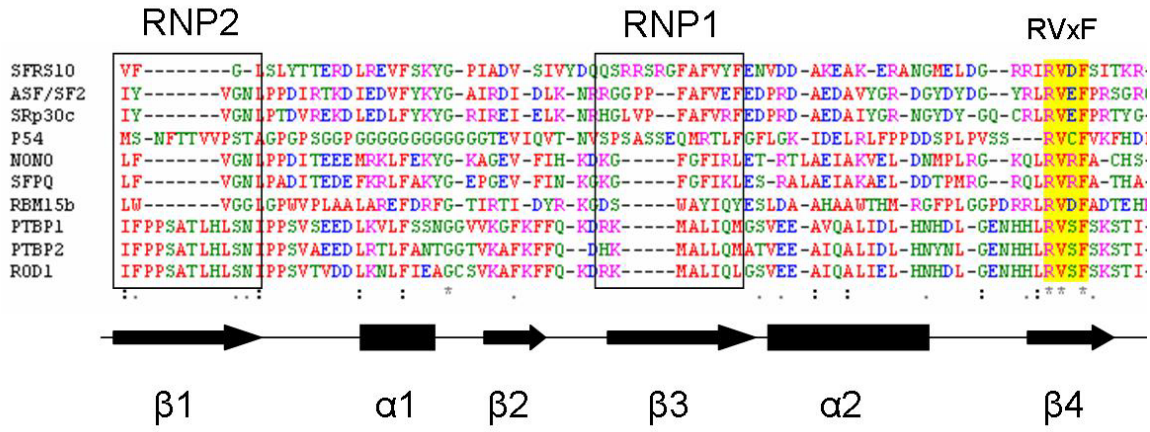
| | | | | | |
|---|---|---|---|---|---|
| SR and SR-RELATED | SRp75_E3 | -0.07 | 90.401 | -0.099 | 0 |
| SR and SR-RELATED | Srrm1 (srm160)_alt | -0.624 | 0 | 0.214 | 6.777 |
| SR and SR-RELATED | Srrm1 (srm160)_com | 1.013 | 0 | -0.557 | 0 |
| SR and SR-RELATED | srrm2 (srm300) | 0.077 | 0 | -0.15 | 0 |
| SR and SR-RELATED | tra2-beta_alt1 | -1.421 | 0 | 1.108 | 0 |
| SR and SR-RELATED | tra2-beta_alt2 | -1.453 | 0 | 1.187 | 0 |
| SR and SR-RELATED | tra2-beta_alt3 | -1.243 | 0 | 0.216 | 0 |
| SR and SR-RELATED | U2af1 | -0.222 | 0 | 0.064 | 8.014 |
| SR and SR-RELATED | U2AF2 | -0.103 | 0 | 0 | 1 |
| targets | Ache-mE2 | -0.426 | 0 | 0 | 1 |
| targets | Aqp4_alt1 | 0 | 1 | 0.012 | 0 |
| targets | Aqp4_alt2 | 0.477 | 0 | 0 | 1 |
| targets | Aqp4_com | 0 | 1 | 0 | 1 |
| targets | ARS2_alt1 | 0.073 | 0 | -0.25 | 19.662 |
| targets | ARS2_com | -0.098 | 9.723 | -0.499 | 0 |
| targets | Ars2-pending | -0.127 | 0 | 0.327 | 0 |
| targets | Ccnl_alt1 | 0.218 | 0 | 0.237 | 0 |
| targets | Ccnl_alt2 | -0.248 | 0 | 0 | 1 |
| targets | Ccnl_com | -0.382 | 0 | 0 | 1 |
| targets | Clcn3_alt1 | 0.157 | 0 | -0.154 | 0 |
| targets | Clcn3_alt2 | 0.068 | 0 | 0.081 | 0 |
| targets | Clcn3_com | -0.172 | 0 | -0.001 | 0 |
| targets | mE1a | 0.465 | 0 | 0.24 | 0 |
| targets | mE1b | 0.072 | 0 | -0.601 | 0 |
| targets | mE1c | 0.296 | 0 | 0.111 | 7.421 |
| targets | mE1d | 0.273 | 6.441 | 0.024 | 72.165 |
| targets | mE1e | 0.021 | 32.128 | 0.159 | 0 |
| targets | mE3 | -0.178 | 18.086 | 0 | 1 |
| targets | mE5 | -0.195 | 0 | 0 | 1 |
| targets | mE6 | -0.042 | 0 | -0.389 | 0 |
| targets | mI4 | 0.035 | 0 | -0.207 | 0 |
| targets | Snca_alt1 | -0.045 | 0 | -0.578 | 0 |
| targets | Snca_alt2 | 0 | 1 | 0 | 1 |
| targets | Snca_com | 0 | 1 | 0 | 1 |
| unknown function | Ayelet 5 | 0.202 | 0 | -0.057 | 0 |
| unknown function | Ayelet3 | 0.094 | 0 | 0 | 1 |
| unknown function | puf60-homolog | -0.037 | 0 | -0.099 | 0 |

## .3.2 PP1 binding site in SR protein

### SNP in aligned Tra2-beta1 protein sequences.




### PP1 binding site in SR and SR related proteins



### Tra2-beta1 PP1 binding site

```
Human                    AFVYFENVDDAKEAKERANGMELDGRRIRVDFSITKRPHTPTPGIYMGRP
Ponpy                    AFVYFENVDDAKEAKERANGMELDGRRIRVDFSITKRPHTPTPGIYMGRP
chimp                    AFVYFERIDDSKEAMERANGMELDGRRIRVDYSITKRAHTPTPGIYMGRP
Mouse                    AFVYFENVDDAKEAKERANGMELDGRRIRVDFSITKRPHTPTPGIYMGRP
Rat                      AFVYFENVDDAKEAKERANGMELDGRRIRVDFSITKRPHTPTPGIYMGRP
Chicken                  AFVYFENVEDAKEAKERANGMELDGRRIRVDFSITKRPHTPTPGIYMGRP
pig                      AFVYFENVDDAKEAKERANGMELDGRRIRVDFSITKRPHTPTPGIYMGRP
cow                      AFVYFENVDDAKEAKERANGMELDGRRIRVDFSITKRPHTPTPGIYMGRP
HoneyBee                 CFVYFESLEDAKVAKEQCAGMEIDGRRMRVDYSITQRAHTPTPGIYLGKP
Dog                      AFVYFENVDDAKEAKERANGMELDGRRIRVDFSITKRPHTPTPGIYMGRP
zebrafish                ALVYFENREDSKEAKERANGMELDGRRIRVDYSITKGPHTPTPGIYMGRP
ricefish                 AFVYFENTPDAKEAKEKANGMELDGRRIRVDFSITKRPHTPTPGIYMGRP
Fruitfly                 CFIYFEKLSDARAAKDSCSGIEVDGRRIRVDFSITQRAHTPTPGVYLGRQ
Housefly                 CFIYYKHLADAEVARDQCCGQEVDGRRIRVAYSITERPHSPTPGVYRGRS
Westernclawedfrog        SFVYFENVDDAKEAKERANGMELDGRRIRVDFSITKRPHTPTPGIYMGRP
Africanclawedfrog        SFVYFENVDDAKEAKERANGMELDGRRIRVDFSITKRPHTPTPGIYMGRP
C.elegans                GFIYFNLIEDATAARDKLCNTDLDGHKIRVDFSLTKRGHSPTPGQYMGDR
Mosquito                 GFVYFKSQAEASIARANCNGLQIHGRRIRVDYSITDQPHPPTPGVYMGRR
malariamosquito          GFVYFESAEDAKVAHDQANGIEIGDRRIRVDFSATNKPHDPTPGVYYGKV
Silkmoth                 CFVYFEDMEDAKIAKNECTGMEIDGRRIRVDYSITQRAHTPTPGIYMGKP
Ustilagomaydis_fungi_    GFITMRSIEDATQCINKLNGFTIHGRNIRVDYSATPKPHDPTPGQYLGPK
                          ::  .   ::  .     .  : .:.:** :* *   * **** * *
```

### ASF/SF2 PP1 binding site

```
Human            PRDAEDAVYGRDGYDYDGYRLRVEFPRSG------RGTGRGGGGGGGG--------GAPR
Ponpy            PRDAEDAVYGRDGYDYDGYRLRVEFPRSG------RGTGRGGGGGGGG--------GAPR
Mouse            PRDAEDAVYGRDGYDYDGYRLRVEFPRSG------RGTGRGGGGGGGG--------GAPR
Rat              PRDAEDAVYGRDGYDYDGYRLRVEFPRSG------RGTGRGGGGGGGG--------GAPR
Chicken          PRDAEDAVYGRDGYDYDGYRLRVEFPRSG------RGTGRGGGGGGGG--------GAPR
Dog              PRDAEDAVYGRDGYDYDGYRLRVEFPRSG------RGTGRGGGGGGGG--------GAPR
Pig              PRDAEDAVYGRDGYDYDGYRLRVEFPRSG------RGTGRGGGGGGGG--------GAPR
Cow              PRDAEDAVYGRDGYDYDGYRLRVEFPRSG------RGTGRGGGGGGGG--------GGFP
Sheep            PRDAEDAVYGRDGYDYDGYRLRVEFPRSG------RGTGRGGGGGGGG--------GAPR
Horse            PRDAEDAVYGRDGYDYDGYRLRVEFPRSG------RGTGRGGGGGGGG--------GAPR
Zebrafish        PRDAEDAVYGRDGYDYDGYRLRVEFPRSG------RGGGRGGGGGGGV--------GAPR
Bonyfish         PRDADDAVYGRDGYDYDGYRLRVEFPRSG------RGS-RGGFGGIG---------GAPR
Fruitfly         ARDADDAVKARDGYDYDGYRLRVEFPRGGGPGS-YRGNR--NDRNNGR--------DGGR
Honeybee         PRDAEDAVHARDGYDYDGYRLRVEFPRGGGPSNNFRGGRGAGDSGRGG--------RGEM
Mosquito         ARDADDAVKARDGYDYDGYRLRVEFPRGGGPGS-YRGSRQGNSDRNSR--------GGDR
Arabidopsis      ARDADDAIYGRDGYDFDGHHLRVELAHGG------RRSSHDARGSYSGRGRGGRGGGDGG
Maize            PRDAEEAIAGRDGYNFDGHRLRVEAAHGG------RGNASSHDRSSGF--------GGGG
Barley           PRDAEDAIQGRDGYNFDGNRLRVELAHGG------RANSSSLPNSH-----------GGG
Rice             PRDAEDAIRGRDGYNFDGNRLRVELAHGG------RGNSSSFNNS-----------GGG
                 .***::*: .****::** :**** .:.*       *
```

### P54 PP1 binding site

```
human            ELRLFP-PDDSPLPVSSRVCFVKFHDPDSAVVAQHLTNTVFVDRALIVVPYAEG--VIPD
chimp            ELRLFP-PDDSPLPVSSRVCFVKFHDPDSAVVAQHLTNTVFVDRALIVVPYAEG--VIPD
rat              ELRLFP-PDDSPLPVSSRVCFVKFHDPDSAVVAQHLTNTVFVDRALIVVPYAEG--VIPD
mouse            ELRLFP-PDDSPLPVSSRVCFVKFHDPDSAVVAQHLTNTVFVDRALIVVPYAEG--VIPD
frog             ELRLFP-PDDSPLPVTSRVCFVKFQDPDSAVVAQHLTNTVFVDRALIVVPYAEG--IIPD
bonyfish         ELRLFP-PDDSPLPVTSRVCFVKFHEPESVGVSQHLTNTVFVDRALIVVPFAEG--VIPD
cow              ---------DSPLPVSSRVCFVKFHDPDSAVVAQHLTNTVFVDRALIVVPYAEG--VIPD
pig              ELRLFP-PDDSPLPVSSRVCFVKFHDPDSAVVAQHLTNTVFVDRALIVVPYAEG--VIPD
honeybee         DIRLYPTIRDVAVPVQSRICYIKFHDQGCVAVAQHMTNTVFIDRALIVIPYQNG--DIPD
mosquito         EIRLYPTIRDVSCPVVSRICYVKYFESSCVAVAQHLTNTVFIDRAVIVIPVANG--VIPD
fly              EIRLYPTIRDVSCPVQSRICYVKYTDTTSVPVAQHLTNTVFIDRALIVIPVL----AIPE
Bloodfluke       EVVVYP-SDDK-EELASKVCYIRYQEPINAEVALHLNNTVFLDRALIVLPLSGDRDAIPD
```

```
Cbriggsae       ELKVYPSEGNINASTLLKTAFIKFDDERCAEVGQHLTNTVLIDRAIVCLPYPNQ--IIPD
Celegans        DLKVYPSEGNITANTLLKTAFIKFDDERCVEVAQHLTNTVVIDCAIVCLPYPNP--VIPD
                         :       : .:::: :   . *. *:.***.:* *:: :*       **:
```

**SRp30c PP1 binding site**

```
human            ----DAEDAIYGRNGYDYGQCRLRVEFPRT----YGGRGGWPRGG--------RNGPPTR
panpy            ERGIDAEDAIYGRNGYDYGQCRLRVEFPRT----YGGRGGWPRGG--------RNGPPTR
mouse            ----DAEDAIYGRNGYDYGQCRLRVEFPRT----YGGRGGWPRGA--------RNGPPTR
pig              ----DAEDAIYGRNGYDYGQCRLRVEFPRT----YGGRGGWPRGG--------RNGPPTR
dog              ----DAEDAIYGRNGYDYGQCRLRVEFPRT----YGSRGGWPRGG--------RNGPPTR
Rat              ----DAEDAIYGRNGYDYGQCRLRVEFPRA----YGGRGGWPRAS--------RNGPPTR
cow              ----DAEDAVYGRNGYDYGQCRLRVEFPRT----YGGRGGWPRGG--------RSGPPTR
frog             ----DAEDAVFGRNGYDFGSCRLRVEFPRS----FRGSGGGGGGGGGYGGSRGRNGPPSR
chicken          ----DAEDAVYGRDGYDYDGYRLRVEFPRSGRGTGRGGGGGGGGG----APRGRYGPPSR
Zebrafish        ----DAEDAVFGRNGYGFGDCKLRVEYPRSSGSKFSGPAGGGGGG-----PRGRFGPPTR
Arabidopsis      ----DADDAIYGRDGYDFDGCRLRVEIAHG------GRRFSPSVDRYSSSYSASR-APSR
Rice             ----DADDAICGRDGYNFDGYRLRVELAHG------GRGQSYSYDRPRSYSSGRRGGVSR
                     **:**: **:**.:.  :**** .:       .                .  :*
```

### .3.3 splicing factors PP1 binding site alignment

# Pre-mRNA-processing factor 39

**RITF**
**KVEF**

# Small nuclear ribonucleoprotein polypeptide A
**U1-A**

### PP1 binding sites: {R/K}-X(0-1)-{V/I}-X-{F/W}

**KISF**

```
Homo_sapien                    FKEVRLVPGRHDIAFVEFDNEVQAGAARDALQGFKITQNNAMKISFAKK-
Pan_troglodytes                FKEVRLVPGRHDIAFVEFDNEVQAGAARDALQGFKITQNNAMKISFAKK-
Monodelphis_domestica          FKEVRLVPGRHDIAFVEFDNEVQAGAARDALQGFKITQNNAMKISFAKK-
Macaca_mulatta                 FKEVRLVPGRHDIAFVEFDNEVQAGAARDALQGFKITQNNAMKISFAKK-
Mus_musculus                   FKEVRLVPGRHDIAFVEFDNEVQAGAARDALQGFKITQNNAMKISFAKK-
Rattus_norvegicus              FKEVRLVPGRHDIAFVEFDNEVQAGAARDALQGFKITQNNAMKISFAKK-
Bos_taurus                     FKEVRLVPGRHDIAFVEFDNEVQAGAARDALQGFKITQNNAMKISFAKK-
Sus_scrofa                     FKEVRLVPGRHDIAFVEFDNEVQAGAARDALQGFKITQNNAMKISFAKK-
Oryctolagus_cuniculus          FKEVRLVPGRHDIAFVEFDNEVQAGAARDSLQGFKITQNNAMKISFAKK-
Canis_familiaris               FKEVRLVPGRHDIAFVEFDNEVQAGAARDALQGFKITQNNAMKISFAKK-
Xenopus_tropicalis             FKEVRLVPGRHDIAFVEFDNEVQAGAARESLQGFKITQSNSMKISFAKK-
Xenopus_laevis                 FKEVRLVPGRHDIAFVEFDNEVQAGAARESLQGFKITQSNSMKISFAKK-
Gallus_gallus                  FKEVRLVPGRHDIAFVEFENECQAGAARDALQGFKITPSHAMKITYAKK-
Danio_rerio                    FKEVRLVPGRHDIAFVEFDNEVQAGAARDALQGFKITQSNAMKISFAKK-
Tetraodon_nigroviridis         FKEVRLVPGRHDIAFVEFDNEVQAAAARDALQGFKITQTNAMKISFAKK-
Aedes_aegypti                  FKEVRLVPNRHDIAFVEFATELQSGAARDALQGFKITPTHAMKISFAKK-
Drosophila_melanogaster        FKEVRLVPNRHDIAFVEFTTELQSNAAKDALQGFKITPTHAMKITFAKK-
Drosophila_pseudoobscura       FKEVRLVPNRHDIAFVEFTTELQSNAAKDALQGFKITPTHAMKITFAKK-
Tribolium_castaneum            FKEVRLVPNRHDIAFVEFENELQSGAAKDALQGFKITPTHAMKISFAKK-
Anopheles_gambiae              FKEVRLVPNRHDIAFVEFATELQSGAARDALQGFKITPTHAMKISFAKK-
Caenorhabditis_elegans         LKDIRMVPNRPGIAFVEFDTDSLAIPARTTLNNFKISAEHTMRVDYAKK-
Caenorhabditis_briggsae        LKDIRMVPNRPGIAFVEFDTDSLAIPARTTLNNFRISAEHVMRVDYAKK-
Arabidopsis_thaliana           FKEVRMIEAKPGIAFVEFADEMQSTVAMQGLQGFKIQ-QNQMLITYAKK-
Oryza_sativaJaponica           FREVRMIEAKPGIAFVEYEDDSQSMVAMQALQGFKITPYNPMAISYAKK-
Oryza_sativaIndica             FREVRMIEAKPGIAFVEYEDDSQSMVAMQALQGFKITPYNPMAISYAKK-
Solanum                        FREVRMIEAKPGIAFVEFDDDVQSSVAMQALQGFKITPQNPMAITYAKK-
Coprinopsis_cinereaokayama7    LYEVRLIPTKKDIAFVEFLDEASSGVAKDALHNFKIDGENKIKVP--LLW
Dictyostelium_discoideum       FKEVHMVESKKGIAFIEFEDEIKSGFAMTNLQHFKVTPEKPMVVSFAAQ-
Neurospora_crassa              FREVRTVPGRSGIAFVEYDAEAGAITAKENTAGMALKNGEKIMKVTYQRQ
Aspergillus                    FREVRLVPGRKGIAFVEYENESGAISAKEATSGMPMGDEGKPIRVTYQRQ
Bigelowiella_natans            FKEVRLVDGKPDIAFIEFNDAQESALAKEGLQNFKITSQNAMKLTFAKQ-
Chaetomium_globosum            FREVRTVPGRSGIAFVEYEAEAGAITAKENTAGMPLKNGEKMMKVTYQRQ
Coccidioides_immitisRS         FKEVRMVPGRKGIAFVEYENETGAISAKEATSGMALGENGKPMRVTYQRQ
Cryptococcus                   LIEIRTIPAKKDIAFVEFADEGAATIAKDALHNFKIDGETKMKVS--VLL
Ustilago_maydis521             YVDVQTIPGKAEIAFVEFADIPSSATARGALNGYNFGAGDKLKASDTLFW
Magnaporthe_grisea7015         FREVRLVPGRRGIAFVEYDGEAGAIAAKENTAGMPLGAEGKPVKVTYQRQ
Neosartorya_fischeriNRRL181    FQEVRLVPGRKGIAFVEYENESGAISAKEATANMPMGDNGKPIRVTYQRQ
Plasmodium                     FVEARIIPQR-NVAFVDFTDETTATFAMKAVQNYELQ-GSKLKISYAKRY
Schizosaccharomyces_pombe97    FQEVRMVPGRRGIAFVEYDSDREATVAKNGTTGMSLSGNQIKVTFARKAS
Tetrahymena_thermophilaSB210   FKEVRLIAPR-KVAFVEFSQEDEATVALNGLQNFQLTPQVFLKLNYAKF-
Trichomonas_vaginalis          FVEVRTLPGKQTIAFVEYKTEEQSAVAIQELNGFEIENHHLT-IQFSK--
                                 : : :  :  :**:::     :  *         .
```

## Small nuclear ribonucleoprotein Sm D1

### PP1 binding sites: {R/K}-X(0-1)-{V/I}-X-{F/W}

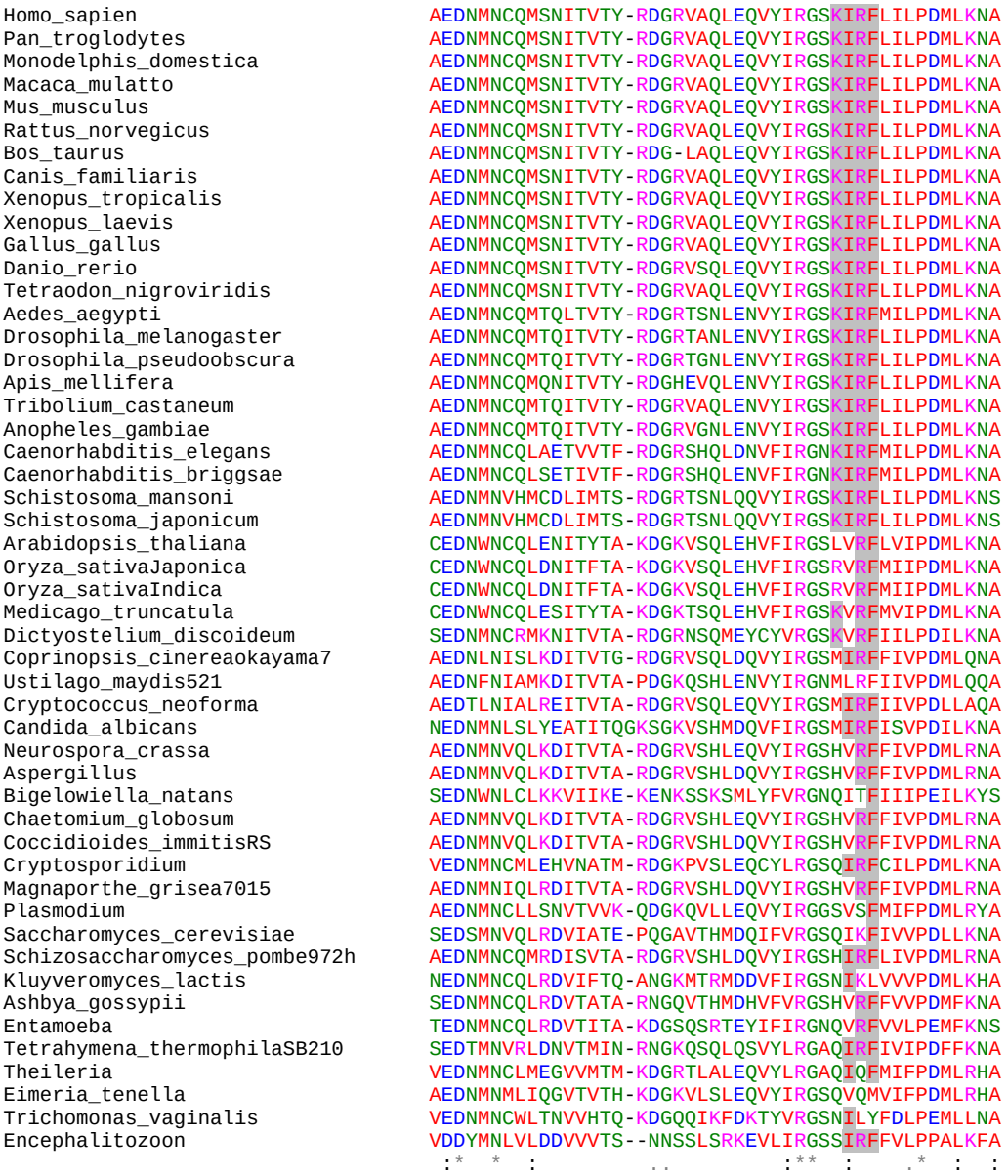
**KLVRF**

```
Homo_sapien                          MKLVRFLMKLSHETVTIELKNGTQVHGTITGVDVSMNTHLKAVK-MTLKN
Pan_troglodytes                      MKLVRFLMKLSHETVTIELKNGTQVHGTITGVDVSMNTHLKAVK-MTLKN
Monodelphis_domestica                MKLVRFLMKLSHETVTIELKNGTQVHGTITGVDVSMNTHLKAVK-MTLKN
Macaca_mulatto                       MKLVRFLMKLSHETVTIELKNGTQVHGTITGVDVSMNTHLKAVK-MTLKN
Mus_musculus                         MKLVRFLMKLSHETVTIELKNGTQVHGTITGVDVSMNTHLKAVK-MTLKN
Rattus_norvegicus                    MKLVRFLMKLSHETVTIELKNGTQVHGTITGVDVSMNTHLKAVK-MTLKN
Bos_taurus                           MKLVRFLMKLSHETVTIELKNGTQVHGTITGVDVSMNTHLKAVK-MTLKN
Canis_familiaris                     MKLVRFLMKLSHETVTIELKNGTQVHGTITGVDVSMNTHLKAVK-MTLKN
Xenopus_tropicalis                   MKLVRFLMKLSHETVTVELKNGTQVHGTITGVDVSMNTHLKAVK-MTVKN
Danio_rerio                          MKLVRFLMKLSHETVTIELKNGTQVHGTITGVDVSMNTHLKAVK-MTLKN
Strongylocentrotus_purpuratus        MKLVRFLMKLSHETVTIELKNGTQVHGTITGVDVSMNTHLKSVK-LTPKN
Aedes_aegypti                        MKLVRFLMKLSHETVTIELKNGTQVHGTITGVDVAMNTHLKAVK-MTIKN
Drosophila_melanogaster              MKLVRFLMKLSHETVTIELKNGTQIHGTITGVDVAMNTHLKSVR-MTIKN
Apis_mellifera                       MKLVRFLMKLSHETVTIELKNGTQVHGTITGVDVAMNTHLKTVK-MTIKN
Bombyx_mori                          MKLVRFLMKLSHETVTIELKNGSVVHGTITGVDVAMNTHLKAVK-VTLKN
Caenorhabditis_elegans               MKLVRFLMKLSHETVNIELKNGTQVSGTIMGVDVAMNTHLRAVS-MTVKN
Caenorhabditis_briggsae              MKLVRFLMKLSHETVNIELKNGTQVSGTIMGVDVAMNTHLRAVS-MTVKN
Schistosoma_japonicum                MKLVRFLMKLSHETVTIELKNGTQVHGSIAGVDVSMNTHMRSVT-LTLKN
Gibberella_zeae                      MKLVRFLMKCANETVTIELKNGTIIHGTISSVSPQMNTALRNVK-MTIKG
Tribolium_castaneum                  MKLVRFLMKLSHETVTMELKNGTQVHGTITGVDVAMNTHLKAVK-VTVKN
Ustilago_maydis521                   MKLVRFLMKLNNESVTIELKNGTVVHGTVTGVDIQMNTHLKTVK-MTVRG
Arabidopsis_thaliana                 MKLVRFLMKLNNETVSIELKNGTIVHGTITGVDVSMNTHLKAVK-LTLKG
Oryza_sativaJaponica                 MKLVRFLMKLNNETVTIELKNGTVVHGTITGVDISMNTHLKTVK-LTLKG
Oryza_sativaIndica                   MKLVRFLMKLNNETVTIELKNGTTVHGTITGVDISMNTHLKTVK-LTLKG
Medicago_truncatula                  MKLVRFLMKLNNETVSIELKNGTIVHGTITGVDISMNTHLKTVK-LTLKG
Brassica_napus                       MKLVRFLMKLNNETVSIELKNGTVVHGTITGVDVSMNTHLKTVK-MTLKG
Coprinopsis_cinereaokayama7          MKLVRFLMKLNNETVTIELKNGAVVHGTITGVDMQMNTFLKTVK-MTMRN
Dictyostelium_discoideum             MKLVRFLMKLHNETVTIELKNGTIVQGSVAGVDVSMNTHLKTVK-LTLKG
Neurospora                           MKLVRFLMKCANETVTIELKNGTIVHGTIASVTPRMDTALRNVK-MTPKG
Aspergillus_terreusNIH2624           MKLVRFLMKCANETVTIELKNGTILHGTITSVSPQMNTALRTVK-MTPKG
Aspergillus_oryzae                   MKLVRFLMKCANETVTIELKNGTILHGTITSVSPQMNTSLRTVK-MTPKG
Aspergillus_fumigatusAf293           MKLVRFLMKCANETVTIELKNGTILHGTITAVSPQMNTSLRTVK-MTPKG
Aspergillus_clavatusNRRL             MKLVRFLMKCANETVTIELKNGTILHGTIISVSPQMNTALRTVK-MTPKG
Aspergillus_nidulansFGSC             MKLVRFLMKCANETVTIELKNGTILHGTITSVSPQMNTSLRTVK-MTPKG
Bigelowiella_natans                  MKLVKFLMRLNNETVTVELKNGTVVQGTISGVDMSMNTHLKIVK-MTLKG
Chaetomium_globosum                  MKLVRFLMKCANETVTIELKNGTIVHGTIASVSPQMNTALRNVK-MTPRG
Coccidioides_immitisRS               MKLVRFLMKCANETVTIELKNGTILHGTIASVSPQMNTALRTVK-MTPKG
Cryptosporidium_parvumIowa           MKLIRFLMKLVNNSVVIELKNGTIIQGTIVTVDMSMNTYLKNVK-MSVKH
Cryptosporidium_hominisTU502         MKLIRFLMKLINNSVVIELKNGTIIQGTIVSVDMSMNTYLKNVK-MSVKH
Magnaporthe_grisea7015               MKLVRFLMKCANETVTIELKNGTIVHGTIASVSPQMNTALRNVK-MTPRG
Karlodinium_micrum                   MRLVRFLMKLSNESVVVELKNSTVVAGTITGVDVSMNTHMKNVK-FTVKG
Neosartorya_fischeriNRRL181          MKLVRFLMKCANETVTIELKNGTILHGTITAVSPQMNTSLRTVK-MTPKG
Plasmodium_bergheiStrainANKA         MKLVTFLMKLTNENVTIELKNGTLISGVITGVDIKMNTHMKNVK-VVIKN
Plasmodium_yoeliiYoeliiStr17XN       MKLVTFLMKLTNENVTIELKNGTLISGVITGVDIKMNTHMKNVK-VVIKN
Plasmodium_chabaudiChabaudi          MKLVTFLMKLTNENVTIELKNGTLISGVITGVDIKMNTHMKNVK-VVIKN
Plasmodium_falciparum3D7             MKLVHFLMKLTNENVTIELKNGTLITGIITAVDIKMNTHMKNVK-VVIKN
Paramecium_tetraurelia               MKLVRFLMKLKNEQVIVELKNGTVVLGTITGVDVRMNTHLSKVK-LTLKG
Saccharomyces_cerevisiae             MKLVNFLKKLRNEQVTIELKNGTTVWGTLQSVSPQMNAILTDVK-LTLPQ
Schizosaccharomyces_pombe972h        MKLVRFLMKLTNETVSIELKNGTIVHGTITSVDMQMNTHLKAVK-MTVKG
Tetrahymena_thermophilaSB210         MKLVRFLMKLKGQTVTIELKNGTQIIGSIIAVDIRMNTHLSKVK-MTIKG
Theileria_annulataStrainAnkara       MKLVRFLMKLANESVTVELKNGTVLTGTVIGIDISMNTHLKNVK-VVNKG
Theileria_parvaStrainMuguga          MKLVRFLMKLANEELTLELKNGTVITGVVIGIDISMNTHLKSVK-VVTKG
Trichomonas_vaginalis                MKLVHFLRKLVRETVTVELKDNTVIKGTVVGVDSAMNTHLRLVH-IKAPG
nucleomorph_GuillardiaTheta          MKIYNFIKQLKGEEIIIELKNNIIIMGVIVNVDKNMNLKVSNISDIDIDG
                                     *::  *: :   : : :***:.  : * :  :   *:  :  :  .
```

# Small nuclear ribonucleoprotein Sm D3

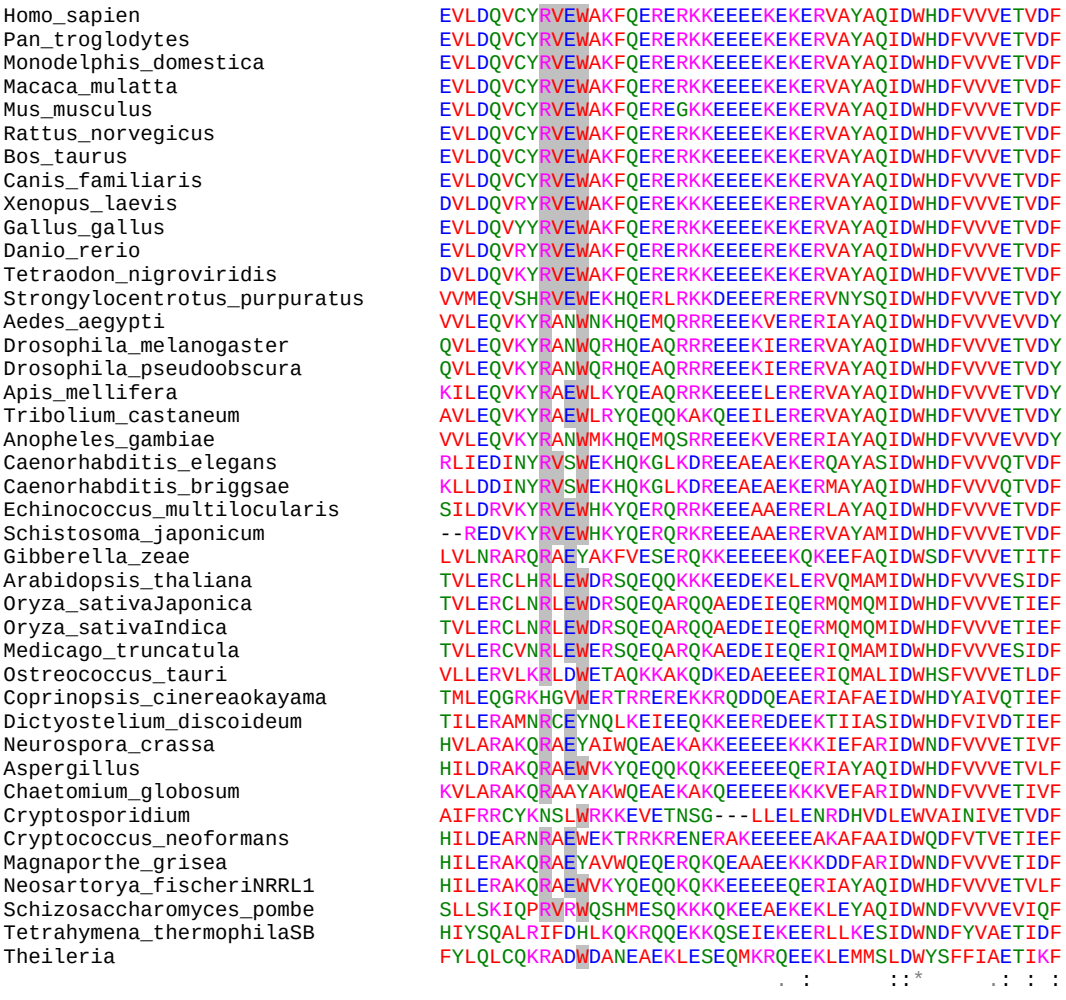
### PP1 binding sites: {R/K}-X(0-1)-{V/I}-X-{F/W}

**KIRF**



# Splicing factor 3 subunit 1

**Splicing factor 3 subunit 1**
**SAP 114**
**SF3a120**
**RVEW**
**KVTW**



# Splicing factor 3 subunit 1

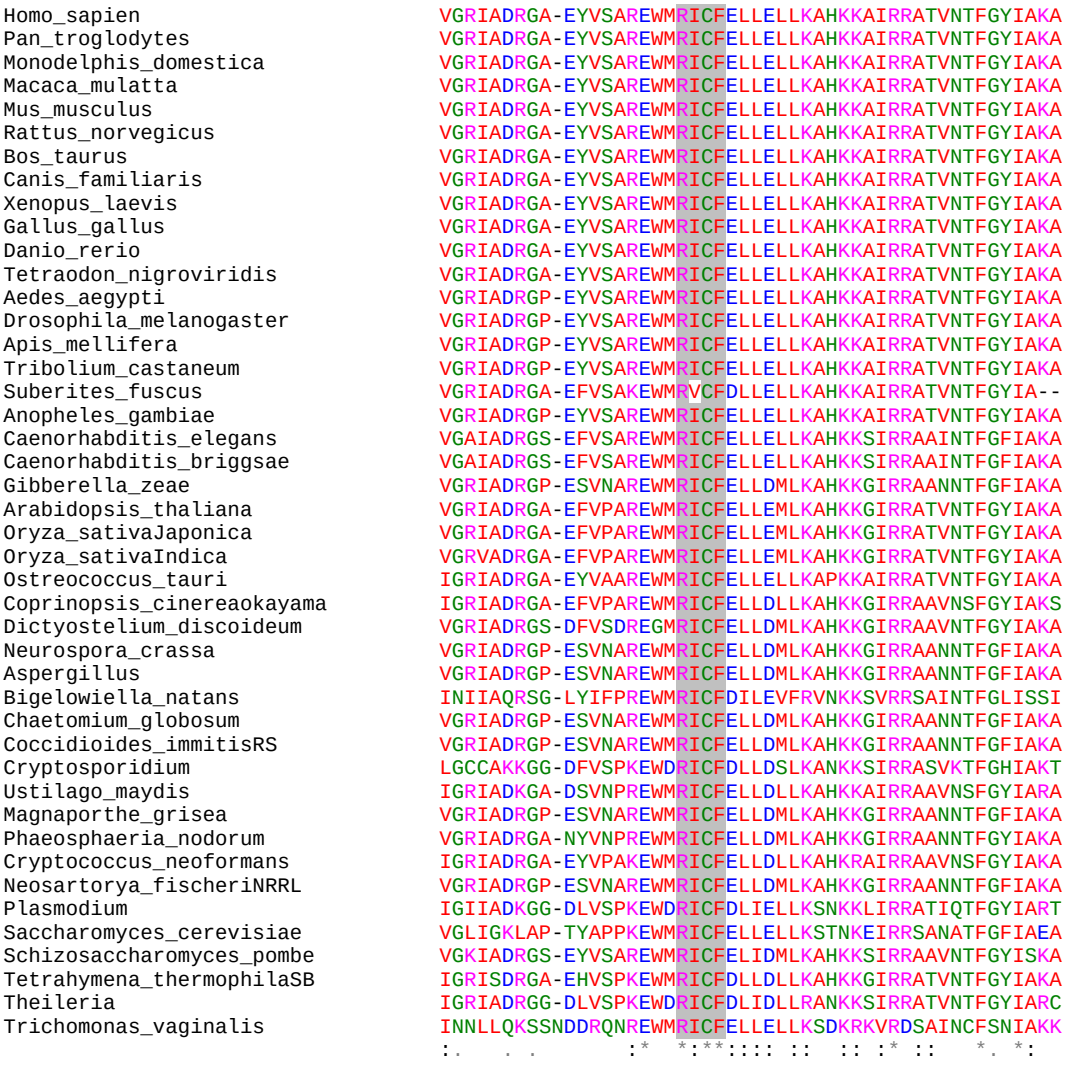
**Splicing factor 3 subunit 1**
**SAP 114**
**SF3a120**
**RVEW**
**KVTW**

```
Homo_sapien                          RTDIFGV--EETAIGKKIGEE--EIQKPEEKVTWDG---HSGSMARTQQA
Pan_troglodytes                      RTDIFGV--EETAIGKKIGEE--EIQKPEEKVTWDG---HSGSMARTQQA
Monodelphis_domestica                RTDIFGV--EETAIGKKIGEE--EIQKPEEKVTWDG---HSGSMARTQQA
Macaca_mulatta                       RTDIFGV--EETAIGKKIGEE--EIQKPEEKVTWDG---HSGSMARTQQA
Mus_musculus                         RTDIFGV--EETAIGKKIGEE--EIQKPEEKVTWDG---HSGSMARTQQA
Rattus_norvegicus                    RTDIFGV--EETAIGKKIGEE--EIQKPEEKVTWDG---HSGSMARTQQA
Bos_taurus                           RTDIFGV--EETAIGKKIGEE--EIQKPEEKVTWDG---HSGSMARTQQA
Canis_familiaris                     RTDIFGV--EETAIGKKIGEE--EIQKPEEKVTWDG---HSGSMARTQQA
Xenopus_laevis                       RTDIFGV--EETAIGKKIGEE--EIQKPEEKVTWDG---HSGSMARTQQA
Gallus_gallus                        RTDIFGV--EETAIGKKIGEE--EIQKPEEKVTWDG---HSGSMARTQQA
Danio_rerio                          RTDIFGV--EETAIGKKIGEE--EIQKPEEKVTWDG---HSGSMARTQQA
Aedes_aegypti                        RTDIFGVGDEEAAIGKKLGEE--ETKK-DDRVTWDG---HTSSVEAATRA
Drosophila_melanogaster              RTDIFGVGDEETVIGKKLGEE--ETKK-DDRVTWDG---HTSSVEAATRA
Drosophila_pseudoobscura             RTDIFGVGDEETVIGKKLGEE--ETKK-DDRVTWDG---HTSSVEAATRA
Apis_mellifera                       RTDIFGVGDEETAIGKKIGEE--DKKK-DDKVTWDG---HTSSVEAATRA
Tribolium_castaneum                  RTDIFGVGDEETAIGKKIGEE--DVRK-DEKVTWDG---HTSSVEAATRA
Anopheles_gambiae                    RTDIFGVGDEEAAIGKKLGEE--EPRK-DDRVTWDG---HTSSVEAATRA
Caenorhabditis_elegans               RTDIFGVGGEQTMIGKKLGEED-NSQQGQNKLIWDG---TEETRDMITRA
Caenorhabditis_briggsae              RTDIFGVGGEQTMIGKKLGEEE-GGQQGQNKLIWDG---TEESRDMITRA
Echinococcus_multilocularis          RSDIFGVGSEETQIGKTPEEAAGGKQTKPDKLIWDG---HAASAEVVAKR
Gibberella_zeae                      RSDVFDPV----TGQAISEDELARRKKAAIHSYDGA----MDAKSQAQLG
Arabidopsis_thaliana                 RPDIFGTTEEEVSNAVKAEIEKKKDEQPKQVIWDGHTG---SIGRTANQA
Oryza_sativaJaponica                 RPDIFGTTEEEVSNAVKAEIEKKKDEQPKQVIWDGHSG---SIGRTATQA
Oryza_sativaIndica                   RPDIFGTTEEEVSNAVKAEIEKKKDEQPKQVIWDGHSG---SIGRTATQA
Medicago_truncatula                  RPDIFGTTEEEVSNAVKAEIEKKNDEQPKQVIWDGHSG---SIGRTANQA
Coprinopsis_cinereaokayama7          RVDIFGTETDEERRKREEEEERLRRRERERKVVWDG----HTASKANTLDK
Dictyostelium_discoideum             RVDIFGETESSKKQDEQPTQAPKVIWDGHSGSIPRVQAAQQAAQLAAQQA
Neurospora_crassa                    RTDVFDAV----TGQPISEEEQARRKKVAMHSYDGN----PEGRSQAHIN
Aspergillus                          RSDVFDSTV---LPETGDPEEEARKKRMAVEGAPGQGPIPPMVGPAGAPA
Chaetomium_globosum                  RSDVFDTT----TGQPISEEELARRKKVALHAFDGN----PDGKSQAHIN
Cryptococcus_neoformans              RTDIFGDDVDEAERKRREEEERQKRRERERKIVWDG----HTASAAKTAET
Magnaporthe_grisea7015               RTDVFDSV----TGQPLSEEEMARRKKAAINSYDGN----PDGKSQAHIA
Neosartorya_fischeriNRRL181          RSDVFDSS----LTAGLDPEEEARKKRMAYENPSGAGPTPPMVGPAGGPP
Schizosaccharomyces_pombe972h        RTDLFDVQ----NGVEISQEEIERRKRAATQSAWGAT---PTNKRR----
Tetrahymena_thermophilaSB210         RPEIFGYAEEQFEKIVENEPSGPSKPIWDGQSATMTR----TTATVAMLA
Theileria                            RPDLFGSADEEVNDHKESNHYSKGNMKYVACDIDTKRKKL----------
```

# Splicing factor 3B subunit 1

**Pre-mRNA-splicing factor SF3b 155 kDa subunit**

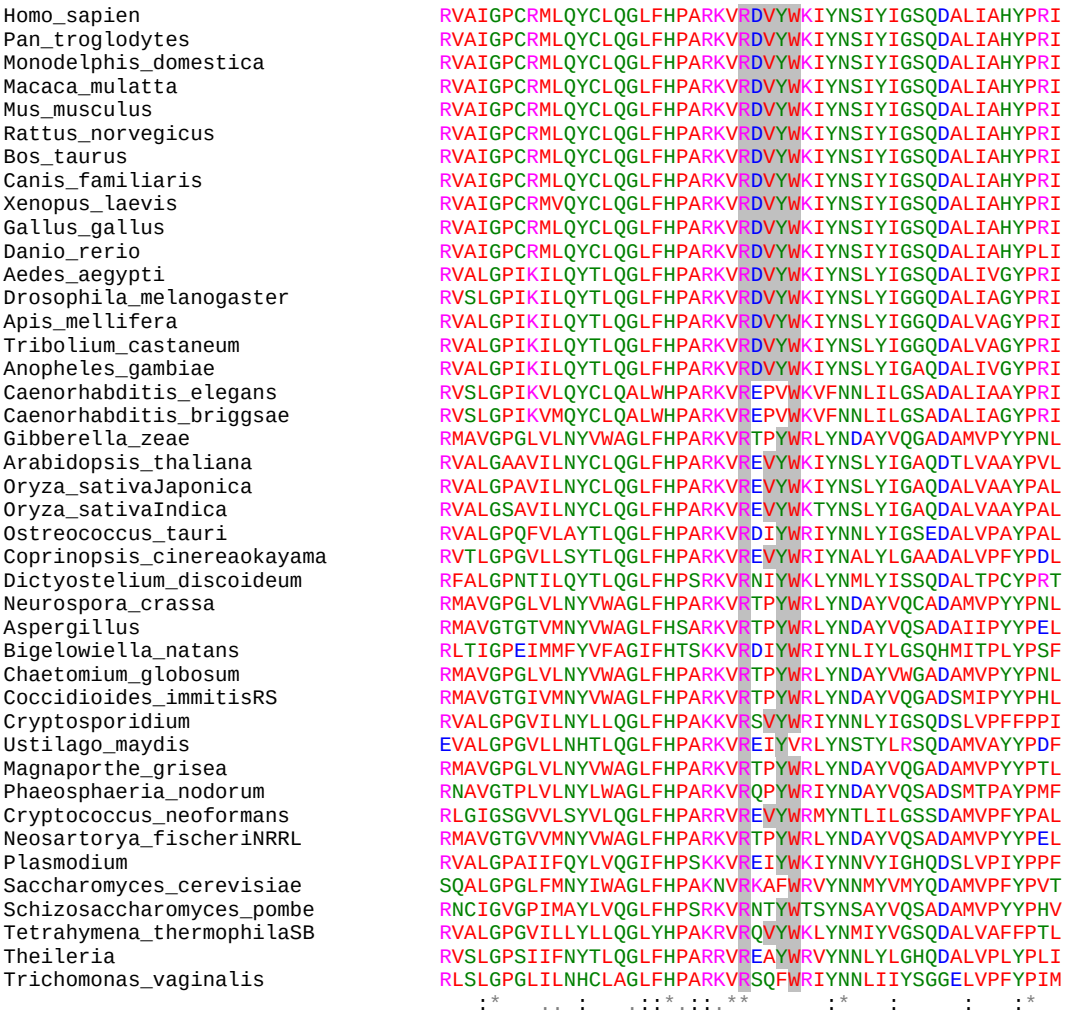
**RICF**
**RDVYW**



# Splicing factor 3B subunit 1

**Pre-mRNA-splicing factor SF3b 155 kDa subunit**

**RICF**
**RDVYW**



# U2-associated SR140 protein

**140 kDa Ser/Arg-rich domain protein**
**U2-associated protein SR140**

**KIMW**
**RMIEF**

```
Homo_sapien                    LGNINPQMNEEMLCQEFGRFGPLASVKIMWPRTDEERARERN-CGFVAFM
Pongo_pygmaeus                 LGNINPQMNEEMLCQEFGRFGPLASVKIMWPRTDEERARERN-CGFVAFM
Pan_troglodytes                LGNINPQMNEEMLCQEFGRFGPLASVKIMWPRTDEERARERN-CGFVAFM
Monodelphis_domestica          LGNINPQMNEEMLCQEFGRFGPLASVKIMWPRTDEERARERN-CGFVAFM
Macaca_mulatta                 LGNINPQMNEEMLCQEFGRFGPLASVKIMWPRTDEERARERN-CGFVAFM
Mus_musculus                   LGNINPQMNEEMLCQEFGRFGPLASVKIMWPRTDEERARERN-CGFVAFM
Rattus_norvegicus              LGNINPQMNEEMLCQEFGRFGPLASVKIMWPRTDEERARERN-CGFVAFM
Bos_taurus                     LGNINPQMNEEMLCQEFGRFGPLASVKIMWPRTDEERARERN-CGFVAFM
Canis_familiaris               LGNINPQMNEEMLCQEFGRFGPLASVKIMWPRTDEERARERN-CGFVAFM
Gallus_gallus                  LGNINPQMNEEMLCQEFGRFGPLASVKIMWPRTDEERARERN-CGFVAFM
Danio_rerio                    LGNINPQMNEEMLCQEFGRYGPLASVKIMWPRTDEERARERN-CGFVAFM
Tetraodon_nigroviridis         LGNINPQMNEEMLCQEFGRYGPLASVKIMWPRTDEERARERN-CGFVAFM
Strongylocentrotus_purpuratus  LGSINPKMNEEMLCKEFVQYGPLASVKIMWPRTDAERQRNRN-CGFVAFM
Aedes_aegypti                  LGNLNPKISEQELMELFGKYGPLASIKIMWPRSEEERARGRN-CGFVAYM
Drosophila_melanogaster        LGNLNPKISEQQLMEIFGRYGPLASIKIMWPRSEEEKQRGRN-CGFVAYM
Drosophila_pseudoobscura       LGNLNPKISEQQLMETFGRYGPLASIKIMWPRSDEEKARGRN-CGFVAYM
Anopheles_gambiae              LGNLNPKISEQALMELFGKYGPLASIKIMWPRSEEEKMRNRN-CGFVAYM
Caenorhabditis_elegans         VSNIPHSVTEDDLLFTFGSFGPLAALKILYPRSEEERRRPHI-CAFVAFM
Caenorhabditis_briggsae        VSNIPHSVTEQDLLFTFGSFGPLAALKILYPRSEEERRRPHI-CAFVAFM
Arabidopsis_thaliana           VGNLSPKVDENFLLRTFGRFGPIASVKIMWPRTDEEKRRQRN-CGFVSFM
Oryza_sativaJaponica           VGNLSPKVDENFLMRTFGRFGPIASVKIMWPRTEEERRRQRN-CGFVAFM
Oryza_sativaIndica             VGNLSPKVDENFLMRTFGRFGPIASVKIMWPRTEEERRRQRN-CGFVAFM
Dictyostelium_discoideum       LGTLSVESNEQVIDELFSTYGIINSIKIITPKNEDDRKRGINYCAIVTYQ
Ustilago_maydis521             ILSLPPNVDERSMGEFFAAWGDVATVKIMWPR-GEQRERLAGLTGFVAFM
Plasmodium                     LGNLSAEVTEEYLCQRFGKFGKVSSVKIMYPRKEEDKKKGRI-SGFVCFE
Theileria                      IGNLSPNVTEDILMSHFSKFGTIVGIRLIPSRTDTP-PDNKQ-TGFVSFM
```

# U2-associated SR140 protein

**140 kDa Ser/Arg-rich domain protein**
**U2-associated protein SR140**

**KIMW**
**RMIEF**

```
Homo_sapien                      VVIPTERNLLALIHRMIEFVVREGPMFEAMIMNREINNPMFRFLFENQTP
Pongo_pygmaeus                   VVIPTERNLLALIHRMIEFVVREGPMFEAMIMNREINNPMFRFLFENQTP
Pan_troglodytes                  VVIPTERNLLALIHRMIEFVVREGPMFEAMIMNREINNPMFRFLFENQTP
Monodelphis_domestica            VVIPTERNLLALIHRMIEFVVREGPMFEAMIMNREINNPMFRFLFENQTP
Macaca_mulatta                   VVIPTERNLLALIHRMIEFVVREGPMFEAMIMNREINNPMFRFLFENQTP
Mus_musculus                     VVIPTERNLLALIHRMIEFVVREGPMFEAMIMNREINNPMFRFLFENQTP
Rattus_norvegicus                VVIPTERNLLALIHRMIEFVVREGPMFEAMIMNREINNPMFRFLFENQTP
Bos_taurus                       VVIPTERNLLALIHRMIEFVVREGPMFEAMIMNREINNPMFRFLFENQTP
Canis_familiaris                 VVIPTERNLLALIHRMIEFVVREGPMFEAMIMNREINNPMFRFLFENQTP
Gallus_gallus                    VVIPTERNLLALIHRMIEFVVREGPMFEAMIMNREINNPMFRFLFENQTP
Danio_rerio                      VVIPTERNLLSLIHRMIEFVVREGPMFEAMIMNREINNPLYRFLFENQSP
Tetraodon_nigroviridis           VVIPTERNLLSLIHRMIEFVVREGPMFEAMIMNREINNPMYRFLFENQSP
Strongylocentrotus_purpuratus    VVIPTERPLLQVIHRMIEFVVNEGPMFEAMVMNREINNPMFRFLFDNQTP
Aedes_aegypti                    VVLPTERHLLMLIHRMVEFVIREGPMFEALIMTREIDNPMYKFLFENESP
Drosophila_melanogaster          VHIPTEKAVLNVIHRMIEFVIREGPMFEALIMIREMENPLFAFLFDNESP
Drosophila_pseudoobscura         VFIPTEKAVLNVIHRMIEFVIREGPLFEASVMIREMENSLFSFLFDNESP
Apis_mellifera                   VVIPTERNLVMLIHRMVEFVIREGPMFEAMIMNRELNNPMFRFLFENYSP
Anopheles_gambiae                VVIPTERPLLMLIHRMVEFVIREGPMFEALIMTKEMDNPMYKFLFENESP
Caenorhabditis_elegans           VVVPPDRKLVRVMDRMAVYVVTEGPQFEAMICAEEFQNPMFQFLWDNTSA
Caenorhabditis_briggsae          VVIPPDRQLVRIMDRMAVYVCTEGPQFEAMICADEYQNPMFQFLWDNTSA
Arabidopsis_thaliana             VVTPEDEHLRHVIDTLALYVLDGECAFEQAIMERGRGNPLFKFMFELGSK
Oryza_sativaJaponica             VAPPDDAHVRHVIDTMALHVLDGGCAFEQAVMERGRGNSLFSFLFDLKSK
Oryza_sativaIndica               VAPPDDAHVRHVIDTMALHVLDGGCAFEQAVMERGRGNSLFSFLFDLKSK
Coprinopsis_cinereaokayama7      KVLDDDDVTDTFIRAVAAEVKGHGSKYEQMLKEREKSNPKYKFMLERTHR
Dictyostelium_discoideum         VQIPQNLFIKSIIDNLAYYVSKEGYPFEKLIQEREYSNMNFQFLFDHQSD
Ustilago_maydis521               VHNQYPEMQRQLIETVASRIRSNGAHFEHILREREAENAQFAFLFEPDSV
Plasmodium                       IILPEDKKVKRIIDLLAKYVTEEGYAFEEIIKKNEKDNPMFNFIFN-TSD
Theileria                        VYVPTPQYKKRIIDLTSKYVSESGKDFEEVIMKNEPRNGLFSFVFDRFTP
                                                          :          .
```

# U5-116

**Elongation factor Tu GTP-binding domain protein 2**
**hSNU114**

## RNVKF

```
Homo_sapien                          DKALLGSVKDSIVQGFQWGTREGPLCDELIRNVKFKILDAVVAQEPLHRG
Pan_troglodytes                      DKALLGSVKDSIVQGFQWGTREGPLCDELIRNVKFKILDAVVAQEPLHRG
Pongo_pygmaeus                       DKALLGSVKDSIVQGFQWGTREGPLCDELIRNVKFKILDAVVAQEPLHRG
Monodelphis_domestica                DKALLGSVKDSIVQGFQWGTREGPLCDELIRNVKFKILDAVVAQEPLHRG
Macaca_mulatta                       DKALLGSVKDSIVQGFQWGTREGPLCDELIRNVKFKILDAVVAQEPLHRG
Mus_musculus                         DKALLGSVKDSIVQGFQWGTREGPLCDELIRNVKFKILDAVVAQEPLHRG
Rattus_norvegicus                    DKALLGSVKDSIVQGFQWGTREGPLCDELIRNVKFKILDAVVAQEPLHRG
Bos_taurus                           DKALLGSVKDSIVQGFQWGTREGPLCDELIRNVKFKILDAVVAQEPLHRG
Canis_familiaris                     DKALLGSVKDSIVQGFQWGTREGPLCDELIRNVKFKILDAVVAQEPLHRG
Xenopus_tropicalis                   DKALLSSVKDSIVQGFQWGTREGPLCDELIRNVKFKILDAVIAQEPLHRG
Xenopus_laevis                       DKALLSSVKDSIVQGFQWGTREGPLCDELIRNVKFKILDAVIAQEPLHRG
Gallus_gallus                        DKALLGSVKDSIVQGFQWGTREGPLCDELIRNVKFKILDAVIAQEPLHRG
Danio_rerio                          DKALLGSVKDSIVQGFQWGTREGPLCDEPIRNVKFKILDAVIAQEPLHRG
Tetraodon_nigroviridis               DKALLGSVKDSIVQGFQWGTREGPLCDEPIRNVKFKILDAVIAQEPLHRG
Strongylocentrotus_purpuratus        DKDLLTSVKDSIVQGFQWGTREGPLCDEPIRNVKFKILDGVIASEPIHRG
Aedes_aegypti                        DKTLLGSVKDSIIQGFQWGTREGPLCEEPIRNVKFKILDAVIAQEPLHRG
Drosophila_melanogaster              DKNLLTAVKDSIVQGFQWGTREGPLCEEPIRNVKFKILDGVIANEALHRG
Drosophila_pseudoobscura             DKNLLTAVKDSIVQGFQWGTREGPLCEEPIRNVKFKILDGVIANEALHRG
Apis_mellifera                       DKTLLNSARDAIIQGFQWGTREGPLCEEPIRNVKFKILDAVIAQEPLHRG
Spodoptera_exigua                    -VQYLNEIKDSVVAGFQWAAKEGVMAEENLRGVRFNIYDVTLHTDAIHRG
Toxoptera_citricida]                 -VQYLNEIKDSVVAGFQWATKEGVLAEENMRAVRFNIYDVTLHADAIHRG
Bombyx_mori                          -VQYLNEIKDSVVAGFQWAAKEGVMAEENLRGVRFNIYDVTLHTDAIHRG
Tribolium_castaneum                  DKGLLSSVKDSIVQGFQWGTREGPLCEEPIRNTKFKILDAVIANEPLHRG
Anopheles_gambiae                    DKTLLGTVKDSIVQGFQWGTREGPLCEEPIRNVKFKILDAVIAPEPLHRG
Caenorhabditis_elegans               DKHLLSTVRESLVQGFQWATREGPLCEEPIRQVKFKLLDAAIATEPLYRG
Caenorhabditis_briggsae              DKHLLSTVRESLVQGFQWATREGPLCEEPIRQVKFKLLDATIASEPLYRG
Schistosoma_japonicum                ---------------FQWGTREGPLCDEPIRNVKFKILDALISGEAHQRG
Gibberella_zeae                      DKKTLNAVRESIRQGFSWATREGPLCEEPIRNTKFKVTDVLLANEAIFRG
Arabidopsis_thaliana                 DRNLMMAVKDSIVQGFQWGAREGPLCDEPIRNVKFKIVDARIAPEPLHRG
Oryza_sativaJaponica                 DKNLLNAVKDSIVQGFQWGAREGPLCDEPIRNVKFKILNANIAPEPLHRG
Oryza_sativaIndica                   DKNLLNAVKDSIVQGFQWGAREGPLCDEPIRNVKFKILNANIAPEPLHRG
Medicago_truncatula                  DKSLLNAVKDSIVQGFQWGAREGPLCDEPIRNVKFKIVDARIAPEPLHRG
Coprinopsis_cinereaokayama7          DKKLLGTVKEHIKQGFQWGAREGPLCDEPMRNVKFRLLDTTLAAEPIFRG
Dictyostelium_discoideum             NRSLLLSISDSVVRGFQWATKEGPLVDEPIRNVKFKLLDATIASEPIQRS
Neurospora_crassa                    DKKRLNTVKDFIRQGFNWAVREGPLCEEPIRNTKFRLIDVSLAQEAIFRG
Aspergillus                          DKKLLGNVRDSITQGFSWGTREGPLCEEPIRNTKFRLTDVSLADQAIYRG
Bigelowiella_natans                  -IQYLDEIKDSCVSAFQDVTKEGILAHENMRGVIFTIVDLELHADSIHRG
Chaetomium_globosum                  DKKRLNTVKESIRQGFSWATREGPLCEEPIRNTKFRLIDIALAQEAIFRG
Coccidioides_immitisRS               DKKLLGTVRDSIRQGFSWGTREGPLCEEPIRNTKFRLTDITLADQAIFRG
Cryptosporidium                      DKKLLYDVKDDIIQGFNWAVKEGPLLEEPIRNVKFKILDVNLSSDKVSRG
Ustilago_maydis521                   DKKLLYAVKESITQGFQWGCREGPLADEPMRNVKFRILDATLAHEPIHRG
Magnaporthe_grisea7015               DKKLLNTVKESIRQGFSWATREGPLCEEPIRNTKFKIMDAALSQEAIYRG
Neosartorya_fischeriNRRL181          DKKLLGNVRDSITQGFSWGTREGPLCEEPIRNAKFRLTDVSLADQAIYRG
Saccharomyces_cerevisiae             -VQYLHEIKDSVVAAFQWATKEGPIFGEEMRSVRVNILDVTLHADAIHRG
Schizosaccharomyces_pombe972h        DKNVLNSVKEYIKQGFQWGTREGPLCDETIRNVNFRLMDVVLAPEQIYRG
Tetrahymena_thermophilaSB210         DKNILRECKEHINQGFCWATREGPLCDEPVRNVKFKLIEANISSEPLYRA
Theileria                            DQNLLNRVKSSVIQGFNWAIKEGPLIEEPIRSVKFRLINCELSNEYINIT
Trichomonas_vaginalis                -AEYLQEVKEHFISSFQWATKLGVLAEEPLRGVRFNVVEVFLHADAAHRN
                                              *    : * :  * :*  .  : :  :  :
```

# Pre-mRNA-processing-splicing factor 8

**KIAF**
**RAVFW**

```
Homo_sapien                          TEYKIAFPYLYNNLP--HHVHLTWYHTPNVVFIK--TEDPDLPAFYF
Monodelphis_domestica                TEYKIAFPYLYNNLP--HHVHLTWYHTPNVVFIK--TEDPDLPAFYF
Macaca_mulatta                       TEYKIAFPYLYNNLP--HHVHLTWYHTPNVVFIK--TEDPDLPAFYF
Mus_musculus                         TEYKIAFPYLYNNLP--HHVHLTWYHTPNVVFIK--TEDPDLPAFYF
Rattus_norvegicus                    TEYKIAFPYLYNNLP--HHVHLTWYHTPNVVFIK--TEDPDLPAFYF
Canis_familiaris                     TEYKIAFPYLYNNLP--HHVHLTWYHTPNVVFIK--TEDPDLPAFYF
Xenopus_laevis                       TEYKIAFPYLYNNLP--HHVHLTWYHTPNVVFIK--TEDPDLPAFYF
Gallus_gallus                        TEYKIAFPYLYNNLP--HHVHLTWYHTPNVVFIK--TEDPDLPAFYF
Danio_rerio                          TEYKIAFPYLYNNLP--HHVHLTWYHTPNVVFIK--TEDPDLPAFYF
Aedes_aegypti                        TEYRIAFPYLYNNMP--HFVHLCWYHTPNVVFIK--TEDPDLPAFYF
Drosophila_melanogaster              TEYRIAFPYLYNNMP--HFVHLSWYHTPNVVYIK--TEDPDLPAFYF
Drosophila_pseudoobscura             TEYRIAFPYLYNNMP--HFVHLSWYHTPNVVYIK--TEDPDLPAFYF
Apis_mellifera                       TEYRIAFPYLYNNMP--HFVHLSWYHAPNVVYIK--TEDPDLPAFYF
Tribolium_castaneum                  TEYRIAFPYLYNNMP--HFVHLSWYHAPNVVYIK--TEDPDLPAFYF
Anopheles_gambiae                    TEYRIAFPYLYNNMP--HFVHLSWYHTPNVVFIK--TEDPDLPAFYF
Caenorhabditis_elegans               TEYRIAFPFMYNNLISSLPVQVSWYHTPSVVFIK--TEDPDLPAFYY
Caenorhabditis_briggsae              TEYRIAFPFMYNNLISSLPVQVSWYHTPSVVFIK--TEDPDLPAFYY
Ashbya_gossypii                      TEYRIAFPHLYNPRP--RSVEISWYHDPVSCLLP-EDDDDAMPAFHF
Gibberella_zeae                      TEFRVAYPYLYNSLP--RSVHLSWHSHPQVVFNR--ADDPDLPTFHF
Arabidopsis_thaliana                 TEYKVAFPHLYNNRP--RKVKLCVYHTPMVMYIK--TEDPDLPAFYY
Oryza_sativaJaponica                 TEYRIAFPHLYNNRP--RKVRLGVYHTPMIMYIK--TEDPDLPAFYY
Oryza_sativaIndica                   TEYRIAFPHLYNNRP--RKVRLGVYHTPMIMYIK--TEDPDLPAFYY
Ostreococcus_tauri                   TEYKVAFPYLYNNRP--RKVALATYHSPMVMFIK--TEDPDLPAYYY
Coprinopsis_cinereaokayama7          TEYKVAFPHLYNSLP--RSVHIAPYHYPKNVYIR--TDDPDLPAFYF
Dictyostelium_discoideum             TEYKIAFPYLYNSRP--RKVKTPTYHTPNNCYIK--NDSPDLPGFYF
Neurospora_crassa                    TECRVAYPHLYNALP--RSVQLSVHSYPQVVYTR--TEDYNLPAFYF
Aspergillus                          TEYRVAFPFLYNTLP--RSVSVSWYSYPQVVYVR--TDDPNLPAFYF
Bigelowiella_natans                  VEYKLDYPYLYNSNVN--LIDDKIYLDQLSHIVK--IEDIQLPIFLE
Chaetomium_globosum                  TECRVSYPYLYNALP--RSVHLSWHSHPQVVYHK--PEDPDLPAFYF
Coccidioides_immitisRS               TEYRVAFPYLYNSLP--RSVKLSWYSHPQVVYVR--TDDPNLPAFYF
Cryptosporidium                      SEYKIAFPHFYNSLP--KFVSTSVYHYIVNIFTK--PENPNSPIFEF
Entamoeba_histolytica                TEYKVAYPNLYANRP--REVHIIPYHYPMQMIVK--PDDPTLPSYYF
Ustilago_maydis521                   TEYRVAFPHLYNSRP--RAVHIGTYHEPKNVYIR--SDEIDQ-AFYF
Yarrowia_lipolytica                  TEHKVAFPHVYNSMP--RKVELVTYSYPQEVVVK--SEDPDLPPYYF
Pichia_stipit                        TEYKVAFPYLYNSFV--KGVHVSWYHDPVKCYVE-SSDP-VATSFTF
Candida_albicans                     TEYKVELPFLYNSFV--KKVSVSPLGAPLDCRSQ-QPQSKGLPAFTF
Magnaporthe_grisea                   NESRVAFPYLYNHLP--RSVQISWYSHPQIVYKQ--PENPDLPAFHF
Debaryomyces_hansenii                TEYKVAFPFLYNSFA--KLIYVGWYHEPLNCFIKNSMEDTDLPAFHF
Phaeosphaeria_nodorum                TEYRVSYPFLYNSLP--RSVKISWYSHPQVVYVR--AEDPSLPAFYF
Neosartorya_fischeriNRRL181          TEYRVAFPFLYNTLP--RSVKVSWYSHPQVVYVR--T-DPNLPAFYF
Plasmodium                           TEYKIAFPYLYNNRP--RKIAVSKYHSPMCVYIK--LEDIDLPPFYF
Saccharomyces_cerevisiae             SEYKVAFPHLYNSRP--RSVRIPWYNNPVSCIIQ-NDEEYDTPALFF
Schizosaccharomyces_pombe972h        TEYRIAFPYLYNSRA--RSVALSEYHQPSNVFVP--PEDPDLPAFFW
Kluyveromyces_lactis                 TEYQIAFPQIYAARP--RSVETLWYHDPLSCTVK-SEMEHSSATFFF
Tetrahymena_thermophilaSB210         TEYKIAFPFLYNSRP--RRVAIAPYHYPACCYIK--QENPDLPAFYF
Paramecium_tetraurelia               SEYKIAFPFLYNSRP--RSVAIAPYHYPANVFIK--QDNPEIPTYNF
Theileria                            TEYKIAFPYLYNSRP--RKVAMTNYHTKLCSYIR--HEDPDLPIFHY
Trichomonas_vaginalis                TEWKIAYPNLYNNRP--RKISIAPYHYPLSCFAK--YNTIITPVFQL
nucleomorph_GuillardiaTheta          VEYKLDYPYLYNSNVN--LIDDKIYLDQLSHIVK--IEDIQLPIFLE
Trypanosoma_cruzi                    TERQIAFPFLYGSVV--EAVTVAPYHYPAEIRVE--NEDPALPCFSW
Leishmania_major                     TERQIAFPFLYGSVV--DGVELAPYHYPATVRVE--NNDPELPCFTF
                                       *     *
```

# Pre-mRNA-processing-splicing factor 8

**KIAF**

**RAVFW**

```
Homo_sapien                          KKCWPRDARMRLMKHDVNLGRAVFWDIKNRLPRSVTTVQWEN
Monodelphis_domestica                KKCWPRDARMRLMKHDVNLGRAVFWDIKNRLPRSVTTVQWEN
Macaca_mulatta                       KKCWPRDARMRLMKHDVNLGRAVFWDIKNRLPRSVTTVQWEN
Mus_musculus                         KKCWPRDARMRLMKHDVNLGRAVFWDIKNRLPRSVTTVQWEN
Rattus_norvegicus                    KKCWPRDARMRLMKHDVNLGRAVFWDIKNRLPRSVTTVQWEN
Canis_familiaris                     KKCWPRDARMRLMKHDVNLGRAVFWDIKNRLPRSVTTVQWEN
Xenopus_laevis                       KKCWPRDARMRLMKHDVNLGRAVFWDIKNRLPRSVTTVQWEN
Gallus_gallus                        KKCWPRDARMRLMKHDVNLGRAVFWDIKNRLPRSVTTVQWEN
Danio_rerio                          KKCWPRDARMRLMKHDVNLGRAVFWDIKNRLPRSVTTIQWEN
Tetraodon_nigroviridis               KKCWPRDARMRLMKHDVNLGRAVFWDIKNRLPRSVTTVQWEN
Strongylocentrotus_purpuratus        KKCWPRDARMRLMKHDVNLGRAVFWDMKNRLPRSITTFMWEQ
Aedes_aegypti                        KKCWPRDARMRLMKHDVNLGRAVFWDIKNRLPRSVTTVQWDN
Drosophila_melanogaster              KKCWPRDARMRLMKHDVNLGRAVFWDIKNRLPRSVTTIGWES
Drosophila_pseudoobscura             KKCWPRDARMRLMKHDVNLGRAVFWDIKNRLPRSVTTIGWES
Apis_mellifera                       KKCWPRDARMRLMKHDVNLGRAVFWDIKNRLPRSTTTIQWEN
Tribolium_castaneum                  KKCWPRDARMRLMKHDVNLGRAVFWDIKNRLPRSVTTIQWEN
Anopheles_gambiae                    KKCWPRDARMRLMKHDVNLGRAVFWDIKNRLPRSVTTVQWDN
Caenorhabditis_elegans               KKCWPRDARMRLMKHDVNLGRAVFWDIKNRLPRSITTVEWEN
Caenorhabditis_briggsae              KKCWPRDARMRLMKHDVNLGRAVFWDIKNRLPRSITTVEWEN
Ashbya_gossypii                      KKCWPRDSRMRLIRQDVNLGRAVFWEIENRVPPSFADILWEN
Gibberella_zeae                      KKCWPRDSRMRLMRHDVNLGRAVFWDLKNRLPRSVTTIDWDD
Arabidopsis_thaliana                 KKCWPRDARMRLMKHDVNLGRSVFWDMKNRLPRSITTLEWEN
Oryza_sativaJaponica                 KKCWPRDARMRLMKHDVNLGRSVFWDMKNRLPRSITTLEWEN
Oryza_sativaIndica                   KKCWPRDARMRLMKHDVNLGRSVFWDMKNRLPRSITTLEWEN
Ostreococcus_tauri                   KKCWPRDARMRLMKHDVNLGRAVFWDIKNRLPRSLTTLEWDS
Coprinopsis_cinereaokayama           KRCWPRDCRMRLIKHDVNLGRAVFWNVKQSLPRSLTTIEWED
Dictyostelium_discoideum             KKCWPRDCRMRLMKHDVNLGRAVFWQIKNRLPRSLTTIDWED
Neurospora_crassa                    KKCWPRDSRMRLMRHDVNLGRAVFWDLKNRLPRSVTTIEWED
Aspergillus                          KKCWPRDCRMRLMRHDVNLGRAAFWDLKNRLPRSITTIDWDD
Chaetomium_globosum                  KKCWPRDSRMRLMRHDVNLGRAVFWDLKNRLPRSVTTVEWED
Coccidioides_immitisRS               KKCWPRDSRMRLMRHDVNLGRAVFWDLKNRLPRSITTIEWDD
Cryptosporidium                      KKCWPRECRMRLVKNDVIIGKSVYWELSNRLPKSITTLEWER
Entamoeba_histolytica                KKCWPKDCRMRLMRHDVNLGRAVFWELKNRLPRSLTTLNWED
Ustilago_maydis                      RKCWPRDARMRLVKHDVNLGRAVFWTVKNSLPRSLTTIEWED
Yarrowia_lipolytica                  KKCWPRDSRMRLMRSDVNLARAAFWEIQNRLPRAMTSIEWSE
Pichia_stipit                        KRCWPKDSRMRLMRHDVNLGRAVFWEIASRIPRSLTSIEWKD
Candida_albicans                     HRCWPRDSRMRLMRHDVNLGRATFWEISGRIPTSLTSIEWED
Magnaporthe_grisea                   KKCWPRDARMRLMRHDVNLGRAVFWDLKNRLPRSITTIEWED
Debaryomyces_hansenii                RRCWPRDSRMRLMRHDVNLGRAVFWEIAGRIPRSLTTIEWED
Phaeosphaeria_nodorum                KKCWPRDSRMRLMRHDVNLGRAVFWDMKNRLPRSVTTIEWDD
Neosartorya_fischeriNRRL             KKCWPRDCRMRLMRHDVNLGRAAFWDLKNRLPRSITTIEWDD
Plasmodium                           KTCWPRDCRMRKMKHDVNLGRATFWEIQNRIPRSLTSLDWDH
Saccharomyces_cerevisiae             RKCWPKDSRMRLIRQDVNLGRAVFWEIQSRVPTSLTSIKWEN
Schizosaccharomyces_pombeh           KNCWPRDARMRLMKHDVNLGRAVFWEIRNRLPRSLTTLEWED
Kluyveromyces_lactis                 KKCWPRDARMRLIRHDVNLGRAVFWELTSRVPKSLVNITWEN
Tetrahymena_thermophilaSB            KKCWPKDCRMRLMKHDVNLGRAVFWEMKNRLPRCLTTMEWEH
Paramecium_tetraurelia               KKCWPKDCRMRLMKHDVNLGRAVFWDIKNRLPRCLTTLAWEH
Theileria                            KNCWPKDCRMRLMKHDVNLGRAAFWEMQSRLPRSITTLEWSD
Trichomonas_vaginalis                KTCWPKDARMRLIKHDVNLGRAVFWDLQNRLPRSLCEVNWNS
Trypanosoma_cruzi                    ANQWPRDARMRLFLNDVNLARAVLWEFRGRLPPSIAEMNESN
Leishmania_major                     AKEWPRDARMRLFLSDVNLARAVLWEFRSRLPPSMATINDAN
                                        **::.***  .  ** :.::. *  .   :* .   .
```

# 5. Discussion

## *minigene in vivo splicing assays*

### 5.1.1 advantage and applications of this method

A minigene contains a genomic fragment including the alternative exon(s) and the surrounding introns as well as the flanking constitutively spliced cloned in a eukaryotic expression vector. All minigenes described in the AEdb minigene database come from the published literature and have been compiled. More than half of the constructions are minigenes containing one cassette exon (45 from 82). Other mechanism are minigenes containing multiple cassette exon, retained intron, incremental combinatorial exons, mutually exclusive exons, alternative 3' splice site, alternative 5' splice site and minigenes containing both alternative 5' and 3' splice site (Table 5.1)

The minigene entries are linked to the appropriate entries of the AEdb sequence collection, which allows the user to quickly identify experimentally useful minigenes by searching the database (Stamm, Riethoven et al. 2006).

**Table 5.1 AEdb-Minigene data statistics – 82 entries**

| | |
|---|---|
| Organism distribution | Human (46); mouse (17); rat (15); drosophila (4); chicken (2); others (3) |
| Splicing mechanism distribution | Cassette exon (45); multiple casette exons (3); incremental combinatorial exons (2); Alternative acceptor sites (9); Alternative donor sites (8); Intron retention (2); Mutually exclusive exons (13) |
| Reported tissue specificity | 55 |
| Known regulatory factors | 32 |
| Enhancer and silencer | 97 |
| Superlinks to AEdb sequence database | 105 |

Using minigenes has two advantages: (i) often exon trap constructs are easier and faster to clone and (ii) if these constructs behave like the endogenous gene, it is clear that the regulatory region is confined to the cloned exon. However, it has been often found that there is a discrapency between alternative exons flanked by their normal or heterologous contexts (Hartmann, Rujescu et al. 2001).

The minigenes can be transfected into the cell line of choice by standard methods, including calcium phosphate, electroporation, and liposome transfer. When studying tissue specific alternative splicing, usually several cell lines are screened to find a cell line that recapitulates the alternative splicing pattern observed in vivo. As a result, minigenes have been analyzed in numerous cell lines, including HeLa, HEK293, primary neurons, HepG2, CHO. Once such a system is established, two major questions can be addressed: which cis-sequences are necessary for the regulation and what trans-acting factors are involved. Cis-acting sequences are usually determined mutagenesis, either at specific sites or through deletion of larger parts. In a number of studies mutations resembling human mutations were analysed by minigenes (Nissim-Rafinia, Chiba-Falek et al. 2000). The role of trans-acting factors is usually studied by increasing their amount through cotransfection. Since splicing factors mostly work in a concentration dependent manner (Smith and Valcarcel 2000), a correlation between alternative exon usage and amount of trans-factor is a good indication for regulation of this exon by that particular factor. However, since indirect effects, e.g. sequestration, influence of mRNA stability or interference with the general splicing machinery will occur, controls and additional experiments have to be performed.

Although in most cases regulatory factors are increased through cotransfections, an increasing number of experiments are reported that use oligonucleotides or RNA interference to decrease the amount of trans-factors. Finally, the modification of trans-acting factors through phosphorylation has been studied by employing the appropriate kinases (Hartmann, Rujescu et al. 2001; Weg-Remers, Ponta et al. 2001; Matter, Herrlich et al. 2002).

By far the most experiments are analyzed by RT-PCR. To achieve reproducibility, it is important that mRNA is not damaged during isolation. In our experience this is best

achieved with commercially available spin column bases kits (e.g. RNeasy, Qiagen). At least one primer is chosen to be specific for the minigene to avoid amplification of the endogenous gene. RT-RCR should be performed with the lowest amount of cycles possible to ensure a linear relation between mRNA isoforms and amplified signals. A frequently occurring problem is the amplification of the minigene if it is short. This can be avoided by shortening the extension time of by adding DpnI into the reverse transcription reaction. DpnI cuts GATC sequences in doublestranded DNA when the adenosine is methylated but does not cut non-methylated ssDNA or cDNA. We found that commercially available preparation of DpnI are essentially RNAse free and do not interfere with the reverse transcription. The primers for amplification are usually chosen in the flanking constitutive exons, which results in two bands corresponding to exon skipping and inclusion. PCR products are separated on agarose or acrylamide gels, the intensity of the bands is quantified and their ratio determined. The detection of the PCR products can be done by ethidium bromide staining, and labeling the primers with $^{32}P$ or a fluorescent oligo nucleotide (Nissim-Rafinia, Chiba-Falek et al. 2000). The detection is performed by UV light, autoradiography or a DNA sequencer, respectively. Numerous alternative spliced mRNAs have been quantified from different tissues using real-time PCR with boundary spanning TaqMan probes or molecular beacons and scorpion primers (Taveau, Stockholm et al. 2002).

Other methods such as RNase protection assays have been used to analyze minigenes. The quantitative nature of this assay is the major advantage. However, the method is much more laborious than RT-PCR. An interesting development is the analysis of alternative splicing by array formats (Yeakley, Fan et al. 2002). Finally, assay systems have been developed that rely on the proteins generated by alternative splicing of minigenes. These different isoforms can be detected by Western blots (Stoilov, Daoud et al. 2004) if specific antibodies are available. Related to this method are chimeric minigenes that express EGFP fusions depending on alternative exon usage. Because splicing events of these constructs can be detected by EGFP fluorescence they are suitable for FACS analysis and genetic screening. A direct measurement of alternative splicing is possible when luciferase reporters are used, which has been used to determine

signal transduction pathways (Weg-Remers, Ponta et al. 2001; Matter, Herrlich et al. 2002) regulating alternative splicing. To account for variations in transfection efficiency, double reporter assays have been developed (Nasim, Chowdhury et al. 2002).

The analysis of alternative splicing with minigenes requires several controls. First, it is important to determine the transfection efficiency in each experiment, which can be easily done using GFP tagged constructs. When assaying the influence of trans-acting factors, Western blots needs to be performed to determine whether an increase of cDNA expression constructs really causes an increase of protein generated. A fist step in determining whether a trans-acting factor is acting directly with the pre-mRNA of interest is to determine whether it immuno precipitates with the pre-mRNA. To account for unspecific effects, the amount of cDNA transfected and the amount of promoter should be maintained constant. This is usually done by adding "empty" parental vector DNA. Usually, minigene analyses are quite robust. However, the changes of alternative splicing in vivo are often relatively small, around 2 to three fold. It is therefore necessary to determine the statistical significance of the experiments by calculating the standard deviations and performing student's test. (check T-test in webpage http://www.physics.csbsju.edu/stats/t-test.html).

Minigenes have been used to determine trans-acting factor binding sites by in vivo SELEX (Cooper 1999). Here, a randomized sequence is cloned into an alternative exon and the complete mixture is transfected with a trans-acting factor. After RT-PCR, the mixture of alternative exons is isolated and recloned for a second round. This method was successfully used to determine A/C-rich splicing enhancer (Coulter, Landree et al. 1997). Other modifications include the use of kinases and DNA or RNA oligonucleotides to either phosphorylate or remove regulatory factors.

### 5.1.2 Tra2beta1 and Clk2 regulated Clk2 exon4 inclusion in minigene assay

CLK2 binds and phosphorylates hTRA2-beta1 (Stoilov et al. 2004). Like almost all pre-mRNAs of splicingregulatory proteins, htra2-beta1 and clk2 pre-mRNAs undergo alternative splicing (Hanes etal. 1994; Nayler et al. 1998a). Skipping of exons 2 and 3 of htra2-beta pre-mRNA generatesan inactive protein, hTRA2-beta3 (Stoilov et al. 2004),

whereas skipping of exon 4 of clk2pre-mRNA generates a frameshift resulting in the inactive variant, clk2tr (Duncan et al.1997).

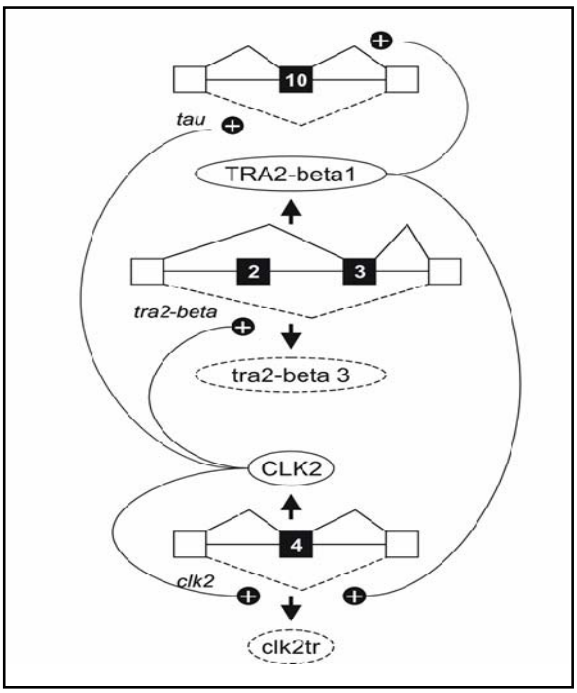



**Figure 5.1 Summary of the splice site regulation of tau exon 10, tra2-beta1 and clk2.** Proteins active in splice site selection are indicated as an oval with a thick line. Inactivevariants are indicated with a dashed line. clk2tr mRNA (tr: truncated) is most likely subject tononsense-mediated decay, therefore, is not translated into protein. The regions of the premRNAs subjected to alternative splicing are schematically indicated, introns as horizontallines, exons as boxes, alternative exons as black boxes. Splicing patterns are indicated byconnecting lines.

Clk2tr mRNA is subject to nonsense-mediated decay and, therefore, is not translated into any protein (Hillman et al. 2004). As shown in Figure 5.1, an increase of hTRA2-beta1 promotes skipping of clk2 exon 4, most likely through binding to a htra2-beta1 motif AAGAGCGA present in the 3' part of the clk2 exon 4. Similar to the situation in clk1 in vivo (Duncan et al. 1997), clk2 promotes skipping of its exon 4 generating a frame shift resulting in an inactive form. We then performed similar assays with the expression constructs encoding the inactive proteins htra2-beta3 and Clk2-KR. Clk2-KR is an inactive CLK2 mutant because it cannot bind ATP. We tested this mutant to determine whether clk2tr protein might influence pre-mRNA splicing if the mRNA escapes nonsense mediated decay under special conditions. We found that these inactive forms have no effect on clk2 pre-mRNA splicing (Figure 5.1). Interestingly, previous research has revealed that CLK2 influences the splicing pattern of tra2-beta by promoting exon skipping and the formation of the inactive protein isoform TRA2-beta3 (Stoilov et

al. 2004).This suggests that the amount of active TRA2-beta1 and CLK2 is controlled through a feedback of alternative splicing decisions.

In this work, our results indicate that CLK2 promotes exon skipping in the tra2-beta, clk2 resulting in the formation of the inactive htra2-beta3 isoform and the inactive clk2. The formation of CLK2 is influenced by hTRA2-beta1, which promotes the formation of the inactive clk2tr isoform, the mRNA of which is subjected to nonsense-mediated decay and is not transcribed into protein (Hillman et al.2004). Tra2beta1 Clk2 and Alzheimer's Disease related gene Tau are involved in this loop and was analyzed by RT-PCR in AD brain tissue which would be shown in 5.2.

## *5.2 Alternative splicing in Alzheimer's Disease*

Tau protein is a cytoskeletal component that is predominantly expressed in axons of mature and growing neurons. The tau gene undergoes extensive alternative splicing leading to the expression of multiple isoforms in the brain (Andreadis 2005). Exon 10 is one of the alternatively spliced exons and encodes a microtuble binding site. The alternative usage of exon 10 in the adult brain is specific for humans, since rodents use this exon alternatively during development, but constitutively at the adult stage. Therefore, in humans, alternative splicing controls the formation of tau proteins containing either 4 or 3 microtuble binding domains (4R, 3R). Adult human neurons have a characteristic ratio of these isoforms and perturbation of this ratio leads to the formation of neurofibrillary tangles (NFTs) and cell death. NFTs are a hallmark of several inherited neurodegenerative diseases that are commonly named "tauopathies". They include progressive supranuclear palsy (PSP), corticobasal degeneration (CBD), Pick's disease (PiD) and frontotemporal dementia with parkinsonism linked to chromosome 17 (FTDP-17). In all of these diseases, the correct ratio between the 3R and 4R isoforms is disturbed by mutation of elements regulating exon 10 alternative splicing (Delacourte and Buee 2000; Spillantini, Van Swieten et al. 2000).

In recent years, several groups have analyzed the regulation of exon 10 alternative splicing (Gao, Memmott et al. 2000; Hartmann, Rujescu et al. 2001; D'Souza and Schellenberg 2002; Jiang, Tang et al. 2003; Wang, Gao et al. 2004). Like other alternative exons, exon 10 is regulated by a finely-tuned balance of sequences and trans-

acting factors. Exon10 contains two enhancers, a GAR (guanidine/adenosine-rich) and an ACE (adenosine/cytosine-enhancer) motif and two weak silencers that are disturbed by mutations occurring in the human population. These elements bind to splicing regulators that either inhibit exon 10 inclusion (SRp30c, SRp55, SRp75, 9G8, U2AF, PTB and hnRNP G) or promote inclusion (hTRA2-beta1, CELF3, CELF4) (Andreadis 2005; Wang, Wang et al. 2005). From these proteins, TRA2-beta1 appears to be of central importance since it binds directly to the exon 10 GAR-type enhancer (Jiang, Tang et al. 2003) and can interact with most of the other exon 10 regulatory proteins (SRp30c, SRp55, SRp75, 9G8, U2AF and hnRNP G) (Nayler, Cap et al. 1998; Wang, Wang et al. 2005). Finally, the usage of tau exon 10 is influenced by phosphorylation of splicing factors. The kinases CLK2 and GSK3ß have been shown to alter the phosphorylation status of splicing factors, which in turn promote skipping of Exon 10 (Hartmann, Rujescu et al. 2001; Hernandez, Perez et al. 2004; Stoilov, Daoud et al. 2004).

In contrast to the inherited tauopathies, there are no obvious mutations in splicing regulatory elements of the tau gene in sporadic tauopathies. However, recent data suggest that the 4R/3R ratio is disturbed in these diseases as well (Umeda, Taniguchi et al. 2004). We, therefore, analyzed post mortem brain tissue for the alternative splicing patterns of tau exon 10, htra2-beta1 exon 3 and the alternative exon of clk2. We discovered that in brain tissues affected by sporadic AD, the fraction of mRNA including exon 10 is increased. Furthermore, we also found that in addition the mRNA isoform ratios of proteins regulating exon 10 usage are altered, suggesting that defects in pre-mRNA processing contribute to sporadic AD (Figure 4.10 in result part).

Our data demonstrate that tau mRNAs containing exon 10, which codes for the 4R protein isoform is relatively increased in the temporal cortex of AD patients. It has been shown that a distortion of the 3R and 4R tau isoform ratio leads to enhanced NFT formation and neurodegeneration (Andreadis 2005). For example, the disruption of the proper balance of the 3R and 4R repeat isoforms through mutations in splicing regulatory regions causes frontotemporal dementia with parkinsonism linked to chromosome-17 (FTPD-17) (Spillantini and Goedert 2000). We find an increase of the 4R tau isoform in the brain regions affected by sporadic AD. This is in agreement with previous data

showing that the 4R tau isoform is upregulated in brain areas from Alzheimer's disease patients with a heavy burden of neurofibrillary tangles (Yasojima et al. 1999). The increase of the 4R tau isoforms is quantitative and therefore semi-quantitative PCR is necessary for its detection. Furthermore, the tau exon 10 ratio varies between individuals and brain areas. This could explain why there is a variability between studies (Umeda et al. 2004) and why some studies did not describe differences in exon 10 usage in Alzheimer's disease (Chambers et al. 1999; Boutajangout et al. 2004). Since previous studies have reported an absence of mutations in tau pre-mRNA associating with AD (Poorkaj et al. 2001) we analyzed the splicing patterns of two pre-mRNAs encoding proteins that regulate tau exon 10 splicing, tra2-beta and clk2 (Hartmann et al. 2001; Jiang et al. 2003; Kondo et al. 2004; Wang et al. 2004; Wang et al. 2005). Results from the investigation revealed a statistically significant change in the ratio of clk2 isoforms. In brain tissue of AD patients, the mRNA for the clk2tr isoform, encoding a non-functional mRNA, is increased in comparison to the full-length isoform encoding the active kinase. These findings point to a defect in splicing regulation associated with AD. A decrease of CLK2 activity favors the inclusion of exon 10 and the formation of 4R tau isoforms and could explain the observed changes in tau exon 10 splicing. However, the activity of CLK2 is not only regulated by its abundance, but also by its phosphorylation state (Nayler et al. 1998). At this point, we can only speculate that the CLK2 activity is reduced in brains affected by AD. We, therefore, analyzed tra2-beta1 as an example of another pre-mRNA that is regulated by CLK2. CLK2 activity promotes the formation of the tra2-beta3 isoform. Our finding that the ratio between tra2-beta3 and tra2-beta1 is increased in temporal cortex affected by AD strongly suggests a decrease of CLK2 activity in this area, since CLK2 promotes tra2-beta3 formation (Figure 4.12 and 4.13 in result part). Considering that it has been well established that tra2-beta1 promotes tau exon 10 inclusion (Jiang et al. 2003; Kondo et al. 2004), it is natural to assume that a relative increase of tra2-beta1 would favor exon 10 inclusion. It is, therefore, possible that an abnormally low activity of CLK2 could contribute to tau exon 10 missplicing, since a low concentration of CLK2 would fail to correctly promote both tau exon 10 and tra2-beta exon 3 inclusion. Surprisingly, we found a down regulation of mRNA encoding active CLK2 kinase also in supplementary motocortex, a brain region not affected by

Alzheimer's disease. For sporadic Alzheimer's disease, aberrant splicing of Presenilin-2 exon 5 has been demonstrated previously (Sato et al. 1999). We therefore determined Presenilin-2 exon 5 usage in our samples and found pronounced changes similar to tau exon 10 and tra2-beta1 exon 3 in temporal cortex, but no changes in supplementary motocortex. Interestingly, the intron upstream of exon 5 harbors a TRA2-beta1 binding site (Figure 4.13 in result part). Thus, so far all alternative exons that are changed in AD contain TRA2-beta1-binding sites which is rich of RAAG motif. It is therefore possible that changes in TRA2-beta isoforms cause abnormal splicing of other pre-mRNAs containing TRA2-beta1-binding sites. Since alternative pre-mRNA splicing is regulated by combinatorial control involving numerous, often antagonistic factors, it is not possible to predict what other mRNAs might be affected. This combinatorial control could also explain why alterations of CLK2 isoforms in supplementary motocortex have no influence on the splicing patterns of tau exon 10 and tra2-beta1, since this brain area might express other factors that compensate the loss of CLK2.

The sequencing of several genomes has underlined the importance of alternative splicing. Indeed, cDNA expression array data indicate that up to 75% of all human genes are subject to alternative splicing (Johnson et al. 2003), which serves as a major mechanism for creating functionally different proteins from a surprisingly small number of genes in humans (Stamm et al. 2004). It is, therefore, interesting to note the increasing number of human diseases suspected to be caused by abnormal regulation of splicing (Stoilov et al. 2002; Faustino and Cooper 2003). Whereas the majority of the diseases associated with missplicing are caused by mutations in regulatory regions, a number of missplicing events occur without any obvious mutations in the affected genes. The most likely reason for these changes are alterations in regulatory factors, suggesting that splicing could be a genetic modifier of disease (Nissim-Rafinia and Kerem 2002). Pre-mRNA splicing modulations are frequent during senescence, which could be caused by age-dependent changes in the splicing machinery (Meshorer and Soreq 2002). Our findings suggest that missplicing events originating from the tra2-beta1 system contribute to the pathogenesis of sporadic AD.

In this work, we investigated whether human postmortem brain tissue of AD patients reveal differences in alternative splicing patterns of the tau, htra2-beta, presenilin 2 and clk2 genes when compared with age-matched controls. We found that the splicing patterns of all four genes are altered in affected brain areas of sporadic AD patients. In these affected areas, the amount of mRNAs of tau isoforms including exon 10, the htra2-beta1 isoform and an inactive form of clk2 are significantly increased. These findings suggest that a misregulation of alternative splicing seems to contribute to sporadic AD.

## 5.3 *Alternative splicing in breast cancer*

Changes in alternative splicing of the CD44 gene are associated with tumor progression and metastasis in breast cancer. A recently published study suggested that specific alterations in the relative concentrations of SR proteins during breast cancer development might be responsible for these effects (Stickeler, Kittrell et al. 1999), but functional data were not determined.

dTra2 (Tra2 in drosophila) was the first member of the SR family of splicing factors shown to be required for exon recognition regulated by exonic enhancer sequences (Burtis 1993). Further experiments showed that dTra2, dTra, and several other SR proteins act synergistically to recognize individual exonic enhancer sequences. In humans, no orthologue for Drosophila dTra has been uncovered, suggesting that human Tra2 proteins may bind to RNA with noticeably different binding partners and RNA specificity. In vitro, GAAGAA-based enhancers were identified as targets for hTra2-ß1 binding, suggesting that hTra2-ß1 binds to G/A-rich exon enhancers. Experiments in vivo showed that hTra2-ß1 binds to a more degenerate sequence GVVGANR, which is partially found in the doublesex gene, where the protein could interact with the dsx-repeat AAAGGACAAAGGACAAA, which is rich in CAA sequences (putative Tra2-beta1 binding sites are italicized). This enhancer could be considered as a version of C/A-rich elements, which we identified earlier as an exon enhancer sequence involved in the specific regulation of alternative CD44 exons v4 and v5 splicing by YB-1 (Stickeler, Fraser et al. 2001).

This background made human Tra2-beta1 a potential candidate to regulate alternative CD44 splicing. We observed Tra2-beta1-dependent activation of inclusion of

two CD44 variable exons that have both C/A-rich and G/A-rich exon enhancer sequences. The ability of both proteins to enhance inclusion, however, was dependent on a single C/A-rich element (CAGACAACCACAAGGA) that resides within exon v4. When this sequence was mutated to CAGAuAAggACuAGGA, inclusion of v4 was eliminated but some level of inclusion of exon v5 remained. Increasing the concentration of Tra2 did not result in increased inclusion of exon v5 in this mutant background despite the presence of the sequence GAAGAA within an enhancer in exon v5. Inclusion of exon v4 was also not restored despite the fact that the created mutation contained two GGA elements similar to the purine-rich enhancer element from the Drosophila doublesex exon that binds dTra2. This observation suggests that human Tra2 recognizes C/A-rich exon enhancers rather than G/A-rich exon enhancers in CD44 and agrees with the binding sequence of Tra2-beta1 found in vivo.

It should be noted that CD44 alternative splicing is often accompanied by inclusion of neighboring variable exons, suggesting interaction between the factors that recognize individual CD44 exons. Exons v4 and v5 may be particularly prone to this type of multiexon recognition. Examination of cytoplasmic RNAs coding for CD44 variable exons has detected not only RNAs that contain both exons v4 and v5 but also RNA species that contain the short intron between them but no other intron. Thus, it is possible that exons v4 and v5 are initially recognized as a single large exon followed by removal of the intervening sequence between them. Several reports have also suggested that the levels of CD44 mRNA containing this intron are increased in tumor cells where inclusion of exons v4, v5, and v6 are also increased (Stickeler, Mobus et al. 1997).

Although CD44 splicing was not dependent on Drosophila dTra, it is possible that other human proteins play the role of Tra in Tra2-mediated recognition of the CD44 exons, such as hnRNP G (Hofmann and Wirth 2002). Two human proteins that influence exon v4 inclusion, YB-1 and the RNA helicase p72, were thus far identified (Stickeler, Fraser et al. 2001). Like Tra2, YB-1 binds to C/A-rich sequences. Therefore, it seems possible that the interaction of Tra2-Tra-SR proteins could be replaced with a Tra2-YB-1-p72 interaction on CD44 exon v4.

The statistically significant induction of Tra2-beta1 expression levels in invasive breast cancer, in contrast to their corresponding pairs of normal tissue controls, is in

accordance to our findings of differential SR protein expression changes in a mouse model of breast cancer development (Stickeler, Kittrell et al. 1999). These findings were conclusive on the RNA as well as protein level and, here, especially morphologically as determined by immunohistochemistry. These results are accompanied by an induction of alternative CD44 splicing in the same paradigm.

In additional cell culture experiments, we analyzed the potential Tra2-beta1 effect on endogenous CD44. In parallel to findings with YB-1, these experiments could not detect an overall induction of alternative splicing of the endogenous CD44 gene (data not shown). YB-1 and Tra2-beta1 are, to date, the strongest splicing activators of CD44. The missing effects on endogenous CD44 might be due to the large internal cassette of variable CD44 exons with large numbers of silencing cis- and trans-acting elements that are deleted in our reporter constructs.

Besides CD44, several other genes show altered RNA processing during breast cancer development, including hormone receptors and genes involved in apoptosis (Orban and Olah 2003). Because splicing changes in these proteins may have a pronounced effect on cellular function, it is important to understand the role of alternative processing in breast cancer. Specific splicing factor expression patterns might be responsible for the metastatic phenotype of certain tumors and thereby represent new targets for intervention in subgroups of patients at high risk.

In this work, we found a significant induction of tra2-beta1 in invasive breast cancer, both on the RNA and protein levels. With respect to our functional data, our findings suggest a specific involvement of human Tra2-beta1 in the recognition of CD44 variable exons and a potential influence in regulating the expression of CD44 isoforms involved in breast cancer progression and metastasis.

## *5.4 Role of tra2-beta1 in translation*

In our previous study, It was shown that TRA2-BETA1 as a SR-like protein shuttles between nucleus and cytoplasm. Under steady state conditions, it is localized in the nucleus, where it was mainly implicated in the recognition of exons (Nayler, Cap et al. 1998; Stoilov, Daoud et al. 2004).

However TRA2-BETA1 accumulates in the cytosol under cellular stress conditions (Daoud, Mies et al. 2002). Yeast two hybrid studies showed that TRA2-BETA1 directly binds to RPL3, a protein of the large ribosomal subunit that plays a role in peptidyltransferase center formation. We confirmed the interaction between TRA2-BETA1 and RPL3 using in vitro pull down assays with recombinant proteins (Fig. 4.21 A and B). This finding raised our interest that whether tra2-beta1 plays role for the translation in cytoplasm.

In sucrose gradient fractionation, the cytoslic TRA2-BETA1 co-sediments with ribosomes and polysome fractions (Fig. 4.22). In the western blot with tra2 antibody, the strongest tra2 signals are closed to the ribosome 80S complex. Furthermore, CLIP (RNA Cross-Linking and ImmunoPrecipitation) of TRA2-BETA1 shows that most of the CLIP targets from cytoslic RNA is ribosomal RNA which were localized mainly on the large subunit of the ribosome, near the RPL3 binding sites in the 28S rRNA (Fig. 4.20 A), a region called sarcin/ricin loop. Interestingly, such area contains several AAG rich sequences in loop structure while in our previous research that tra2-beta1 found to bind with single strand RAAG rich motif. It indicates that tra2-beta1 might be linked to a role in translation initiation and the complex's recruitment. Taken together, the interaction between HTRA2-BETA1, RPL3 and rRNA indicates a previously unknown function of the SR like protein HTRA2-BETA1 in translational regulation.

The retained questions are, where and when does TRA2-BETA1 bind with RPL3? We tried to co-localize both endogenous TRA2-BETA1 and over expressed RPL3 with EGFP tag. No clear overlap were found between these two proteins while RPL3 mainly storage in the nuclei. The Pietropaolo's group found that RPL3 involved alternative splicing and NMD for the gene expression (Cuccurese, Russo et al. 2005). Our sucrose gradient experiment shows that RPL3 co-sediments with TRA2-BETA1 not only in the ribosome and polysome fractions but also in the beginning of the gradient where the mRNP complex localized. Is it possible that TRA2-BETA1 with RPL3 together play roles in splicing complex? RPL3 is believed to be assembled in the nuclei and then exported to cytoplasm. How does RPL3 departure from TRA2-BETA1 or even further that might TRA2-BETA1 play potential role in the beginning of the ribosome assembly? These mechanism is unclear.

Recently, more and more shuttling SR proteins are shown to associate with translating ribosomes. For example, ASF/SF2 enhances translation of reporter mRNAs both in vivo and in vitro (Sanford, Ellis et al. 2005). SRp20 functions in internal ribosome entry site (IRES)-mediated translation of a viral RNA (Bedard, Daijogo et al. 2007). 9G8 has recently been proposed to function in translation to unspliced RNA (Swartz, Bor et al. 2007). In transcription and splicing complex, these SR proteins often directly bind together or tightly related. How these factors are exported (with/without mRNA), assembled to translation machinery and whether they are tightly recruited together in cytoplasm as in nuclear will be of clear interest to further examine.

To investigate the functional relevance of RPL3 and TRA2-BETA1 interaction, we fused a Nuclear Export Signal (NES) to HTRA2-BETA1 and tested its effect on translation. When the tra2-beta1-NES mutant is cotransfected with firefly luciferase construct which is inserted of TRA2-BETA1 binding site, we observed a strong increase of luciferase activity. This effect was less pronounced when the PP1 binding properties of tra2-beta1 (Novoyatleva, Heinrich et al. 2007) were abolished in the tra2-beta1-NES-RATA mutant. Silencing of TRA2-BETA1 by siRNA decreased the translation of luciferase activity (Fig. 4.23 C). Surprisingly, the over expression of TRA2-BETA1 without NES has very limited stimulation to translation. The reson might be that the concentration of cytomic TRA2-BETA1 is strictly controlled by export since under the steady state conditions it is mainly localized in the nucleus.

In conclusion, we suggest a model for tra2-beta1 function in nuclear and cytoplasm (Fig. 5.2). In the nucleus, hyperphosphorylated TRA2-BETA1 can bind to exons containing looped RAAG motifs and regulate splice site selection. CLK2 and other kinases involves in TRA2-BETA1's richment in nuclear speckles where the splicing regulation happens. Dephosphorylation of TRA2-BETA1 controls shuttling of the protein and the dephosphorylation site might be in the second RS domain. After splicing, it probably remains with regulated mRNA and recruit RPL3 with ribosomal large subunit and transport outside to cytoplasm together. In cytoplasm, TRA2-BETA1 binds with PP1 and stimulates translation. Phosphorylated by SRPK or other kinases, TRA2-BETA1 shuttles back to the nucleus for a new regulation cycle.

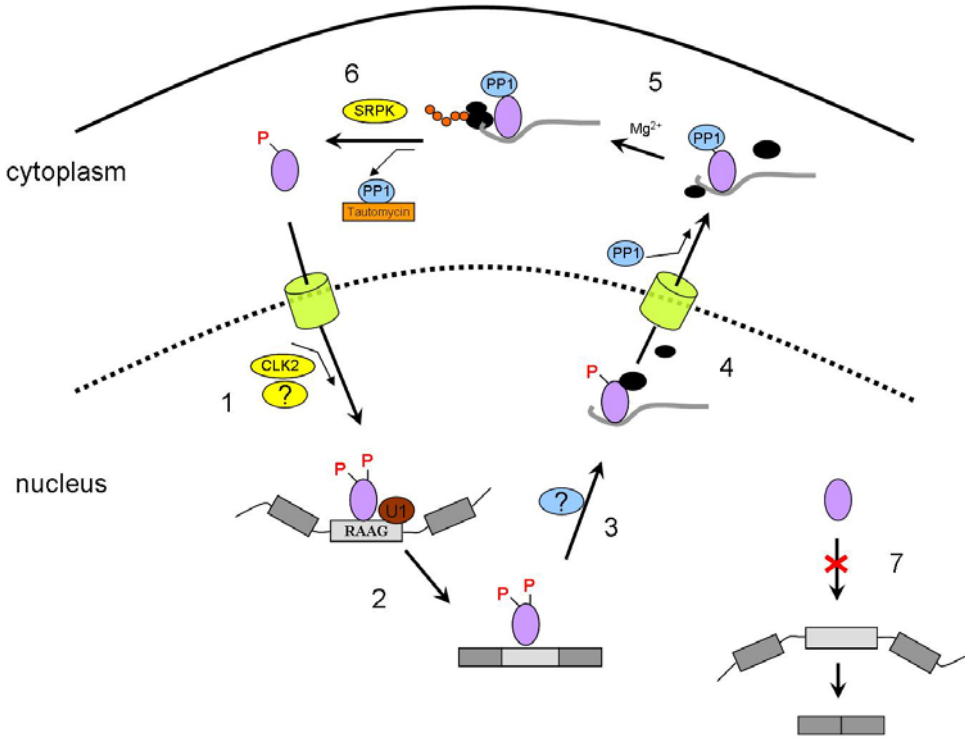



**Figure 5.2. Shuttling model of tra2-beta1.** TRA2-BETA1 is indicated by purple ovals. Kinases acting on TRA2-BETA1 are shown as yellow ovals. Phosphotases are shown in light blue ovals and ribosome subunits are shown in black ovals. In the nucleus, CLK2 and other kinases involves in TRA2-BETA1's richment in nuclear speckles **(step 1)**. Hyperphosphorylated TRA2-BETA1 binds to looped RAAG motif in pre-mRNA and associates with splisosome (brown oval) to enhance the exon's splicing, it forms an exon inclusion variant **(step 2)**. May involves the phosphotase' activity, TRA2BETA1 binds with RPL3 or associated ribosomal subunit which blocks its second RS domain to be further phosphorylated **(step 3)**. TRA2-BETA1 is dephosphorylated by PP1 and forms mRNP complex to be exported though nulear pore **(step 4)**. In cytosol, the ribosomal small subunit is recruited on mRNA with released large subunit and forms translation machinary **(step 5)**. TRA2-BETA1 exists in several phosphorylation states that are not clearly defined. In cytoplasm, SRPK might phosphorylates TRA2-BETA1 on its RS domain which facilitates nuclear import **(step 6)**. In nucleus, dephosphorylated TRA2-BETA1 causes changes in its properties and lead to exon skipping **(step 7)**.

# 6 Reference


Andreadis, A. (2005). "Tau gene alternative splicing: expression patterns, regulation and modulation of function in normal brain and neurodegenerative diseases." Biochem. Biophys. Acta 1739: 91-103.

Ballatore, C., V. M.-Y. Lee, et al. (2007). "Tau-mediated neurodegeneration in Alzheimer's disease and related disorders." 8(9): 663-672.

Bedard, K. M., S. Daijogo, et al. (2007). "A nucleo-cytoplasmic SR protein functions in viral IRES-mediated translation initiation." Embo J 26(2): 459-67.

Beil, B., G. Screaton, et al. (1997). "Molecular cloning of htra2-beta-1 and htra2-beta-2, two human homologs of tra-2 generated by alternative splicing." DNA Cell Biol 16(6): 679-90.

Bentley, D. (2002). "The mRNA assembly line: transcription and processing machines in the same factory." Current Opinion in Cell Biology 14(3): 336-342.

Black, D. L. (2003). "MECHANISMS OF ALTERNATIVE PRE-MESSENGER RNA SPLICING." Annual Review of Biochemistry 72(1): 291-336.

Blencowe, B., J. Nickerson, et al. (1994). "Association of nuclear matrix antigens with exon-containing splicing complexes" J. Cell Biol. 127(3): 593-607.

Blennow, K., M. J. de Leon, et al. "Alzheimer's disease." The Lancet 368(9533): 387-403.

Bourgeois, C. F., F. Lejeune, et al. (2004). "Broad specificity of SR (serine/arginine) proteins in the regulation of alternative splicing of pre-messenger RNA." Prog Nucleic Acid Res Mol Biol 78: 37-88.

Bradford, M. (1976). "A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding." Anal Biochem. 72: 248-54.

Burset, M., I. A. Seledtsov, et al. (2001). "SpliceDB: database of canonical and non-canonical mammalian splice sites." Nucleic Acids Res 29(1): 255-9.

Burtis, K. C. (1993). "The regulation of sex determination and sexually dimorphic differentiation in Drosophila." Curr Opin Cell Biol 5(6): 1006-14.

Caceres, J., S. Stamm, et al. (1994). "Regulation of alternative splicing in vivo by overexpression of antagonistic splicing factors." Science 265(5179): 1706-1709.

Caceres, J. F., G. R. Screaton, et al. (1998). "A specific subset of SR proteins shuttles continuously between the nucleus and the cytoplasm." Genes Dev. 12(1): 55-66.

Celotto, A. M. and B. R. Graveley (2001). "Alternative Splicing of the Drosophila Dscam Pre-mRNA Is Both Temporally and Spatially Regulated." Genetics 159(2): 599-608.

Ciccarone, V. C., D. A. Polayes, et al. (1997). Generation of Recombinant Baculovirus DNA in E. coli Using a Baculovirus Shuttle Vector. Molecular Diagnosis of Infectious Diseases. U. Reischl. Totowa, Humana Press. 13: 213-236.

Cooper, T. A. (1999). "In vivo SELEX in vertebrate cells." Methods Mol Biol 118: 405-17.

Coulter, L., M. Landree, et al. (1997). "Identification of a new class of exonic splicing enhancers by in vivo selection [published erratum appears in Mol Cell Biol 1997 Jun;17(6):3468]." Mol. Cell. Biol. 17(4): 2143-2150.

Craven, R. J., W. G. Cance, et al. (1995). "The Nuclear Tyrosine Kinase Rak Associates with the Retinoblastoma Protein pRb." Cancer Res 55(18): 3969-3972.

Cruts, M., L. Hendriks, et al. (1996). "The presenilin genes: a new gene family involved in Alzheimer disease pathology." Hum. Mol. Genet. 5(suppl_1): 1449-1455.

Cuccurese, M., G. Russo, et al. (2005). "Alternative splicing and nonsense-mediated mRNA decay regulate mammalian ribosomal gene expression." Nucl. Acids Res. 33(18): 5965-5977.

Daoud, R., M. Da Penha Berzaghi, et al. (1999). "Activity-dependent regulation of alternative splicing patterns in the rat brain." Eur J Neurosci 11(3): 788-802.

Daoud, R., G. Mies, et al. (2002). "Ischemia induces a translocation of the splicing factor tra2-beta 1 and changes alternative splicing patterns in the brain." J Neurosci 22(14): 5889-99.

Dauwalder, B., F. Amaya-Manzanares, et al. (1996). "A human homologue of the Drosophila sex determination factor transformer-2 has conserved splicing regulatory functions." Proc Natl Acad Sci U S A 93(17): 9004-9.

Debnath, J., M. Chamorro, et al. (1999). "rlk/TXK Encodes Two Forms of a Novel Cysteine String Tyrosine Kinase Activated by Src Family Kinases." Mol. Cell. Biol. 19(2): 1498-1507.

Del Gatto-Konczak, F., C. F. Bourgeois, et al. (2000). "The RNA-Binding Protein TIA-1 Is a Novel Mammalian Splicing Regulator Acting through Intron Sequences Adjacent to a 5' Splice Site." Mol. Cell. Biol. 20(17): 6287-6299.

Delacourte, A. and L. Buee (2000). "Tau pathology: a marker of neurodegenerative disorders." Curr Opin Neurol 13(4): 371-6.

Derry, J. J., S. Richard, et al. (2000). "Sik (BRK) Phosphorylates Sam68 in the Nucleus and Negatively Regulates Its RNA Binding Ability." Mol. Cell. Biol. 20(16): 6114-6126.

D'Souza, I. and G. D. Schellenberg (2002). "tau Exon 10 expression involves a bipartite intron 10 regulatory sequence and weak 5' and 3' splice sites." J Biol Chem 277(29): 26587-99.

Duncan, P., D. Stojdl, et al. (1997). "In vivo regulation of alternative pre-mRNA splicing by the Clk1 protein kinase." Mol. Cell. Biol. 17(10): 5996-6001.

Elliott, D. J., J. P. Venables, et al. (2000). "An evolutionarily conserved germ cell-specific hnRNP is encoded by a retrotransposed gene." Hum Mol Genet 9(14): 2117-24.

Fasken, M. B. and A. H. Corbett (2005). "Process or perish: quality control in mRNA biogenesis." 12(6): 482-488.

Gao, Q. S., J. Memmott, et al. (2000). "Complex regulation of tau exon 10, whose missplicing causes frontotemporal dementia." J Neurochem 74(2): 490-500.

Glatz, D. C., D. Rujescu, et al. (2006). "The alternative splicing of tau exon 10 and its regulatory proteins CLK2 and TRA2-BETA1 changes in sporadic Alzheimer's disease." J Neurochem 96(3): 635-44.

Goedert, M., B. Ghetti, et al. (2000). "Tau Gene Mutations in Frontotemporal Dementia and Parkinsonism Linked to Chromosome 17 (FTDP-17): Their Relevance for Understanding the Neurogenerative Process." Ann NY Acad Sci 920(1): 74-83.

Graveley, B. R. (2000). "Sorting out the complexity of SR protein functions." RNA 6(9): 1197-1211.

Gu, M. and C. D. Lima (2005). "Processing the message: structural insights into capping and decapping mRNA." Current Opinion in Structural Biology. Folding and binding / Protein-nucleic acid interactions 15(1): 99-106.

Gunderson, S. I., S. Vagner, et al. (1997). "Involvement of the carboxyl terminus of vertebrate poly(A) polymerase in U1A autoregulation and in the coupling of splicing and polyadenylation." Genes Dev. 11(6): 761-773.

Hao, Q., D. Ferris, et al. (1991). "Nuclear and cytoplasmic location of the FER tyrosine kinase." Mol. Cell. Biol. 11(2): 1180-1183.

Hartmann, A. M., O. Nayler, et al. (1999). "The interaction and colocalization of Sam68 with the splicing-associated factor YT521-B in nuclear dots is regulated by the Src family kinase p59(fyn)." Mol Biol Cell 10(11): 3909-26.

Hartmann, A. M., D. Rujescu, et al. (2001). "Regulation of alternative splicing of human tau exon 10 by phosphorylation of splicing factors." Mol Cell Neurosci 18(1): 80-90.

Hernandez, F., M. Perez, et al. (2004). "Glycogen synthase kinase-3 plays a crucial role in tau exon 10 splicing and intranuclear distribution of SC35. Implications for Alzheimer's disease." J Biol Chem 279(5): 3801-6.

Hieronymus, H. and P. A. Silver (2003). "Genome-wide analysis of RNA-protein interactions illustrates specificity of the mRNA export machinery." 33(2): 155-161.

Higashide, S., K. Morikawa, et al. (2004). "Identification of regulatory cis-acting elements for alternative splicing of presenilin 2 exon 5 under hypoxic stress conditions." Journal of Neurochemistry 91(5): 1191-1198.

Hillman, R. T., R. Green, et al. (2004). "An unappreciated role for RNA surveillance." Genome Biology 5(2): R8.

Hofmann, Y. and B. Wirth (2002). "hnRNP-G promotes exon 7 inclusion of survival motor neuron (SMN) via direct interaction with Htra2-{beta}1." Hum. Mol. Genet. 11(17): 2037-2049.

Huang, Y. and J. A. Steitz (2005). "SRprises along a Messenger's Journey." Molecular Cell 17(5): 613-615.

Isono, K.-i., Y.-i. Fujimura, et al. (2005). "Mammalian Polyhomeotic Homologues Phc2 and Phc1 Act in Synergy To Mediate Polycomb Repression of Hox Genes." Mol. Cell. Biol. 25(15): 6694-6706.

Jemal, A., T. Murray, et al. (2005). "Cancer Statistics, 2005." CA Cancer J Clin 55(1): 10-30.

Jensen, K. B., B. K. Dredge, et al. (2000). "Nova-1 Regulates Neuron-Specific Alternative Splicing and Is Essential for Neuronal Viability." Neuron 25(2): 359-371.

Jiang, Z., H. Tang, et al. (2003). "Mutations in Tau Gene Exon 10 Associated with FTDP-17 Alter the Activity of an Exonic Splicing Enhancer to Interact with Tra2{beta}." J. Biol. Chem. 278(21): 18997-19007.

Jiang, Z., H. Tang, et al. (2003). "Mutations in tau gene exon 10 associated with FTDP-17 alter the activity of an exonic splicing enhancer to interact with Tra2-beta1." J. Biol. Chem. 278: 18997-19007.

Julian P. Venables (2006). "Unbalanced alternative splicing and its significance in cancer." BioEssays 28(4): 378-386.

Jumaa, H., J. Guenet, et al. (1997). "Regulated expression and RNA processing of transcripts from the Srp20 splicing factor gene during the cell cycle." Mol. Cell. Biol. 17(6): 3116-3124.

Jurica, M. S. and M. J. Moore (2003). "Pre-mRNA Splicing: Awash in a Sea of Proteins." Molecular Cell 12(1): 5-14.

Kapranov, P., A. T. Willingham, et al. (2007). "Genome-wide transcription and the implications for genomic organization." 8(6): 413-423.

Keene, J. D. (2007). "RNA regulons: coordination of post-transcriptional events." 8(7): 533-543.

Kim Guisbert, K., K. Duncan, et al. (2005). "Functional specificity of shuttling hnRNPs revealed by genome-wide analysis of their RNA binding profiles." RNA 11(4): 383-393.

Kohler, A. and E. Hurt (2007). "Exporting RNA from the nucleus to the cytoplasm." advanced online publication.

Kornblihtt, A. R., M. de la Mata, et al. (2004). "Multiple links between transcription and splicing." RNA 10(10): 1489-98.

Krecic, A. M. and M. S. Swanson (1999). "hnRNP complexes: composition, structure, and function." Current Opinion in Cell Biology 11(3): 363-371.

Kunkel, T., J. Roberts, et al. (1987). "Rapid and efficient site-specific mutagenesis without phenotypic selection." Methods Enzymol. 154: 367-82.

Ladd, A. N. and T. A. Cooper (2002). "Finding signals that regulate alternative splicing in the post-genomic era." Genome Biol 3(11): REVIEWS0008.

Laemmli, U. (1970). "Cleavage of structural proteins during the assembly of the head of bacteriophage T4." Nature 227(5259): 680-5.

Lander ES, e. a. (2001). "Initial sequencing and analysis of the human genome." 409(6822): 860-921.

Lorson, C. L., E. Hahnen, et al. (1999). "A single nucleotide in the SMN gene regulates splicing and is responsible for spinal muscular atrophy." PNAS 96(11): 6307-6311.

Luckow, V. A., S. C. Lee, et al. (1993). "Efficient generation of infectious recombinant baculoviruses by site-specific transposon-mediated insertion of foreign genes into a baculovirus genome propagated in Escherichia coli." J. Virol. 67(8): 4566-4579.

Luo, M.-j. and R. Reed (1999). "From the Cover: Splicing is required for rapid and efficient mRNA export in metazoans." PNAS 96(26): 14937-14942.

Maquat, L. E. (2005). "Nonsense-mediated mRNA decay in mammals." J Cell Sci 118(9): 1773-1776.

Matter, N., P. Herrlich, et al. (2002). "Signal-dependent regulation of splicing via phosphorylation of Sam68." Nature 420(6916): 691-695.

Modrek, B., A. Resch, et al. (2001). "Genome-wide detection of alternative splicing in expressed sequences of human genes." Nucl. Acids Res. 29(13): 2850-2859.

Nasim, M. T., H. M. Chowdhury, et al. (2002). "A double reporter assay for detecting changes in the ratio of spliced and unspliced mRNA in mammalian cells." Nucl. Acids Res. 30(20): e109-.

Nayler, O., C. Cap, et al. (1998). "Human transformer-2-beta gene (SFRS10): complete nucleotide sequence, chromosomal localization, and generation of a tissue-specific isoform." Genomics 53(2): 191-202.

Nayler, O., F. Schnorrer, et al. (1998). "The cellular localization of the murine serine/arginine-rich protein kinase CLK2 is regulated by serine 141 autophosphorylation." J Biol Chem 273(51): 34341-8.

Nayler, O., F. Schnorrer, et al. (1998). "The Cellular Localization of the Murine Serine/Arginine-rich Protein Kinase CLK2 Is Regulated by Serine 141 燗utophosphorylation." J. Biol. Chem. 273(51): 34341-34348.
Nayler, O., W. Stratling, et al. (1998). "SAF-B protein couples transcription and pre-mRNA splicing to SAR/MAR elements." Nucl. Acids Res. 26(15): 3542-3549.
Nayler, O., W. Stratling, et al. (1998). "SAF-B protein couples transcription and pre-mRNA splicing to SAR/MAR elements." Nucleic Acids Res 26(15): 3542-9.
Nishikura, K. (2006). "Editor meets silencer: crosstalk between RNA editing and RNA interference." 7(12): 919-931.
Nissim-Rafinia, M., O. Chiba-Falek, et al. (2000). "Cellular and viral splicing factors can modify the splicing pattern of CFTR transcripts carrying splicing mutations." Hum. Mol. Genet. 9(12): 1771-1778.
Niwa, M., C. C. MacDonald, et al. (1992). "Are vertebrate exons scanned during splice-site selection?" 360(6401): 277-280.
Novoyatleva, T., B. Heinrich, et al. (2007). "Protein phosphatase 1 binds to the RNA recognition motif of several splicing factors and regulates alternative pre-mRNA processing." Hum. Mol. Genet.: ddm284.
Novoyatleva, T., Y. Tang, et al. (2006). "Pre-mRNA missplicing as a cause of human disease." Prog Mol Subcell Biol 44: 27-46.
Orban, T. I. and E. Olah (2003). "Emerging roles of BRCA1 alternative splicing." Mol Pathol 56(4): 191-197.
Pajares, M. J., T. Ezponda, et al. (2007). "Alternative splicing: an emerging topic in molecular and clinical oncology." The Lancet Oncology 8(4): 349-357.
Peltz, S. W., A. B. Hammell, et al. (1999). "Ribosomal Protein L3 Mutants Alter Translational Fidelity and Promote Rapid Loss of the Yeast Killer Virus." Mol. Cell. Biol. 19(1): 384-391.
Proudfoot, N. J., A. Furger, et al. (2002). "Integrating mRNA Processing with Transcription." Cell 108(4): 501-512.
Rafalska, I., Z. Zhang, et al. (2004). "The intranuclear localization and function of YT521-B is regulated by tyrosine phosphorylation." Hum. Mol. Genet. 13(15): 1535-1549.
Renz, A. and F. Fackelmayer (1996). "Purification and molecular cloning of the scaffold attachment factor B (SAF-B), a novel human nuclear protein that specifically binds to S/MAR- DNA." Nucl. Acids Res. 24(5): 843-849.
Rosenthal, J. A., H. Chen, et al. (1999). "The Epsins Define a Family of Proteins That Interact with Components of the Clathrin Coat and Contain a New Protein Module." J. Biol. Chem. 274(48): 33959-33965.
Saguez, C., J. R. Olesen, et al. (2005). "Formation of export-competent mRNP: escaping nuclear destruction." Current Opinion in Cell Biology
Nucleus and gene expression 17(3): 287-293.
Sanford, J. R., J. Ellis, et al. (2005). "Multiple roles of arginine/serine-rich splicing factors in RNA processing." Biochem Soc Trans 33(Pt 3): 443-6.
Sanford, J. R., J. D. Ellis, et al. (2005). "Reversible phosphorylation differentially affects nuclear and cytoplasmic functions of splicing factor 2/alternative splicing factor." Proc Natl Acad Sci U S A 102(42): 15042-7.

Sanford, J. R., N. K. Gray, et al. (2004). "A novel role for shuttling SR proteins in mRNA translation." Genes Dev. 18(7): 755-768.
Shatkin, A. J. and J. L. Manley (2000). "The ends of the affair: Capping and polyadenylation." 7(10): 838-842.
Shen, H. and M. R. Green (2004). "A pathway of sequential arginine-serine-rich domain-splicing signal interactions during mammalian spliceosome assembly." Mol Cell 16(3): 363-73.
Shen, J. and R. J. Kelleher, III (2007). "The presenilin hypothesis of Alzheimer's disease: Evidence for a loss-of-function pathogenic mechanism." Proceedings of the National Academy of Sciences 104(2): 403-409.
Sitz, J. H., M. Tigges, et al. (2004). "Dyrk1A Potentiates Steroid Hormone-Induced Transcription via the Chromatin Remodeling Factor Arip4." Mol. Cell. Biol. 24(13): 5821-5834.
Smith, C. W. J. and J. Valcarcel (2000). "Alternative pre-mRNA splicing: the logic of combinatorial control." Trends in Biochemical Sciences 25(8): 381-388.
Spillantini, M. G., J. C. Van Swieten, et al. (2000). "Tau gene mutations in frontotemporal dementia and parkinsonism linked to chromosome 17 (FTDP-17)." Neurogenetics 2: 193-205.
Stamm, S. (2002). "Signals and their transduction pathways regulating alternative splicing: a new dimension of the human genome." Hum Mol Genet 11(20): 2409-16.
Stamm, S., S. Ben-Ari, et al. (2005). "Function of alternative splicing." Gene 344C: 1-20.
Stamm, S., J. J. Riethoven, et al. (2006). "ASD: a bioinformatics resource on alternative splicing." Nucleic Acids Res 34(Database issue): D46-55.
Stamm, S., J.-J. Riethoven, et al. (2006). "ASD: a bioinformatics resource on alternative splicing." Nucl. Acids Res. 34(suppl_1): D46-55.
Stewart, M. (2007). "Ratcheting mRNA out of the Nucleus." Molecular Cell 25(3): 327-330.
Stickeler, E., S. D. Fraser, et al. (2001). "The RNA binding protein YB-1 binds A/C-rich exon enhancers and stimulates splicing of the CD44 alternative exon v4." Embo J 20(14): 3821-30.
Stickeler, E., F. Kittrell, et al. (1999). "Stage-specific changes in SR splicing factors and alternative splicing in mammary tumorigenesis." Oncogene 18(24): 3574-82.
Stickeler, E., V. J. Mobus, et al. (1997). "Intron 9 retention in gene transcripts suggests involvement of CD44 in the tumorigenesis of ovarian cancer." Anticancer Res 17(6D): 4395-8.
Stoilov, P., R. Daoud, et al. (2004). "Human tra2-beta1 autoregulates its protein concentration by influencing alternative splicing of its pre-mRNA." Hum Mol Genet 13: 509-524.
Stoilov, P., E. Meshorer, et al. (2002). "Defects in pre-mRNA processing as causes of and predisposition to diseases." DNA Cell Biol 21: 803-818.
Stoss, O., T. Novoyatleva, et al. (2004). "p59(fyn)-mediated phosphorylation regulates the activity of the tissue-specific splicing factor rSLM-1." Mol Cell Neurosci 27(1): 8-21.
Stoss, O., P. Stoilov, et al. (1999). "The in vivo minigene approach to analyze tissue-specific splicing." Brain Res Brain Res Protoc 4(3): 383-94.

Stoss, O., P. Stoilov, et al. (1999). "The in vivo minigene approach to analyze tissue-specific splicing." Brain Research Protocols 4: 383-394.

Swartz, J. E., Y.-C. Bor, et al. (2007). "The Shuttling SR Protein 9G8 Plays a Role in Translation of Unspliced mRNA Containing a Constitutive Transport Element." J. Biol. Chem. 282(27): 19844-19853.

Tacke, R., M. Tohyama, et al. (1998). "Human Tra2 Proteins Are Sequence-Specific Activators of Pre-mRNA Splicing." Cell 93(1): 139-148.

Tanahashi, H. and T. Tabira (2001). "Three novel alternatively spliced isoforms of the human beta-site amyloid precursor protein cleaving enzyme (BACE) and their effect on amyloid beta-peptide production." Neuroscience Letters 307(1): 9-12.

Tang, Y., T. Novoyatleva , et al. (2005). Analysis of alternative splicing in vivo using minigenes. Handbook of RNA Biochemistry. Hartmann, Wiley-VCH: 755-782.

Taveau, M., D. Stockholm, et al. (2002). "Quantification of Splice Variants Using Molecular Beacon or Scorpion Primers." Analytical Biochemistry 305(2): 227-235.

Thanaraj, T. A., S. Stamm, et al. (2004). "ASD: the Alternative Splicing Database." Nucl. Acids Res. 32(Database issue): D64-69.

Timothy W. Nilsen (2003). "The spliceosome: the most complex macromolecular machine in the cell?" BioEssays 25(12): 1147-1149.

Ule, J., K. Jensen, et al. (2005). "CLIP: A method for identifying protein-RNA interaction sites in living cells." Methods. Post-transcriptional Regulation of Gene Expression 37(4): 376-386.

Ule, J., K. B. Jensen, et al. (2003). "CLIP identifies Nova-regulated RNA networks in the brain." Science 302(5648): 1212-5.

Umeda, Y., S. Taniguchi, et al. (2004). "Alterations in human tau transcripts correlate with those of neurofilament in sporadic tauopathies." Neurosci Lett 359(3): 151-4.

Wang, J., Q.-S. Gao, et al. (2004). "Tau Exon 10, Whose Missplicing Causes Frontotemporal Dementia, is Regulated by an Intricate Interplay of Cis Elements and Trans Factors." J.Neurochem. 88: 1078-1090.

Wang, Y., J. Wang, et al. (2005). "Tau exons 2 and 10, which are misregulated in neurodegenerative diseases, are partly regulated by silencers which bind a complex comprised of SRp30c and SRp55 that either recruits or antagonizes htra2beta 1." J Biol Chem: 14230-9.

Warmuth, M., M. Bergmann, et al. (1997). "The Src Family Kinase Hck Interacts with Bcr-Abl by a Kinase-independent Mechanism and Phosphorylates the Grb2-binding Site of Bcr." J. Biol. Chem. 272(52): 33260-33270.

Watermann, D. O., Y. Tang, et al. (2006). "Splicing factor Tra2-beta1 is specifically induced in breast cancer and regulates alternative splicing of the CD44 gene." Cancer Res 66(9): 4774-80.

Weg-Remers, S., H. Ponta, et al. (2001). "Regulation of alternative pre-mRNA splicing by the ERK MAP-kinase pathway." EMBO J. 20: 4194?203.

Will, C. L. and R. Luhrmann (2001). "Spliceosomal UsnRNP biogenesis, structure and function." Curr Opin Cell Biol 13(3): 290-301.

Wong, B. R., D. Besser, et al. (1999). "TRANCE, a TNF Family Member, Activates Akt/PKB through a Signaling Complex Involving TRAF6 and c-Src." Molecular Cell 4(6): 1041-1049.

Xia, C., Z. Bao, et al. (2000). "GCIP, a Novel Human Grap2 and Cyclin D Interacting Protein, Regulates E2F-mediated Transcriptional Activity." J. Biol. Chem. 275(27): 20942-20948.

Yang, W., C. G. Lo, et al. (2001). "The Cdc42 Target ACK2 Directly Interacts with Clathrin and Influences Clathrin Assembly." J. Biol. Chem. 276(20): 17468-17473.

Yeakley, J. M., J. B. Fan, et al. (2002). "Profiling alternative splicing on fiber-optic arrays." Nat Biotechnol 20(4): 353-8.

Zhang, J., E. Berenstein, et al. (1996). "Transfection of Syk protein tyrosine kinase reconstitutes high affinity IgE receptor-mediated degranulation in a Syk-negative variant of rat basophilic leukemia RBL-2H3 cells." J. Exp. Med. 184(1): 71-79.

Zhang, Z. and A. R. Krainer (2004). "Involvement of SR Proteins in mRNA Surveillance." Molecular Cell 16(4): 597-607.

Zheng, Z. M. (2004). "Regulation of alternative RNA splicing by exon definition and exon sequences in viral and mammalian gene expression." J Biomed Sci 11(3): 278-94.

Zhou, Z., L. J. Licklider, et al. (2002). "Comprehensive proteomic analysis of the human spliceosome." Nature 419(6903): 182-5.

# Curriculum Vitae

**Name:** Yesheng Tang
**Date and Place of Birth:** 12th Feb 1976.
**Place of Birth:** Gansu, P.R.China

**Education and professional experience:**

[1991-1992] Xiangtan University Affilated Middle School,
Human, P.R.China
[1992-1994] Shanghai JiaoTong University Affiliated Middle School,
Shanghai,P.R.China
[1994-1998] Shanghai JiaoTong University, Shanghai, P.R.China
Bachelor of Science, Biological Technology, July 1998
[1998-2000] National Center for Gene Research, Chinese Academy of Sciences, Shanghai, P.R.China
Research Assistant
[2000-2002] Joint MBA program by University of the Sunshine Coast, Queensland, Australia and Shanghai JiaoTong University, Shanghai, P.R.China
Master of Business Administration, March 2002
[2000-2003] Shanghai JiaoTong University, Shanghai, P.R.China
Master of Science, Biochemistry and Molecular Biology,
March 2003
[2003- ] Friedrich-Alexander-University, Erlangen, Germany
PhD Student

**Awards received**

May,2002 Dongshi Dongfang Awards for graduate students,
Shanghai Jiaotong University, Shanghai, P.R.China

**Publications**


- **Tang, Y.,** Kishore, S., Stamm, S. Shuttling Splicing factor Tra2beta1 binds with Ribosomal Protein L3 and regulates translation. Preparing
- Novoyatleva T, Heinrich B, **Tang Y,** Benderska N, Ben-Dov C, Bracco L, Bollen M, Stamm S. Protein phosphatase 1 regulates the usage of tra2-beta1 dependent alternative exons. Human Molecular Genetics. 2007 Oct 2
- Novoyatleva T, **Tang Y,** Rafalska I, Stamm S. Pre-mRNA Missplicing as a Cause of Human Disease. Prog Mol Subcell Biol. 2006;44:27-46.
- Watermann DO, **Tang Y,** Zur Hausen A, Jager M, Stamm S, Stickeler E. (2006) Splicing factor Tra2-beta1 is specifically induced in breast cancer and regulates alternative splicing of the CD44 gene. Cancer Res. May 1;66(9):4774-80.
- Stamm S, Riethoven JJ, Le Texier V, Gopalakrishnan C, Kumanduri V, **Tang Y,** Barbosa-Morais NL, Thanaraj TA. (2006) ASD: a bioinformatics resource on alternative splicing. Nucleic Acids Res. Jan 1;34(Database issue):D46-55.
- Glatz DC, Rujescu D, **Tang Y,** Berendt FJ, Hartmann AM, Faltraco F, Rosenberg C, Hulette C, Jellinger K, Hampel H, Riederer P, Moller HJ, Andreadis A, Henkel K, Stamm S. (2005) The

alternative splicing of tau exon 10 and its regulatory proteins CLK2 and TRA2-BETA1 changes in sporadic Alzheimer's disease. J Neurochem. Feb; 96(3):635-44. Epub 2005 Dec 20.

- Stamm, S., Ben-Ari, S., Rafalska, I.,**Tang, Y.,** Zhang, Z., Toiber, D., Thanaraj, T.A., Soreq, H. Function of alternative splicing. Gene. 2005 Jan 3;344:1-20
- **Tang, Y.,** Novoyatleva , T., Benderska, N., Kishore, S., Thanaraj, T.A. and Stamm, S. (2004) Analysis of alternative splicing in vivo using minigenes. In Westhof, Bindereif, Schön and Hartmann (eds.), Handbook of RNA Biochemistry. 2/ 755–782 Wiley-VCH.
- Liu, X., Wang, H., Li, Y., **Tang, Y.,** Liu, Y., Hu, X., Jia, P., Ying, K., Feng, Q., Guan, J., Jin, C., Zhang, L., Lou, L., Zhou, Z., Han, B. (2004) Preparation of single rice chromosome for construction of a DNA library using a laser microbeam trap. J Biotechnol. 109(3):217-26.
- Feng, Q., Zhang, Y., Hao, P., Wang, S., Fu, G., Huang, Y., Li, Y., Zhu, J., Liu, Y., Hu, X., Jia, P., Zhang, Y., Zhao, Q., Ying, K., Yu, S., **Tang, Y.,** Weng, Q., Zhang, L., Lu, Y., Mu, J., Lu, Y., Zhang, LS., Yu, Z., Fan, D., Liu, X., Lu, T., Li, C., Wu, Y., Sun, T., Lei, H., Li, T., Hu, H., Guan, J., Wu, M., Zhang, R., Zhou, B., Chen, Z., Chen, L., Jin, Z., Wang, R., Yin, H., Cai, Z., Ren, S., Lv, G., Gu, W., Zhu, G., Tu, Y., Jia, J., Zhang, Y., Chen, J., Kang, H., Chen, X., Shao, C., Sun, Y., Hu, Q., Zhang, X., Zhang, W., Wang, L., Ding, C., Sheng, H., Gu, J., Chen, S., Ni, L., Zhu, F., Chen, W., Lan, L., Lai, Y., Cheng, Z., Gu, M., Jiang, J., Li, J., Hong, G., Xue, Y., Han B. (2002) Sequence and analysis of rice chromosome 4. Nature. 420(6913):316-20.
- Zhao, Q., Zhang, Y., Cheng, Z., Chen, M., Wang, S., Feng, Q., Huang, Y., Li, Y., **Tang, Y.,** Zhou, B., Chen, Z., Yu, S., Zhu, J., Hu, X., Mu, J., Ying, K., Hao, P., Zhang, L., Lu, Y., Zhang, LS., Liu, Y., Yu, Z., Fan, D., Weng, Q., Chen, L., Lu, T., Liu, X., Jia, P., Sun, T., Wu, Y., Zhang, Y., Lu, Y., Li, C., Wang, R., Lei, H., Li, T., Hu, H., Wu, M., Zhang, R., Guan, J., Zhu, J., Fu, G., Gu, M., Hong, G., Xue, Y., Wing, R., Jiang, J., Han,B. (2002) A fine physical map of the rice chromosome 4. Genome Res.12(5):817-23.
- **Tang, YS.,** Li, Y., Zhu, JJ., Hu, X., Lin, ZX., Han, B., Hong, GF. (2002) Colony PCR apply to the rice genome sequencing. Prog Biochem Biophys 29 (2): 316-318 (in Chinese)
- **Tang, YS.,** Tang, YL., Gong, GY. (2001) Regulatory effects of Enterococcus Faecium JT1701 on intestinal bacteria of humans, Journal of shanghai jiaotong university.35,1066-1070 (in Chinese)