The RNA-binding protein YT521-B plays a role in the alternative splice site selection and is regulated by tyrosine phosphorylation.

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

vorgelegt von
Ilona Rafalska
aus Kraków, Polen

2005

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

Tag der mündlichen Prüfung:	17. Juni 2005
Vorsitzender der Prüfungskomission:	Prof. Dr. Donat-Peter Häder
Erstberichterstatter:	PD Dr. Fritz Titgemeyer
Zweitberichterstatter:	Prof. Dr. Michael Wegner

To my Mom
Mojej Mamie

ACKNOWLEDGMENTS

The work presented here was performed in the Institute of Biochemistry at Friedrich-Alexander-University Erlangen-Nürnberg and it was supported by the Deutsche Forschungsgemeinschaft grants Sta399/7-1 and Sta399/7-2.

I would like to thank Prof. Stefan Stamm for giving me the opportunity to work in his lab and for his support during my Ph.D. research.

I am grateful to Prof. Michael Wegner and Prof. Lars Nitschke for being part of my committee.

Special thanks to Dr. Fritz Titgemeyer for agreeing to be my *Doktorvater* at *Nat. Fak. II* and for his help and guidance during preparations for the defense.

Many thanks to former and present lab members of the Prof. Stefan Stamm's group, Annette M. Hartmann, Oliver Nayler, Marieta Gencheva, Peter Stoilov, Bettina Heinrich, Zhaiyi Zhang, Tatyana Novoyatleva, Nataliya Benderskaya, Yesheng Tang, and Shivendra Kishore, for their help and for creating a good working atmosphere.

Ich möchte mich beim gesamten Lehrstuhl I and Lehrstuhl II von Institut für Biochemie dafür bedanken, dass ich ihre Geräte benutzen konnte.

I would like to thank Dr. Horst Wolff and Prof. Ruth Brack-Werner from the Institue of Molecular Virology, GSF in Neuherberg for performing protein shuttling assay.

I would also like to thank Loyola D'Silva, Dr. Paweł Śmiałowski and Dr. Tad Holak from the Max-Planck-Institute for Biochemistry in Martiensried for performing NMR analyses.

My gratitude to Dominique Olbert for reading and correcting the manuscript.

Dziękuję mojej Rodzinie i moim Przyjaciołom za niezmienne trwanie przy mnie i za wspieranie mnie w najtrudniejszych chwilach. Bez Was nie mogłabym istnieć.

Thank you, Timothy.

PUBLICATIONS

Parts of this work are included in the following publications:

Stoilov, P., Rafalska, I. and Stamm, S. (2002) YTH: a new domain in nuclear proteins. *Trends in Biochemical Sciences*, **27**, 495-497

Rafalska, I., Zhang, Z., Benderska, N., Wolff, H., Hartmann, A.M., Brack-Werner, R. and Stamm, S. (2004) The intranuclear localization and function of YT521-B is regulated by tyrosine phosphorylation. *Hum. Mol. Genet.*, **13**, 1535-1549

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

CONTENTS

ABBREVIATIONS

ActD	actinomycin D
ASD	alternative splicing database
ASePCR	alternative splicing electronic RT-PCR
ASF	alternative splicing factor
ATP	adenosine 5'-triphosphate
bp	base pairs
BSA	bovine serum albumin
CBs	Cajal bodies
CBP	CREB binding protein
cDNA	complementary DNA
CFTR	cystic fibrosis transmembrane conductance regulator ATP-binding cassette subfamily C member 7
CLK	CDC2-like kinase
CT	chromosome teritory
CTD	carboxyterminal domain (of RNA polymerase II)
DFC	dense fibrillar component
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
9G8	splicing factor, arginine/serine-rich 7
GC	granular component
GH	growth hormone
GnRH	gonadotrophin releasing hormone
HEK	human embryonic kidney
hnRNP	heterogenous nuclear ribonucleoprotein
ICD	interchromatin domain

ICH-1	interleukin-1β converting enzyme homologue 1
IGC	interchromatin granule cluster
IL-4	interleukin-4
IPTG	isopropylβ-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
RNA	ribonucleic acid
RNase	ribonuclease
rpm	revolutions per minute
RRM	RNA recognition motif
RT-PCR	reverse transcription followed by polymerase chain reaction
RUST	regulated unproductive splicing and translation
SAF	scaffold attachment factor (A or B)
Sam68	Src associated in mitosis 68kDa
SC35	splicing component, 35 kDa; splicing factor, arginine/serine-rich 2
SDS	sodium dodecyl sulfate
SF	splicing factor (1 or 2)
SFRS14	splicing factor, arginine/serine-rich 14
SH	Src homology domain (2 or 3)
SLM	Sam68 like molecule (1 or 2)
SMA	spinal muscular atrophy
SMN	survival motor neuron gene (1 or 2)
SNB	Sam68 nuclear body
snoRNP	small nucleolar ribonucleoprotein
snRNP	small nuclear ribonucleoprotein paricle
SR-protein	serine-arginine- rich protein
STAR	signal transduction and activation of RNA
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
YTH domain	YT521-B homology domain

ZUSAMMENFASSUNG

Die Sequenzierung des menschlichen Genoms hat gezeigt, dass Menschen eine unerwartet geringe Anzahl von Genen besitzen. Einer der Hauptmechanismen, der an der Herstellung der Proteinvielfalt aus der überraschend geringen Anzahl von Genen beteiligt ist, ist alternatives Spleißen. Daten von Mikroarray-Analysen zeigen, dass etwa 75% der menschlichen Gene alternativ gespleißt werden. Somit ist die Kontrolle des alternativen Spleißens eine wichtige Komponente der Genexpressions-Regulation. Bislang sind insbesondere die mit Signaltransduktionwegen assoziierten Mechanismen der Spleiß-Regulation kaum untersucht und verstanden.

In dieser Arbeit wurde ein Beitrag zum Verständnis des Mechanismus der Regulation des alternativen Spleißens erbracht. Die Ergebnisse konzentrieren sich auf das Protein YT521-B, das mit verschiedenen Spleißfaktoren interagiert und selbst Spleißvorgänge beeinflussen kann. YT521-B ist ein Kernprotein, das in einer neuen Kern-Substruktur, den *YT-bodies*, vorliegt. YT521-B gehört zu keiner bekannten Proteinfamilie und hat keine vorher beschriebenen RNA-Bindungsdomänen.

Im ersten Teil der Arbeit wurde eine neue Proteindomäne in YT521-B identifiziert, die phylogenetisch sehr stark konserviert ist. Diese Domäne wurde YTH, für YT521-B Homologie, genannt. Ihre vorhergesagte Sekundärstruktur besteht aus vier α-Helices und sechs β-Faltblättern. Die Konservierung von aromatischen Aminosäuren in den β-Faltblättern ist vergleichbar mit dem RNA Erkennungsmotiv (RRM). EMSA-Analysen zeigen, dass die YTH Domäne purinreiche RNA binden kann. Analysen von YTH Mutanten zeigen auch, dass mindestens 7 aus 14 konservierten Aminosäuren wichtig für die Bildung von *YT-bodies* sind.

Im zweiten Teil der Arbeit wurde gezeigt, dass Lokalisation und Funktion von YT521-B durch Tyrosin-Phosphorylierung reguliert sind. YT521-B wird durch einige spezifische Tyrosinkinasen phosphoryliert. Obwohl YT521-B im *steady state* im Zellkern lokalisiert ist, kann es durch membranständige Tyrosinkinasen phosphoryliert werden, weil YT521-B sich sehr schnell zwischen Cytosol und Zellkern bewegt. Die Tyrosin-Phosphorylierung löst *YT-bodies* auf und sorgt für eine Assoziation zwischen YT521-B in der unlöslichen Kernmatrix. Durch diese Sequestrierung reguliert die Tyrosin-Phosphorylierung den Effekt von YT521-B auf die Auswahl alternativer Spleißstellen.

ABSTRACT

Sequencing of the human genome has revealed that humans have an unexpectedly small number of genes. Alternative splicing emerges to be one of the major mechanisms involved in generating protein diversity from the relatively limited number of genes. Microarray data now show that nearly 75% of human genes produce transcripts that are alternatively spliced. In consequence, the control of splicing becomes an important component of gene regulation. To date, the mechanisms that regulate splicing, in particular the associated signal-transduction pathways, are not well understood.

The focus of this work is to understand the mechanism of the alternative splicing regulation. The results presented here concentrate on the YT521-B protein, which has been identified as a protein interacting with several splicing factors, such as hnRNP-G, Sam68, and rSAF-B. YT521-B is present in the nucleus, where it defines a novel compartment, the YT bodies. It has also been shown that YT521-B can change alternative splice site usage in a concentration-dependent manner. YT521-B does not belong to any of the known protein families and does not have in its structure any previously known RNA-binding domain.

The first part of this work shows the presence of a domain, which is conserved among proteins of several species. It is named YT521-B Homology (YTH) domain. Its predicted secondary structure consists of four α-helices and six β-strands. The conservation of aromatic residues in the β-sheets is reminiscent of the RNA recognition motif (RRM). The performed electrophoretic mobility shift (EMSA) experiments with YT521-B deletion mutant containing YTH domain show that the protein binds purine-rich RNA. In addition, mutational analysis shows that at least 7 of the 14 conserved residues are necessary for YT bodies' formation.

The second part of this work demonstrates that localization and function of YT521-B is regulated by tyrosine phosphorylation. It is shown that YT521-B is phosphorylated by specific nuclear non-receptor tyrosine kinases. It is also phosphorylated by membrane-bound kinases, as the protein shuttles between nucleus and cytosol. Tyrosine phosphorylation causes dispersion of YT521-B from YT bodies to the nucleoplasm and moves the phosphorylated protein into insoluble nuclear fraction. Finally, tyrosine phosphorylation of YT521-B changes the protein effect on alternative splice site selection of several reporter genes.

1. INTRODUCTION

The initial primary transcript synthesized by RNA polymerase II undergoes several processing steps before a mature mRNA is produced. Eukaryotic messenger RNA processing includes several major events, such as 5' capping, 3' polyadenylation (Shatkin and Manley, 2000), splicing (Neugebauer, 2002; Kornblihtt et al., 2004), and RNA editing (Wedekind et al., 2003). These processes are crucial for eukaryotic gene expression occurring in the specialized compartments of the nucleus (Dundr and Misteli, 2001). Regulation of gene expression at the posttranscriptional level is mainly achieved by RNA-binding proteins (Perez-Canadillas and Varani, 2001). This chapter introduces the process of splicing (sections 1.1.1.–1.1.3.) and focuses on regulation (section 1.1.4.) and function (section 1.1.5.) of the alternative splicing. In addition, the basic knowledge of RNA interacting domains (section 1.2.) and functional architecture of the cell nucleus (section 1.3.) is presented.

1.1. Pre-mRNA splicing

1.1.1. Basic splicing reaction and spliceosome assembly

Almost all eukaryotic pre-mRNAs are transcribed as precursors and the formation of mature mRNAs from primary transcripts requires excision of intervening intron sequences and the subsequent joining (splicing) of exons.

Introns are marked by weakly conserved elements that are essential for intron recognition and its removal (Table 1). Introns with GT-AG terimini are called U2-type introns.

Table 1. Sequence elements marking introns.

Element	Consensus sequence*
5' (donor) splice site	YRG/GURAGU
3' (acceptor) splice site preceded by a polypyrimidine stretch	Y_{12}NYAG/
Branch point located 18-200 nucleotides 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.

There is a second minor class of introns found in higher eukaryotes. These, so-called U12-type, introns have AT-AC termini. Their consensus sequences are more highly

conserved than U2-type GT-AG introns (5' splice site: /AUAUCCUU, 3' splice site: YCCAC/, branch point: TCCTTAAC) (Hall and Padgett, 1994) and they are spliced by a distinct spliceosome (Tarn and Steitz, 1996). Analysis of canonical and non-canonical splice sites in mammalian genomes shows that the 99.24% of splice site pairs has GT-AG termini, 0.69% GC-AG, 0.05% AT-AC, and finally only 0.02% consists of other types of non-canonical splice sites (Burset et al., 2000; Burset et al., 2001).

Splicing proceeds through two sequential transesterification reactions (Figure 1). The first reaction generates a 2', 5'-phosphodiester bond at the branch point upstream of the 3' splice site and a free 3'-hydroxyl group on the upstream exon. In the second reaction, the newly formed 3'-OH group attacks the phosphodiester bond at the 3' splice site. In consequence, the lariat intron with a 3'-hydroxyl group is released and two exons are joined together.

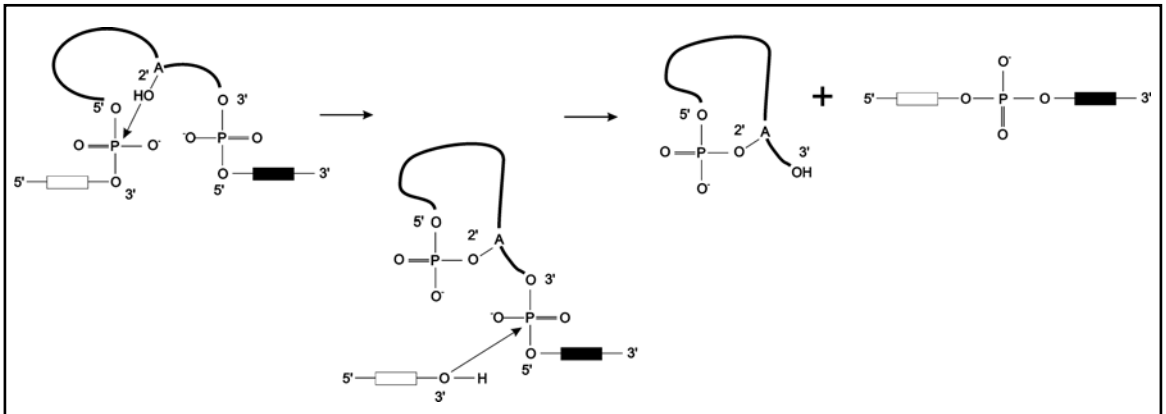


Figure 1. Pre-mRNA splicing basic reaction.

The splicing reaction is catalyzed by a large and highly dynamic ribonucleoprotein complex, the spliceosome. In size and complexity, spliceosomes are comparable to ribosomes. The spliceosome consists of the five small nuclear Ribonucleoprotein Particles (snRNPs) U1, U2, U4, U5, and U6 and 50-100 non-snRNP splicing factors (reviewed in Kramer, 1996; Burge et al., 1999). Spliceosome assembly is a dynamic process involving the recruitment of both snRNP and non-snRNP splicing factors to the pre-mRNA (reviewed in Reed and Palandjian, 1997; Burge et al., 1999; Reed, 2000; Hastings and Krainer, 2001). The 5' splice site is recognized by the U1 snRNP when 5' region of the U1 snRNA basepairs with the consensus sequence of 5' splice site (Zhuang and Weiner, 1986). The polypyrimidine tract and the AG dinucleotide

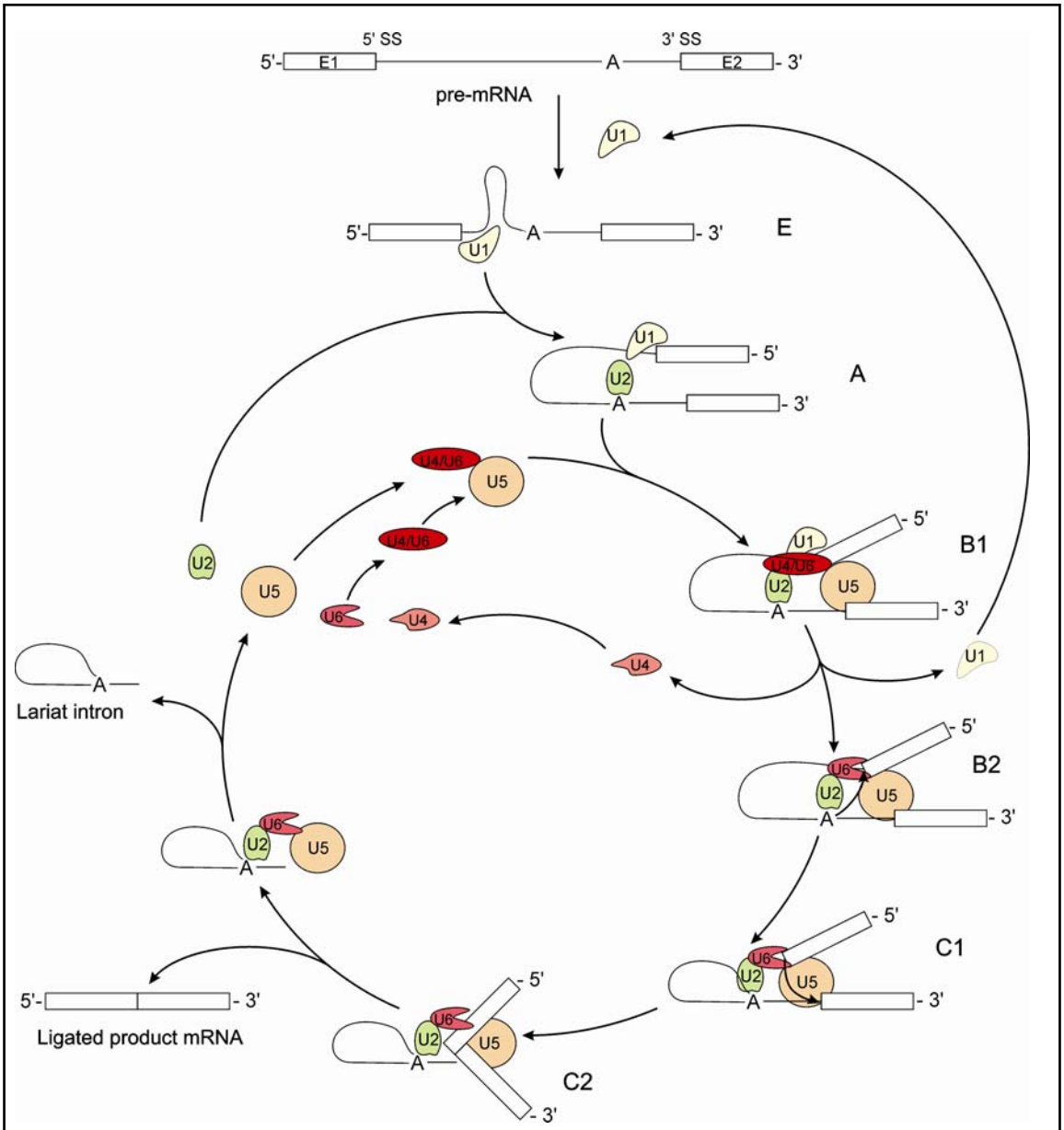


Figure 2. The spliceosome cycle. The processing of a pre-mRNA containing two exons and one intron into the ligated exon product and lariat intron is shown. Small nuclear ribonucleoprotein particles are indicated by different colors. Arrows represent transitions between complexes and/or recycling of snRNP components. Macromolecular complexes characteristic for mammalian systems are marked with capital letters. Occurrence of two transesterification reactions is indicated with thick arrows. [Adapted from (Burge et al., 1999)]

at the 3' splice site are recognized by U2 snRNP Auxiliary Factors U2AF^{65} and U2AF^{35}, respectively (Zhang et al., 1992; Wu et al., 1999). The 3' and 5' splice site recognition is facilitated by exon bridging interactions between U2AF bound to the 3' splice site and U1 snRNP bound to the downstream 5' splice site (Hoffman and Grabowski, 1992; Wu and Maniatis, 1993). Splicing Factor 1 (SF1) binds the branch point and interacts with U2AF,

which is crucial for the branch point recognition (Kramer and Utans, 1991). All these interactions allow assembly of an early (commitment) E complex. There are also alternative spliceosome assembly pathways described, such as U1- (Crispino et al., 1994), U2AF- (MacMillan et al., 1997; Kent et al., 2005), and U2AF RRM3-independent (Banerjee et al., 2004) pathways, showing that the spliceosome assembly is flexible and different pre-mRNAs have distinct rate-limiting steps. Spliceosomal A complex is formed when the U2 snRNP is recruited to the branch point sequence. Association of the U4/U6·U5 tri-snRNP with the pre-spliceosome generates the pre-catalytic B complex. The mature spliceosome is formed through extensive rearrangements of snRNP components. U1 and U4 snRNPs are released from the spliceosome and U6 snRNP replaces U1 snRNP, forming the RNA:RNA duplex with the 5' splice site. In addition, the 5' splice site is brought to the close proximity of the branch point and the 3' splice site by base-pairing between U6 and U2 snRNAs and interaction of U5 snRNP with both upstream and downstream exons. At this point, the first transesterification occurs and the intron lariat is formed. Finally, in C complex the second transesterification reaction takes place and two exons are joined while intron lariat is released. snRNPs leave the complex and are recycled for the next splicing reaction. The schematic spliceosomal cycle is presented in Figure 2 (Burge et al., 1999).

1.1.2. Exon recognition and intron bridging

Splicing regulatory elements, such as 5' and 3' splice sites and branch point, are degenerated and described only by loosely followed, short consensus sequences. It is therefore difficult to predict exons in genomic sequences, and current computer programs cannot accurately predict splice sites from genomic DNA. Approximately 60% of the exons can be identified properly when the programs are benchmarked with newly determined and experimentally annotated genomic regions (Thanaraj, 2000). This contrasts with the high accuracy and fidelity of splice site selection *in vivo*. One reason for the splicing efficiency observed *in vivo* is the presence of additional regulatory enhancer or silencer elements, which can be located within exons or introns. These elements are again characterized by loose consensus sequences and this degeneracy prevents them from interfering with the coding capacity of the exons. Several types of

enhancer/silencer elements are described, $(GAR)_n$ purine-rich (Dominski and Kole, 1994; Elrick et al., 1998), pyrimidine-rich (Pret et al., 1999) and A/C-rich (ACEs) (Coulter et al., 1997).

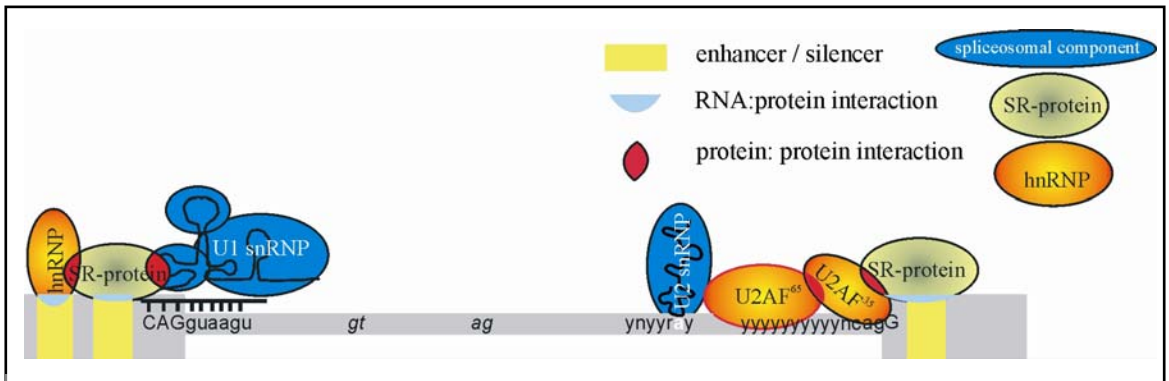


Figure 3. Elements involved in splice site selection. Exons are indicated as boxes, the intron as a thick line. Exonic splicing regulatory elements (enhancers or silencers) are shown as yellow boxes. The 5' splice-site (CAGguaagu) and 3' splice-site $(y)_{10}$ncagG, as well as the branch point (ynyyray), 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 (yellow), bind to splicing regulatory elements; the protein:RNA interaction is shown in light blue. This protein complex assembles around an exon enhancer, stabilizing binding of the U1 snRNP close to the 5' splice-site, for example due to protein:protein interaction between an SR protein and the RS domain of U170K (shown in red). This allows hybridization (thick black line with stripes) of the U1 snRNA (black) with the 5' splice-site. The formation of the multi-protein:RNA complex allows discrimination between proper splice-sites indicated at exon:intron borders and cryptic splice-sites (small *gt ag*) that are frequent in pre-mRNA sequences. Factors at the 3' splice-site include U2AF, which facilitates binding to U2 snRNP to the branchpoint sequence. (Stamm et al., 2004)

Proteins binding to regulatory sequence elements can be subdivided into two groups, serine/arginine-rich (SR) proteins (Manley and Tacke, 1996; Tacke and Manley, 1999) and heterogeneous nuclear Ribonucleoproteins (hnRNPs) (Weighardt et al., 1996; Krecic and Swanson, 1999). RNA recognition occurs through the specialized domains, such as RRM or KH motif (section 1.2.). Since the interaction between individual factor and regulatory sequence is weak and not highly specific, different SR and SR-like proteins can act through the same enhancers and influence the same splice sites. In addition, SR and SR-like proteins bridge introns by interactions among each other and through interactions with the core spliceosome components, forming this way a stable complex that marks the exon (Wu and Maniatis, 1993; Tacke and Manley, 1999; Hertel et al., 1997) (Figure 3). SR and SR-like proteins participate in the recruitment of $U2AF^{65}$ and $U2AF^{35}$ to the polypyrimidine tract and the adjacent 3' splice site (Graveley et al., 2001), as well as in the recruitment of U1 snRNP to the 5' splice site (Eperon et al., 1993;

Eperon et al., 2000) and U2 snRNP to the branch point (Tarn and Steitz, 1995; Boukis et al., 2004). These protein:protein interactions are mediated by RS domains, rich in serine and arginine residues. hnRNPs can bind to enhancers and silencers in a sequence-dependent manner (Caputi and Zahler, 2001) and can antagonize SR and SR-like proteins function. For example, hnRNP G that possibly binds to the specific AAGT sequence (Nasim et al., 2003) interacts with tra2-beta and is capable of inhibiting tra2-beta-dependent splicing (Venables et al., 2000).

The importance of the splicing regulatory sequences and their interacting proteins is emphasized by the increasing number of pathologies associated with them (reviewed in Stoss et al., 2000; Stoilov et al., 2002a; Faustino and Cooper, 2003). For example, Spinal Muscular Atrophy (SMA) is caused by loss or mutations in Survival Motor Neuron 1 gene (SMN1). A nearly identical SMN2 gene fails to compensate for the loss of SMN1 because of the translationally silent C→T conversion in exon 7. This change disrupts an exonic splicing enhancer, what in consequence leads to exon 7 skipping and generation of a truncated, less stable SMN2 protein (reviewed in Ogino and Wilson, 2002). In Frontotemporal Dementia with Parkinsonism linked to chromosome 17 (FTDP-17) tau mutations can either increase or decrease exon 10 splicing. Regulation of exon 10 splicing is complex and involves at least three different cis-acting elements. Described tau mutations can enhance an existing splicing enhancer element, create a new exon splicing enhancer element or destroy an exon splicing silencing element (D'Souza et al., 1999).

1.1.3. Alternative splicing

The sequencing of the human genome has demonstrated that humans contain only about 20,000 – 25,000 genes, which is far less then previously expected (International Human Genome Sequencing Consortium, 2004). Therefore, the generation of protein isoforms through regulated alternative splicing is an important step in gene expression of complex organisms. An average human gene contains a mean of 8.8 exons with a mean internal size of 145 bp. The mean intron length is 3365 bp and the 5' and 3' UTR are 300 and 770 bp, respectively. As a result, a "standard" gene spans about 27 kb. After pre-mRNA processing the average mRNA exported into the cytosol consists of 1340 bp coding sequence, 1070 bp untranslated regions and a poly (A) tail (Lander et al., 2001).

These numbers show that more than 90% of the pre-mRNA is removed as introns and only about 10% of the average pre-mRNA is joined as exonic sequences by pre-mRNA splicing. Human cells are not only capable of accurately recognizing the small exons within the larger intron context, but are also able to recognize exons alternatively. In this process, an exon is either incorporated into the mRNA, or is excised as a part of an intron. Initially, this was thought to be only a minor processing pathway affecting about 5% of all genes (Sharp, 1994), but over time it became clear that it is highly abundant. Computational analysis shows that 59% of the 245 genes present on chromosome 22 are alternatively spliced (Lander et al., 2001) and DNA microarray experiments indicate that 74% of all human genes are alternatively spliced (Johnson et al., 2003). The high frequency of alternative splicing in humans is also supported by EST-based database analysis indicating that 35-60% of all human gene products are alternatively spliced (Mironov et al., 1999; Brett et al., 2000; Kan et al., 2001; Lander et al., 2001; Modrek et al., 2001; Modrek and Lee, 2002), suggesting that alternative splicing of human genes is the rule, not the exception.

On average, a human gene generates 2-3 transcripts. However, extreme cases exist, 3 human neurexin genes can potentially form over 1000 transcripts due to alternative splicing and six alternative promoters (Missler and Sudhof, 1998). In *Drosophila*, the Down syndrome cell adhesion molecule (Dscam) can potentially generate 38,016 isoforms due to alternative splicing (Celotto and Graveley, 2001). This number is larger than the total number of genes present in *Drosophila*. Alternative splicing is observed in all tissues, but tissue-specific splicing is most prevalent in brain cells (Stamm et al., 2000; Xu et al., 2002). EST data comparison strongly indicates that similar levels of alternative splicing occur in evolutionarily distinct species such as human, mouse, *Drosophila* and *C.elegans*, emphasizing the importance of alternative splicing throughout evolution (Brett et al., 2002).

Most alternative splicing events can be classified into five basic splicing patterns: cassette exons, alternative 5' splice sites, alternative 3' splice sites, mutually exclusive cassette exons and retained introns (Figure 4). An estimated 75% of all alternative splicing patterns change the coding sequence (Kan et al., 2001; Okazaki et al., 2002;

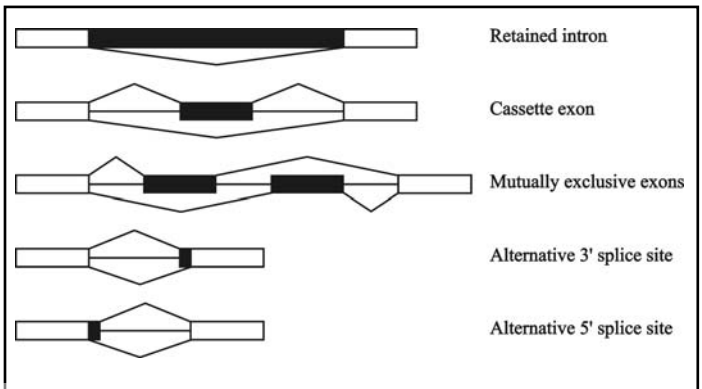


Figure 4. Types of alternative exons. Exons are indicated as boxes, introns as horizontal lines. Black color indicates alternatively spliced exons, flanking constitutive exons are shown in white.

Zavolan et al., 2003). The alternative usage of internal cassette exons is the most prominent splicing pattern. Database analysis revealed that cassette exons represent around 53% of human alternative exons (Stamm et al., 2000). More complicated alternative splicing patterns consist of combined basic patterns and are frequently observed, e.g. in the simultaneous skipping of multiple exons in the CD44 gene (Screaton et al., 1992), combination of intron retention in cassette exons of the splicing factors 9G8 (Popielarz et al., 1995) and SFRS14 (Sampson and Hewitt, 2003) and multiple alternative 3'splice sites in the gene encoding SC35 (Sureau and Perbal, 1994).

1.1.4. Regulation of the alternative splicing

Alternative splicing is tightly regulated in a tissue- or developmental stage-specific fashion through the activity of tissue-specific factors (Buckanovich et al., 1993; Elliott et al., 1996; Grabowski, 1998; Venables et al., 1999; Di Fruscio et al., 1999). In addition, changes in splice site selection may occur in response to extracellular stimuli, such as stress (Borsi et al., 1995; Denegri et al., 2001; Daoud et al., 2002), growth factors (Shifrin and Neel, 1993; Smith et al., 1997), hormones (Chalfant et al., 1998; Xie and McCobb, 1998), cytokines (McKay et al., 1994; Eissa et al., 1996), drugs (Yao et al., 1996), membrane depolarization in neurons (Zacharias and Strehler, 1996; Xie and Black, 2001) or antigenic stimulation of the T-cell receptor (Konig et al., 1998; Weg-Remers et al., 2001), and are a part of a normal adaptation process of the cell (reviewed in Stamm, 2002; Shin and Manley, 2004). These extracellular cues act through the activation of signaling pathways, such as MKK3/6-p38 or Ras-Raf-MEK-ERK, that transduce the information to the nucleus and in effect change splicing factors activation (Lynch and Weiss, 2000; van der Houven van Oordt et al., 2000; Weg-Remers et al.,

2001; Matter et al., 2002). However, in many cases signaling components between the cell surface and the nuclear splicing machinery are not yet identified. Ischemia induced in the brain effects in the hyperphosphorylation of tra2-beta1 and its translocation from the nucleus to the cytosol (Daoud et al., 2002). In addition, some of tra2-beta1 interacting proteins, like Sam68 and serine/arginine-rich proteins, also accumulate in the cytosol. Concomitant with this relocalization is a change in the alternative splice site usage of the Interleukin-1β Converting enzyme Homologue 1 (ICH-1) gene. A well studied example of hormonal influence on the alternative splicing is the incorporation of the exon βII of the Protein Kinase C (PKC) due to activation of the insulin receptor by insulin (Chalfant et al., 1998; Patel et al., 2001).

1.1.4.1. Phosphorylation-dependent control of the pre-mRNA splicing machinery

Phosphorylation and dephosphorylation of spliceosomal components are main mechanisms providing precise control during the pre-mRNA splicing process. The phosphorylation status of proteins is known to play a role in the assembly of the spliceosome, regulation of the splice site selection, and subcellular localization of splicing factors (reviewed in Soret and Tazi, 2003). Sequential phosphorylation and dephosphorylation of SR proteins mark particular steps in the splicing reaction (Cao et al., 1997; Xiao and Manley, 1997) and experiments using phosphatase inhibitors demonstrate that dephosphorylation is important for single splicing reactions (Mermoud et al., 1992). Moreover, changes in the phosphorylation level, leading to hyper- or hypophosphorylation of SR proteins, were shown to inhibit splicing (Prasad et al., 1999). Phosphorylation of serine residues recruits SR proteins from storage compartment, nuclear speckles to the sites of active transcription, resulting also in diffusion of speckles (Misteli et al., 1998; Wang et al., 1998). It was demonstrated that phosphorylation of specific splicing factors by CDC2-Like Kinases (CLK1-4) promotes exclusion of Tau exon 10 (Hartmann et al., 2001). Exon 10 of human Tau microtubule-associated protein is known to be associated with frontotemporal dementia and Parkinsonism linked to chromosome 17 (FTDP-17) (Hutton et al., 1998; Poorkaj et al., 1998; Spillantini and Goedert, 1998; Varani et al., 1999; Jiang et al., 2000; Kowalska et al., 2002). Mutations increasing the inclusion of exon 10 effect in the production of Tau protein, which

contains additional microtubule-binding sites, which can result in protein aggregation and formation of filamentous lesions (Hasegawa et al., 1999). Another example of phosphorylation-dependent regulation of splice site selection is formation of variant CD44 isoforms during immune response (Weg-Remers et al., 2001). After T-cell receptor stimulation, the Ras-Raf-MEK-ERK signaling cascade is activated and exon v5 inclusion is observed. The similar effect of the enhanced ERK-mediated exon v5 inclusion was also observed after forced expression of Sam68 and phorbol-ester stimulation (Matter et al., 2002). Tyrosine phosphorylation of Sam68-Like Mammalian proteins (SLM-1 and SLM-2) inhibits their RNA-binding activities (Haegebarth et al., 2004) and influences their splicing properties (Stoss et al., 2004).

1.1.5. Function of alternative splicing

Gene regulation through alternative splicing is more versatile than regulation through promoter activity. Variant transcripts generated through alternative splicing, similar to those initiated from distinct promoters, are often tissue and/or developmentaly specific, resulting in effects seen only in certain cells or developmental stages. However, changes in promoter activity alter predominantly the expression levels of the mRNA. In contrast, changes in alternative splicing can modulate transcript expression levels by subjecting mRNAs to nonsense-mediated decay and alter the structure of the gene product by inserting, or deleting, novel protein parts. The structural changes fall into three categories: introduction of stop codons, changes of the protein structure and changes in the 5' or 3' untranslated region. The effects caused by alternative splicing range from a complete loss of function to subtle effects that are difficult to detect. Data on functional aspects of alternative splicing is collected in the Alternative Splicing Database (ASD) at http://www.ebi.ac.uk/asd/ and reviewed in (Stamm et al., 2005). Several examples of the functional importance of alternative splicing are presented in Table 2 and described below.

Table 2. Examples of functional effects of alternative splicing.

Functional effect of alternative splicing	Example	Reference
Receptor insensitivity to a ligand	TSH receptor	(Ando et al., 2001)
Changed ligand specificity of the receptor	FGFR-2	(Miki et al., 1992)
Modified affinity of a ligand to the receptor	GnRH receptor	(Wang et al., 2001)
Differential binding between proteins	Tenascin	(Chiquet-Ehrismann et al., 1991)
Generation of soluble receptor isoforms	IL-4 receptor	(Blum et al., 1996)
Modulation of the ligand-dependent internalization of the receptor	μ-Opioid receptor	(Koch et al., 2001)
Sequestration of the receptor in membrane-enclosed compartments	mGluR1	(Chan et al., 2001)
Changed sublocalization within an organelle	MEK5	(English et al., 1995)
Regulation of the enzymatic activity	dNOS	(Stasiv et al., 2001)
Changed protein stability	FosB	(Nestler et al., 1999)
Insertion of posttranslational modification sites	Kv4.3	(Po et al., 2001)
Modulation of ion-channel properties	KCNQ2	(Pan et al., 2001)

Abbreviations: dNOS, *Drosophila* nitric-oxide synthase; FGFR, fibroblast growth factor receptor; GnRH, gonadotrophin releasing hormone; IL-4, interleukin-4; KCNQ2, potassium voltage-gated channel, KQT-like subfamily, member 2; Kv4.3, potassium voltage gated channel, Shal-related family, member 3; MEK5, MAP/ERK kinase 5; mGluR1, metabotropic glutamate receptor 1; TSH, thyroid stimulating hormone.

1.1.5.1. Introduction of stop codons

mRNAs containing premature stop codons can be degraded by Nonsense-Mediated Decay (NMD). During pre-mRNA splicing, exon-exon junctions are marked with a protein complex that influences the subsequent mRNA translation. In the absence of translation, mRNA is not subject to NMD, even when premature stop codons fulfill the NMD criteria (Maquat, 2004). About 25-35% of alternative exons introduce frameshifts or stop codons into the pre-mRNA (Stamm et al., 2000; Lewis et al., 2003). Since approximately 75% of these exons are predicted to be subject to nonsense-mediated decay, an estimated 18-25% of transcripts are switched off by stop codons introduced in alternative splicing and nonsense mediated decay (Lewis et al., 2003). This process, which has been termed RUST for Regulated Unproductive Splicing and Translation, currently represents the function of alternative splicing with the most obvious biological consequences. The exact number of genes affected by RUST is only a crude estimate, as mRNAs undergoing nonsense-mediated decay are unstable and underrepresented in cDNA libraries.

1.1.5.2. Changes of the protein structure

Approximately 75% of alternative splicing events occur in the translated regions of mRNAs and will affect the protein coding region (Okazaki et al., 2002; Zavolan et al., 2003). Changes in the protein primary structure can alter the binding properties of proteins, influence their intracellular localization and modify their enzymatic activity and/or protein stability by diverse mechanisms. One commonly found mechanism is the introduction of protein domains that are subject to posttranslational modification, such as phosphorylation.

1.1.5.2.1. Binding properties

Protein isoforms generated by alternative splicing differ in their binding properties both to small molecular weight ligands (e.g. hormones) and to macromolecules such as proteins or nucleic acids. Binding activity can be completely abolished due to alternative splicing as it happens in the TSH (Thyroid Stimulating Hormone) receptor case. Alternative variants of the TSH receptor, occurring in TSH-secreting tumors, are unable to bind TSH and cause insensitivity to TSH (Ando et al., 2001). Alternative splicing can also determine the ligand specificity of a receptor. A well-studied example is that of the Fibroblast Growth Factor Receptor gene FGFR-2, which creates two isoforms that differ by 49 amino acids in the extracellular domain. Depending on the presence of this domain, the receptor binds to both fibroblast and keratinocyte growth factor or only to fibroblast growth factor (Miki et al., 1992). Moreover, the affinity between the modified protein and its ligand can be altered. For example, binding of GnRH (Gonadotrophin Releasing Hormone) to shorter splice variants of the GnRH receptor is reduced 4-10 fold (Wang et al., 2001), which abolishes signaling. Similarly, interaction of transcription factors with DNA can be modified by alternative splicing, which contributes to transcriptional regulation (Lopez, 1995). Alternative splice variants can inhibit transactivation in a dominant negative way or can modulate transactivation. Frequently, alternative splicing does not directly affect DNA binding, but modulates the formation of complexes between various transcription factors, which in consequence regulates the affinity between transcription factor complexes and DNA (Kozmik et al., 1993; Ormondroyd et al., 1995).

1.1.5.2.2. Intracellular localization

Alternative splicing determines the intracellular localization of numerous proteins, usually by influencing localization signals or regulating the interaction of proteins with membranes. By deleting or interrupting transmembrane or membrane-association domains, soluble isoforms are generated by alternative splicing. Non-membrane bound isoforms can lose the ability to transduce signals (Kestler et al., 1995; Tone et al., 2001), they can be less stable (Garrison et al., 2001) or have a different effect on immune system modulation (Riteau et al., 2001). If the soluble isoform retains the ability to bind a ligand, it can regulate the concentration and bioactivity of that ligand, which indirectly interferes with the function of the membrane-bound form. This type of regulation has been described for the IL-4 (interleukin-4) receptor and the growth hormone binding protein (Blum et al., 1996; Rosenfeld, 1994). By deleting the transmembrane region, constitutively active molecules that do not require the activation by proteases can also be formed (Lyman et al., 1995).

Alternative splicing can determine the localization of proteins in various subcellular sites and organelles. Proteins can be sequestered into compartments, where they perform no function. This mechanism is widely used for receptor molecules and alternative splicing can regulate their retention in membrane-enclosed compartments. For example, the inclusion of an endoplasmatic reticulum retention signal in the metabotropic Glutamate Receptor mGluR1 reduces the cell surface expression of this receptor and restricts its trafficking (Chan et al., 2001). Also sublocalization of proteins within organelle can be regulated by alternative splicing. For example, in the nucleus proteins can be present in different nuclear substructures, the nucleoplasm and speckles, due to alternative splicing events (Nishizawa et al., 2001).

1.1.5.2.3. Protein stability

Inclusion of alternate protein domains can regulate the half-life of proteins. Protein stability can be altered due to autophosphorylation that signals the degradation of receptor molecules. For the interleukin-1 receptor-associated kinase, this autophosphorylation-dependent degradation is isoform-specific, leading to a molecule that is not down-regulated by its ligand (Jensen and Whitehead, 2001). The effect of

alternative splicing-dependent protein stability has been studied for the *fosB* gene, which generates a shorter isoform that is more stable than the full-length protein. This isoform accumulates in brain in a region-specific manner in response to many types of chronic behavioral changes (Nestler et al., 1999).

1.1.5.2.4. Posttranslational modifications

Posttranslational modifications can be dictated by alternative splicing, usually by generating consensus sites for phosphorylation, glycosylation, palmitoylation or sulfatation. For instance, isoform-dependent phosphorylation of the potassium channel Kv4.3 allows the modulation of outward currents by the alpha-adrenergic system via protein kinase C (Po et al., 2001). In addition, binding between a kinase and its substrate can be regulated by alternative splicing, e.g. binding of the SR-protein kinase 1 to scaffold attachment factor B (Nikolakaki et al., 2001).

1.1.5.3. Changes of the mRNA properties

Some examples have been described where alternative splicing changes the properties of the mRNA. Alternative splicing events occuring in 5' and 3' UTRs may change the stability of the RNA. For example, alternative exons in the 5' UTR of the HIV-1 virus can either promote or inhibit the nuclear degradation of their mRNA, which regulates HIV-1 gene expression (Krummheuer et al., 2001).

1.2. RNA interacting domains

Mature messenger RNA is produced as a consequence of the series of processing reactions, which include 5' end capping, pre-mRNA splicing, 3' end polyadenylation, and RNA editing (Shatkin and Manley, 2000; Zorio and Bentley, 2004; Wedekind et al., 2003; Kornblihtt et al., 2004). All these processes are mediated by RNA-binding proteins and are necessary for mRNA export to the cytoplasm and for its subsequent targeting to the ribosome. Recent studies identify an increasing number of conserved motifs responsible for RNA-binding activity. Currently, there are around 80 different RNA-binding domains deposited in Pfam database. Table 3 presents several examples of them. The best characterized are RNA Recognition Motif (RRM), hnRNP K Homology (KH) motif, and double-stranded RNA-Binding Domain (dsRBD).

Table 3. RNA-binding domains

Motif	Pfam accession number	Examples of proteins	Function	Reference
RRM	PF00076	hnRNP A1, hnRNP C1/C2, U2AF^{65}, SC35, SF2/ASF	• mRNA biogenesis • pre-mRNA splicing • polyadenylation	(Birney et al., 1993)
dsRBD	PF00035	Staufen, PKR kinase, ADAR	• translation regulation • RNA editing	(St Johnston et al., 1992)
KH motif	PF00013	hnRNP K, FMR-1, ribosomal protein S3	• pre-mRNA splicing • mRNA biogenesis	(Musco et al., 1996)
GSG domain	-	GRP33, Sam68, GLD-1	• pre-mRNA splicing • localization to SNBs • KH motif is a part of GSG domain	(Chen et al., 1999; Chen et al., 2001)
RGG box	-	Fibrillarin, hnRNP A2/B1, hnRNP R, hnRNP U	• pre-mRNA splicing • subcellular localization	(Cartegni et al., 1996; Nichols et al., 2000)
G-patch	PF01585	hnRNPA/B, CHERP (DAN26), SPF45	• pre-mRNA splicing	(Aravind and Koonin, 1999a)
LSM (Sm) domain	PF01423	snRNP Sm proteins (B/B', D1, D2, D3, E, F and G)	• pre-mRNA splicing	(Hermann et al., 1995)
CPSF A subunit region	PF03178	CPSF A (160 kDa) subunit	• splicing of single-intron pre-mRNAs	(Li et al., 2001)
PRP38 family	PF03371	Prp38	• pre-mRNA splicing	(Xie et al., 1998)
S4 domain	PF01479	ribosomal protein S4	• translation regulation	(Aravind and Koonin, 1999b)
Surp module (SWAP domain)	PF01805	CHERP (DAN26)	• pre-mRNA splicing	(Denhez and Lafyatis, 1994)
CAT RNA binding domain	PF03123	Antiterminator proteins SacY, LicT, BglG	• transcription antitermination	(Declerck et al., 1999)
Nop10p family	PF04135	Nop10p	• 18S rRNA production • rRNA pseudouridylation	(Henras et al., 1998)
S1 RNA binding domain	PF00575	ribosomal protein S1, RNase E, RNase II, NusA	• translation regulation	(Bycroft et al., 1997)
ANTAR	PF03861	AmiR	• transcription antitermination	(Shu and Zhulin, 2002)
Tudor domain	PF00567	Tudor, SMN	• transport of mRNA during oogenesis in *Drosophila*	(Ponting, 1997)
Zinc finger	PF02008	Nab2p	• poly(A) RNA transport	(Marfatia et al., 2003)
Rho termination factor RNA binding domain	PF07497	Rho protein	• transcription termination	(Bogden et al., 1999)

Abbreviations: ANTAR, AmiR and NasR transcription antitermination regulators; CAT, Co-AntiTerminator; CPSF, cleavage and polyadenylation specificity factor; GSG domain, GRP33–Sam68–GLD-1 domain; LSM domain, Sm-like domain; Nop10p, nucleolar protein 10kDa; SWAP, suppressor-of-white-apricot

1.2.1. RNA recognition motif (RRM)

RNA recognition motif is also called the Ribonucleoprotein (RNP) domain. It is one of the most common eukaryotic RNA-binding domains (Varani and Nagai, 1998; Rubin et al., 2000). One or more RRMs are found in many RNA interacting proteins, such as hnRNPs, SR proteins, snRNP-associated proteins, pre-rRNA processing factors, and poly(A)-binding proteins. RRM is 90-100 amino acids long and it follows a general pattern of β_1-α_1-β_2-β_3-α_2-β_4 (Nagai et al., 1990; Hoffman et al., 1991). It forms the four-stranded antiparallel β sheet packed against two perpendicular α helices (Figure 5A). The β_1 and β_3 strands are involved in RNA binding (Gorlach et al., 1992) and contain conserved solvent-exposed aromatic residues that provide stacking interactions with RNA (Oubridge et al., 1994).

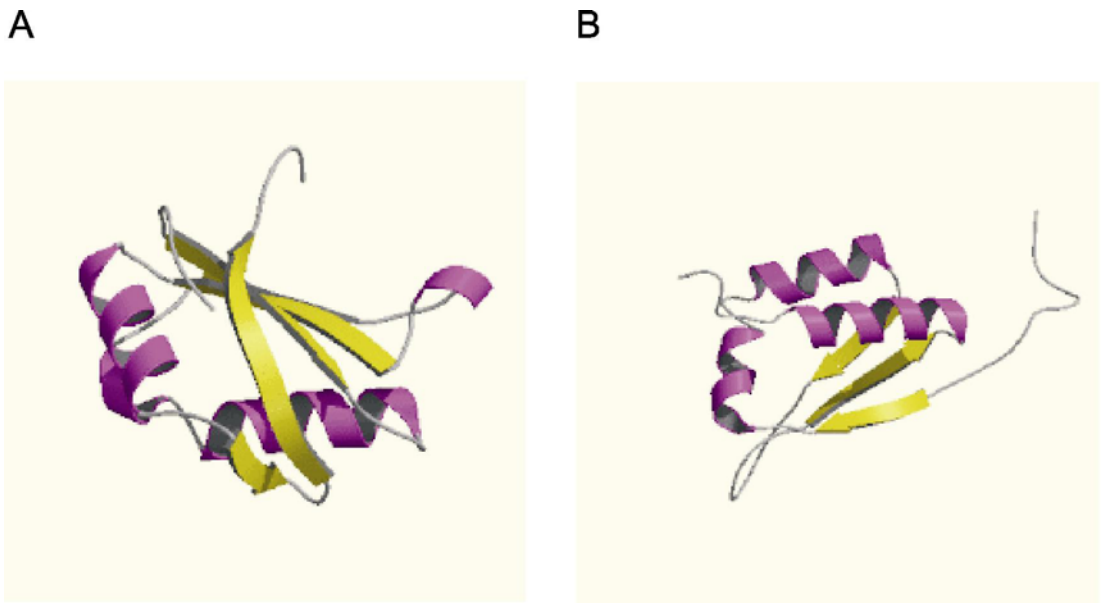


Figure 5. RNA interacting domains. (A) RNA recognition motif of U1 snRNP A, PDB code 1OIA. **(B)** KH domain of hnRNP K, PDB code 1KHM. Purple color marks α helices, yellow – β strands. Both structures were downloaded from CATH Protein Structure Classification Database UCL (www.biochem.ucl.ac.uk/bsm/cath).

1.2.2. K-homology (KH) domain

The KH domain was first identified in hnRNP K (Siomi et al., 1993) and is also present in other RNA interacting proteins, such as hnRNP E1, FMR-1, Nova, and STAR proteins (signal transduction and activation of RNA). Most KH proteins possess more

than one copy of the KH motif. The KH domain spans around 70 residues and consists of three antiparallel β strands packed against three α helices (β_1-α_1-α_2-β_2-β_3-α_3) (Figure 5B) (Musco et al., 1996; Lewis et al., 1999). The invariant Gly-X-X-Gly segment (where X represents lysine, arginine or glycine) in the loop connecting helices α_1 and α_2 and the variable loop connecting β_2 and β_3 play a central role in RNA recognition (Lewis et al., 2000).

1.2.3. Double-stranded RNA-binding domain (dsRBD)

The double-stranded RNA-binding domain was first described for the *X.laevis* XlrbpA and *Drosophila* Staufen proteins (St Johnston et al., 1992). The functions of dsRBD proteins vary and include RNA localization, RNA editing, and translational repression. Most of these proteins have multiple, tandemly arranged dsRBDs, although several single copy dsRBD proteins have also been described. dsRBDs are approximately 70 amino acids in length and have common α_1-β_1-β_2-β_3-α_2 structure, where a three-stranded antiparallel β sheet is packed against two helices (Figure 6) (Bycroft et al., 1995). The recognition of dsRNA and discrimination against dsDNA is mediated by two highly conserved basic loops contacting a sugar-phosphate backbone without any direct base contact. Loop 2 interacts with several 2'-OH groups in the minor groove of the dsRNA type-A helix and loop 4 interacts with the phosphodiester backbone across the major groove from the sites of loop 2 contacts (Ryter and Schultz, 1998).

Figure 6. The double-stranded RNA binding domain. dsRBD of XlrbpA, PDB code 1DI2, was downloaded from CATH Protein Structure Classification Database UCL. Purple color marks α helices, yellow – β strands.

1.3. Nuclear organization and subnuclear bodies

The cell nucleus was one of the first intracellular structures identified in microscopic observations in the 19th century. It was originally described by Franz Bauer in 1802 and later popularized by Robert Brown (1833) (Harris, 1999). Initially nucleus was believed to have very little organization. The most recent studies show high compartmentalization within the nucleus and growing number of specialized subnuclear domains (Lamond and Earnshaw, 1998; Spector, 2001; Dundr and Misteli, 2001; Platani and Lamond, 2004) (Figure 7).

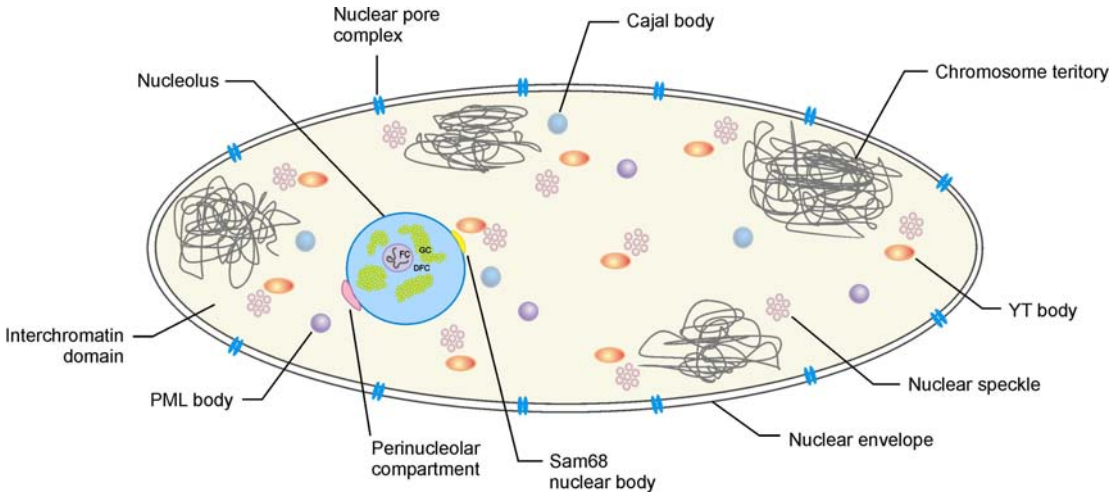


Figure 7. Diagram of subnuclear compartments. The diagram shows the nucleus surrounded by the nuclear envelope. The relative locations of chromosome territories and the principal classes of nuclear bodies are indicated. [Adapted from (Platani and Lamond, 2004) and (Spector, 2001)]

The nucleus is surrounded by a double membrane, of which the outer membrane is contiguous with a rough endoplasmic reticulum. Transport of molecules between the nucleus and cytoplasm occurs through protein complexes located in the nuclear envelope, Nuclear Pore Complexes (NPCs) (Stoffler et al., 1999; Fahrenkrog et al., 2004).

Within the nucleoplasm, the chromosomes are arranged into large-scale domains, termed Chromosome Territories (CTs) (Cremer et al., 2000; Cremer and Cremer, 2001). They occupy distinct regions of the nucleus and have a modular and dynamic organization. Three distinguishable parts can be described for CTs, an "open" higher-order chromatin compartment (euchromatin) with chromatin domains containing active

genes, a "closed" chromatin compartment (heterochromatin) comprising inactive, silenced genes, and an Interchromatin Domain (ICD) compartment that contains macromolecular complexes for transcription, splicing, DNA replication, and repair.

The best-characterized example of a subnuclear body is the nucleolus (Scheer and Hock, 1999). It is the site of ribosome subunit biogenesis. The nucleolus is differentiated into three distinct morphological regions, Fibrillar Centers (FCs), Dense Fibrillar Component (DFC), and Granular Component (GC). The FCs are surrounded by DFCs, and the GCs emerge from DFC regions. The rDNA genes reside at the FCs and fibrillar centers are thought to be the interphase equivalent of the Nucleolar Organizer Regions (NORs) of chromosomes. Transcription and processing of rRNA is thought to occur in DFCs. All the processing and modification events are carried out by a complex array of nucleolar factors, including the small nucleolar Ribonucleoproteins (snoRNPs) (Filipowicz and Pogacic, 2002). The GC region is made up of pre-ribosomal particles at different stages of maturation, as well as large and small ribosomal subunits.

The Perinucleolar Compartment (PNC) and the Sam68 Nuclear Body (SNB) are structures detected at the periphery of the nucleolus (Huang, 2000). Despite their different nuclear localizations, PNCs and SNBs share some common characteristics. They are predominantly observed in transformed cells and are rarely found in normal primary cells. They both are enriched with RNA binding proteins and nucleic acids and are thought to play a role in RNA metabolism. HnRNP proteins, splicing factors, and small RNAs transcribed by RNAPol III are present in PNCs (Ghetti et al., 1992; Matera et al., 1995; Wang et al., 2003), whereas STAR proteins, such as Sam68, SLM-1, and T-STAR (human SLM-2), are localized to SNBs (Chen et al., 1999).

PML bodies, also termed PODs (PML Oncogenic Domains), ND10 (Nuclear Domain 10), or Kr bodies, are discrete interchromosomal accumulations of several proteins, including PML, Sp100, SUMO1 and CBP (Maul et al., 2000). It has been suggested PML bodies play a role in aspects of transcriptional regulation and appear to be targets of viral infections.

Pre-mRNA splicing factors are localized in a pattern of 25–50 nuclear speckles as well as being diffusely distributed throughout the nucleoplasm (Spector, 1990; Lamond and Spector, 2003). Many of the larger speckles correspond to Interchromatin Granule

Clusters (IGCs) that are involved in the assembly and/or modification of pre-mRNA splicing factors. Nuclear speckles are thought to be sites of storage of mRNA splicing factors, from which splicing factors are recruited to sites of transcription (Misteli et al., 1997).

Some of pre-mRNA splicing factors are also localized in Cajal Bodies (CBs), previously called coiled bodies (Gall, 2000). These dynamic nuclear bodies are thought to play a role in snRNP biogenesis and in the trafficking of snRNPs and snoRNPs, which move through the Cajal body en route to splicing speckles or nucleoli, respectively.

Gems (gemini of Cajal bodies) are found in the nucleoplasm and are coincident with or adjacent to Cajal bodies. They are characterized by their association with the survival of motor neurons gene product (SMN) and an associated factor, Gemin2 (Matera, 1999). SMN and Gemin2 / SIP1 form a complex that plays an essential role in cytoplasmic snRNP biogenesis (Fischer et al., 1997).

YT bodies containing splicing factor YT521-B (section 1.4.) are dot-like structures located in close contact to speckles and Sam68 nuclear bodies (Nayler et al., 2000). They contain focal sites of transcription and disperse upon actinomycin D treatment accumulating protein in the unsoluble nuclear fraction. The effect of actinomycin D on YT bodies differentiates them from nuclear speckles, which round up and increase in size after transcription inhibition with actinomycin D (Misteli et al., 1997).

1.4. YT521-B as a putative splicing factor

YT521 was first isolated using rat tra2-beta1 (RA301) as a bait in a yeast two-hybrid screen (Imai et al., 1998), but the interaction between tra2-beta1 and YT521 was never shown *in vivo*. The level of YT521 mRNA was shown to be upregulated in astrocytes exposed to hypoxia / reoxygenation. At the same time, an alternatively spliced variant, YT521-B, was isolated via its interaction with hnRNP G and STAR proteins, Sam68, SLM-1, SLM-2 (Hartmann et al., 1999). The interaction of YT521-B with Sam68 was shown to be regulated by tyrosine phosphorylation emanating from SRC family kinases. YT521-B has modular structure and consists of several distinguishable domains: a glutamic acid-rich region, a proline-rich domain containing putative SH3 domain

binding sites, a glutamic acid/arginine-rich region involved in protein:protein interaction, and four putative nuclear localization signals (Figure 8). YT521-B is ubiquitously

Figure 8. Domain structure of YT521-B. The four nuclear localization signals (NLS) are indicated with numbers. The E-rich region, the proline-rich region, and the ER-rich region are indicated by different shading. The inserted A indicates the location of the alternative exons of YT521-B.

expressed and is localized in the nucleoplasm. It defines a novel compartment in the nucleus, the YT bodies (Figure 9) (Nayler et al., 2000). YT bodies are dynamic structures that disappear during mitosis and are most pronounced in G1. YT bodies contain sites of active transcription and are in close proximity to speckles and Sam68 nuclear bodies, which contain proteins interacting with YT521-B. When all three RNA polymerases are blocked by high concentrations of actinomycin D, YT bodies disperse throughout the nucleoplasm and accumulate in an unsoluble nuclear fraction (Nayler et al., 2000). Moreover, it was shown that YT521-B can change splice site selection of htra2-beta and SRp20 mRNAs in a concentration-dependent manner (Hartmann et al., 1999). Finally, it was shown recently that YT521-B interacts with emerin, a nuclear membrane protein associated with Emery-Dreifuss muscular dystrophy (EDMD) (Wilkinson et al., 2003). YT521-B-dependent influence on splice site selection of CD44 exon v5 is inhibited in the presence of emerin, suggesting a possible role of YT521-B in EDMD pathogenesis.

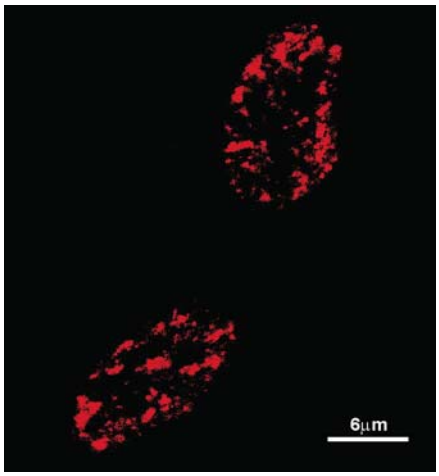


Figure 9. YT521-B localizes to the subnuclear compartment, YT bodies. Endogenous YT521-B was immunostained with specific PK2 antibodies (Nayler et al., 2000) and visualized with CY3-coupled anti-rabbit secondary antibodies.

2. RESEARCH OVERVIEW

Pre-mRNA processing requires proteins containing RNA-binding motifs in their structures. Most of the splicing factors include in their structures at least one RNA-binding domain. YT521-B was shown previously to be involved in the splice site selection, but it does not belong to any of the known RNA-binding protein families. The focus of the first part of this work was the characterization of a novel domain present in YT521-B. This domain is possibly involved in the interaction with RNA. Both the computational analysis and experimental data are presented here.

The second part of this work concerns the regulation of pre-mRNA splicing. Mechanisms controlling the alternative splice site selection are so far not very well described. It is therefore important to investigate the effect signaling pathways have on splicing factors. The goal of this work was to determine the effect of tyrosine phosphorylation on splicing factor YT521-B. It was shown previously that YT521-B is a target for tyrosine kinases. In studies presented here, it was tested which kinases can phosphorylate YT521-B. The interaction between YT521-B and kinases was analyzed and the influence of phosphorylation on both subnuclear localization and biological activity of YT521-B was determined.

3. MATERIALS AND METHODS

3.1. Materials

3.1.1. Chemicals

Product	Supplier	Product	Supplier
30% Acrylamide/Bis solution	Sigma	β-Mercaptoethanol	Merck
Agar	GibcoBRL	Methanol	Roth
UltraPure agarose	Invitrogen	Ni-NTA Agarose	Qiagen
Ampicilin	Sigma	Nonidet P-40 / Igepal CA-630	Sigma
Aprotinin	SIgma	dNTPs	Invitrogen
[γ-^{32}P]-ATP	Amersham	Paraformaldehyde	Merck
Benzonase	Sigma	PEG 3500	Sigma
Boric acid	Roth	Perhydrol 30% H_2O_2	Merck
Bradford reagent (BioRad Protein Assay)	BioRad	Phenol: Chloroform: Isoamyl alcohol	Sigma
Brilliant Blue R 250	Sigma	PMSF	Sigma
Bromophenol blue	Merck	Poly[C]/[U]/[G]/[A] Agarose Beads	Sigma
Chloramphenicol		Potassium chloride	Merck
Chloroform: Isoamyl alcohol	Sigma	2-Propanol	Roth
Deoxycholic acid	Sigma	Protease Inhibitor Cocktail	Sigma
ssDNA/dsDNA Cellulose	Sigma	Protein A Sepharose	Amersham
Dextrose	Sigma	RNase Inhibitor	Roche
DMSO	Sigma	SDS	Sigma
DTT	Merck	Sepharose CL-4B	Pharmacia
EDTA	Merck	Silver Stain Plus	BioRad
Ethanol	Roth	Sodium acetate	Merck
Ethidium bromide	Sigma	Sodium chloride	Roth
Ficoll 400	Fluka	Sodium dihydrogen phosphate	Merck
Gelatin	Sigma	Sodium fluoride	Sigma
Glycerol	Sigma	Sodium hydroxide	Merck
Glycerol 2-phosphate	Sigma	Sodium orthovanadate	Sigma
Glycin	Roth	Sodium pyrophosphate	Merck
HEPES	Sigma	di-Sodiumhydrogen phosphate	Merck
Imidazole	Roth	TEMED	Sigma
p-Iodophenol	Sigma	Tris base	Aldrich
IPTG	Sigma	TRIzol	Sigma
Kanamycin	Sigma	Triton X-100	Sigma
Luminol	Sigma	Tryptone	Sigma
Lysozyme	Sigma	Tween 20	Sigma
Magnesium chloride	Merck	Yeast Extract	Sigma
Magnesium sulfate	Sigma	Xylene cyanole FF	Sigma

3.1.2. Enzymes

Product	Supplier	Product	Supplier
EcoRI	NEB	AmpliTaq DNA polymerase	Roche
NotI	NEB	Platinum Pfx polymerase	Invitrogen
SacII	NEB	T4 PNK	NEB
DpnI	NEB	T7 DNA Polymerase	NEB
SuperScript II	Invitrogen	FastLink T4 DNA Ligase	Biozym

3.1.3. Cell lines and media

Name	Description	ATCC number
COS-7	African green monkey kidney SV40 transformed	CRL-1651
HEK293	Human embryonic kidney transformed with adenovirus 5 DNA	CRL-1573
Neuro-2a	Mouse neuroblastoma	CCL-131

COS-7, HEK293 and Neuro-2a cells were maintained in DMEM supplemented with 10% fetal calf serum (both from GibcoBRL). For subculturing cells, 1 × Trypsin/EDTA (GibcoBRL) was used.

3.1.4. Bacterial strains and media

Strain	Genotype	Reference
BL21(DE3)-RIL	*ompT hsdS*(r_B m_B) *dcm*$^+$ Tet^r *gal* λ(DE3) endA Hte [*argU ileY leuW* Cam^r]	(Studier et al., 1990)
XL1-Blue MRF'	Δ(*mcrA*)183 Δ(*mcrCB-hsdSMR-mrr*) 173 *endA1 supE44 thi-1 recA1 gyrA96 relA1 lac* [F′ *proAB lacI*^{q}*ZΔM15* Tn*10* (Tet^r)]	(Bullock et al., 1987)
CJ236	F′ *cat (pCJ105 = pOX38::cat= FΔ(HindIII)::cat [Tra*$^+$ *Pil*$^+$ *Cam*^{R}*]/ ung-1 relA1 dut-1 thi-1 spoT1*	(Kunkel et al., 1987)

LB MEDIUM (1L):
10 g NaCl
10 g tryptone
5 g yeast extract

LB AGAR (1L):
10 g NaCl
10 g tryptone
5 g yeast extract
20 g agar

3.1.5. Antibiotics

Antibiotic	Stock solution		Working concentration	
	concentration	storage	liquid culture	plates
Ampicilin	50 mg/ml in H_2O	4°C	100 µg/ml	100 µg/ml
Chloramphenicol	30 mg/ml in ethanol	–20°C	15 µg/ml	30 µg/ml
Kanamycin	20 mg/ml in H_2O	4°C	20 µg/ml	20 µg/ml

3.1.6. Antibodies

Antibody	Organism	Dilution used for western blotting	Supplier
anti Abl	Mouse	1:2000	Santa Cruz
anti FLAG M2	Mouse	1:1000	Sigma
anti GFP	Mouse	1:4000	Roche
anti pTyr (PY20)	Mouse	1:5000	Santa Cruz
anti YT521-B (PK2)	Rabbit	1:3000	custom made *
anti YTH	Rabbit	1:1000	custom made **

* Antiserum was raised against a mixture of two YT521-B peptides: P1 RSARSVILIFSVRESGKFQCG and P2 KDGELNVLDDILTEVPEQDDECG (Nayler et al., 2000)

** Antiserum was raised against a YT521-B peptide: CVRESGKFQGFARLSSE

3.1.7. Plasmids

3.1.7.1. Clones from the lab collection or outside sources

Name	Backbone	Description	Reference
pEGFP-C2	pEGFP-C2	CMV promoter, Kan^r/Neo^r, f1 ori	Clontech
pht6-Fl-C2	pEGFP-C2	YT521-B EGFP-tagged	(Hartmann et al., 1999)
pht6-Fl-FLAG	pcDNA3	YT521-B FLAG-tagged	O. Nayler
c-src wt	pcDNA3.1	c-Src kinase	(Wong et al., 1999)
pRK5-abl	pRK5	c-Abl kinase	O. Nayler
Sik YF	pcDNA3	Constitutively active Sik kinase	(Derry et al., 2000)
pRK5-fyn	pRK5	Fyn kinase	O. Nayler
pSVL-Syk	pSVL	Syk kinase	(Zhang et al., 1996)
CSK	pcDNA3	CSK kinase	O. Nayler
AUG1(pcDNA3-Rlk)	pcDNA3	Rlk kinase	(Debnath et al., 1999)
pUHG10-3(FER)	pUHD10-3	FerH kinase	(Hao et al., 1991)
pRK5-c-src-KA	pRK5	Catalytic inactive c-Src kinase	O. Nayler
pRK5-fyn-KA	pRK5	Catalytic inactive Fyn kinase	O. Nayler
pCR3.1 MGtra	pCR3.1TA	Tra2-beta minigene	(Stoilov et al., 2004)
pXB (X16)	pCRneo	SRp20 minigene	(Jumaa and Nielsen, 1997)
pET v5	Exontrap	CD44 v5 minigene	(Konig et al., 1998)
pMTE1A	pMT	Adenoviral E1A minigene	(Caceres et al., 1994)
IL4R Lng MG	Exontrap	IL-4R long minigene	(Rafalska et al., 2004)
SV9/10L/11	Exontrap	Tau minigene	(Gao et al., 2000)
pXGH5 wt ISEm1 ISEm2	pXGH5	GH1 wt minigene GH1 ISE mutant minigeneGH1 ISE mutant minigene	(Ryther et al., 2003)
7T3 p3#10 p3#6	pSI	CFTR exon 9 minigene (12GT/7T) CFTR exon 9 minigene (11GT/5T) CFTR exon 9 minigene (10GT/9T)	(Nissim-Rafinia et al., 2000)

3.1.7.2. Newly made clones

Name	Backbone	Description	Tag
pET-DelKpn	pET28a(+)	YT521-B TH ΔNLS4 deletion mutant	HIS
pht6-HTa	pFastBac HTa	Full length YT521-B in the insect expression vector	HIS
K364L pht6	pEGFP-C2	YT521-B mutant (K→L) within YTH domain	EGFP
S365A pht6	pEGFP-C2	YT521-B mutant (S→A) within YTH domain	EGFP
W380D pht6	pEGFP-C2	YT521-B mutant (W→D) within YTH domain	EGFP
L390E pht6	pEGFP-C2	YT521-B mutant (L→E) within YTH domain	EGFP
L402E pht6	pEGFP-C2	YT521-B mutant (L→E) within YTH domain	EGFP
F404D pht6	pEGFP-C2	YT521-B mutant (F→D) within YTH domain	EGFP
S405A pht6	pEGFP-C2	YT521-B mutant (S→A) within YTH domain	EGFP
F412D pht6	pEGFP-C2	YT521-B mutant (F→D) within YTH domain	EGFP
G414I pht6	pEGFP-C2	YT521-B mutant (G→I) within YTH domain	EGFP
W431D pht6	pEGFP-C2	YT521-B mutant (W→D) within YTH domain	EGFP
W450D pht6	pEGFP-C2	YT521-B mutant (W→D) within YTH domain	EGFP
N469D pht6	pEGFP-C2	YT521-B mutant (N→D) within YTH domain	EGFP
R478N pht6	pEGFP-C2	YT521-B mutant (R→N) within YTH domain	EGFP
D479K pht6	pEGFP-C2	YT521-B mutant (D→K) within YTH domain	EGFP
S54A pht6	pEGFP-C2	YT521-B mutant (S→A)	EGFP
S54E pht6	pEGFP-C2	YT521-B mutant (S→E)	EGFP
P133G pht6	pEGFP-C2	YT521-B mutant (P→G)	EGFP
S140A pht6	pEGFP-C2	YT521-B mutant (S→A)	EGFP
S140E pht6	pEGFP-C2	YT521-B mutant (S→E)	EGFP
Y253F pht6	pEGFP-C2	YT521-B mutant (Y→F)	EGFP
S287A pht6	pEGFP-C2	YT521-B mutant (S→A)	EGFP
S287E pht6	pEGFP-C2	YT521-B mutant (S→E)	EGFP
S311A pht6	pEGFP-C2	YT521-B mutant (S→A)	EGFP
S311E pht6	pEGFP-C2	YT521-B mutant (S→E)	EGFP

3.1.8. Primers

3.1.8.1. Primers used for huYT521-B gene

Name	Orientation	Sequence
F1-Exon1	sense	ATGGCGGCTGACAGTCGGG
R1-Exon3	antisense	CTCAGAACCATCTGGCGTAGGAG
F2-Exon2	sense	ATGGAGAACTTAATGTTCTGGATG
R2-Exon5	antisense	CTGCATATGACTCTGATGCAGAG
F3-Exon4	sense	AGAATTGGGCTTGAAGTGGATAG
R3-Exon7	antisense	CTTCGCTTTGGCAAGAGACAC
F4-Exon6	sense	GTTCAGAAAAGAAGCATGAGAAATT
R4-Exon9	antisense	CTGCAAATCCAGTCAATTTTAAAGA
F5-Exon8	sense	GGTGTATGGTCCACGCTCCCT
R5-Exon11	antisense	CCTTCCCACATCCCGGACTG
F6-Exon10	sense	GCGTGAATTACCCTTCACTAAGTC
R6-Exon13	antisense	CTGTCCACTTCCTGGTATCGTG

Name	Orientation	Sequence
F7-Exon12	sense	GCGTCGACCAGAAGATTATGATAT
R7-Exon15	antisense	CATTCCTTGCCAAGGTGGTG
F8-Exon14	sense	TTTCACAAATCTTATTCCCAACAG
R8-Exon 17	antisense	TCTTCTATATCGACCTCTCTCCCC

3.1.8.2. Primers used for cloning

Name	Orientation	Sequence	Name of the newly made clones
EcoRI-DelKpn	sense	CCGAATTCGAATATGAACAGGATG	pET-DelKpn
DelKpn-STOP-NotI	antisense	GCGGCCGCTCAGGTACCGGATGGTG	
EcoRI-pht6-HTa	sense	GAATTCGCCACCATGGCGG	pht6-HTa
pht6-NotI-HTa	antisense	GCGGCCGCTTATCTTCGATAACGACCTCTTTCCCC	

3.1.8.3. Primers used for mutagenesis

Name	Sequence	Introduced mutation	Name of the generated clone
K364L	TTTTTCCTCATACTGAGTAACAACCAT	aag → ctg	K364L pht6
S365A	TTCCTCATAAAGGCTAACAACCATGAG	agt → gct	S365A pht6
W380D	GCAAAGGGTGTAGATTCCACATTACCT	tgg → gat	W380D pht6
L390E	AATGAGAAGAAAGAAAATCTTGCGTTT	tta → gaa	L390E pht6
L402E	AGGAGTGTTATAGAAATATTTTCTGTC	tta → gaa	L402E pht6
F404D	GTTATATTAATAGATTCTGTCAGGGAA	ttt → gat	F404D pht6
S405A	ATATTAATATTTGCTGTCAGGGAAAGT	tct → gct	S405A pht6
F412D	GAAAGTGGAAAGGATCAAGGTTTTGCC	ttt → gat	F412D pht6
G414I	GGAAAGTTTCAAATTTTTGCCAGATTG	ggt → att	G414I pht6
W431D	TCTCCTATACATGATGTGCTTCCAGCA	tgg → gat	W431D pht6
W450D	TTTAAAATTGACGATATTTGCAGGCGT	tgg → gat	W450D pht6
N469D	ACCAATCCCTGGGATGAACATAAGCCA	aat → gat	N469D pht6
R478N	GTAAAGATTGGAAATGATGGACAGGAA	cgt → aat	R478N pht6
D479K	AAGATTGGACGTAAAGGACAGGAAATT	gat → aaa	D479K pht6
S54A	GAAAGAATGGAAGCTATTGACACCAAG	tct → gct	S54A pht6
S54E	GAAAGAATGGAAGAAATTGACACCAAG	tct → gaa	S54E pht6
P133G	CAACCAGAAAAAGGTTGTCTACGGAAA	cct → ggt	P133G pht6
S140A	CGGAAAAGGGATGCTGAAAGAAGGGCC	tct → gct	S140A pht6
S140E	CGGAAAAGGGATGAAGAAAGAAGGGCC	tct → gaa	S140E pht6
Y253F	GAGGAAGAAGAATTTGAACAGGATGAG	tat → ttt	Y253F pht6
S287A	TTCACAGATGGAGCTGTCAGGTCTGGT	tct → gct	S287A pht6
S287E	TTCACAGATGGAGAAGTCAGGTCTGGT	tct → gaa	S287E pht6
S311A	GCTCGAGGCATAGCTCCCATTGTCTTT	tca → gct	S311A pht6

Name	Sequence	Introduced mutation	Name of the generated clone
S311E	GCTCGAGGCATAGAACCCATTGTCTTT	tca → gaa	S311E pht6

3.1.8.4. Primers used for RT-PCR

Name	Orientation	Sequence	Target
N3Ins N5Ins	antisense sense	CTCCCGGGCCACCTCCAGTGCC GAGGGATCCGCTTCCTGCCCC	CD44 v5 and IL4R Lng minigenes
IL4R-Ex9 Rev IL4R-intron Rev IL4R-Ex7 For	antisense antisense sense	TTGGTAATGCTGAAGTAACAGAACA CAGCACAATGACTGTGACCTC TACAAGGAACCCAGGCTGA	Endogenous IL4R
X16R T7	antisense sense	CCTGGTCGACACTCTAGATTTCCTTTCATTTGACC TAATACGACTCACTATAGGG	SRp20 minigene
INS3 INS1	antisense sense	CACCTCCAGTGCCAAGGTCTGAAGGTCACC CAGCTACAGTCGGAAACCATCAGCAAGCAG	Tau minigene
8Ri5 F10Rx3	antisense sense	TGCATTAATGCTATTCTGATTC TTGGCATGCTTTGATGACGC	CFTR minigenes
hGH-RT2 Exon5 Rev Exon2 For	antisense antisense sense	GGACAAGGCTGGTGGGCACTGG CACAGCTGCCCTCCACAGAGC CCATCGTCTGCACCAGCTGGC	GH1 minigenes
E1A Rev E1A For	antisense sense	CTCAGGCTCAGGTTCAGACACAGG GTTTTCTCCTCCGAGCCGCTCCGA	E1A minigene
pCR3.1-RT MGTra-Xho MGTra-Bam	antisense antisense sense	GCCCTCTAGACTCGAGCTCGA GGGCTCGAGTACCCGATTCCCAACATGACG GGGCCAGTTGGGCGACCGGCGCGTCGTGCGGGG	Tra2-beta minigene

3.2. Methods

3.2.1. 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 allow completion of amplification of all products.

3.2.2. Plasmid DNA isolation

Plasmid DNA was isolated from small-scale bacterial cultures by alkaline lysis with SDS first described by Birnboim and Doly (Birnboim and Doly, 1979). Briefly, bacterial cells carrying the plasmid were cultured overnight at 37°C with vigorous shaking in 3 ml of LB medium containing the appropriate antibiotic. The cells were harvested by centrifugation for 30 sec at 14000 rpm in a microfuge. The pellet was resuspended in 150 µl of buffer P1 and the equal volume of lysis buffer P2 was added. The lysis was performed for 5 min at RT and then 150 µl of neutralization buffer P2 were added. After centrifugation for 10 min at 14000 rpm the resulting supernatant was precipitated by adding 1 volume of isopropanol. DNA was pelleted by centrifugation for 10 min at 14000 rpm, washed with 70 % ethanol, air-dried and dissolved in 30 µl of buffer TE.

Large amounts of plasmid DNA were prepared with use of the Qiagen Plasmid Maxi Kit according to the manufacturer's protocol.

BUFFER P1:
50 mM Tris-HCl, pH 8.0
10 mM EDTA
100 µg/ml RNase A

BUFFER P2:
200 mM NaOH
1% SDS

BUFFER P3:
3 M Potassium acetate, pH 5.5

BUFFER TE:
10 mM Tris-HCl, pH 8.0
1 mM EDTA

3.2.3. Determination of nucleic acid concentration

The concentration of nucleic acids in solution was estimated using a spectrophotometer (Eppendorf BioPhotometer 6131). The absorbance of the solution was measured at 260 nm and concentration was calculated using following formulas:

1 OD_{260} = 50 µg/ml for double stranded DNA

1 OD_{260} = 37 µg/ml for single stranded DNA

1 OD_{260} = 40 µg/ml for RNA

3.2.4. 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, λ = 260 nm.

1 × TBE:
90 mM Tris-borate
20 mM EDTA

6 × GEL–LOADING BUFFER:
0.25% bromophenol blue
0.25% xylene cyanol FF
15% Ficoll 400 in dH_2O

3.2.5. 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 × CRYSTAL VIOLET GEL–LOADING BUFFER:
0.25% crystal violet
15% Ficoll 400 in dH_2O

3.2.6. Site-directed mutagenesis of DNA

Site-directed mutagenesis was performed according to the method described by Kunkel (Kunkel, 1985). The DNA of interest was cloned into a vector carrying the f1 phage origin of replication and thus capable of existing in both single- and double-stranded forms. The recombinant plasmid was transformed into *E.coli* strain CJ236 deficient in dUTPase (*dut*) and uracil N-glycosylase (*ung*). These mutations result in a number of uracils being substituted for thymine in the nascent DNA. After transformation bacteria were grown on plates containing chloramphenicol in addition to the plasmid-specific antibiotic, to ensure the presence of the F' episome necessary for production of helper phage. To isolate single-stranded DNA from the plasmid of interest, colonies were grown in 5 ml of LB medium for 90 min and then 5 x 10^8 pfu of helper phage M13KO7 (New England BioLabs) was added. The culture was grown overnight at 37°C and single-stranded DNA was isolated with the Qiagen M13 kit according to the manufacturer's protocol. This uracil-containing DNA was used as a template in the *in vitro* mutagenesis reaction. Phosphorylated oligonucleotides containing desired mutations were annealed to

the template at a molar ratio of 20:1 in 10 µl of 1 x T7 DNA Polymerase buffer. The DNA was denatured for 5 min at 94°C and then the temperature was gradually decreased from 70°C to 37°C at a rate of 1°C per minute. The extension of the annealed primer was carried out in 20 µl by adding to the same tube 1 µl of 10 x T7 DNA Polymerase buffer, 0.8 µl of 10 mM dNTPs, 1.5 µl of 10 mM ATP, 3 U T7 DNA Polymerase and 2 U FastLink T4 DNA Ligase. The reaction was incubated at 37°C for 45 min. The ligase was inactivated by incubation at 65°C for 20 min. The mutagenesis reaction was transformed into competent XL1Blue *E.coli* cells. Replication of the plasmid in this strain leads to repair of the template strand and consequently to production of plasmid carrying the desired mutation. All mutant plasmids were verified by sequencing.

3.2.7. 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 (OD_{600} = 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 were stored at -80°C afterwards.

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

3.2.8. 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 brought with water up 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 × KCM BUFFER:
500 mM KCl
150 mM $CaCl_2$
250 mM $MgCl_2$

3.2.9. Expression and purification of HIS-tagged proteins in *E.coli*

To express HIS-tagged TH ΔNLS4 YT521-B deletion mutant, pET-DelKpn construct was transformed into BL21(DE3)-RIL *E.coli* strain. After the transformation cells were plated on LB agar plate containing both kanamycin (to select plasmid containing bacteria) and chloramphenicol (to maintain pACYC plasmid coding for additional *argU*, *ileY*, and *leuW* tRNAs). Single colony was then inoculated into 100 ml of LB medium and grown overnight. The next day this culture was inoculated into 500 ml of fresh LB medium to OD_{600} around 0.4. When OD_{600} reached 0.7 1 M IPTG was added to final concentration of 1 mM and cultured was grown for another 2 hr at 37°C with vigorous shaking. After the induction cells were harvested by centrifugation for 30 min at 4000 rpm. The pellet was resuspended in 10 ml of lysis buffer and lysozyme was added to final concentration of 1 mg/ml. After the 30-minute incubation cells were sonicated with six 10-second bursts at 300 W with a 10 sec cooling period between each burst and later sheared 10 times with 1.10 × 30 mm needle. The cellular debris was pelleted by centrifugation for 60 min at 4000 rpm and supernatant was then pushed through 0.45 μm filter. 0.5 ml of equilibrated with the lysis buffer Ni-NTA Agarose was added to the bacterial lysate followed by the incubation on the rotating wheel for 1 hour at room temperature. The agarose beads were spun down for 7 min at 800*g* and then washed 3 times with the washing buffer 80, 3 times with the washing buffer 100 and subsequently protein was eluted twice with the elution buffer.

LYSIS BUFFER:
50 mM Tris, pH 8.0
500 mM NaCl
0.05% Triton X-100
20 mM Imidazole

WASHING BUFFER 80:
50 mM Tris, pH 8.0
500 mM NaCl
0.05% Triton X-100
80 mM Imidazole

WASHING BUFFER 100:
50 mM Tris, pH 8.0
500 mM NaCl
0.05% Triton X-100
100 mM imidazole

ELUTION BUFFER:
50 mM Tris, pH 8.0
500 mM NaCl
0.05% Triton X-100
500 mM Imidazole

3.2.10. Determination of protein concentration

Protein concentration was estimated using BioRad Protein Assay Kit based on Bradford method. Protein in 800 µl of distilled water was mixed with 200 µl of 1 × Dye Reagent and incubated for 5 min at RT. Absorbance of the solution was measured in a spectrophotometer at λ = 595 nm. Concentration of samples was read from the standard curve where OD_{595} was plotted versus concentration of BSA standards.

3.2.11. Nucleic acid binding assay

Purified TH ΔNLS4 YT521-B deletion mutant was subjected to binding with ribonucleotide homopolymers, double-stranded DNA and single-stranded DNA (Carmichael, 1975; Kiledjian and Dreyfuss, 1992). Ribohomopolymers attached to agarose beads or deoxyribonucleic acid-cellulose were added to 1 µg of protein in 1 × binding buffer, 5 mM DTT, 2.5 mM $MgCl_2$. The reaction was incubated for 15 minutes at 4°C, the beads were then pelleted by centrifugation and washed five times in binding buffer. Bound protein was eluted from the beads by boiling in SDS–PAGE sample buffer for 5 minutes and run on a 10% SDS–PAGE gel (section 3.2.21.). Resulting bands were visualized by silver staining (section 3.2.22.) or Western Blot using anti YTH antibodies (section 3.2.23.).

5 × BINDING BUFFER:
50 mM Hepes, pH 7.5
500 mM KCl
0.1% Triton X-100
50% Glycerol

3.2.12. RNA electrophoretic mobility shift assay (RNA gel shift)

RNA gel shift assay was performed according to the protocols reviewed previously (Black et al., 1998; Thomson et al., 1999). 1 µg of purified TH ΔNLS4 recombinant protein was pre-incubated for 8 minutes at room temperature in 1 × binding

buffer (section 3.2.11.), 150 µg/ml BSA, 5 mM DTT, and 0.6 µl of 50 mM $MgCl_2$ in a final volume of 15 µl. After radiolabeled RNA oligonucleotides were added, the mixture was incubated for 15 more minutes. 15 µl of 10% glycerol as a sample buffer were added and RNA:protein complexes were separated on 4% native polyacrylamide gel in 0.5 × Tris / Glycine buffer at 4°C. The gel was pre-run for 30 minutes at 4°C before loading the samples.

5 × TRIS / GLYCINE BUFFER (1L):
30.28 g Tris base
142.7 g Glycine
3.92 g EDTA

4% NATIVE PA GEL FOR GEL SHIFT (30 ml):
4.02 ml 30% AA/Bis 37.5:1
3 ml 5 × Tris / Glycine buffer
1.5 ml 50% Glycerol
20 µl TEMED
300 µl 10% APS
21.16 ml dH_2O

3.2.12.1. Radiolabeling of RNA oligonucleotides

100 pmol of RNA oligonucleotides were mixed with 5 µl of 10 mCi/ml, 6000 Ci/mmol [γ-^{32}P]-ATP in 1 × PNK buffer. 10 units of T4 polynucleotide kinase were added and the reaction was incubated at 37°C for 30 minutes. Radiolabeled oligos were purified using Qiagen QIAquick Nucleotide Removal Kit according to the manufacturer's protocol.

RNA oligonucleotide	Sequence
Purine-rich	GAGGAGGAAAGAGGAGAGAGAAAGGAGG
Pyrimidine-rich	CUCCUCCUUUCUCCUCUCUCUUUCCUCCUU
TraBS3	GUGGGGAAGAUGAAAGAAGUCAGAAUUUG

3.2.13. 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 × 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.

3.2.14. 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 applictions 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 dropwise, 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 × HBS:
280 mM NaCl
10 mM KCl
1.5 mM $Na_2HPO_4 \times 2\ H_2O$
12 mM Dextrose
50 mM Hepes
pH 6.95

3.2.15. Fixing attached eukaryotic cells on cover slips

Cells grown on cover slips and transfected with pEGFP-C2 constructs were fixed with 4% formaldehyde in 1 x PBS, pH 7.4 for 20 min at 4°C. Cells were washed 3 times in PBS prior to mounting on microscope slides with Gel/Mount (Biomeda). The cells were examined by confocal laser scanning microscopy (Leica).

3.2.16. Immunostaining

Cells were fixed in 4% formaldehyde and PBS for 20 min at 4°C, washed three times in PBS and 0.1% Triton X-100, and blocked in PBS, 0.1% Triton X-100, and 3% NGS for 2 hours at room temperature. Cells were then incubated with anti Abl antibody (diluted 1:1000 in PBS, 0.1% Triton X-100, and 3% bovine serum albumin) overnight at 4°C and washed three times in PBS and 0.1% Triton X-100. Cells were then incubated with CY3-coupled secondary antibody (Dianova), diluted 1:1500 in PBS and 0.1% Triton X-100, for 2 hours at room temperature, washed three times in PBS and 0.1% Triton X-100, and mounted on glass slides with Gel/Mount. Finally the cells were examined by confocal laser scanning microscopy.

3.2.17. Quantification of cell staining

Cells were analyzed with confocal microscopy and single images were analyzed using Photoshop. First, the cell nucleus was outlined. Next, YT bodies were marked through visual inspection and the mean of their integrated signals „–“ standard deviations was calculated as YT bodies. The black areas were marked by visual inspection and the mean of their integrated distribution „+” standard deviation was designated as areas void of YT521-B. The remaining area was defined as dispersed YT521-B staining.

3.2.18. Immunoprecipitation of proteins

24 h after the transfection (section 3.2.14.), cells were washed with 1 × PBS and lysed for 20 min at 4°C in 200 µl of RIPA buffer. The lysates were collected in Eppendorf tubes and cleared by centrifugation for 1 min at 12000 rpm. The supernatant was diluted with 3 volumes of RIPA rescue buffer. The antibody recognizing the expressed protein or the tag was added to the lysate and incubated at 4°C on the rotating wheel. After 90 min of incubation 50 µl of Protein A Sepharose / Sepharose CL-4B (1:1) were added to the solution and the incubation continued overnight under the same conditions. The Sepharose beads were pelleted by centrifugation for 1 min at 1000 rpm in a microcentrifuge followed by 3 washes with 500 µl of HNTG buffer. Finally, 30 µl of 3 × SDS sample buffer were added to the pellet and boiled for 5 min at 96°C to denature the proteins. The beads were spun down and the supernatant loaded on SDS-

polyacrylamide gel (section 3.2.21.). Proteins resolved on the gel were transfered to nitrocellulose membrane and analyzed by Western blot and ECL using the appropriate antibodies (section 3.2.23.).

Protein A Sepharose / Sepharose CL-4B was made as follows: Protein A Sepharose beads were washed in 15 ml of distilled H_2O and pelleted at 500 rpm for 2 min. After a second wash with dH_2O equal volume of Sepharose CL-4B was added and the beads were washed two more times in RIPA rescue buffer and kept at 4°C.

RIPA:
1% NP40
1% Sodium deoxycholate
0.1% SDS
150 mM NaCl
10 mM Na-phosphate, pH 7.2
2 mM EDTA
50 mM NaF
5 mM β-glycerophophate
freshly added:
4 mM Sodium orthovanadate
1 mM DTT
1 mM PMSF
20 µg/ml Aprotinin
100 U/ml Benzonase

RIPA RESCUE:
20 mM NaCl
10 mM Na-phosphate, pH 7.2
1 mM NaF
5 mM β-glycerophosphate
freshly added:
2 mM Sodium orthovanadate
1 mM DTT
1 mM PMSF
20 µg/ml Aprotinin

HNTG WASH:
50 mM HEPES, pH 7.5
150 mM NaCl
1 mM EDTA
10% Glycerol
0.1% Triton X-100
freshly added:
2 mM Sodium orthovanadate
100 mM NaF
1 mM PMSF
20 µg/ml Aprotinin

3.2.19. Solubility assay

24 hours after the transfection (section 3.2.14.), cells were lysed for 20 min on ice in 200 µl of HNTG buffer. The lysates were centrifuged at 13000*g* for 15 min at 4°C, 100 µl 3 x SDS sample buffer was added to the supernatant, and 300 µl 3x SDS sample buffer was added to the remaining pellet. The probes were mixed, boiled for 5 min, and 30 µl of each fraction loaded onto 7.5% SDS-polyacrylamide gels. Alternatively, cells were lysed for 20 min on ice in 200 µl of RIPA buffer. Lysates were centrifuged for 15 min at 4°C, 100 µl 3 x SDS sample buffer was added to the supernatant, and 300 µl 3 x SDS sample buffer was added to the pellet. 30 µl of the fractions was loaded in each lane and analyzed

on 7.5% SDS-polyacrylamide gels. The proteins were analyzed by Western blotting and ECL using anti GFP and PK2 antibodies.

HNTG LYSIS BUFFER:
50 mM Hepes, pH 7.5
150 mM NaCl
1% Triton X-100
10% Glycerol
1 mM EDTA
20 mM Sodium pyrophosphate
2 mM Sodium orthovanadate
100 mM NaF
5 mM ß-glycerophosphate
1 mM PMSF
1 µg/ml Aprotinin
100 U/ml Benzonase

RIPA LYSIS BUFFER:
0.01 M Sodium phosphate
1% NP-40
1% Sodium deoxycholate
0.1% SDS
0.15 M NaCl
2 mM EDTA
1 mM NaF
4 mM Sodium orthovanadate
5 mM ß-glycerolphosphate
1 mM PMSF
1 µg/ml Aprotinin
100 U/ml Benzonase

3.2.20. Isolation of nuclear protein 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 12000*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, pushed through a narrow-gauge needle and centrifuged. Immunoprecipitation with anti GFP antibody and Protein A Sepharose was performed overnight at 4°C on the supernatant, 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.

BUFFER A:
10 mM HEPES, pH 8.0
1.5 mM $MgCl_2$
10 mM KCl
1 mM DTT

BUFFER C:
20 mM HEPES, pH 8.0
25% (v/v) Glycerol
420 mM NaCl
0.2 mM EDTA, pH 8.0
1 mM DTT
0.5 mM PMSF

NET-TRITON:
150 mM NaCl
50 mM Tris-HCl, pH 7.4
0.1% Triton X-100
Protease inhibitors

3.2.21. Electrophoresis of proteins

Proteins were resolved on denaturing SDS polyacrylamide gels. Typically, the BioRad gel electrophoresis system was used, to make 10 cm × 7.5 cm × 0.5 mm gels. The separating gel was 7.5-15%, depending on the molecular weight of the proteins, and the stacking gel was 4%. The proteins were mixed with sample loading buffer, denatured at 96°C for 5 min and loaded on the gel. Electrophoresis was carried out at 100 V for 2 hours in SDS gel running buffer.

Separating gel (10 ml)	**7.5%**	**10%**	**15%**	**Stacking gel (10 ml)**	**4%**
dH_2O	4.85 ml	4.1 ml	2.5 ml	**dH_2O**	6.1 ml
1.5 M Tris-HCl, pH 8.8	2.5 ml	2.5 ml	2.5 ml	**0.5 M Tris-HCl, pH 6.8**	2.5 ml
10% SDS	100 µl	100 µl	100 µl	**10% SDS**	100 µl
30% Acrylamide/Bis	2.5 ml	3.3 ml	5.0 ml	**30% Acrylamide/Bis**	1.3 ml
10% Ammonium Persulfate	100 µl	100 µl	100 µl	**10% Ammonium Persulfate**	100 µl
TEMED	10 µl	10 µl	10 µl	**TEMED**	10 µl

3 x SDS SAMPLE BUFFER:
150 mM Tris-HCl, pH 6.8
6% SDS
30% Glycerol
3% β-Mercaptoethanol
0.3% Bromophenol blue

SDS GEL RUNNING BUFFER:
250 mM Glycine, pH 8.3
25 mM Tris
0.1% SDS

3.2.22. Staining of protein gels

Coomassie staining was used to detect proteins in SDS polyacrylamide gels. After electrophoresis the gel was placed in staining solution (2.5% Coomassie Brilliant Blue R250, 45% Methanol, 10% Acetic acid) for 2-3 h at RT. After this incubation the gel was washed 2-3 times for 30 min in 50% Methanol/10% Acetic acid and 2-3 more times in 20% Methanol/10% Acetic acid. Alternatively, polyacrylamide gels were stained with Silver Stain Plus solutions according to the manufacturer's protocol.

3.2.23. Western blotting

Proteins resolved on SDS polyacrylamide gels were transferred to nitrocellulose membrane (Schleicher & Schuell) in transfer buffer, for 45 min at 120 V. The membrane was blocked for 1 hour in 1 x NET-gelatine buffer at 4°C, then primary antibody was added and the incubation was allowed to proceed overnight at 4°C. The membrane was washed three times for 20 min in 1 x NET-gelatine and incubated with a secondary antibody coupled to horseradish peroxidase for 1 hour. The membrane was subsequently washed three times for 20 min in 1 x NET-gelatine and the bound antibodies were detected by the ECL system. Equal amounts of solutions ECL1 and ECL2 were mixed and added to the membrane for 5 min. The membrane was then exposed to an X-ray film (Fuji SuperRX) and developed in a Kodak developing machine.

TRANSFER BUFFER:
192 mM Glycine
25 mM Tris
20 % Methanol

NET-GELATINE:
150 mM NaCl
5 mM EDTA
50 mM Tris-HCl, pH 7.5
0.05 % Triton X-100
0.25 % Gelatine

ECL1:
4.5 mM Luminol
4.3 mM p-Iodophenol
100 mM Tris, pH 9.5

ECL2:
0.003 % H_2O_2
100 mM Tris, pH 9.5

3.2.24. *In vivo* splicing assay

To determine the influence of a protein on the splicing of selected minigenes, *in vivo* splicing was performed as described (Stoss et al., 1999; Tang et al., 2005). 1 to 2 µg of the minigene plasmid were transfected in eukaryotic cells together with an expression construct for the protein. Usually a concentration-dependent effect was assessed. The protein was transfected in increasing amounts, in the range of 0 to 3 µg. To avoid 'squelching' effects, the 'empty' parental expression plasmid containing the promotor was added in decreasing amounts, to ensure a constant amount of transfected DNA. Cells were plated in 6-well plates and transfection (section 3.2.14.) was done 24 hours after plating. After incubation for 14-17 hours at 3% CO_2 total RNA was isolated from the cells (section 3.2.23.).

400 ng of RNA were used in a reverse transcription reaction (section 3.2.24.). The reverse primer used for RT was specific for the vector in which the minigene was cloned, to suppress reverse transcription of the endogenous RNA. To avoid the problem of the amplification of minigene DNA, DpnI restriction enzyme was added into the reverse transcription reaction. DpnI cuts GATC sequence in double-stranded DNA when the adenosine is methylated but does not cut non-methylated single-stranded DNA or cDNA. A control reaction with dH_2O instead of RNA was included.

1/8 of the reverse transcription reactions were used for PCR with minigene-specific primers (section 3.2.24.). The primers were selected to amplify alternatively spliced minigene products. A control reaction with no template (RNA instead of cDNA) was included in the PCR. The PCR programs were optimized for each minigene in trial experiments.

Minigene	PCR conditions
Tra MG	94°C 2 min; 33 cycles – 94°C 20 sec, 65°C 20 sec, 72°C 40 sec; 72°C 2 min
CD44 v5 MG	94°C 5 min; 30 cycles – 94°C 20 sec, 72°C 50 sec; 72°C 7 min
SRp20 MG	94°C 2 min; 30 cycles – 94°C 30 sec, 55°C 1 min, 72°C 1 min; 72°C 20 min
Tau MG	94°C 2 min; 30 cycles – 94°C 1 min, 60°C 1 min, 72°C 48 sec; 72°C 10 min
CFTR MGs	94°C 3 min; 35 cycles – 94°C 1 min, 55°C 30 sec, 72°C 1 min; 72°C 7 min
GH1 MGs	94°C 3 min; 25 cycles – 94°C 45 sec, 72°C 90 sec; 72°C 8 min
IL-4R Lng MG	94°C 5 min; 30 cycles – 94°C 20 sec, 72°C 90 sec; 72°C 7 min
pMTE1A*	94°C 2 min; 20 cycles – 94°C 30 sec, 65°C* 30 sec, 72°C 1 min; 10 cycles – 94°C 30 sec, 55°C 30 sec, 72°C 45 sec; 72°C 20 sec

* Touchdown PCR – annealing temperature was decreasing from 65°C to 55°C, 0.5°C every cycle.

PCR reactions were resolved on a 0.3-0.4 cm thick 2 % agarose TBE gel and the image was analyzed using ImageJ analysis software (http://rsb.info.nih.gov/ij/).

3.2.25. 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 40 µl of RNase-free dH_2O.

Alternatively, in RNA immunoprecipitation procedure RNA was isolated from Sepharose beads using TRIzol reagent according to the manufacturer's protocol. After ethanol precipitation, the RNA pellet was dissolved in 20 µl of RNase-free dH_2O.

3.2.26. 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 held in 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 (see section 3.2.22. for conditions of amplifying minigene products from *in vivo* splicing assays).

RT BUFFER:
300 µl 5 × First strand synthesis buffer (Invitrogen)
150 µl 0.1 M DTT (Invitrogen)
75 µl 10 mM dNTPs
475 µl dH_2O

3.3. Databases and computational tools

Database / software	URL	Description	Reference
ASD	http://www.ebi.ac.uk/asd	The alternative splicing database	(Thanaraj et al., 2004)
ASePCR	http://genome.ewha.ac.kr/ASePCR/	Web-based application emulating the RT-PCR in various tissues	(Kim et al., 2005)
ClustalW	http://www.ebi.ac.uk/clustalw/index.html	Multiple sequence alignment program for DNA or proteins	(Thompson et al., 1994)
TIGR EGO	http://www.tigr.org/tdb/tgi/ego/orth_search.shtml	Database for orthologous genes in eukaryotes	(Lee et al., 2002)
Human BLAT Search	http://www.genome.ucsc.edu/cgi-bin/hgBlat	Sequence alignment tool similar to BLAST	(Kent, 2002; Kent et al., 2002)
MyHits Motif Scan	http://myhits.isb-sib.ch/cgi-bin/motif_scan	Scans protein sequences for the occurrence of patterns, profiles and motifs	(Pagni et al., 2004)
NCBI BLAST and PSI-BLAST	http://www3.ncbi.nlm.nih.gov/BLAST/	Finds regions of sequence similarity	(Altschul et al., 1990; Altschul et al., 1997)
NCBI HomoloGene	http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=homologene	Automated detection of homologs among the annotated genes of several completely sequenced eukaryotic genomes	(Wheeler et al., 2005)

Database / software	URL	Description	Reference
PSORT II	http://psort.nibb.ac.jp/form2.html	Prediction of subcellular localization of proteins	(Nakai and Kanehisa, 1992)
Scansite Motif Scan	http://scansite.mit.edu/motifscan_seq.phtml	Searches for motifs within proteins that are likely to be phosphorylated by specific protein kinases	(Obenauer et al., 2003)
SMART	http://smart.embl-heidelberg.de/	Allows identification of genetically mobile domains and analysis of domain architectures	(Schultz et al., 1998; Letunic et al., 2002)

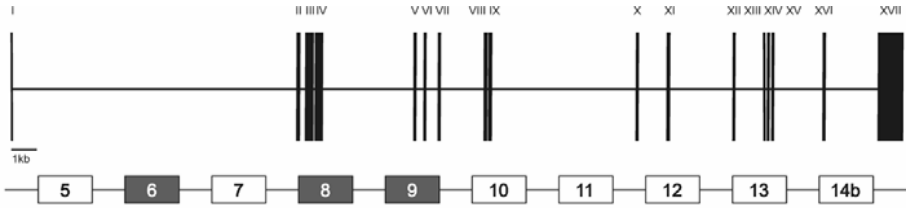
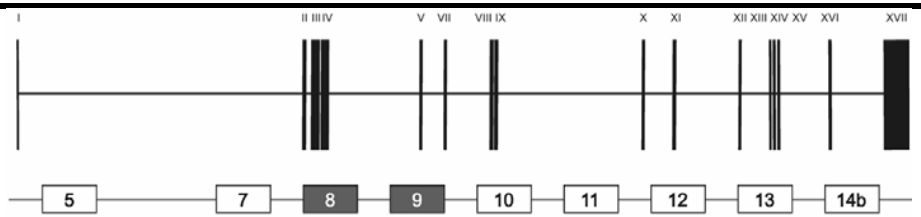
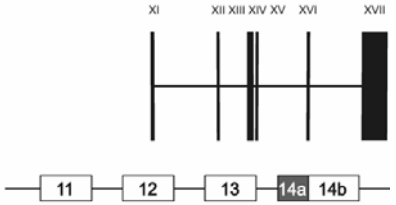
4. RESULTS

The transcripts of most metazoan protein-coding genes are alternatively spliced. However, mechanisms that are involved in the control of splicing are not well understood. YT521-B is known to be involved in alternative splice site selection. It is also a tyrosine phosphorylated nuclear protein. It has therefore a potential of being a target for molecules involved in signaling pathways. Studies presented in this chapter show that tyrosine phosphorylation emanating from non-receptor tyrosine kinases can alter both intranuclear localization and function of YT521-B. In addition, the existence of a novel, possibly RNA-binding motif is reported here. Finally, based of phylogenetic conservation profile, it is suggested that YT521-B might be a vertebrate-specific splicing factor.

4.1. Human YT521-B gene structure

Rat YT521-B was described as the protein interacting with several factors implicated in splicing, such as hnRNP-G, Sam68, and rSAF-B. It was also shown to be involved in splice selection of different mRNAs (Hartmann et al., 1999). Since finishing the Human Genome Project, a complete genome sequence is available and thus it is possible to search for human proteins homologous to previously characterized factors. To determine human YT521-B gene structure, the GenBank non-redundant nucleotide database was searched with the rat YT521-B sequence. The human transcript is predicted to be 3.3 kb with a single open reading frame of 2184 nt. The GenBank mRNA sequence was subsequently inspected using Human BLAT Search (Kent, 2002). The search revealed that the gene encoding for YT521-B protein is located on chromosome 4q13.3, on the reverse complement strand in positions 69182418 – 69219187. This analysis also showed that two YT521-B splice variants are deposited in GenBank. One consists of 17 exons, whereas the other one consists of 16 exons with the exon VI being skipped. In addition, the Alternative Splicing Database (ASD) (Thanaraj et al., 2004) was searched. There are two more isoforms deposited in the ASD. One with 24-nucleotide insertion within exon XIV is confirmed by five ESTs and the second one with exons VIII and IX being skipped is confirmed by one EST. All YT521-B isoforms together with their gene

Table 4. YT521-B isoforms.

Description	Event	Confirming EST / mRNA sequences	Gene structure based on EST sequence
ISOFORM 1 • 17 exons	-	complete mRNA sequence BC041119	
ISOFORM 2 • skipped exon 6	Cassette exon	complete mRNA sequence BC053863	
ISOFORM 3 • exon 14 with 24-nt inclusion	Alternative 3' splice site	BU508352 BU509020 CF144386 BU785242 BU785378	
ISOFORM 4 • exons 8 and 9 skipped	Cassette exon	BU659487	

Exons are indicated as boxes or vertical lines. Introns are shown as horizontal lines. Gene structure based on EST sequences is drawn to scale. Enlargement of alternative regions with numbering based on the actual exon position is given.

structures are shown in Table 4. Subsequently, the alternative splicing (AS) electronic RT-PCR (ASePCR) with primers flanking putative alternative exons was performed. ASePCR is a web-based application that emulates RT-PCR in various tissues (Kim et al., 2005). Results of this analysis are given in Table 5. ASePCR shows that alternatively spliced isoforms of human YT521-B, similarly to rat protein, can be expressed in various tissues.

Table 5. Results of ASePCR performed with human primers flanking putative alternative exons.

Splicing event	Primers	Product size (bp)*	Tissue information**
Skipping of exon 6	Exon 4 Forward AGAATTGGGCTTGAAGTGGATAG Exon 7 Reverse CTTCGCTTTGGCAAGAGACAC	**663** 609	bladder, blood, brain, cervix, eye, liver, lung, muscle, pancreas, pituitary gland, placenta, prostate, skin
24-nt inclusion within exon 14	Exon 13 Forward GGTATTTAAAGGATCCACGATACCA Exon 16 Reverse TACTCGTTTATCTCTGTATCTTGCTTCA	299 **275**	Blood, bone**,** bone marrow, brain, breast, cervix, colon, eye, kidney, larynx, liver, lung, muscle, pancreas, prostate, skin, stomach, testis
Skipping of exons 8 and 9	Exon 7 Forward ATCAAACCAGTAAACTCAAATATGTGCT Exon 10 Reverse CTGTCCATCACGTCCGATCTTTA	**407** 180	blood, bone, brain, breast, cervix, colon, eye, kidney, liver, lung, muscle, pancreas, pituitary gland, prostate, skin

* More abundant, not alternatively spliced isoforms are shown in bold.
** Tissues where both isoforms are present are underlined, tissues without underlining represent those where only more abundant isoforms are present.

Exon-intron boundaries of human YT521-B gene are shown in Table 6. Length of the exons range from ~ 40 bp (exons XIII and XIV) to 1000 bp (exon XVII) and length of the introns range from 70 bp (intron M) to 11341 bp (intron A). Potential splice sites were individually compared with a splice site consensus sequence and scored for the degree of sequence identity (Clark and Thanaraj, 2002) (Table 6). Splice sites of exon VI deviate significantly from the consensus. They have values of 3.1 and 4.5 for 3' and 5' splice sites, respectively. The average scores are 9.3 ± 3.3 for 3' splice sites and 8.2 ± 2.3 for 5' splice sites. The alternative 3' splice site generated by 24-nucleotide inclusion within exon XIV also deviates strongly from the consensus and has a value of 1.8, whereas 3' splice site for the exon XIV without the inclusion has a value of 7.0. This result is in agreement with the earlier observations indicating that the weakest 3' splice sites are found in retained introns and alternative 3' splice sites (Stamm et al., 1994;

Thanaraj and Stamm, 2003). In addition, a number of splice sites, both constitutive and possibly alternative, are also found to be below average strength. However, it is known that the splice sites utilized in alternative gene structures are in general weaker than average values and that differential splice strength is seen to be a fundamental factor in alternative splicing (Clark and Thanaraj, 2002).

Table 6. Compilation of exon-intron boundaries.

exon	5 S	exon sequence	intron sequence	exon sequence	3 S	exon
			(5 UTR) ...agcgtaaacagcgcagggca	tttcgggagcgggattgttt...		exon I 365 nt
I	8.8	..ACAGTCGGGAGGAGAAAG	gttagtgcggagaaa (A, 11341 nt) ...atttttaattttatttgcag	ATGGAGAACTTAATGTTC..	8.2	exon II 102 nt
II	7.3	..ATAAAAATGAGAAAAAGG	gtatgtaagagttct (B, 382 nt) ...ttttacattgtattttaaag	GATCAAAAAGAAAAAGTG..	4.5	exon III 330 nt
III	3.4	..ACGCCAGATGGTTCTGAG	gtagtaaatattatt (C, 121 nt) ...ccttttttctttaaacatcag	AGAATTGGGCTTGAAGTG..	1.2	exon IV 424 nt
IV	4.3	..CAGGCACAGATGGATCAG	gtactacttttttata (D, 3629 nt) ...aaaactcttgttctttttag	ATGAGAAAAAGAAGGAAA..	8.8	exon V 90 nt
V	4.3	..CATCAGAGTCATATGCAG	gtattctcatttttgt (E, 460 nt) ...**tgtactttggctaaatacag**	**GTTCAGAAAAGAAGCATG**..	**3.1**	**exon VI 54 nt**
VI	**4.5**	**..TTCGTGCTGTCCGAAAAG**	**gtatcatttaaattt (F, 596 nt)** ...tgggtttctttttttttcag	ATCAAACCAGTAAACTCA..	10.4	exon VII 95 nt
VII	4.9	..TCTCTTGCCAAAGCGAAG	gtattagcatcagtt (G, 1792 nt) ...**tctttatgttatttatgtag**	**GGTGTATGGTCCACGCTC**..	**4.0**	**exon VIII 109 nt**
VIII	**10.8**	**..AGAGTGGAAAATTTCAAG**	**gtaagaataaaaata (H, 82 nt)** ...**ttatgatatgattctcttag**	**GGTTTGCAAGACTTTCTT**..	**5.8**	**exon IX 118 nt**
IX	**8.3**	**..AATTGACTGGATTTGCAG**	**gtaaattatacactt (I, 5778 nt)** ...tgatattactttgtatttag	GCGTGAATTACCCTTCAC..	4.9	exon X 85 nt
X	6.2	..ATCGGACGTGATGGACAG	gtttgtggccttaat (J, 1223 nt) ...ttgctactttcctattctag	GAAATTGAACTTGAATGT..	6.6	exon XI 167 nt
XI	6.5	..AGTCCGGGATGTGGGAAG	gttagaaacgttctt (K, 2543 nt) ...aatgtgtctaatttgtgtag	GCGTCGACCAGAAGATTA..	3.6	exon XII 83 nt
XII	5.6	..AGTTTCACCAGAGACCAG	gttaatattttataa (L, 1260 nt) ...aatgcttgatgtttgaacag	GGTATTTAAAGGATCCAC..	3.0	exon XIII 37 nt
XIII	12.7	..ATACCAGGAAGTGGACAG	gtaagtcgggttcac *(M, 70 nt)* ...**tatgaattacctgaatttag**	**TTTCACAAATCTTATTCC**..	**1.8**	**exon XIV 64 nt**
XIII	12.7	..ATACCAGGAAGTGGACAG	gtaagtcgggttcac (M, 96 nt) ...acaaatcttattcccaacag	ACGATTTTCAGGAGTTCG..	7.0	exon XIV 40 nt
XIV	8.0	..GATGTGTTTTTAAATGGG	gtaagcatttcattt (N, 108 nt) ...attatcatctttttattttag	TCCTACAATGATTATGTG..	7.0	exon XV 63 nt
XV	3.1	..CCACCTTGGCAAGGAATG	gtttgtattattcta (O, 2070 nt) ...aaatggttttccttttttag	CCCCCTTACCCAGGAATG..	8.2	exon XVI 135 nt
XVI	6.6	..TACAGAGATAAACGAGTA	gtaagtgcacaagga (P, 1990 nt) ...acatttctttccttttctag	CATGATTATGATATGAGG..	8.7	exon XVII 1004 nt

Coding exon sequences are indicated by uppercase letters, and intron and untranslated exon sequences are indicated by lowercase letters. Alternative splice sites are shown in bold. 5' and 3' scores are given for each splice site (Stamm et al., 1994). A 100% match to the mammalian 3' splice site ttttccctccag|G would have a score of 14.2. A perfect 5' splice site AAG|gtaagt would have a score of 12.6. The mean score of the 3' splice site in constitutive exons is 7.9 and 8.1 for constitutive 5' splice sites. The analysis was performed using Alternative Splicing Database (ASD) Workbench.

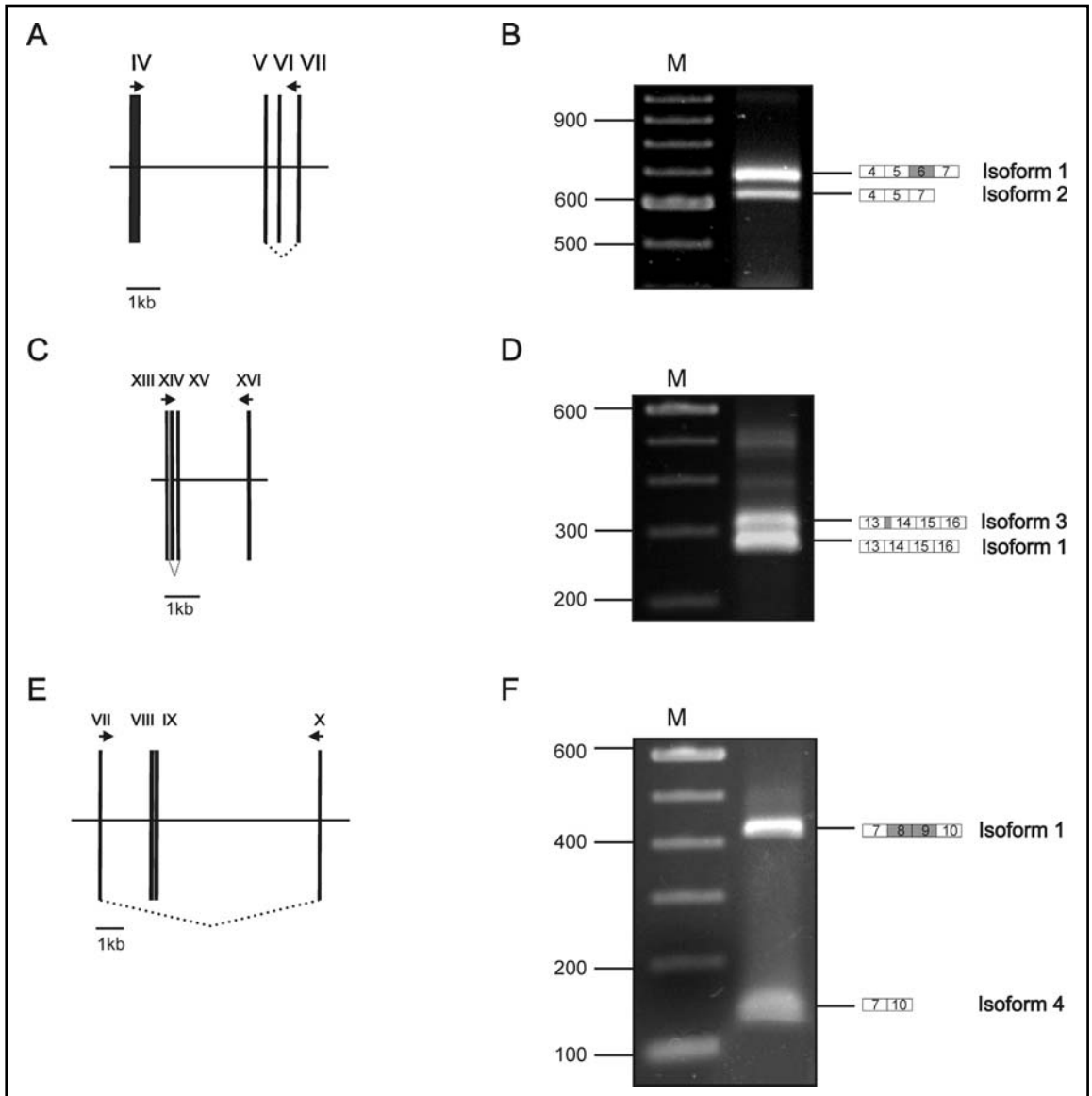
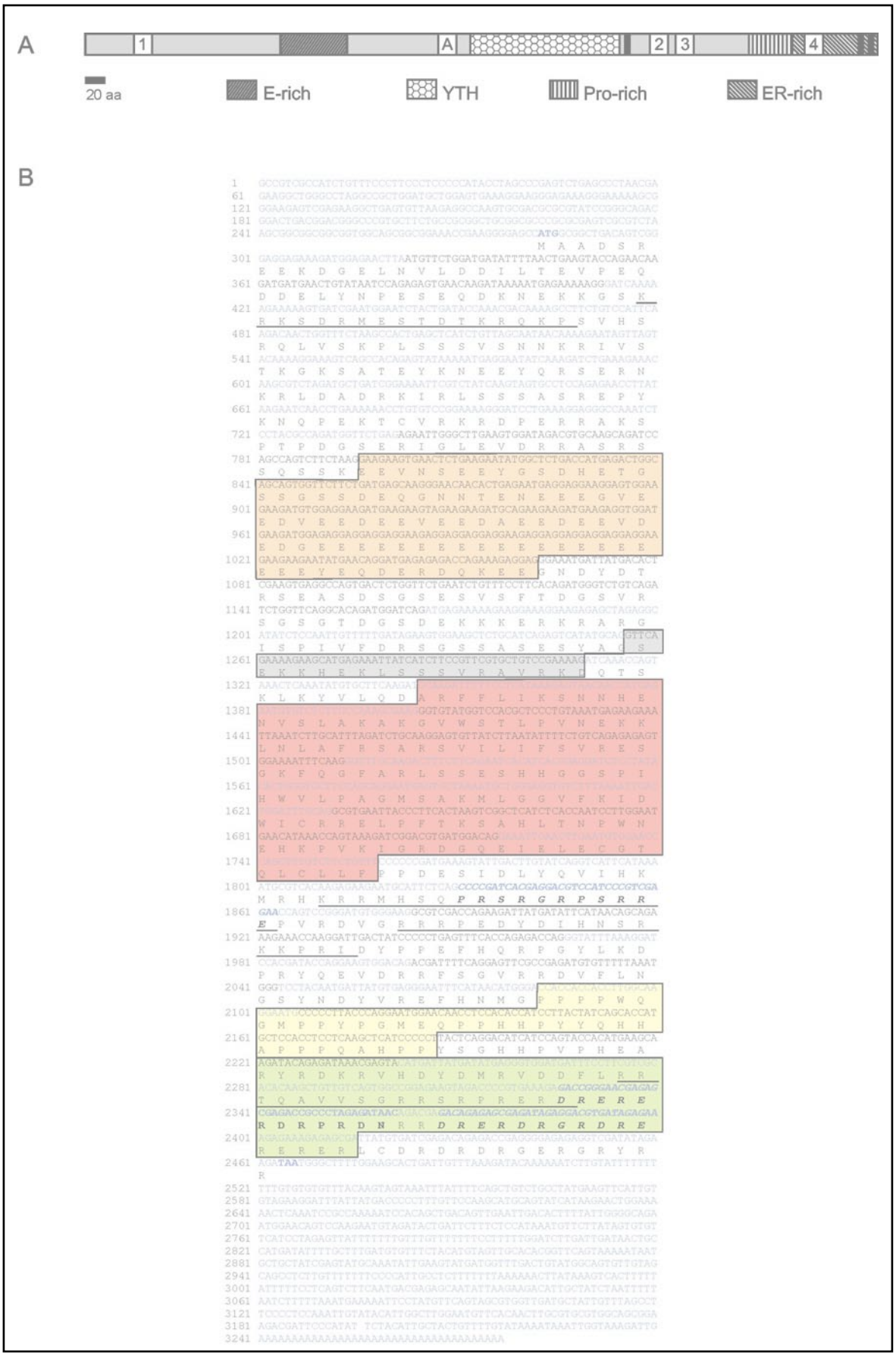


Figure 10. Alternative variants of human YT521-B. (**A**), (**B**) Detection of the isoform 2 with exon 6 being skipped. (**C**), (**D**) Detection of the isoform 3 with alternative 3' splice site resulting in 24-nt insertion within exon 14. (**E**), (**F**) Detection of the isoform 4 with exons 8 an 9 being skipped. (**A**), (**C**), (**E**) Part of the exon-intron structure containing alternatively spliced regions is drawn to scale; exons are indicated as vertical lines and numbered with Roman numerals; introns are shown as horizontal lines. The alternative splicing patterns are indicated with dotted lines. Positions of primers are indicated. (**B**), (**D**), (**F**) RT-PCR was performed with RNA isolated from adult human brain using primers flanking putative alternative exons. The structures of the PCR products are indicated on the right. M, 100 bp ladder

To confirm the existence of alternatively spliced human YT521-B isoforms, RT-PCR with RNA isolated from adult human brain was performed. As shown in Figure 10, all of the putative isoforms (Table 4, Table 5) were detected using primers flanking alternatively spliced exons, the same as those used in ASePCR (Table 5). The size of PCR products is in agreement with earlier electronic analysis. Transcripts of isoforms 1, 2 and 3 are present in human brain in similar amounts (Figure 10B, D), whereas isoform 4 is not as pronounced as isoform 1 (Figure 10F). Skipping of exons VIII and IX (Table 4, Figure 10E) leads to a frameshift. It results in the introduction of the termination codon at the beginning of exon X. Since the introduced termination codon is followed by an exon-exon junction more than 50-55 nucleotides downstream, it might be recognized as a Premature Termination Codon (PTC) (Maquat, 2004). Therefore, mRNA of isoform 4 is predicted to be a subject to nonsense-mediated decay (NMD). It is known that NMD is not only an mRNA quality-control mechanism that degrades abnormal mRNAs arising because of mistakes in gene expression or being product of mutated genes, but it also can target naturally occurring substrates. Together, these data show the gene structure of human YT521-B and indicate alternative splicing events resulting in the occurrence of four possible protein isoforms.

4.2. YT521-B protein has a modular structure

To gain an insight into human YT521-B function, its modular structure was inspected. MyHits Motif Scan program (Pagni et al., 2004) available at ExPASy Proteomics Server was used for this purpose. As shown in Figure 11, human YT521-B shows the same features as the rat protein (Hartmann et al., 1999). There are four putative bipartite nuclear localization signals (underlined sequences). At the N-terminus, there is a glutamic acid-rich region of unknown function (E-rich, shaded in orange). A putative nucleic acid binding YTH domain (section 4.3.) is located in the center of the protein (Stoilov et al., 2002b). At the C-terminus, there is a proline-rich stretch (Pro-rich, shaded in yellow) possibly responsible for binding to SH3 domain containing proteins (Kay et al., 2000) and a glutamic acid/arginine-rich region (ER-rich, shaded in green) involved in the possible protein:protein interactions (Hartmann et al., 1999; Yanagisawa et al., 2000). The conserved DR/ER-rich region is also thought to play a role in the binding of nucleic


A
1
A
2
3
4
20 aa
E-rich
YTH
Pro-rich
ER-rich
B

acids, though this has not been shown experimentally (Levi-Strauss et al., 1988; Speiser and White, 1989; Surowy et al., 1990; Assier et al., 1999). Database searches show that over 1000 proteins contain a domain with alternating ER or DR repeats. Several examples with known function are presented in Table 7. Some of them, such as RD RNA-binding protein (Surowy et al., 1990), negative elongation factor E (Yamaguchi et al., 1999), U1 70K (Spritz et al., 1990), and several putative DEAD helicases (Ono et al., 1994), are known to be involved in RNA metabolism. In addition, the RED protein is predicted to bind chromatin (Assier et al., 1999). Moreover, there are two AT-hook-like motifs (Pfam02178) predicted to be located within the ER-rich region of YT521-B (Figure 11). The AT-hook is identified to be a motif preferentially binding to AT-rich DNA. It is a small motif with a typical 9-amino acid signature sequence, in which a glycine-arginine-proline (GRP) tripeptide is the center of the DNA-binding domain. It was described for high mobility group (HMG) proteins (Reeves and Nissen, 1990), transcription factor AKNA (Siddiqa et al., 2001), MAR binding protein1 AHM1 (Morisawa et al., 2000), and some other proteins (Aravind and Landsman, 1998). In total, there are three AT-hook motifs predicted for YT521-B. They loosely follow the consensus sequence and it is not clear whether they are functional. As shown in Figure 12, the protein domain composition of YT521-B is reflected in its gene structure. Glutamic acid-rich region, proline-rich region and glutamic acid/arginine-rich region are encoded by one exon (IV, XVI, and XVII, respectively), and the YTH domain is encoded by four exons (VII–X).

Figure 11. Modular structure of human YT521-B. (**A**) Domain structure of YT521-B. The four nuclear localization signals (NLS) are indicated with numbers. The E-rich region, the proline-rich region, the ER-rich region, and YTH domain are indicated by different shading. AT-hooks are shown as black boxes. The inserted A indicates the location of the alternative exon VI. (**B**) cDNA and protein sequence of human YT521-B. The cDNA sequence of human YT521-B is shown. Exons are indicated by alternating text color. Start and stop codons are marked in bold. The protein sequence is shown underneath the cDNA sequence in one-letter code. The E-rich, Pro-rich, and ER-rich regions are shown as orange, yellow, and green-shadowed boxes, respectively. A putative RNA binding YTH domain is shown as a red-shadowed box. Four NLS sequences are underlined. AT-hook sequences are in bold italics. Alternatively spliced exon VI is indicated as a gray-shadowed box.

Table 7. Compilation of proteins containing DR/ER-rich low-complexity regions.

Protein*	Species	Comment
ATP-dependent helicase DHX8	Human	(Ono et al., 1994)
DEAD box RNA helicase	*A.thaliana*	gi:15226155
Translation initiation factor 3 subunit 10	Human	(Johnson et al., 1997)
RED protein	Human	(Assier et al., 1999)
MHC HLA-RD protein	Human	(Speiser and White, 1989)
Negative elongation factor E (NELF-E)	Human	(Yamaguchi et al., 1999)
Polyglutamine binding protein 1 (PQBP-1)	Human	(Waragai et al., 1999)
Transcription co-repressor SMRT	Human	(Chen and Evans, 1995)
U5 small nuclear ribonucleoprotein RNA helicase	*A.thaliana*	gi:20259792
RD RNA-binding protein	Human	(Surowy et al., 1990)
U1 70K	Human	(Spritz et al., 1990)
U2AF	*N.plumbaginifolia*	(Domon et al., 1998)
YT521-B putative splicing factor	Rat	(Hartmann et al., 1999)

*Proteins were identified in protein-protein BLAST search for short, nearly exact matches with the inclusion threshold of 0.005.

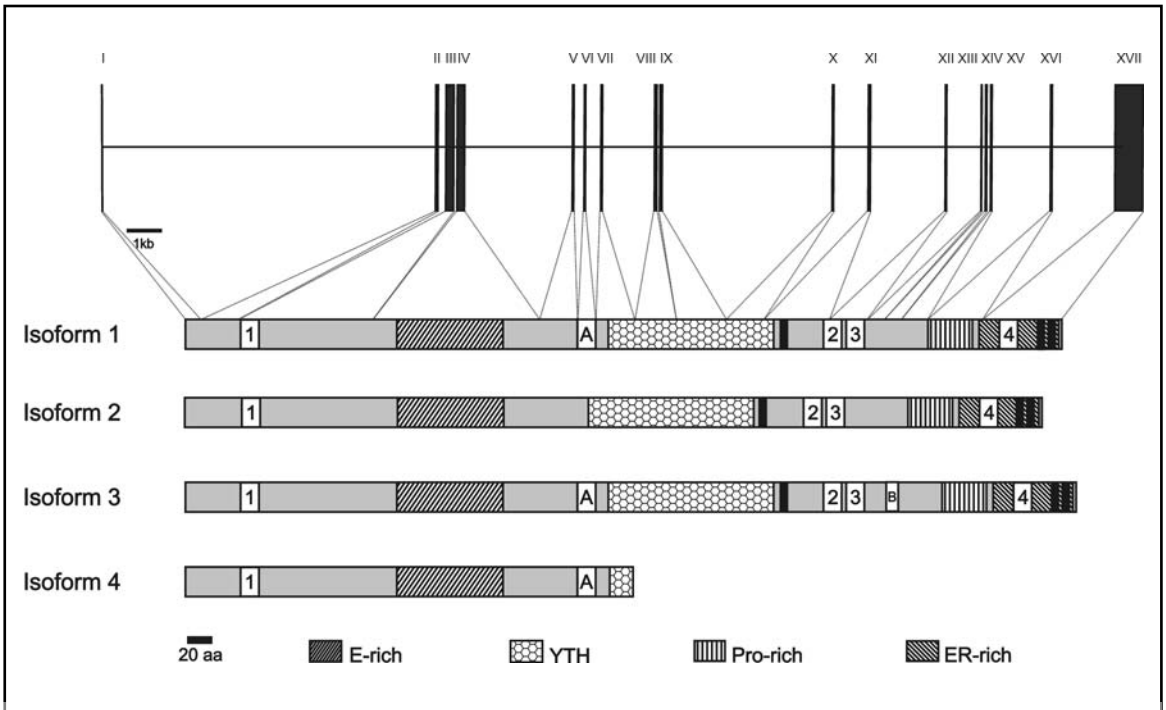


Figure 12. Comparison of the exon-intron organization of human YT521-B and the domain structure of its putative open reading frames. Exon-intron and protein structures are drawn to scale. Exons are indicated as black boxes or vertical lines and numbered with Roman numerals. Introns are shown as horizontal lines. Gene structures and alternative splicing events for isoforms 1 – 4 are given in Table 1. Protein structures of isoforms 1 and 2 are based on complete mRNA sequences, whereas structures of isoforms 3 and 4 are only a prediction based on partial EST sequences. Protein domains are indicated by different shading. AT-hooks are shown as black boxes. The four nuclear localization signals (NLS) are indicated with numbers. The alternative exons of YT521-B are indicated with capital letters, A and B.

Table 8. YT521-B orthologs*.

GenBank identifier	Species		Name	Protein		YTH domain	
				Length (aa)	% sequence identity to human YT521-B	Length (aa)	% sequence identity to human YTH
gi|21166355	Human	*Homo sapiens*	YT521-B	727	-	138	-
gi|19173798	Rat	*Rattus norvegicus*	YT521-B	738	96%	138	100%
gi|51711035	Mouse	*Mus musculus*	YT521-B	660	83%	170	100%
gi|50511209	Mouse	*Mus musculus*	mKIAA1966 protein	507	93%	138	100%
gi|55622698	Chimpanzee	*Pan troglodytes*	Predicted: similar to YT521	487	94%	138	100%
gi|55729205	Orangutan	*Pongo pygmaeus*	Hypothetical protein	482	96%	138	100%
gi|57095582	Dog	*Canis familiaris*	Predicted: similar to YT521	727	92%	138	98%
gi|54038191	African clawed frog	*Xenopus laevis*	Unknown: protein for MGC:83568	704	72%	138	94%
gi|51703808	Western clawed frog	*Silurana tropicalis*	MGC89461 protein	673	69%	138	94%
gi|55250288	Zebrafish	*Danio rerio*	Zgc:101592	679	62%	138	93%

*Orthologs were identified using Eukaryotic Gene Orthologs (EGO) and HomoloGene databases.

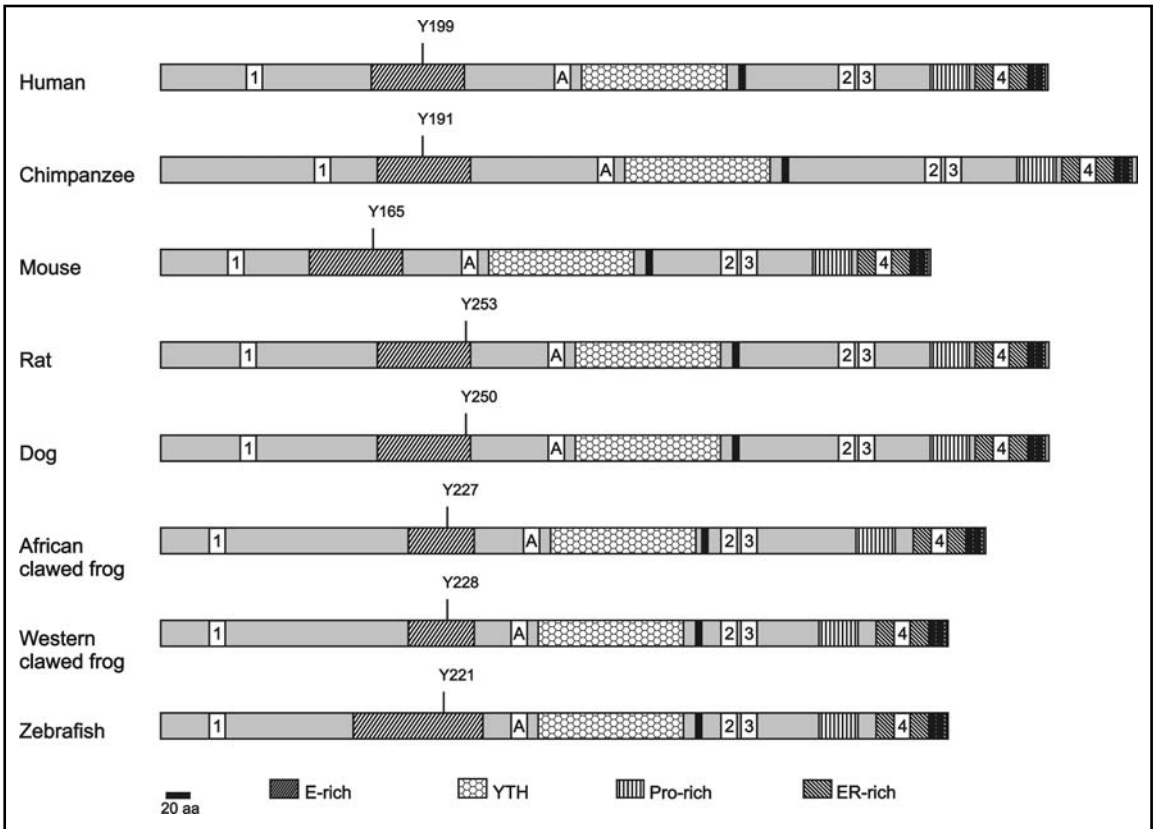


Figure 13. Domain structure of YT521-B orthologs. The four nuclear localization signals (NLS) are indicated with numbers. The protein domains are indicated by different shading. AT-hooks are shown as black boxes. The position of the alternatively spliced exon VI is indicated with letter A. Position of the conserved tyrosine is indicated.

The same conserved modular structure is shared by all known YT521-B orthologs from zebrafish to rodents to primates (Table 8, Figure 13). There was neither yeast nor bacterial YT521-B orthologs found in performed database searches. Moreover, there are no YT521-B orthologs present in *C.elegans* and *D.melanogaster*. Altogether, data

obtained in the performed searches indicate that YT521-B may be a vertebrate-specific protein. The percentage of sequence identity among the YT521-B orthologs varies from 62 to 96% (Table 8). The part of the protein showing the largest conservation among species is the YTH domain (Table 8, Figure 14). These characteristics indicate the importance of the protein, whose structure and function was retained throughout the course of evolution.

```
X.laevis        ARFFLIKSNNHENVSLAKAKGVWSTLPVNEKKLNAAFRSARSVILVFSVRESGKFQGFAR 60
X.tropicalis    ARFFLIKSNNHENVSLAKAKGVWSTLPVNEKKLNAAFRSARSVILVFSVRESGKFQGFAR 60
H.sapiens       ARFFLIKSNNHENVSLAKAKGVWSTLPVNEKKLNLAFRSARSVILIFSVRESGKFQGFAR 60
P.troglodytes   ARFFLIKSNNHENVSLAKAKGVWSTLPVNEKKLNLAFRSARSVILIFSVRESGKFQGFAR 60
M.musculus      ARFFLIKSNNHENVSLAKAKGVWSTLPVNEKKLNLAFRSARSVILIFSVRESGKFQGFAR 60
R.norvegicus    ARFFLIKSNNHENVSLAKAKGVWSTLPVNEKKLNLAFRSARSVILIFSVRESGKFQGFAR 60
C.familiaris    ARFFLIKSNNHENVSLAKAKGVWSTLPVNEKKLNAAFRSARSVILIFSVRESGKFQGFAR 60
D.rerio         ARFFLIKSNNHENVSLAKAKGVWSTLPVNEKKLNAAFRAARSVVLIFSVRESGKFQGFAR 60
                ******************************** ***:****:*:**************

X.laevis        LSSESHHGGSPIHWVLPAGMNAKMLGGVFKIDWICRRELSFTKCVHLTNPWNEHKPVKIG 120
X.tropicalis    LSSESHHGGSPIHWVLPAGMNAKMLGGVFKIDWICRRELPFTKCVHLTNPWNEHKPVKIG 120
H.sapiens       LSSESHHGGSPIHWVLPAGMSAKMLGGVFKIDWICRRELPFTKSAHLTNPWNEHKPVKIG 120
P.troglodytes   LSSESHHGGSPIHWVLPAGMSAKMLGGVFKIDWICRRELPFTKSAHLTNPWNEHKPVKIG 120
M.musculus      LSSESHHGGSPIHWVLPAGMSAKMLGGVFKIDWICRRELPFTKSAHLTNPWNEHKPVKIG 120
R.norvegicus    LSSESHHGGSPIHWVLPAGMSAKMLGGVFKIDWICRRELPFTKSAHLTNPWNEHKPVKIG 120
C.familiaris    LSSESHHGGSPIHWVLPAGMNAKMLGGVFKIDWICRRELPFTKSAHLTNPWNEHKPVKIG 120
D.rerio         LSSESHHGGSPIHWVLPAGMNAKMLGGVFKIDWICRRELPFTKTAHLANPWNEHKPVKIG 120
                ********************.******************.*** .**:************

X.laevis        RDGQEIEPDCGTQLCLLF 138
X.tropicalis    RDGQEIEPDCGTQLCLLF 138
H.sapiens       RDGQEIELECGTQLCLLF 138
P.troglodytes   RDGQEIELECGTQLCLLF 138
M.musculus      RDGQEIELECGTQLCLLF 138
R.norvegicus    RDGQEIELECGTQLCLLF 138
C.familiaris    RDGQEIELECGTQLCLLF 138
D.rerio         RDGQEIEPECGTELCMLF 138
                ******* :***:**:**
```

Figure 14. Alignment of YTH domains from YT521-B orthologs (see Table 8). Asterisk marks the identical residues or nucleotides, colon marks the conserved substitutions, and dot marks the semi-conserved substitutions.

4.3. YTH: a new domain in nuclear proteins

During BLAST searches for YT521-B homologs, a conserved part between residues 356 and 499 of the rat YT521-B protein was identified. A PSI-BLAST search with this portion of the protein and the inclusion threshold of 0.005 was subsequently performed. After four iterations, multiple proteins from *Arabidopsis thaliana*, *Oryza sativa*, *Homo sapiens*, *Mus musculus*, *Drosophila melanogaster*, *Plasmodium falciparum*, and *Saccharomyces cerevisiae* were found to have a similar domain. The reliability of the

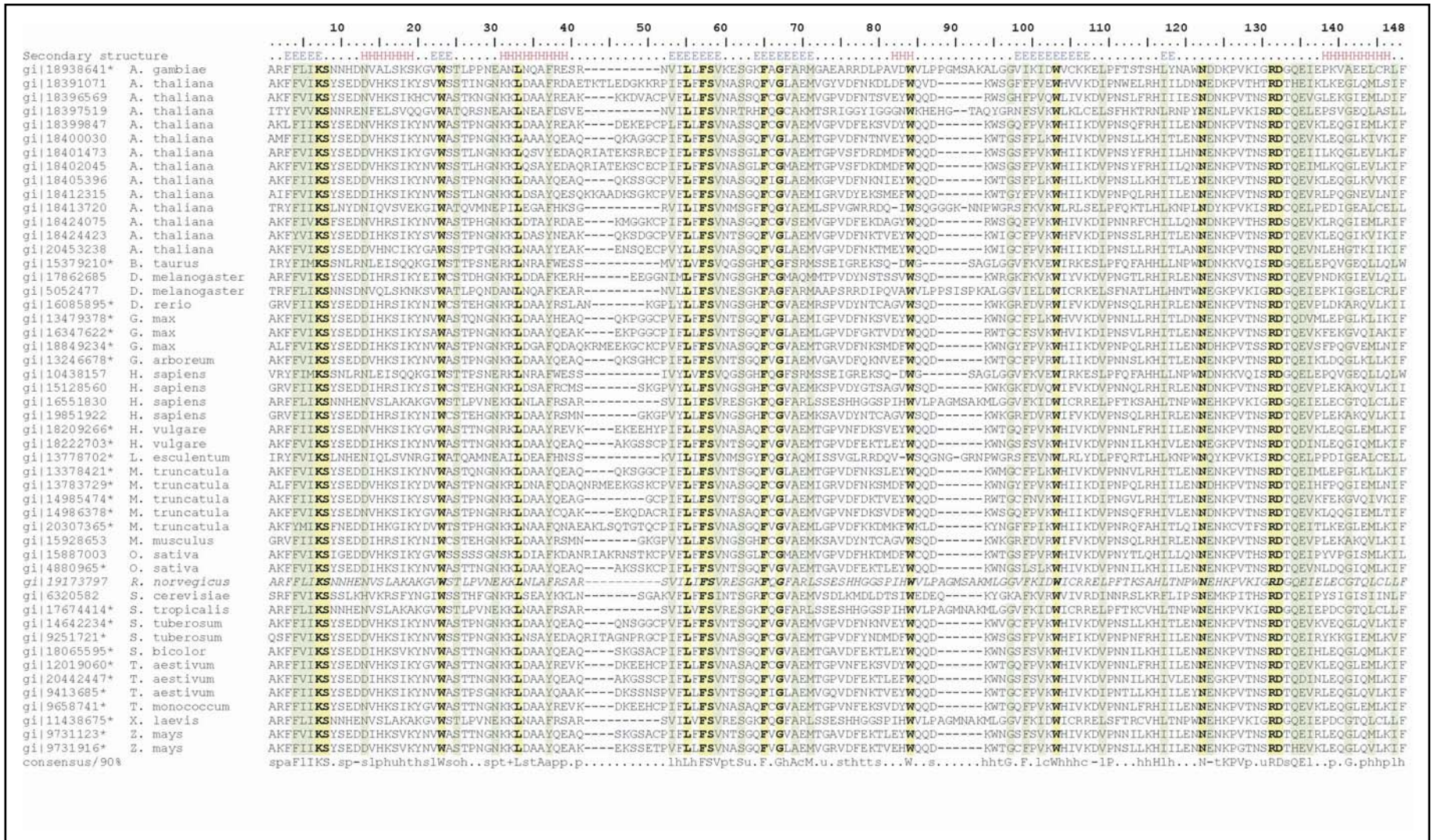


Figure 15. Multiple alignment of sequences containing YT homology (YTH) domains. The sequences are denoted with the GenBank identifier of the corresponding nucleotide sequence followed by the species name. The GenBank identifier for the *Saccharomyces cerevisiae* refers to protein entry in the database. Translations of partial EST sequences are indicated with an asterisk. The sequences were identified using searches of the translated EST and the non-redundant nucleotide databases at NCBI using rat YT521-B (shown in italic). Wherever available, the conceptual translation provided in the database was used for the alignment. If no conceptual translation was provided (i.e. for the EST sequences), the reading frame identified by the BLAST search was used if it did not contain stop codons or unknown residues. Redundant sequences originating from the same species were omitted. The similar residues that are present in all sequences are shaded (green) using the PAM250 matrix. The identical residues are shaded in yellow. The secondary structure, as predicted by the PHD program, is shown on top. E denotes extended (β-strands) structure and H denotes the predicted α-helices.

scores was judged by their EXPECT (E) values. The EXPECT value is a statistical significance threshold for reporting matches against database sequences. More stringent matches have lower EXPECT thresholds with fewer chance matches being reported. The typical threshold for a reliable E value from a BLAST search is 10^{-5} or lower. Found sequences had EXPECT values ranging from 10^{-11} to 10^{-60}. Many of these protein sequences were derived from automated gene prediction and were not confirmed by mRNA or EST sequences. To identify sequences of existing proteins, BLAST searches with the conserved region against the non-redundant CDS translation database at NCBI (filtered using "biomol_mrna[PROP]" as a keyword) and the translated EST database were performed. The sequences from these searches, aligned to the full length of the query and having E values below 10^{-6}, were aligned using ClustalW, after the redundancies had been removed (Figure 15). Additional BLAST searches against the genome databases confirmed that the conserved region is present exclusively in eukaryotic genomes. Most of the proteins identified in the BLAST searches originate from plants, with 13 distinct sequences coming from a single species (*A. thaliana*). It is unclear whether this protein family is more widespread in plants, or whether the observed species distribution is due to bias in the plant databases.

The conserved region, which was named the YT521-B homology (YTH) domain (Stoilov et al., 2002b), defines a new domain in these proteins. The YTH domain is usually located in the middle of the protein or at its C-terminus. It shows remarkable conservation across a wide range of species with 14 invariant and 19 highly conserved residues. The proteins present in the alignment do not share significant similarity outside the YTH domain, with the exception of closely related vertebrate orthologs of YT521-B.

To investigate the evolutionary relationship among members of the YTH family, a phylogenetic tree was constructed using the neighbour joining method (Saitou and Nei, 1987) with Kimura correction of distances (Thompson et al., 1994). Kimura's distance correction applies to multiple substitutions, which have occured at a site where only one difference is observed. It has the effect of stretching long branches in trees while leaving short ones relatively untouched. As a result, distances become proportional to the time passed since divergence (Kimura, 1983). As shown in Figure 16, the YTH family diverged from one ancestral domain and also indicates the occurence of several

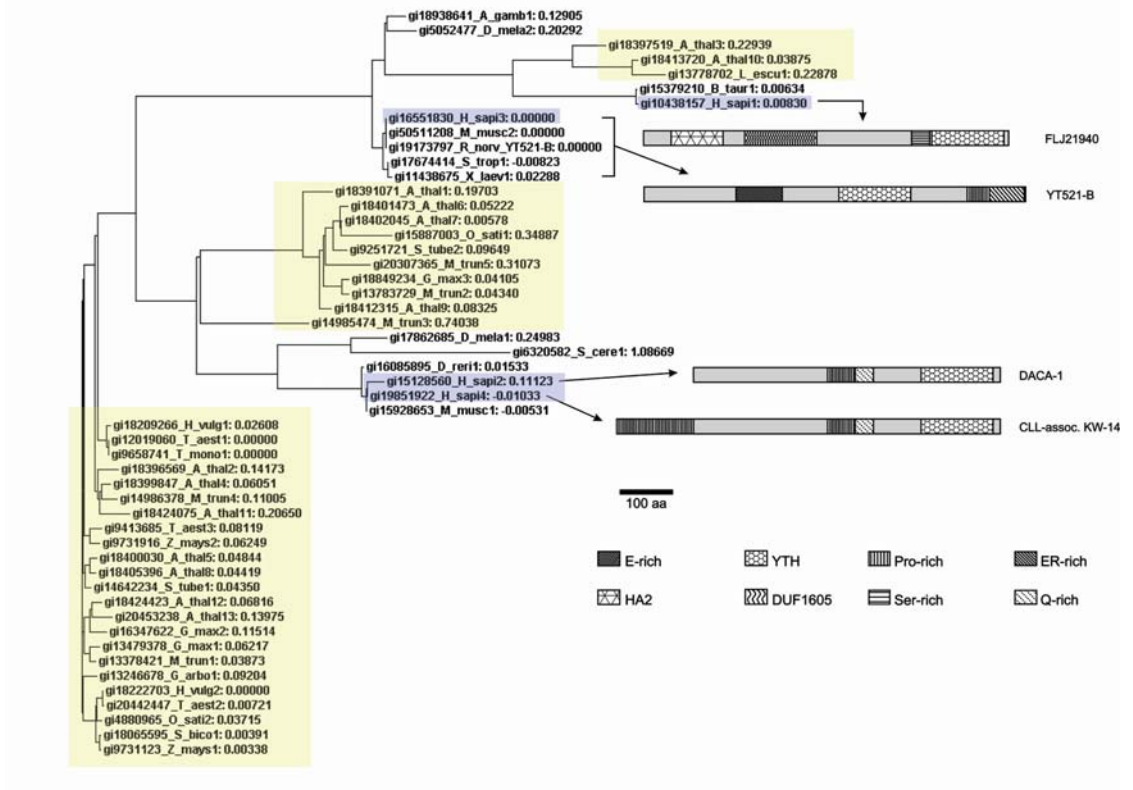


Figure 16. Phylogenetic tree of YTH domain. The phylogeny was derived by the neighbor-joining method from the multiple sequence alignment shown in Figure 5. Plant sequences are shaded in yellow, human sequences – in blue. Schematic representation of human proteins domain structure is shown. The protein domains are indicated by different shading.

independent duplication events. Most of the plant YTH domains, including 6 sequences from *A.thaliana*, diverged early and cluster together, occupying a distinct phylogenetic branch. Rat, mouse and human YTH domains present in YT521-B orthologs show 100% conservation (Table 8). On the other hand, 4 human YTH sequences do not cluster together and show divergence among them (Figure 16, Table 9). The sequence identity

Table 9. YTH domain containing human proteins.

Sequence (Figure 16)	Protein	Predicted localization
gi10438157_H_sapi1	FLJ21940 fis, clone HEP04512 FLJ16598 fis, clone TESTI4006473 (gi47077414) YTH domain containing 2 (gi:38505213)	nuclear
gi15128560_H_sapi2	Dermatomyositis associated with cancer putative autoantigen-1, DACA-1	nuclear
gi16551830_H_sapi3	YT521-B	nuclear
gi19851922_H_sapi4	CLL-associated antigen KW-14	mitochondrial / nuclear

for them ranges from 30% to 90% (Table 10, Figure 17). Human proteins containing YTH domain do not share similar domain organization (Figure 16) and do not have any known function. It is possible that such a divergence among them results in different expression patterns and/or acquiring unique functions or binding specificities by each of these proteins. For example, proteins FLJ21940, FLJ16598, and YTHDC2 (Table 9) contain domains characteristic for helicases suggesting their role in nucleic acid metabolism (Figure 18, Table 11).

Table 10. Sequence identity among human YTH domains.

Sequence 1	Sequence 2	Sequence identity
FLJ21940	DACA-1	35%
FLJ21940	YT521-B	49%
FLJ21940	KW-14	33%
DACA-1	YT521-B	32%
DACA-1	KW-14	90%
YT521-B	KW-14	31%

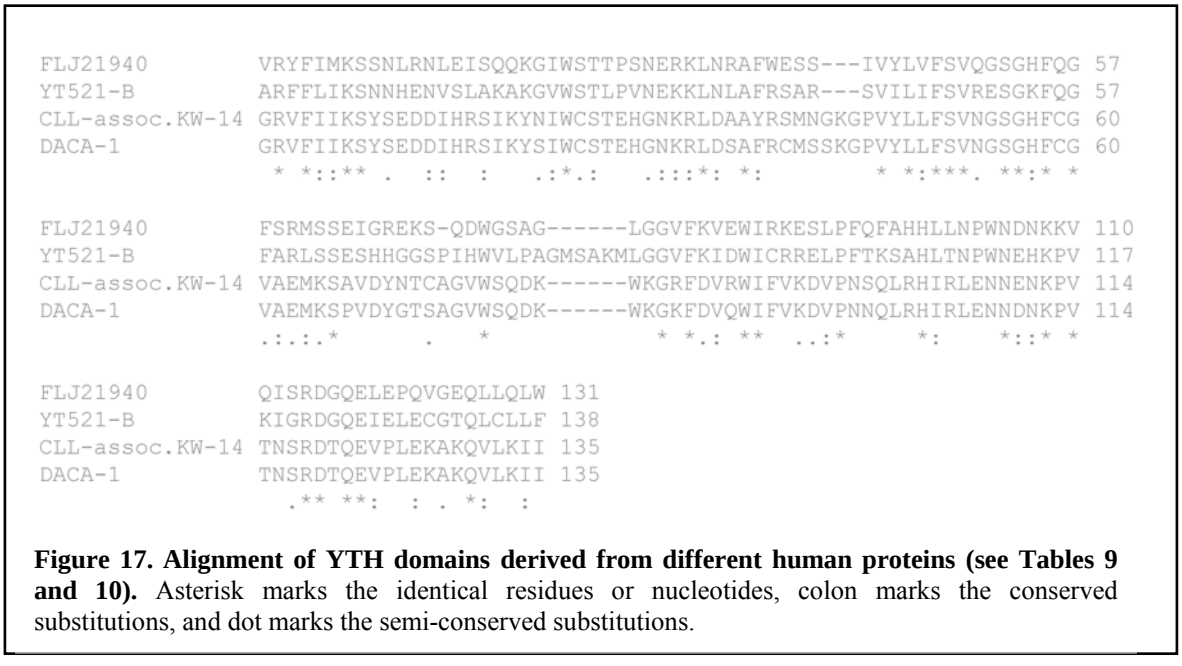


Figure 17. Alignment of YTH domains derived from different human proteins (see Tables 9 and 10). Asterisk marks the identical residues or nucleotides, colon marks the conserved substitutions, and dot marks the semi-conserved substitutions.

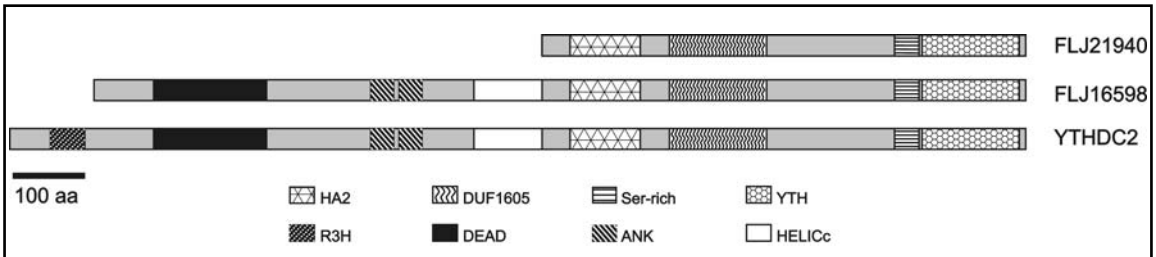


Figure 18. Human proteins containing YTH domain and domains associated with helicase activity. Proteins were identified in SMART searches. The protein domains are indicated by different shading. Brief description of Pfam domains is given in Table 11.

Table 11. Helicase-activity-associated domains and other functional domains present in YTH domain containing proteins.

Name	Domain	Pfam accession number	Function
DEAD	DEAD/DEAH box helicase	PF00270	The DEAD box helicases are involved in various aspects of RNA metabolism, including nuclear transcription, pre-mRNA splicing, ribosome biogenesis, nucleocytoplasmic transport, translation, RNA decay and organellar gene expression
HELICc	Helicase conserved C-terminal domain	PF00271	Restricted to DEAD/H helicases, it may be that this is not an autonomously folding unit, but an integral part of the helicase
HA2	Helicase associated domain	PF04408	Found in a diverse set of RNA helicases, it seems likely to be involved in nucleic acid binding
DUF1605	Domain of unknown function	PF07717	Usually found towards the C-terminus of the DEAD-box helicases where it is associated with helicase associated HA2 domain
R3H	R3H motif	PF01424	Predicted to bind single-stranded nucleic acids
ANK	Ankyrin repeat	PF00023	Provides protein-protein interaction

To test whether any additional known domains are present in the proteins containing YTH domain, SMART and PFAM databases were searched using SMART sequence analysis available at EMBL server. The search did not identify any known domains, except the previously mentioned helicase-associated domains and three C–C–C–H-type zinc finger domains in gi:18397519, an *A. thaliana* predicted protein of unknown function.

The putative secondary structure of the YTH domain was determined using the PHD program (Rost, 1996). The domain is predicted to have a mixed αβ-fold, with four α-helices and six β-strands. The conservation pattern follows the predicted secondary structure, with three blocks of conserved sequence separated by loops of variable size.

Notable features of the domain are the highly conserved aromatic residues located within the β-sheet and being reminiscent of the aromatic residues conservation within the RNA recognition motif. In the RRM domain, conserved aromatic residues located in the β-sheet are crucial for RNA binding (Hoffman et al., 1991). Based on this observation, it is predicted that the biological function of the YTH domain is to bind nucleic acids, probably to bind RNA.

4.4. Antiserum specific for YTH domain

To develop the tool for further investigations, serum against YTH domain was generated. A rabbit polyclonal antiserum was raised against peptide CVRESGKFQGFARLSSE fused to keyhole limpet haemocyanin (KLH). Hydrophobicity plot shows that this peptide has a hydrophilic character and may possibly be antigenic (Figure 19). The obtained serum was tested in Western blot for its specificity. As shown in Figure 20, it can recognize recombinant TH ΔNLS4 YT521-B deletion mutant expressed and purified from *E.coli* (section 4.5.). It cannot however recognize the crude uninduced lysate of BL21(DE3)-RIL bacterial strain (Figure 20 C).

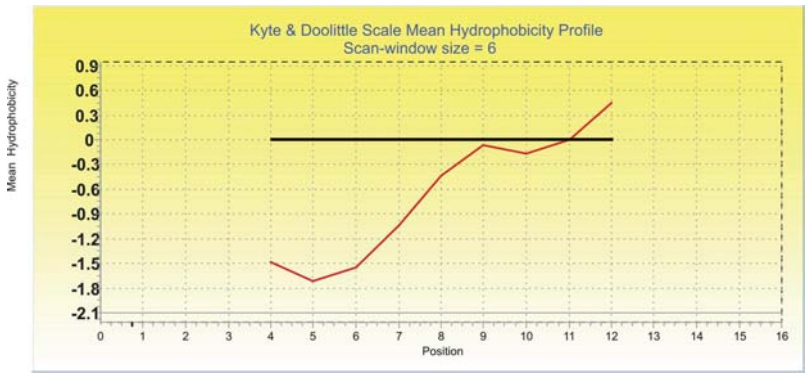


Figure 19. Kyte-Doolittle scale hydrophobicity plot of the peptide within YTH domain. Regions with values above 0 are hydrophobic in character. Window size of 5-7 is used for finding hydrophilic regions that are likely exposed on the surface.

4.5. TH ΔNLS4 YT521-B deletion mutant binds to purine-rich RNA

YT521-B is known to be involved in splice site selection (Hartmann et al., 1999; Nayler et al., 2000). However, its role in the process and potential targets remain unknown. Computational analysis of the protein sequence revealed the possibility that

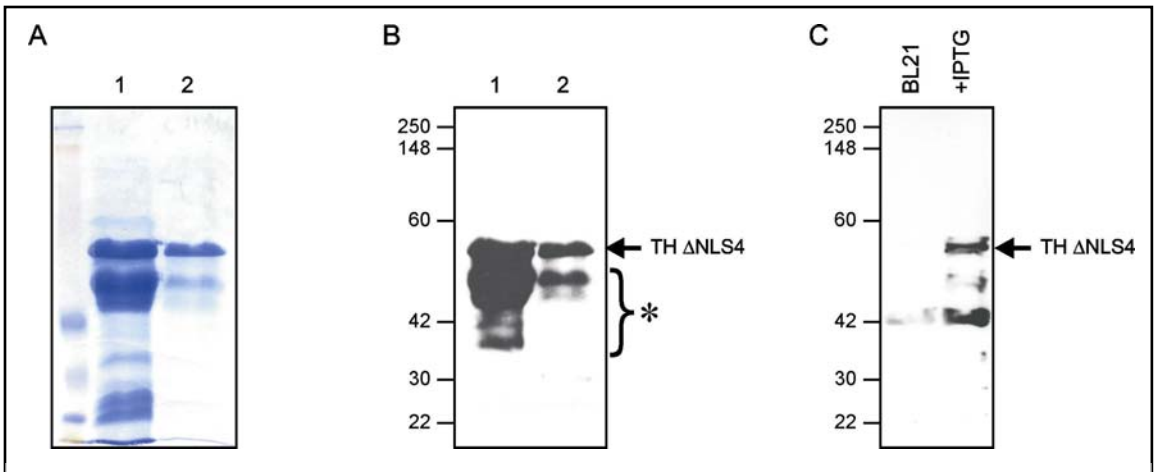


Figure 20. Specificity of the rabbit antiserum generated against YTH domain. (**A**) TH ΔNLS4 deletion mutant purified from *E.coli* visualized on polyacrylamide gel by Coomassie staining. (**B**) TH ΔNLS4 deletion mutant purified from *E.coli* detected on polyacrylamide gel with anti YTH antibody. Numbers 1 and 2 above lanes indicate different purification experiments. Asterisk marks impurities and degradation products, which are purified together with the recombinant protein. (**C**) Crude lysates of BL21(DE3)-RIL bacterial strain visualized on polyacrylamide gel with anti YTH antibody. The lanes are described above the picture; "BL21" indicates the crude lysate of BL21(DE3)-RIL bacterial strain; "+IPTG" indicates the crude lysate of bacterial cells after the induction with IPTG. Pointed arrows show the localization of TH ΔNLS4 band.

YT521-B can directly interact with RNA and that its YTH domain can be characterized as another RNA-binding domain. To test this hypothesis, nucleic acid binding was investigated by several methods. The full-length protein and mutants containing ER-rich region are not expressed in bacteria, as the repetitive glutamic acid/arginine dipeptide seems to be toxic for *E.coli* cells. Experiments were therefore performed with TH ΔNLS4 YT521-B deletion mutant lacking the first 253 residues from the N-terminus and ER-rich region on C-terminus (Figure 21A). This construct contains YTH domain and can be used for nucleic acid binding experiments. The absence of ER-rich region allows successful protein expression in *E.coli*. The gene coding for TH ΔNLS4 YT521-B deletion mutant was cloned into pET-28a(+) expression vector, which introduces a HIS-tag. The presence of 6×His-tag allows efficient purification of the protein on Ni-NTA resin. TH ΔNLS4 YT521-B deletion mutant was purified under native conditions and eluted from Ni-NTA agarose with 500 mM imidazole (Figure 21B). Such a high concentration of imidazole was chosen after performing the elution profile for the protein. TH ΔNLS4 YT521-B deletion mutant is eluted with imidazole concentrations ranging from 150 mM to 500 mM, with a peak at 500 mM. In order to reduce the binding of contaminating proteins to resin, higher than usual concentrations of imidazole in washing buffers (80 and 100 mM)

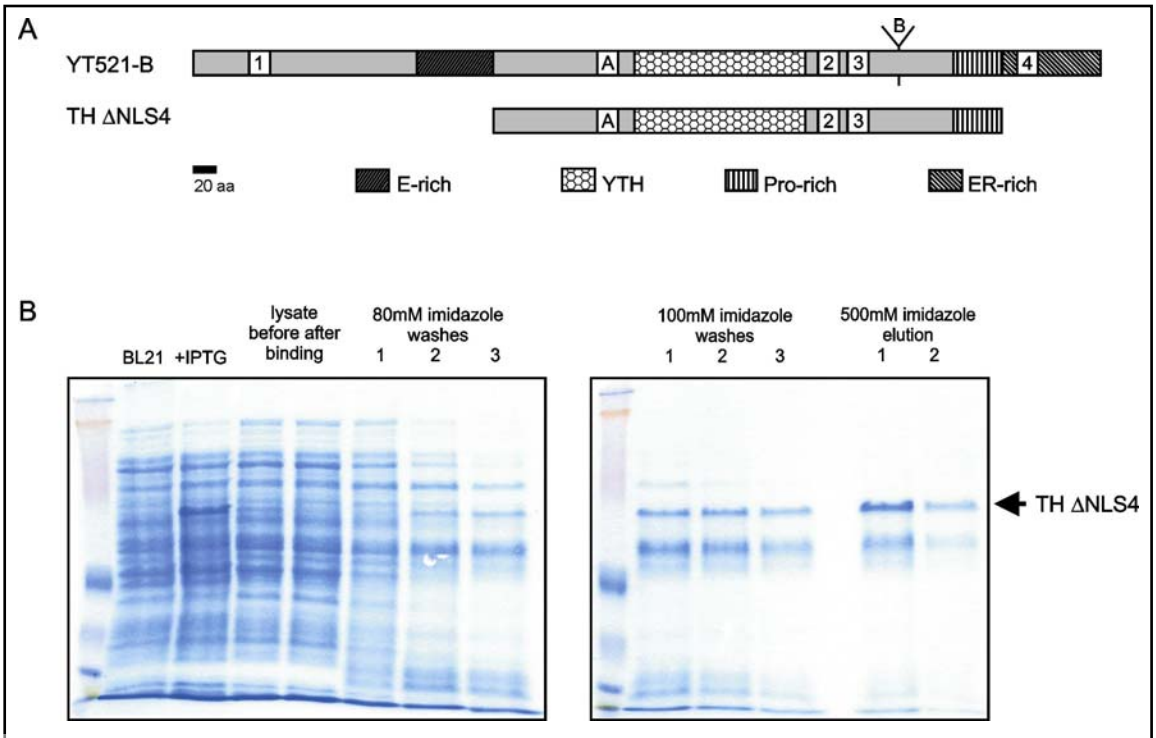


Figure 21. Expression and purification of TH ΔNLS4 deletion mutant. (**A**) Domain structure of full-length YT521-B and TH ΔNLS4 deletion mutant. Nuclear localization signals are indicated with numbers. The E-rich, the proline-rich region, the ER-rich region, and YTH domain are indicated by different shading. The inserted A and B indicate the location of the alternative exons of YT521-B. (**B**) TH ΔNLS4 YT521-B deletion mutant was purified on Ni-NTA agarose beads after the expression in *E.coli* BL21(DE3)-RIL strain. Purified recombinant protein was visualized on polyacrylamide gels by Coomassie staining. The lanes are described above the picture; "BL21" indicates the crude lysate of BL21(DE3)-RIL bacterial strain; "+IPTG" indicates the crude lysate of bacterial cells after the induction with IPTG. A pointed arrow shows the localization of TH ΔNLS4 band.

were used. Purified recombinant TH ΔNLS4 was subsequently incubated with different RNA homopolymer coupled to agarose beads, as well as with single- and double-stranded DNA coupled to cellulose beads (Figure 22A). The bound protein was detected on polyacrylamide gels by silver staining and Western Blot, using specific anti YTH antibodies (section 4.4.). As shown in Figure 22B, C, TH ΔNLS4 binds to poly (G) RNA homopolymers and to DNA beads. Next, an RNA electrophoretic mobility shift assay (R-EMSA) with purine- (R-rich, TraBS3) and pyrimidine-rich (Y-rich) RNA oligonucleotides was performed. This experiment confirmed the possibility that YT521-B deletion mutant interacts with purine-rich sequences (Figure 23). It also revealed that the recombinant protein is unstable and that it precipitates from the solution during the experiment. The extent of protein folding was subsequently analyzed by 1D proton NMR

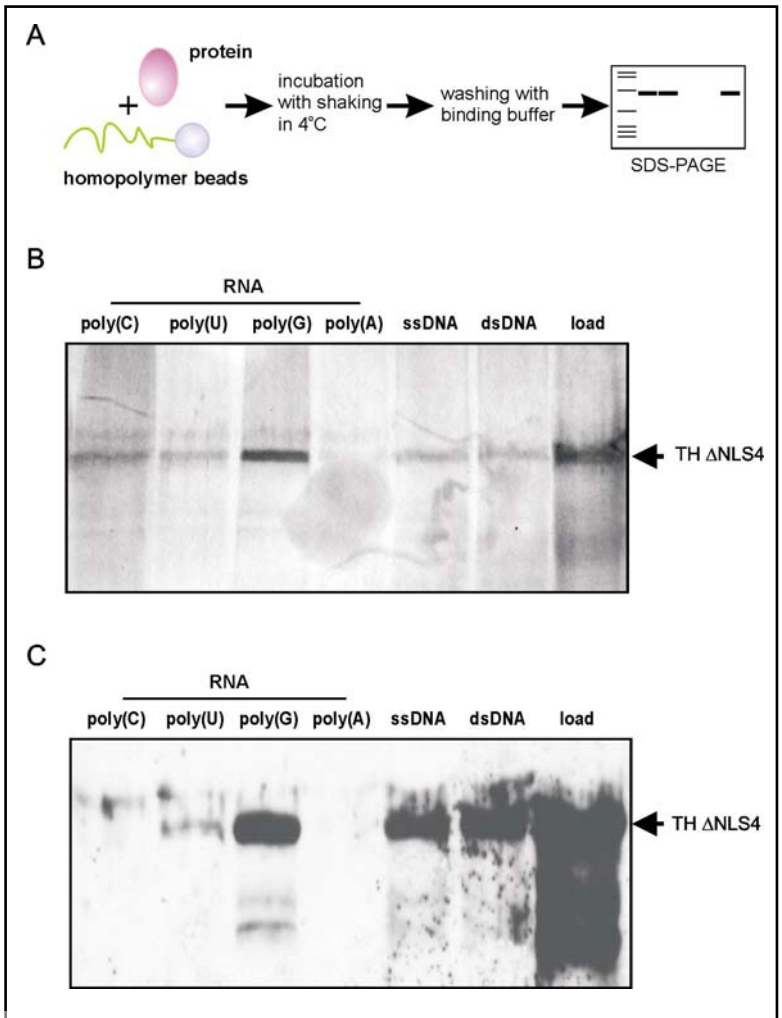


Figure 22. TH ΔNLS4 binds to poly(G) RNA homopolymer beads. (**A**) Schematic representation of the experimental procedure. Purified protein is incubated with homopolymers immobilized on agarose or cellulose beads. Then the resin is washed to remove the unbound protein. Finally, the bound protein is analyzed on gel. (**B**) TH ΔNLS4 bound to homopolymer beads visualized on polyacrylamide gel by silver staining. (**C**) TH ΔNLS4 bound to homopolymer beads detected on polyacrylamide gel with anti YTH antibody. A pointed arrow indicates the localization of TH ΔNLS4 band.

spectra (Weber et al., 2000; Rehm et al., 2002). The analysis was performed in collaboration with Dr. Tad Holak from Max-Planck-Institut for Biochemistry in Martinsried. Inspection of 1D proton NMR spectra yields semi-quantitative information on folding in partially structured proteins or their domains (Rehm et al., 2002). NMR performed for TH ΔNLS4 YT521-B deletion mutant showed that the protein is only partially, in around 30% folded. This suggests that the protein may need a binding partner, such as RNA or DNA, to become stably folded. Another possible interpretation of the result is that in the NMR spectrum a mixture of folded and unfolded protein can be seen. In such a situation, the observed protein activity in binding assays comes only from the folded protein. To get a functional, properly folded protein, this mixture should be separated in the affinity chromatography. Nevertheless, the obtained results indicate the ability of YTH domain containing YT521-B deletion mutant to bind nucleic acids.

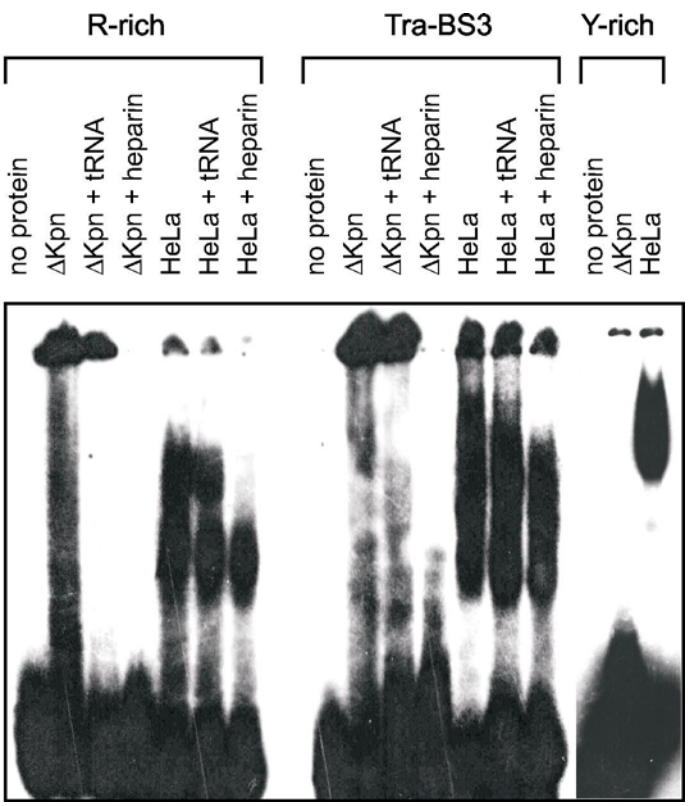


Figure 23. TH ΔNLS4 interacts with purine-rich RNA oligonucleotides. TH ΔNLS4 recombinant protein was bound to different ^{32}P-labelled RNA oligonucleotides. Formed complexes were separated in RNA gel shit experiment and visualized by autoradiography. tRNA and heparin were used as non-specific competitors.

4.6. Mutations within YTH domain changes localization of YT521-B

It is well established that the highly conserved aromatic residues provide stacking interaction between the RRM domain and RNA (Hoffman et al., 1991). Mutagenesis studies bring more information about molecular contacts between the protein and its target RNA (Jessen et al., 1991). Therefore, this method was used in the context of YTH domain characterization. All 14 invariant residues within YTH domain were mutated (Figure 24). The EGFP-tagged mutants were then transfected into HEK293 cells and their intranuclear localization was analyzed. It is known from previous studies that YT521-B is present in the novel subnuclear compartment, in the YT bodies (Nayler et al., 2000). YT bodies disperse upon actinomycin D treatment (Nayler et al., 2000) and after tyrosine phosphorylation (Rafalska et al., 2004) (section 4.12.). As shown in Figure 25, some of

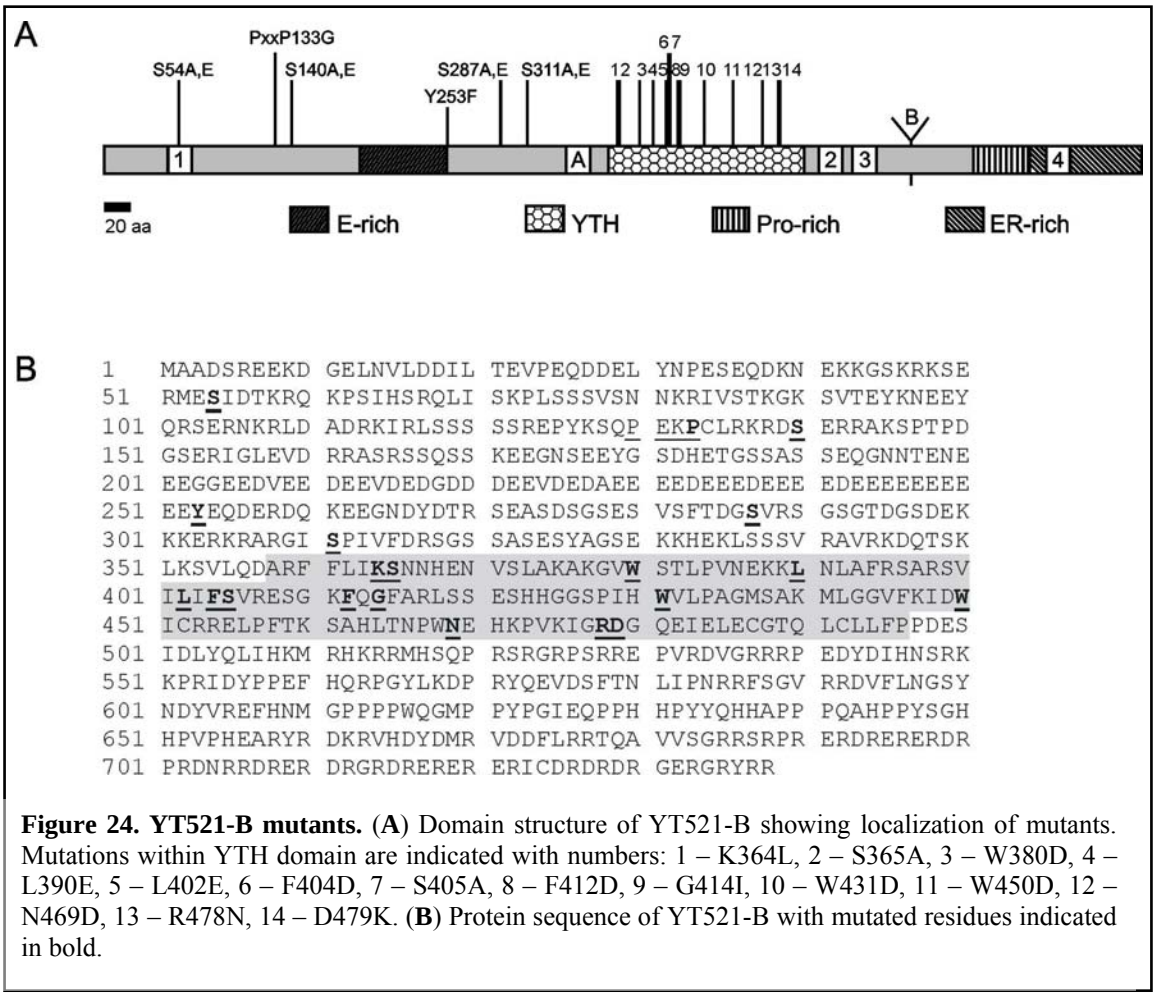


Figure 24. YT521-B mutants. (**A**) Domain structure of YT521-B showing localization of mutants. Mutations within YTH domain are indicated with numbers: 1 – K364L, 2 – S365A, 3 – W380D, 4 – L390E, 5 – L402E, 6 – F404D, 7 – S405A, 8 – F412D, 9 – G414I, 10 – W431D, 11 – W450D, 12 – N469D, 13 – R478N, 14 – D479K. (**B**) Protein sequence of YT521-B with mutated residues indicated in bold.

the invariant residues are necessary for YT bodies' formation and at least 7 mutations, W380D, L390E, L402E, F404D, F412D, G414I, and W450D, cause dispersion of YT bodies (Table 12).

Table 12. Intranuclear localization of YT521-B mutants.

Mutant	YT bodies	Mutant	YT bodies
K364L	Yes	S405A	Yes
S365A	Yes	F412D	Dispersed
W380D	Dispersed	G414I	Dispersed
L390E	Dispersed	W450D	Dispersed
L402E	Dispersed	N469D	Yes
F404D	Dispersed		

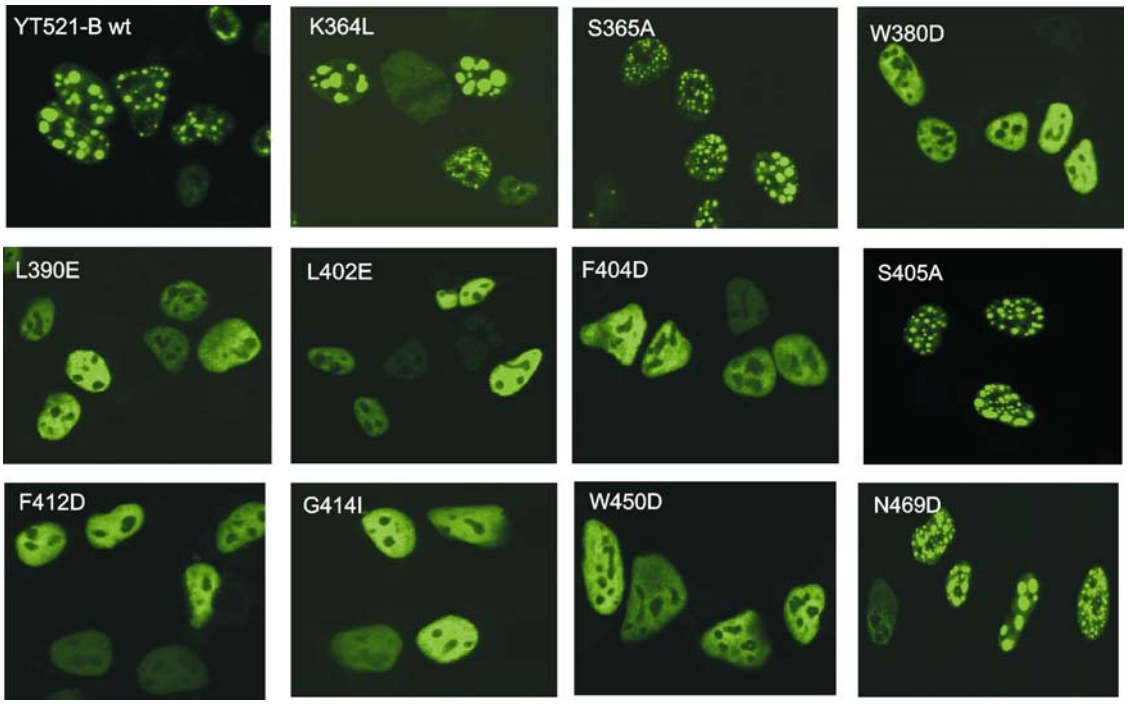


Figure 25. Mutations within YTH domain change localization of YT521-B. EGFP-YT521-B mutants were expressed in HEK293 cells. After 18 h, the formation of YT bodies was analyzed by confocal microscopy.

4.7. Mutation within YTH domain change ability of YT521-B to influence CD44 alternative pre-mRNA splicing

It has been shown previously that YT521-B is able to modulate alternative splice site selection in a concentration-dependent manner (Hartmann et al., 1999). The influence of YTH mutants on CD44 exon v5 minigene (Konig et al., 1998) splicing was analyzed using the i*n vivo* splicing assay approach (Stoss et al., 1999; Tang et al., 2005) (Figure 26). This method involves transient cotransfection of a reporter minigene and a splicing regulator. After the isolation of RNA, the splicing pattern of the minigene is analyzed by RT-PCR using minigene-specific primers. Two mutants were chosen for this experiment, K364L, which is present in the nuclear YT bodies and F412D showing a dispersed staining in the nucleus (Figure 25). As shown in Figure 27, increasing concentration of YT521-B promotes skipping of the alternative exon v5. This effect is no longer observed when F412D YT521-B mutant is cotransfected with the minigene DNA. The difference is statistically significant when comparing the values for 1 and 2 µg of transfected EGFP-

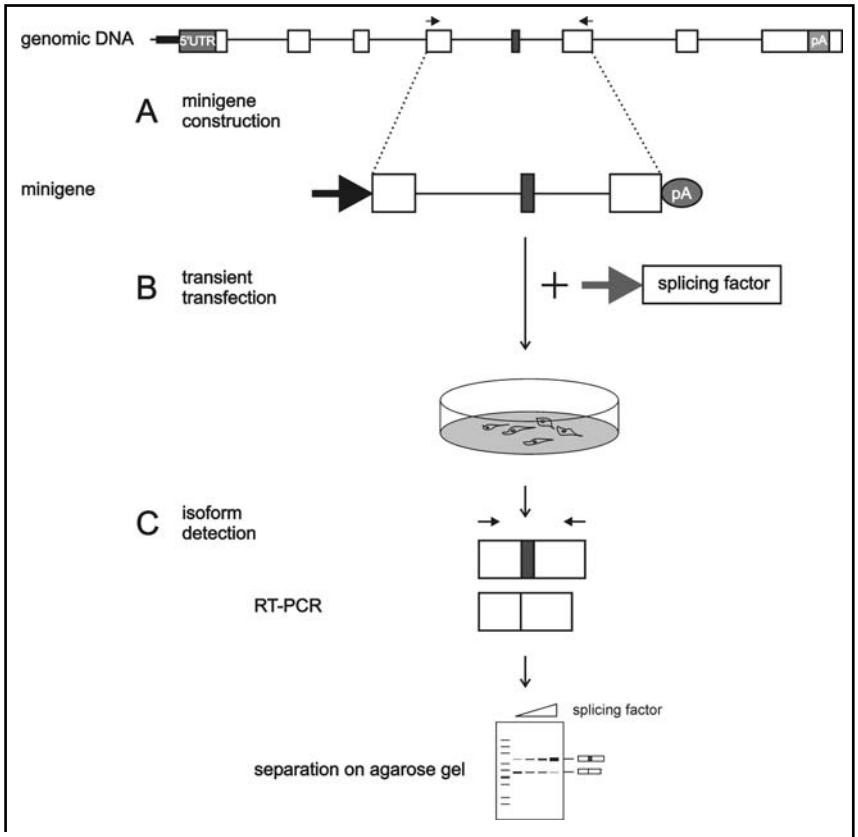


Figure 26. Overview of *in vivo* splicing analysis using minigenes.
(**A**) Minigenes are constructed by long-range PCR amplification of a part of genomic DNA containing the alternative exon (grey box) and its flanking constitutive exons (white boxes). The minigene is cloned into an expression vector containing promoter (black pointed arrow) and polyadenylation site (pA). (**B**) The minigene is analyzed by transfecting it into eukaryotic cells together with the putative splicing factor. (**C**) The different isoforms are determined by RT-PCR using minigene-specific primers and later visualized on the agarose gel stained with ethidium bromide.

YT521-B with corresponding values for EGFP-F412D mutant (*t*-test: $P = 0.0001$ and $P < 0.0001$, respectively). Moreover, when comparing the values for 1 and 2 µg of transfected EGFP-F412D mutant with the value for the transfected pEGFP-C2 empty vector, these differences have no statistical significance and can be considered within the error margin. The effect of K364L YT521-B mutant is the same as that of the wild-type protein. This result shows that YT521-B protein present in nuclear YT bodies is able to influence the alternative splicing of CD44 v5 pre-mRNA, whereas the protein showing the dispersed pattern is not active in this process.

4.8. YT521-B is phosphorylated on tyrosine residues

It has previously been shown that protein:protein interaction between YT521-B and Sam68 is regulated by tyrosine phosphorylation and that $p59^{fyn}$ kinase causes phosphorylation of both proteins (Hartmann et al., 1999). Since YT521-B is detected only in the nucleus (Hartmann et al., 1999), the question arose whether other non-

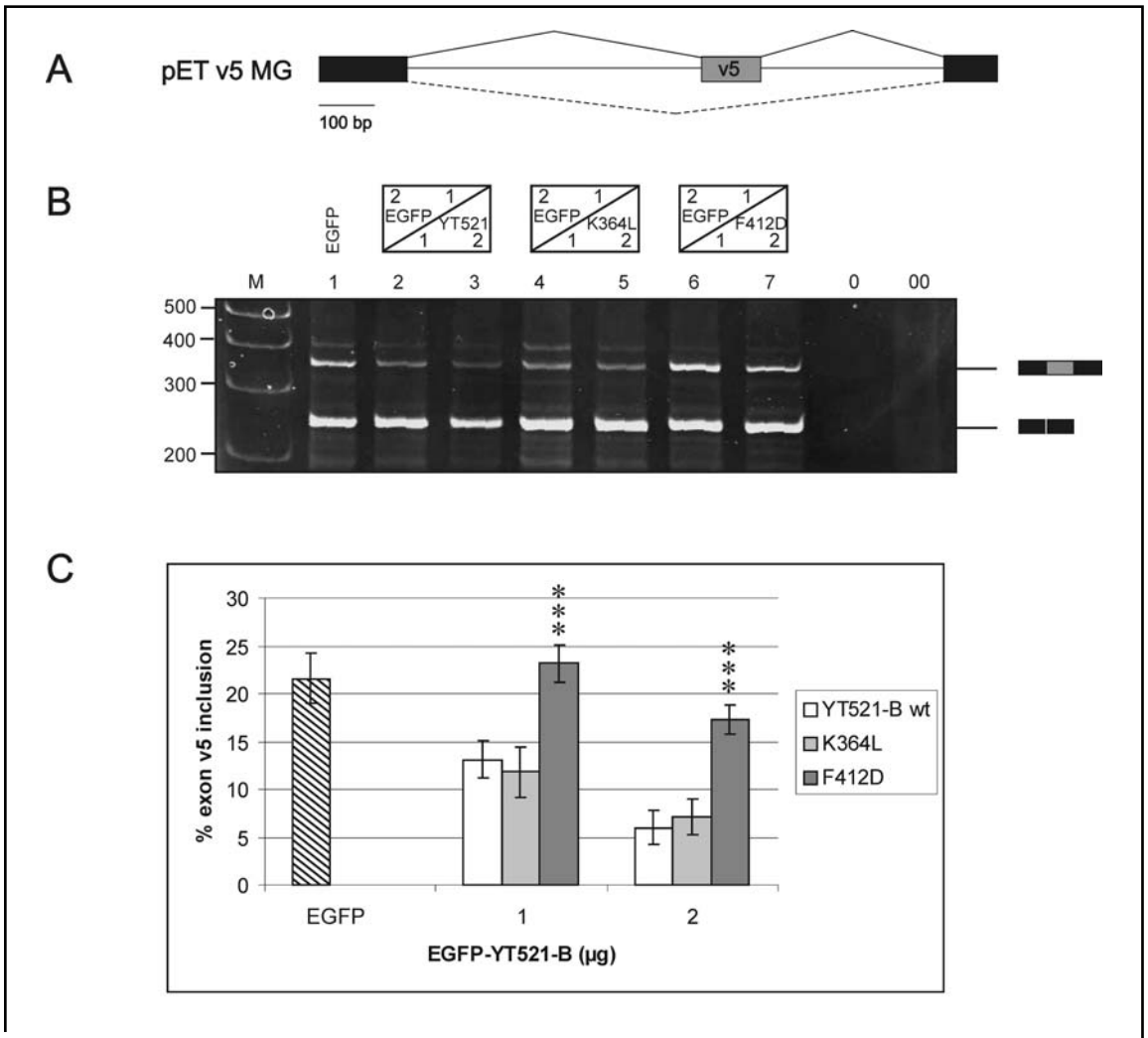


Figure 27. Mutation within YTH domain of YT521-B changes its influence on CD44 exon v5 splice site selection. (**A**) Structure of the CD44 v5 minigene used (Konig et al., 1998). Exons are shown as boxes, introns as lines. The flanking insulin exons are shown in black, the alternative CD44 exon v5 is indicated in gray. The splicing patterns are shown by different line styles. (**B**) Influence of introduced mutations on YT521-B mediated changes in CD44 exon v5 splice site selection. HEK293 cells were transiently transfected with increasing amounts of EGFP-YT521-B in the presence of the CD44 v5 minigene (lanes 2 and 3). The amount of transfected EGFP-YT521-B is indicated and normalized using pEGFP-C2. In lanes 4 and 5 increasing amounts of EGFP-K364L were transfected, in lanes 6 and 7 – EGFP-F412D. The RNA was analyzed by RT–PCR. 0, PCR without RT reaction; 00, PCR without template. The structure of the amplified products is indicated on the right; the alternative exon is shown in gray. M, 100 bp ladder. (**C**) Statistical evaluation of the RT–PCR results. The ratio between the signal corresponding to exon inclusion and all products was determined from at least three different independent experiments.

receptor tyrosine kinases, whose expression is often not only confined to the cell membrane (Pendergast, 1996), can cause its phosphorylation. The sequencing of the human genome has demonstrated that humans possess 90 unique tyrosine kinases. Thirty-two of them are non-receptor type kinases, which can be subdivided into ten subfamilies

(Robinson et al., 2000). In unstimulated cells, the activity of endogenous kinases can only be detected using phosphorylation-specific antibodies against specific target sites. To test the result of kinase activation, kinase activity was increased by transfecting cDNAs expressing the kinases. A member of each subfamily was cotransfected with EGFP-YT521-B. After immunoprecipitation, tyrosine phosphorylation of YT521-B was detected by Western Blot, using phosphotyrosine specific antibody PY20. As shown in Figure 28A, the kinases c-Abl, Rlk, $p59^{fyn}$ and c-Src, as members of the ABL, TEC and SRC families, respectively, phosphorylate YT521-B. Constitutively active YF mutant of Sik from FRK family shows a slight effect in phosphorylating YT521-B. However, in later studies with Sik wild-type kinase no tyrosine phosphorylation was observed (Rafalska et al., 2004). The kinases Syk, Csk, and FerH, representing the SYK, CSK, and FES families have no effect at all. Western Blot analysis of cells crude lysates used for immunoprecipitation (Figure 28B) demonstrates the presence of YT521-B in all experiments.

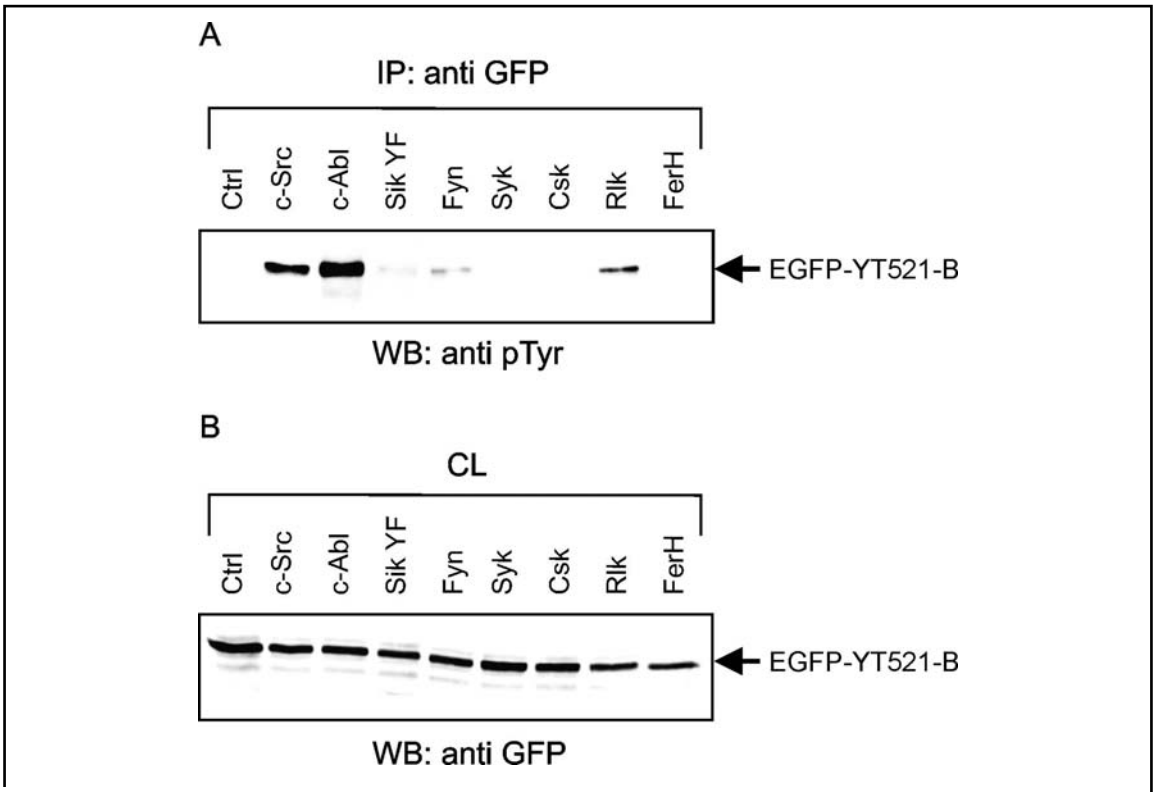


Figure 28. Several non-receptor tyrosine kinases phosphorylate YT521-B. (**A**) EGFP-YT521-B fusion protein was co-expressed in HEK293 cells with indicated tyrosine kinases. Protein was precipitated with anti GFP antibodies and tyrosine phosphorylation was detected with the anti phosphotyrosine antibody PY20. (**B**) Crude lysates of cells were analyzed for EGFP-YT521-B expression with anti GFP antibodies. CL, crude lysate; IP, immunoprecipitation.

In Figure 29, the phosphorylation of YT521-B by Bcr-Abl kinase is shown. Bcr-Abl is a constitutively active tyrosine kinase with transforming capacity for hematopoietic cells. It is a product of fusion between chromosome 9 and 22, called the Philadelphia translocation t(9;22), which is characteristic for chronic myeloid leukemia (CML) (Ben-Neriah et al., 1986; Daley et al., 1990). The fusion between Bcr protein and Abl kinase results in the improper activation of the catalytic activity of Abl. The experiment was performed as described above. In the presence of Bcr-Abl, two bands were detected with anti phosphotyrosine antibody (Figure 29, left). The lower one corresponds to EGFP-YT521-B what was shown after the reblot with anti GFP antibody (Figure 29, right). The upper band most probably represents autophosphorylated Bcr-Abl kinase. Together, these data show that YT521-B is phosphorylated by specific non-receptor tyrosine kinases.

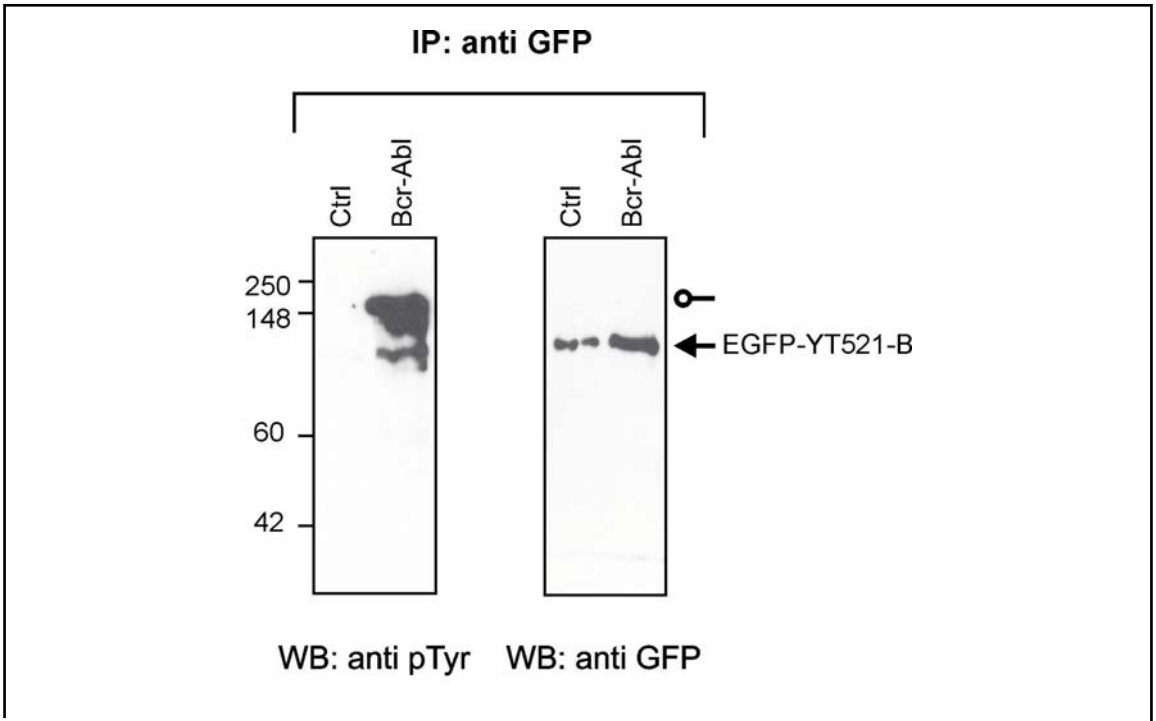


Figure 29. Bcr-Abl tyrosine kinase phosphorylates YT521-B. EGFP-YT521-B fusion protein was co-expressed in HEK293 cells with Bcr-Abl tyrosine kinase. Protein was precipitated with anti GFP antibodies and tyrosine phosphorylation was detected with the anti phosphotyrosine antibody PY20. The reblot was performed with anti GFP to demonstrate the successful immunoprecipitation. A pointed arrow indicates the localization of EGFP-YT521-B, an open round arrowhead shows the possible localization of Bcr-Abl. The molecular mass is given in kilodaltons on the left.

4.9. The proline-rich region is required for YT521-B phosphorylation

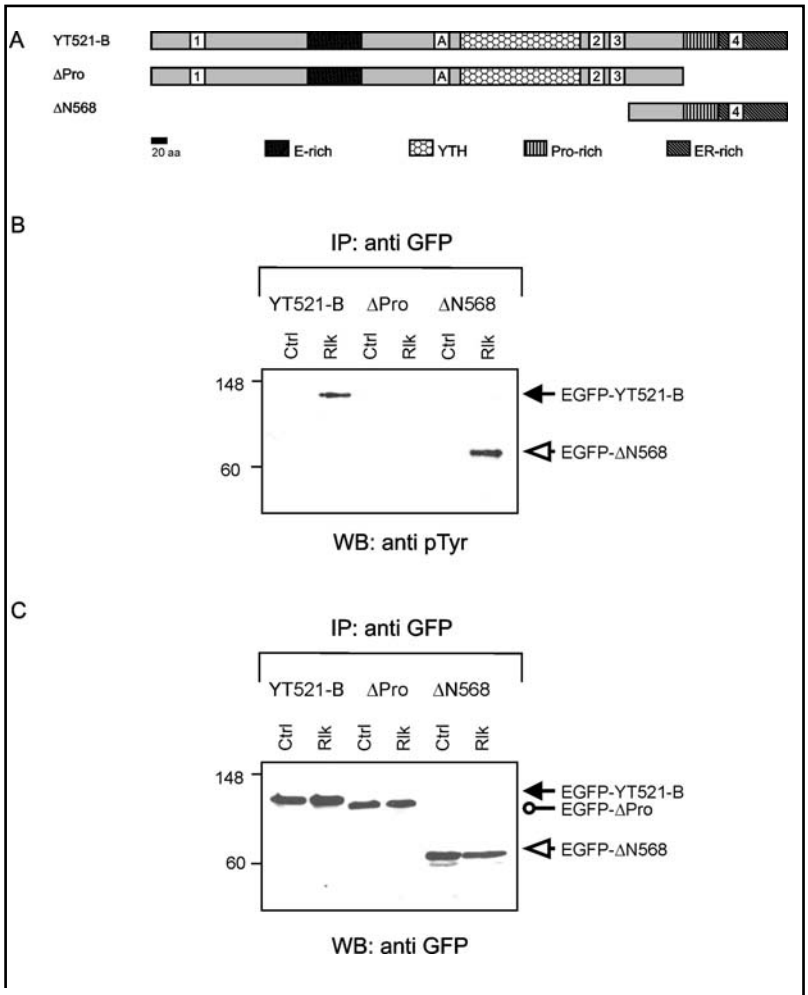


Figure 30. Proline-rich region is necessary for YT521-B phosphorylation. (**A**) Domain structure of full-length YT521-B, ΔPro, and ΔN568 deletion mutants. Nuclear localization signals are indicated with numbers. The E-rich, the proline-rich region, the ER-rich region, and YTH domain are indicated by different shading. The inserted A and B indicate the location of the alternative exons of YT521-B. (**B**) EGFP-tagged full-length YT521-B, ΔPro, and ΔN568 deletion mutants were co-expressed with or without a Rlk expression construct in HEK293 cells. Proteins were immunoprecipitated with an anti GFP antibody and tyrosine phosphorylation was detected with the anti phosphotyrosine antibody PY20. The reblot in (**C**) was performed with anti GFP to demonstrate the successful immunoprecipitation. A pointed arrow indicates the localization of EGFP-YT521-B, an open pointed arrow shows EGFP-ΔN568, and an open round arrowhead shows the localization of EGFP-ΔPro. The molecular mass is given in kilodaltons on the left.

The TEC kinases are structurally similar to SRC family kinases, having a modular organization, Src Homology (SH) 3 and SH2 protein-interaction domains are followed by the tyrosine-kinase catalytic domain (Schaeffer and Schwartzberg, 2000). Autophosphorylation of tyrosine in the SH3 domain of TEC kinases may change the affinity of the SH3 domain for certain binding partners (Morrogh et al., 1999). The splicing factors Sam68 and hnRNP K can interact with the SH3 domain through their proline-rich regions and are found among many proteins interacting with the TEC family kinases (Bunnell et al., 1996;

Guinamard et al., 1997). The Rlk tyrosine kinase, which belongs to the TEC family, was shown to phosphorylate YT521-B (Figure 28A). Subsequently, it was investigated whether Rlk can phosphorylate YT521-B deletion mutant lacking proline-rich region as well. EGFP-tagged full-length YT521-B, ΔPro, and ΔN568 were expressed in the presence of Rlk and then immunoprecipitated using their EGFP tags (Figure 30). The YT521-B ΔN568 deletion mutant contains only the carboxy-terminal domains, the proline-rich region, NLS4, and the glutamic acid/arginine-rich region, whereas ΔPro lacks the entire C-terminus, including the proline-rich region (Figure 30A). Both of these proteins are exclusively nuclear and are not present in YT bodies. ΔPro is also detected in nucleoli, whereas ΔN568, similarly to full-length YT521-B, is excluded from nucleoli (Hartmann et al., 1999). Tyrosine phosphorylation was detected by Western Blot, using the phosphotyrosine specific PY20 antibody. As shown in Figure 30B, full-length and ΔN568 YT521-B, both containing a proline-rich region are phosphorylated by Rlk. The phosphorylation signal was not detected for ΔPro deletion mutant lacking a C-terminus. Autophosphorylated Rlk isoforms, usually migrating at 59 and 53 kDa (Debnath et al., 1999), were not detected in this experiment. It is possible that Rlk construct used in the experiment does not contain the autophosphorylation site. The reblot using anti GFP antibody demonstrates the presence of expressed proteins (Figure 30C). Together, these data show that Rlk tyrosine kinase can phosphorylate YT521-B, as well as its deletion mutant containing a proline-rich region. It is however unclear whether Rlk and YT521-B interact directly through SH3 domain and PXXP motif.

4.10. YT521-B binds to c-Abl kinase

The c-Abl tyrosine kinase phosphorylates YT521-B (Figure 28A) and can be located in the nucleus (Taagepera et al., 1998). Therefore, a direct molecular interaction between YT521-B and c-Abl was investigated. It was tested whether YT521-B would coimmunoprecipitate with c-Abl in the overexpression experiment. FLAG-YT521-B was expressed in the presence of c-Abl and then immunoprecipitated using its FLAG tag (Figure 31, left). As expected, in the presence of c-Abl, two bands are tyrosine phosphorylated (Figure 31 middle). The comparison with the Western Blot probed against YT521-B demonstrates that the lower band corresponds to FLAG-YT521-B

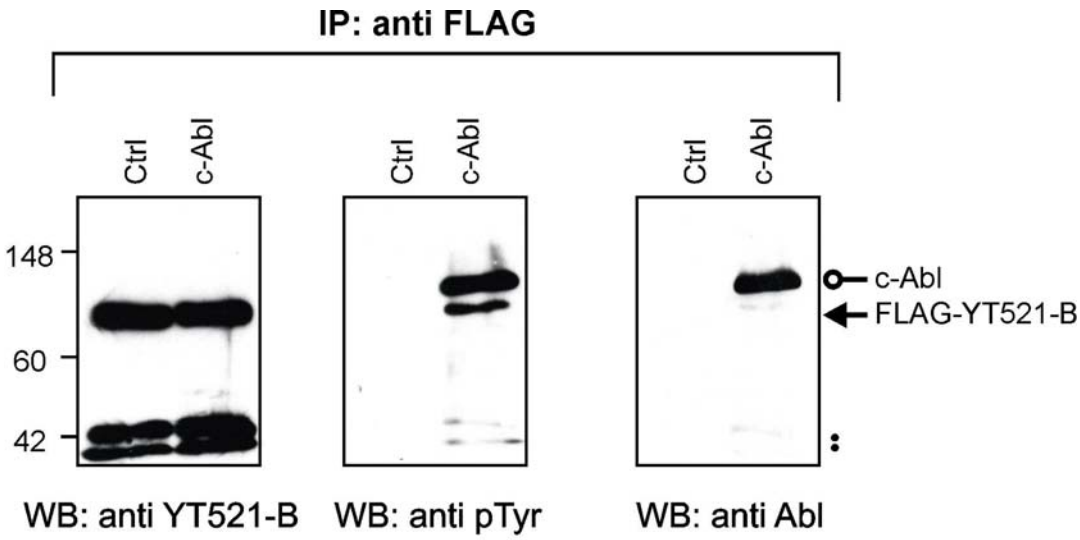


Figure 31. YT521-B binds to c-Abl tyrosine kinase. FLAG-YT521-B was co-expressed with or without a c-Abl expression construct in HEK293 cells. Proteins were immunoprecipitated with an anti FLAG antibody. Immunoprecipitates were analyzed with antisera against YT521-B (PK2) (left), phosphotyrosine (PY20) (middle) and c-Abl (right). A pointed arrow indicates the localization of FLAG-YT521-B, an open round arrowhead shows the localization of c-Abl. Dots indicate the immunoglobulin bands. The molecular mass is given in kilodaltons on the left.

(Figure 31, left). The reblot using an antibody against c-Abl shows that the upper band represents phosphorylated c-Abl (Figure 31, right). This result shows that YT521-B binds c-Abl in a direct or indirect manner, and that they can form a complex under *in vivo* conditions.

4.11. YT521-B colocalizes with c-Abl in the nucleus

Since c-Abl coimmuoprecipitates with YT521-B (Figure 31), it was determined whether it colocalizes with YT521-B in the cell. In the absence of overexpressed c-Abl kinase, EGFP-YT521-B is present in 5-20 nuclear dots, the YT bodies (Figure 32A) (Nayler et al., 2000; Hartmann et al., 1999). After cotransfection of EGFP-YT521-B with c-Abl, these dots disperse and colocalize with c-Abl in the nucleus (Figure 32C, D). As expected, c-Abl is also present in the cytosol (Figure 32B). However, YT521-B is completely absent from the cytosol, demonstrating that overexpression of c-Abl does not induce a translocation of YT521-B into the cytosol (Figure 32D). Together with the coimmunoprecipitation results (section 4.10.), these data suggest that YT521-B is a target of c-Abl in the nucleus.

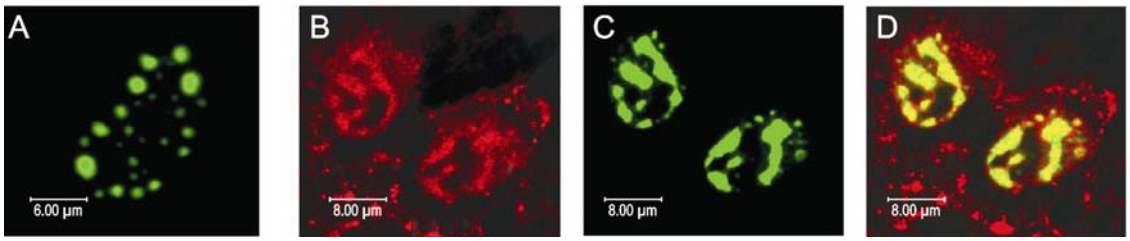


Figure 32. YT521-B colocalizes with c-Abl in the nucleus. EGFP-YT521-B was co-expressed together with c-Abl in COS7 cells. EGFP-YT521-B is shown in green and c-Abl is stained in red. EGFP-YT521-B forms characteristic YT bodies (**A**). Cotransfected c-Abl localizes to the nucleus and can also be detected in the cytosol (**B**). In the presence of transfected c-Abl, YT521-B signal is spread throughout the nucleoplasm (**C**), excluding the nucleoli. Superimposition of the images in (B) and (C) shows nuclear co-localization between EGFP-YT521-B and c-Abl (**D**).

4.12. YT521-B shuttles between nucleus and cytosol

YT521-B is phosphorylated by the membrane bound kinase c-Src (Figure 28A), to which it binds in immunoprecipitation (Rafalska et al., 2004). Since YT521-B is detected exclusively in the nucleus of cells, a phosphorylation by membrane bound kinase would only be possible during mitosis when the nuclear structure disintegrates or would require shuttling of YT521-B between nucleus and cytosol that has been reported for several proteins implicated in splice site selection (Caceres et al., 1998). To test these possibilities, the capability of EGFP-YT521-B to shuttle was assessed using a previously described cell fusion assay. The experiment was performed in collaboration with Dr Ruth Brack-Werner from GSF Institute of Molecular Virology in Neuherberg. The assay monitors accumulation of a fluorescent protein in acceptor nuclei of a newly formed polykaryon (Neumann et al., 2001; Lee et al., 1999). Transfected HeLa cells expressing EGFP-YT521-B were fused with an excess of untransfected HeLa cells and accumulation of EGFP-YT521-B in acceptor nuclei was monitored by time-lapse imaging. The donor cells were marked with red-fluorescent protein to allow distinction from acceptor cells. As shown in Figure 33, EGFP-YT521-B fluorescence is visible in the acceptor nuclei 28 minutes after cell fusion and increases by 307 minutes. Concomitantly, fluorescence intensity of the donor nucleus decreases. No shuttling is observed for GFP-tagged B23 (Neumann et al., 2001) or histone 2B (data not shown) within 360 and 720 minutes, respectively, in this assay. These results indicate that YT521-B shuttles between nucleus and cytosol, where it can interact with membrane bound SRC family kinases.

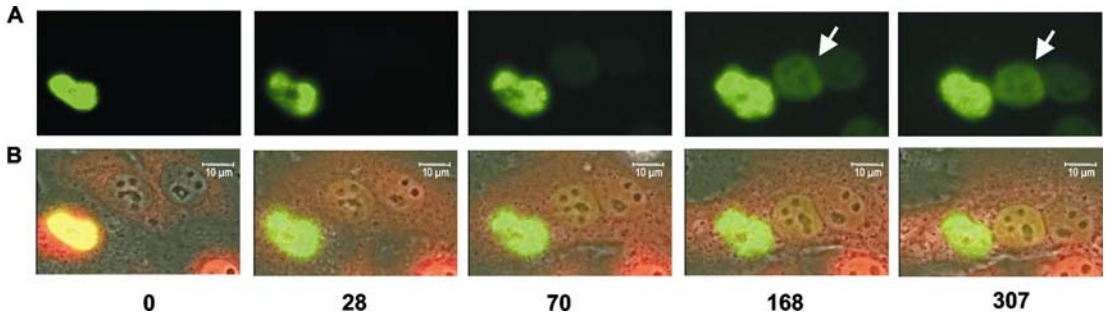


Figure 33. YT521-B shuttles between nucleus and cytosol. HeLa cells transfected with EGFP-YT521-B were fused to an excess of untransfected cells. (**A**) EGFP fluorescence at the time points in minutes after fusion indicated at the bottom. (**B**) The phase contrast images. An arrow indicates acceptor nuclei showing accumulation of EGFP-YT521-B. Numbers indicate the time after fusion in minutes.

4.13. YT bodies disperse after tyrosine phosphorylation

The effect of c-Abl on YT521-B is reminiscent of SR-protein kinases causing dispersion of nuclear speckles containing their target SR-proteins (Colwill et al., 1996a). Therefore, it was tested whether other tyrosine kinases phosphorylating YT521-B have a similar effect. EGFP-YT521-B was cotransfected with c-Src, Rlk, and $p59^{fyn}$ kinases (Figure 28A). As shown in Figure 34A, coexpression of YT521-B with these kinases causes dispersion of YT bodies. As in the situation with c-Abl, YT521-B does not leave the nucleus or enter space occupied by nucleoli. Catalytic inactive mutants of c-Src and $p59^{fyn}$ (SrcKA, FynKA) have no effect, what demonstrates that YT bodies dispersion occurs due to tyrosine phosphorylation. In further studies, it was also shown that activation of the Ras-pathway that ultimately leads to phosporylation of Sam68 (Weg-Remers et al., 2001) has no effect. Similarly, Csk that does not phosphorylate YT521-B does not alter YT body formation (Rafalska et al., 2004). Altogether, these results indicate that dispersion of YT bodies is dependent on YT521-B phosphorylation. Some differences in nuclear staining after phosphorylation can be observed, as small dot-like structures were still visible after phosphorylation with Rlk. Therefore the staining pattern was quantified and the nuclear space, containing YT521-B bodies, diffuse YT521-B staining or no YT521-B signal, was determined. YT bodies were defined as an accumulation of green signal stronger than the double standard deviation of the total signal distribution, which corresponds well with the visual inspection. As shown in Figure 34B, in untreated cells, about 10% of the nuclear space is occupied by YT bodies.

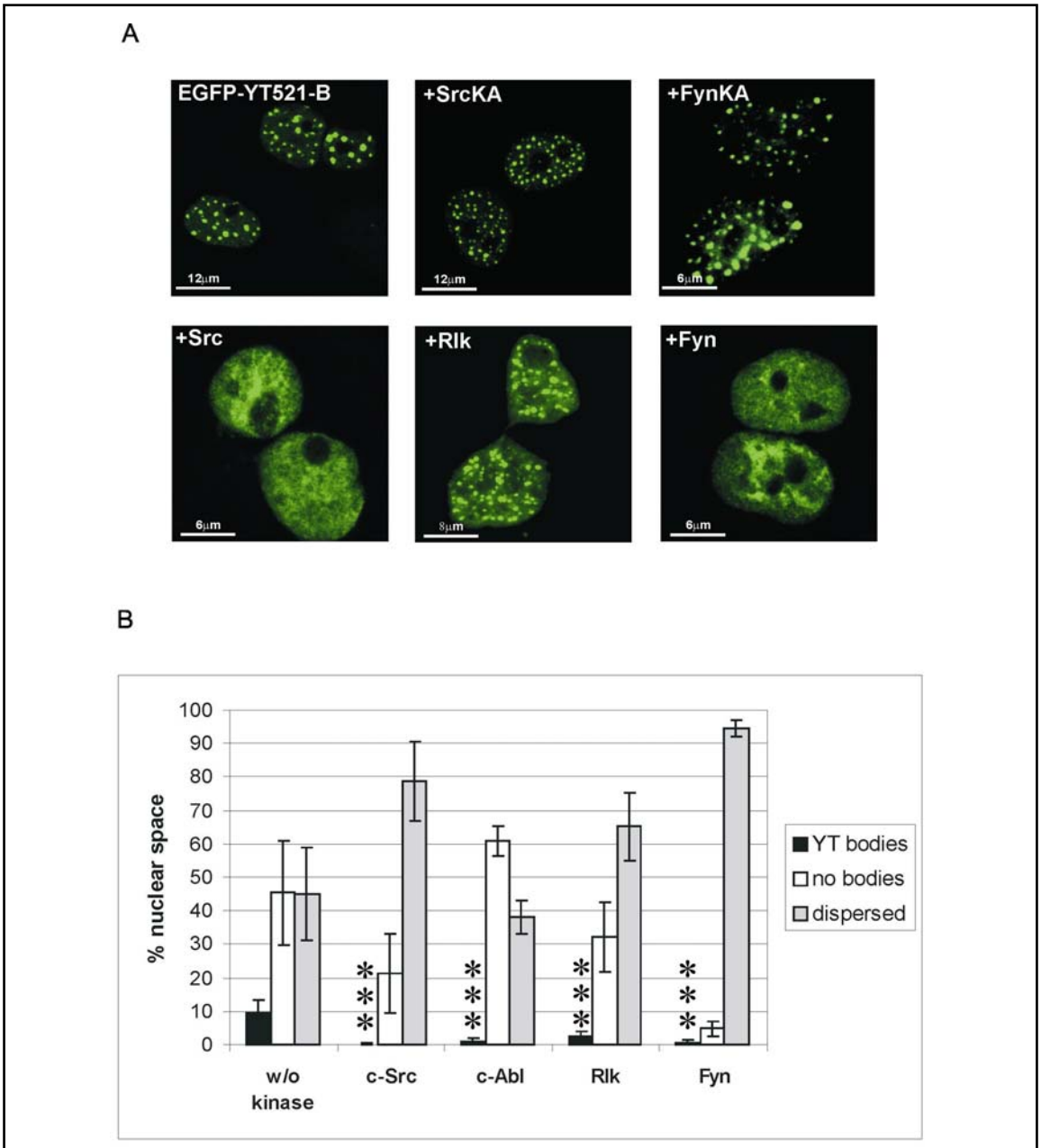


Figure 34. Phosphorylation of YT521-B dissolves YT bodies. EGFP-YT521-B was co-expressed in COS7 cells with indicated kinases. After 18 h, the formation of YT bodies was analyzed by confocal microscopy. (**A**) YT bodies formed by overexpressed EGFP-YT521-B dissolve after phosphorylation. (**B**) Statistical evaluation of YT body formation in the presence of kinases. The nuclear areas containing YT bodies, diffuse nuclear staining and which were void of YT521-B signal were determined from 100 cells. Error bars indicate the standard deviations. The differences for YT body formation were statistically significant. The Student's *t*-test gave values of $P<0.0001$ for all kinases, one-tailed *t*-values were 11.1, 10.0, 10.3 and 10.4 for c-Src, c-Abl, Rlk and $p59^{fyn}$, respectively.

The presence of kinases reduces this number to 0-2.7%. The student's *t*-test shows that these differences are statistically significant ($p < 0.0001$). The dispersion of YT bodies is

concomitant with a change of the cellular areas where YT521-B is not present. It increases from 35-50% in untreated cells to 65-90% in Rlk, c-Src and $p59^{fyn}$ transfected cells. In contrast, in c-Abl transfected cells, the area free of YT521-B slightly increases to 60%. It is possible that this is an effect of c-Abl overexpression, which directs YT521-B to nuclear areas containing AT-rich DNA preferred by c-Abl for binding (Van Etten, 1999). These experiments show that the localization of YT521-B in YT bodies is regulated by tyrosine phosphorylation, similar to the localization of SR-proteins in nuclear speckles that is influenced by serine/threonine phoshorylation.

4.14. The solubility of YT521-B is regulated by phosphorylation

In the nucleus, YT521-B is present in two biochemically distinguishable compartments (Nayler et al., 2000). The first compartment is YT bodies, which are soluble under non denaturing conditions in Triton X-100 containing buffers. The remaining YT521-B protein resides diffusely in the nucleoplasm and is insoluble in Triton X-100 containing buffers. It was previously demonstrated that YT521-B can translocate during the cell cycle or after actinomycin D treatment between these two compartments (Nayler et al., 2000). High concentration of actinomycin D (50µg/ml) that block all three RNA polymerases causes complete dispersion of YT bodies and accumulation of YT521-B in the Triton X-100 insoluble nuclear fraction (Nayler et al., 2000). This characteristic differentiates YT bodies from speckles, which round up and grow after transcription inhibition with actinomycin D (Misteli et al., 1997). Since the picture of cells after actinomycin D treatment (Nayler et al., 2000) is reminiscent of those of cells containing tyrosine phosphorylated YT521-B (Figure 34A) it was therefore tested whether tyrosine phosphorylation of EGFP-YT521-B would also result in the accumulation of the protein in an insoluble nuclear fraction. EGFP-YT521-B was transfected into HEK293 cells and its solubility in a 1% Triton X-100-based cell lysis buffer (HNTG) was analyzed. As shown in Figure 35, in the presence of $p59^{fyn}$, c-Abl or c-Src EGFP-YT521-B is present in the Triton insoluble pellet (P) fraction (Figure 35A). In contrast, when a RIPA/benzonase buffer was used, the protein was present predominantly in the soluble (S) fraction (Figure 35B), demonstrating that the protein is not covalently bound. In later studies, it was proven that this effect is dependent on

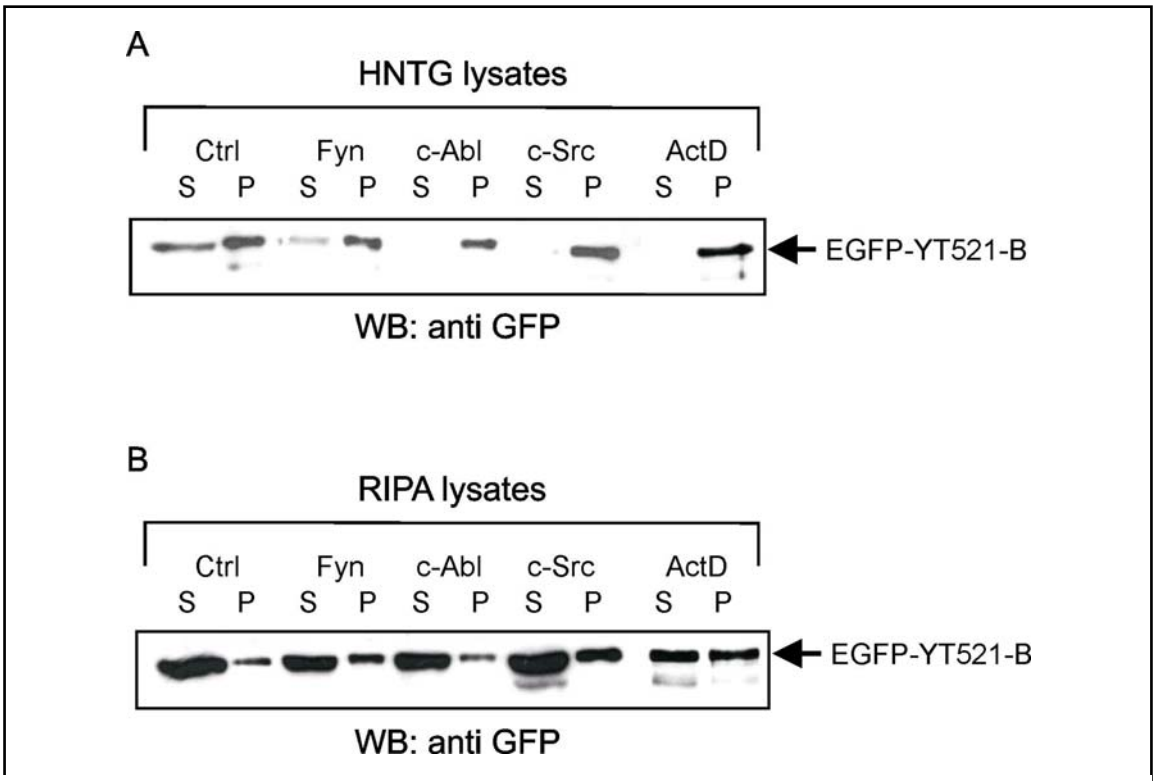


Figure 35. Phosphorylation of YT521-B influences its accumulation in a Triton X-100 insoluble nuclear compartment. The solubility of overexpressed EGFP-YT521 B was determined in the presence of Fyn, c-Abl, c-Src and actinomycin D in HNTG (**A**) and RIPA (**B**) buffers. EGFP-YT521-B was co-expressed with $p59^{fyn}$, c-Abl, and c-Src tyrosine kinases or EGFP-YT521-B transfected cells were treated with 50 µg/ml actinomycin D for 3 h. Cells were lysed in HNTG buffer (A) or RIPA buffer containing benzonase (B). Equal amounts of the soluble fraction (S) or redissolved pellet fraction (P) were analyzed in 10% SDS–PAGE gels. Proteins were analyzed by western blot and ECL using the anti GFP antibody.

YT521-B phosphorylation, as the presence of Csk kinase, which does not phosphorylate YT521-B has no effect (Rafalska et al., 2004). Moreover, all these results were confirmed with experiments conducted on endogenous YT521-B in HEK293 cells (Rafalska et al., 2004). Based on these results, it can be concluded that tyrosine phosphorylation of YT521-B causes its association with insoluble nuclear structures.

4.15. YT521-B regulates the splicing of several minigenes in a concentration-dependent manner

It was described previously that YT521-B is capable of influencing alternative splice site selection (Hartmann et al., 1999). As a part of this work, several minigenes were employed in *in vivo* splicing assays (Stoss et al., 1999; Tang et al., 2005) to test which pre-mRNA can be a target for YT521-B (Table 13).

Table 13. Minigenes tested for YT521-B-dependent splice site selection and changes occurring after tyrosine phosphorylation.

Minigene	Description	YT521-B-dependent change in splicing	Phosphorylation-dependent change in splicing	Reference
pCR MG Tra	hTra2-beta exons 1, 2, 3 and 4	Skipping of exon 2	No change	(Nayler et al., 1998b; Stoilov et al., 2004)
pET v5	CD44 exon v5	Skipping of exon v5	Prevention of exon v5 skipping	(Konig et al., 1998)
pXB	SRp20 (X16) exons 3, 4, 5 and 6	Skipping of exon 4	Prevention of exon 4 skipping	(Jumaa and Nielsen, 1997)
SV9/10L/11	Microtubule-associated protein tau exons 9, 10 and 11	Skipping of exon 10	n.d.	(Gao et al., 2000)
7T3 p3#10 p3#6	Human CFTR exons 8, 9 and 10 with mutations within polypyrimidine tract	No effect	n.d.	(Nissim-Rafinia et al., 2000)
pXGH5 wt ISEm1 ISEm2	Human GH1 exons 2, 3, 4 and 5 with mutations within ISE	No effect	n.d.	(Ryther et al., 2003)
IL4R Lng	Mouse IL4R exons 7, 8 and 9	Retention of intron 9	Prevention of intron 9 retention	(Rafalska et al., 2004)
pMTE1A	Adenovirus E1A gene	No effect	No effect	(Zerler et al., 1986; Caceres et al., 1994)

As shown in Figure 36, increasing amounts of YT521-B expressed with hTra2-beta minigene (Stoilov et al., 2004; Nayler et al., 1998b) result in the blocking of beta4 isoform formation in favor of beta1 isoform. hTra2-beta4 isoform contains all four exons included in the minigene, whereas hTra2-beta1 consists of exons one, three, and four. The observed effect is statistically significant for 1, 1.5, and 2 μg transfected EGFP-YT521-B (*t*-test: $P = 0.005$, $P = 0.0001$, and $P < 0.0001$, respectively). A similar result was obtained for SRp20 minigene (Jumaa and Nielsen, 1997) where the decreasing of exon 4 incorporation is observed due to increasing concentration of YT521-B (Figure 37). This effect is also statistically significant with $P = 0.001$ and $P = 0.0003$ for 1 and 2 μg of transfected DNA, respectively. The effect of YT521-B on the splicing of tau minigene (Gao et al., 2000) is shown in Figure 38. Increasing amounts of YT521-B promote skipping of tau exon 10 ($P < 0.0001$ for 1 and 2 μg of transfected DNA). These results indicate that the relative concentration of YT521-B can influence specific alternative splicing events.

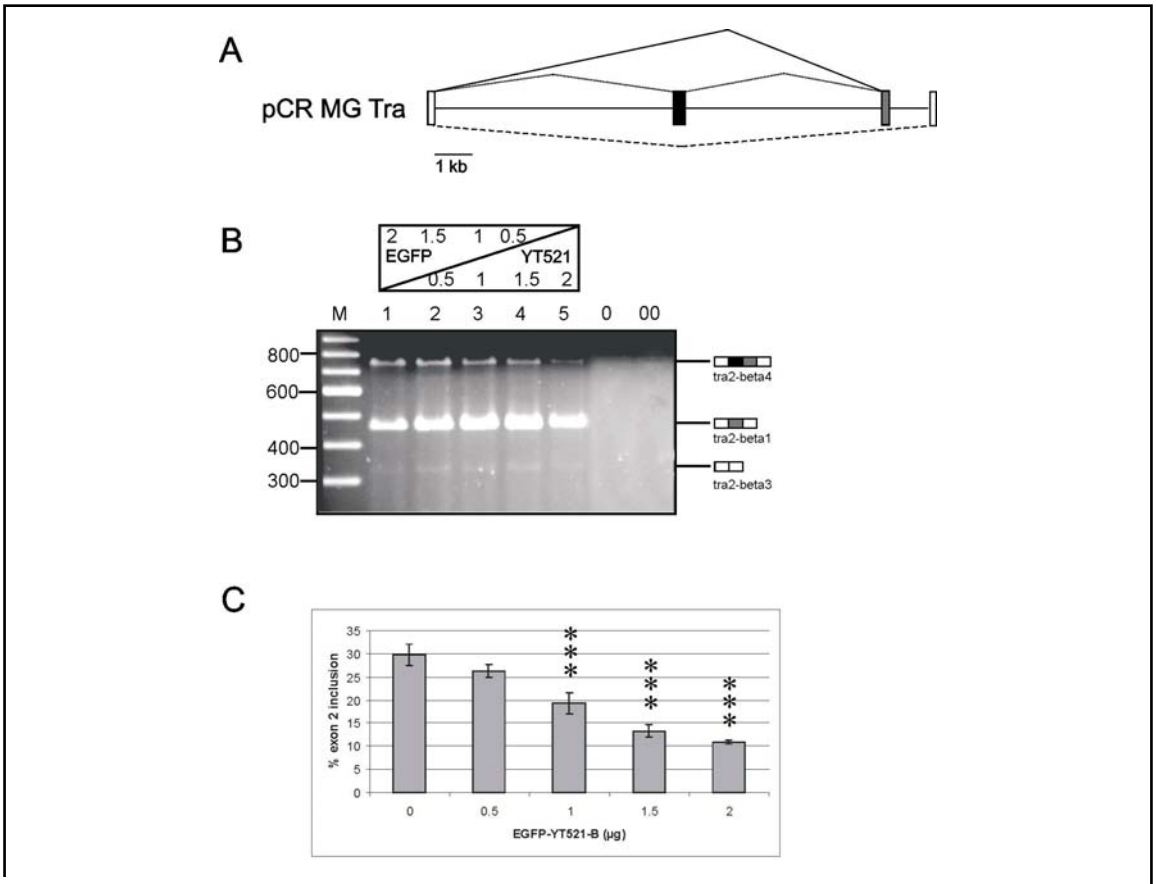


Figure 36. YT521-B influences hTra2-beta splice site selection. (**A**) Structure of the pCR MG Tra minigene used (pCR3.1 MGtra (Stoilov et al., 2004)). Exons are shown as boxes, introns as lines. Constitutive hTra2-beta exons 1 and 4 are shown in white, the alternative exons 2 and 3 are indicated in black and gray, respectively. The splicing patterns are shown by different line styles. (**B**) HEK293 cells were transiently transfected with increasing amounts of EGFP-YT521-B in the presence of the hTra2-beta minigene. The amount of transfected EGFP-YT521-B is indicated and normalized using pEGFP-C2. The RNA was analyzed by RT–PCR. 0, PCR without RT reaction; 00, PCR without template. The structure of the amplified products is indicated on the right. M, 100 bp ladder. (**C**) Statistical evaluation of the RT–PCR results. The ratio between the signal corresponding to exon 2 inclusion and all products was determined from at least three different independent experiments.

4.16. YT521-B acts on the interleukin-4 receptor pre-mRNA

YT521-B has been shown to influence alternative splice site selection *in vivo* (section 4.15.). In addition, it has been shown that tyrosine phosphorylation can regulate physical properties of the protein (sections 4.13. and 4.14.) (Hartmann et al., 1999). It was therefore investigated whether tyrosine phosphorylation can alter the ability of YT521-B to change splice site selection. Alternative splicing of the interleukin-4 receptor is regulated at least in part by tyrosine phosphorylation events emanating from stimulation of the IL-4 receptor by IL-4 (Blum et al., 1996). Due to alternative splicing

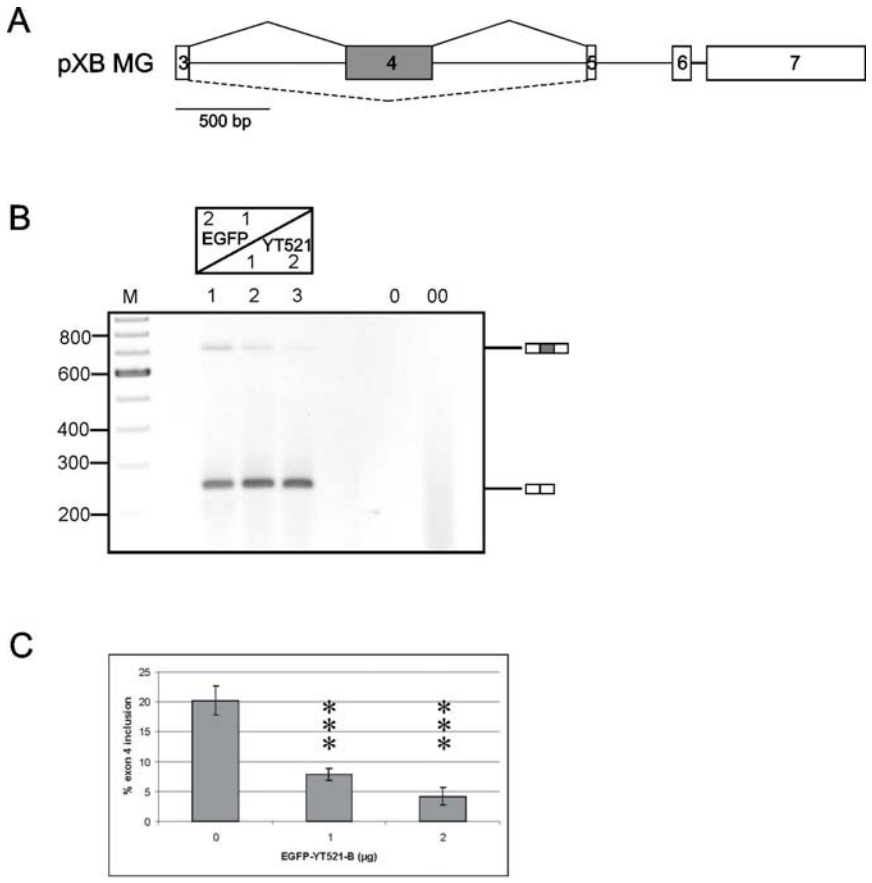


Figure 37. YT521-B influences SRp20 exon 4 splice site selection. (**A**) Structure of the SRp20 minigene used (pXB MG (Jumaa and Nielsen, 1997)). Exons are shown as boxes, introns as lines. The alternative exon 4 is indicated in gray. The splicing patterns are shown by different line styles. (**B**) HEK293 cells were transiently transfected with increasing amounts of EGFP-YT521-B in the presence of the SRp20 minigene. The amount of transfected EGFP-YT521-B is indicated and normalized using pEGFP-C2. The RNA was analyzed by RT–PCR. 0, PCR without RT reaction; 00, PCR without template. The structure of the amplified products is indicated on the right. M, 100 bp ladder. (**C**) Statistical evaluation of the RT–PCR results. The ratio between the signal corresponding to exon inclusion and all products was determined from at least three different independent experiments.

and incorporation of an in-frame stop codon, a soluble and a membrane bound form of the IL-4 receptor are formed. The soluble form has no effect in mediating the IL-4 signal to the cell. An IL-4 receptor minigene comprising of the alternative exon 8 of the mouse IL-4 receptor flanked by the constitutive exons 7 and 9 (Wrighton et al., 1992) was constructed using standard methods (Stoss et al., 1999; Tang et al., 2005). The minigene cassette was cloned between the insulin exons of an exon-trap vector (Figure 39A). First it was investigated whether YT521-B interacts with the IL-4 receptor mRNA *in vivo*. The interaction between YT521-B and pre-mRNA is possible, because YT521-B binds to

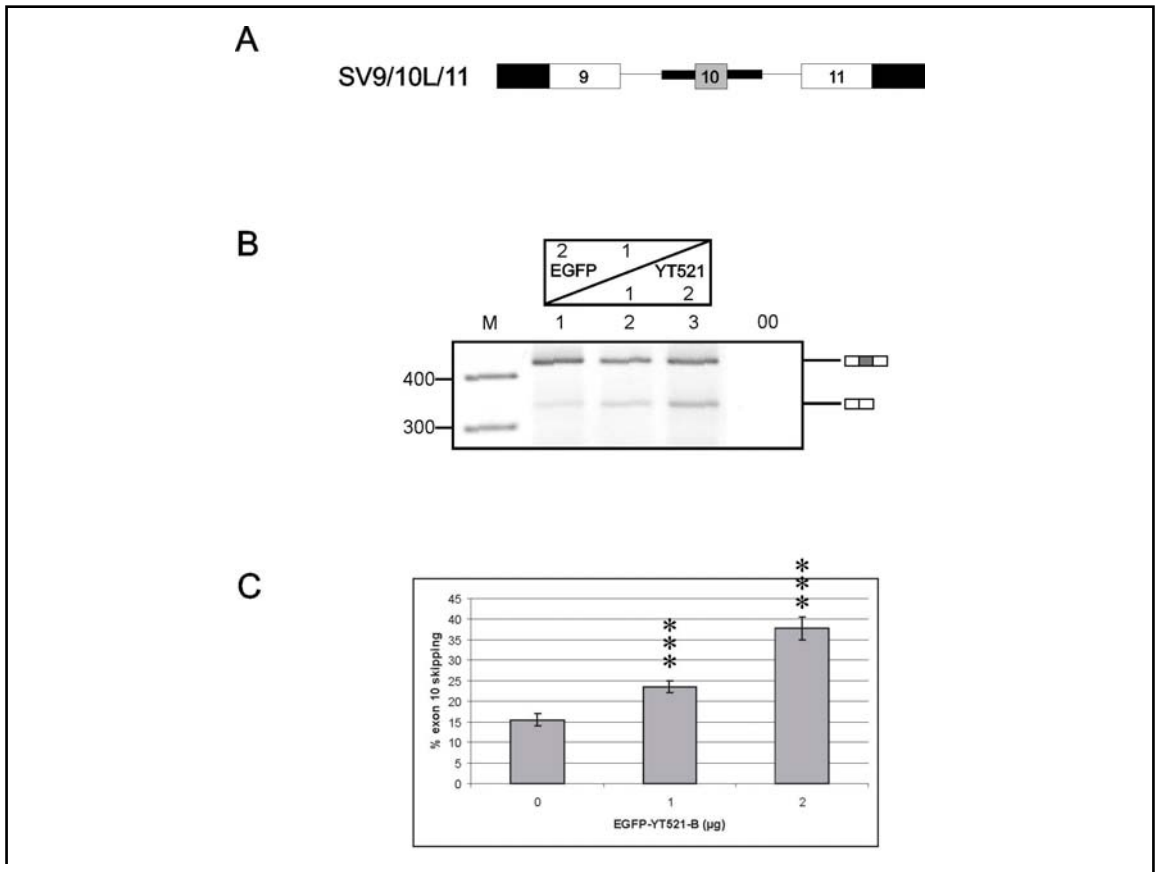


Figure 38. YT521-B influences Tau exon 10 splice site selection. (**A**) Structure of the Tau minigene used (SV9/10L/11 (Gao et al., 2000)). Exons are shown as boxes, introns as lines. The flanking insulin exons are shown in black, the alternative exon 10 is indicated in gray, proximal portions of exon 10 flanking introns (471 bp upstream and 408 bp downstream) are shown as thick lines. (**B**) HEK293 cells were transiently transfected with increasing amounts of EGFP-YT521-B in the presence of the Tau minigene. The amount of transfected EGFP-YT521-B is indicated and normalized using pEGFP-C2. The RNA was analyzed by RT–PCR. 0, PCR without RT reaction; 00, PCR without template. The structure of the amplified products is indicated on the right. M, 100 bp ladder. (**C**) Statistical evaluation of the RT–PCR results. The ratio between the signal corresponding to exon exclusion and all products was determined from at least three different independent experiments. Stars indicate statistical significant differences with $P<0.0001$.

SAM68, rSLM-2, hnRNP G and SAF-B *in vivo*. All these proteins have been implicated in splice site selection (Hartmann et al., 1999). Furthermore, YT521-B can possibly bind nucleic acids via its YTH domain (Stoilov et al., 2002b). EGFP-YT521-B was immunoprecipitated in the presence of the IL-4R minigene and RNA was detected in the precipitates by RT-PCR (Buckanovich and Darnell, 1997) using IL-4 receptor minigene specific primers. The amplification of IL-4 receptor variants demonstrates that YT521-B associates with this mRNA and can possibly regulate it (Figure 39B). The presence of c-Abl reduced the amount of RNA found in the immunoprecipitates (Figure 39B, lanes 1

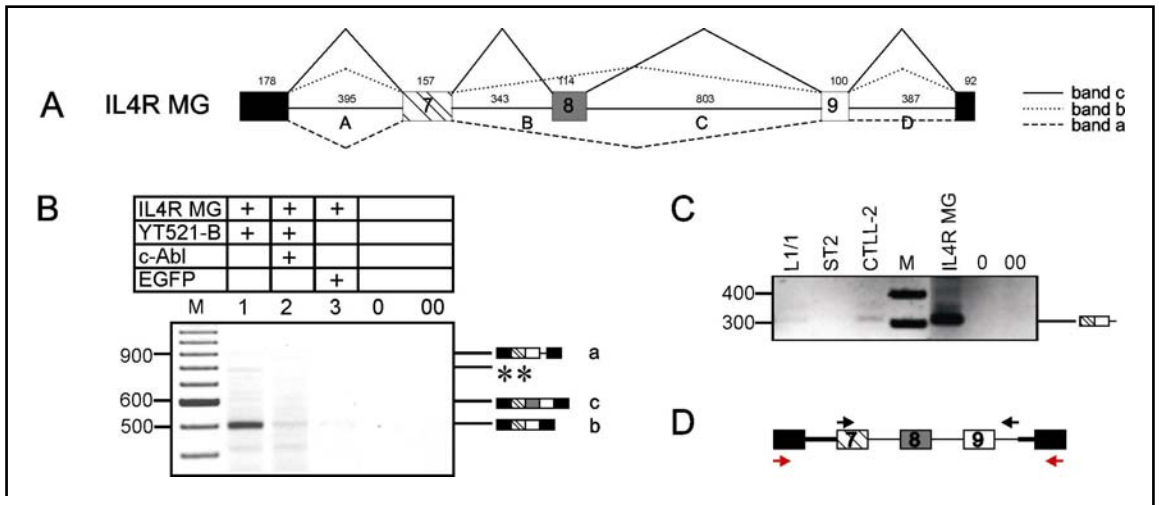


Figure 39. YT521-B is associated with the interleukin-4 receptor pre-mRNA. (**A**) Structure of the IL-4R minigene used. Exons are shown as boxes, introns as lines. The flanking insulin exons are shown in black, constitutive IL-4 receptor exons are shown in white or white-hatched boxes, the alternative IL-4R exon is indicated in gray. Numbers indicate the length of the introns and exons, letters indicate introns of the minigene construct. The splicing patterns are shown by different line styles. Bands a, b and c correspond to the bands in (B). (**B**) YT521-B immunoprecipitates contain interleukin-4 receptor pre-mRNA intermediates. The IL-4R minigene was cotransfected with EGFP-YT521-B. Protein–RNA complexes were isolated by immunoprecipitation with anti GFP in the absence of RNases. In lane 2, the c-Abl was cotransfected. Interleukin-4 receptor mRNA was detected by RT–PCR using minigene specific primers. (**C**) Retention of intron D exists *in vivo*. mRNAs from the lymphocyte cell lines L1/1, ST2 and CTLL-2 were analyzed by RT–PCR using primers in intron D and exon 7 (**D**), which amplify the intron retention event (corresponding to band a) observed in the RNA-IPs. 0: empty PCR control (water instead of RNA) 00: immunprecipitate without reverse transcriptase.

and 2), although similar amounts of EGFP-YT521-B were present in the precipitates (data not shown). Sequencing of the products revealed that they were intermediates of the IL-4 receptor pre-mRNA processing (Figure 39B). Interestingly, one of the products, band a, was generated by retention of intron D (Figure 39A). To determine whether such an intron retention event occurs *in vivo*, RNA from various lymphocytes expressing the IL-4 receptor were amplified using an intron-specific primer (Figure 39D). As shown in Figure 39C, mRNA with intron D retention is expressed in L1/1 and CTLL-2 cells (Schnare et al., 1998). This data shows that YT521-B interacts with the IL-4R pre-mRNA.

4.17. Phosphorylation changes the ability of YT521-B to influence splice site selection of IL-4R pre-mRNA

Next, it was determined whether YT521-B influences splice site selection of the IL-4R pre-mRNA. IL-4R alternative splicing pattern dependency on YT521-B

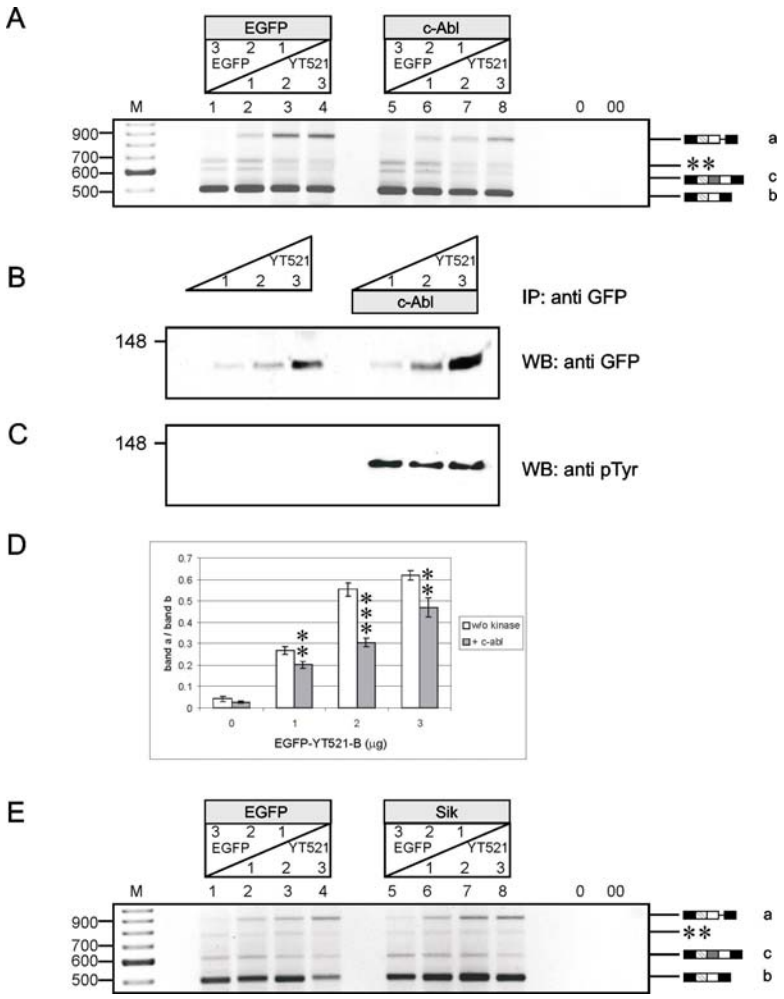


Figure 40. Phosphorylation of YT521-B changes its influence on interleukin 4 receptor splice site selection. (**A**) Influence of c-Abl on YT521-B mediated changes in splice site selection. HEK293 cells were transiently transfected with increasing amounts of EGFP-YT521-B in the presence of the IL-4R minigene (lanes 1–4). The amount of transfected EGFP-YT521-B is indicated and normalized using pEGFP-C2. An additional 0.5 µg EGFP-C2 was present in lanes 1–4 to allow comparison with the c-Abl co-transfections in lanes 5–8. In lanes 5–8, 0.5 µg of c-Abl expression plasmid was co-transfected with an increasing amount of EGFP-YT521-B. The RNA was analyzed by RT-PCR. The structure of the amplified products is indicated on the right; the alternative exon is shown in gray. (**B**) The increase of YT521-B in the transfection assays was determined by immunoprecipitating EGFP-YT521-B with anti GFP. The protein was detected with anti-GFP. (**C**) Phosphorylation of YT521-B was detected with western blot, using an anti phosphotyrosine antibody on the immunoprecipitates of (B). (**D**) Statistical evaluation of the RT–PCR results. The ratio between the signal containing all constitutive exons as well as intron D to the signal of the band containing only constitutive exons [marked a and b in (A)] was determined from at least three different independent experiments. (**E**) Sik does not alter YT521-B mediated splice site selection. The experiment was performed as described in (A). Instead of c-Abl, an expression clone for Sik was employed. M, 100 bp ladder. 0, PCR without RT reaction; 00, PCR without template.

concentration was analyzed by cotransfection and RT-PCR (Stoss et al., 1999; Tang et al., 2005), using primers in the flanking constitutive insulin exons. Transfection

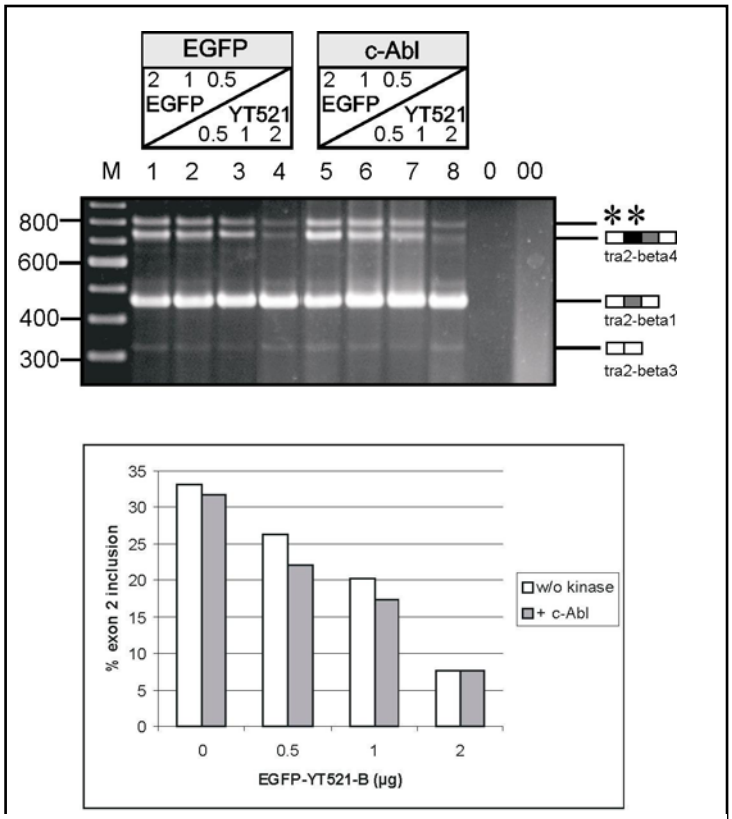


Figure 41. Phosphorylation of YT521-B does not change splice site selection of hTra-2beta. (**A**) HEK293 cells were transiently transfected with increasing amounts of EGFP-YT521-B in the presence of the hTra2-beta minigenes (lanes 1–4). The amount of transfected EGFP-YT521-B is indicated and normalized using pEGFP-C2. An additional 0.5 µg EGFP-C2 was present in lanes 1–4 to allow comparison with the c-Abl co-transfections in lanes 5–8. In lanes 5–8, 0.5 µg of c-Abl expression plasmid was co-transfected with an increasing amount of EGFP-YT521-B. The RNA was analyzed by RT–PCR. 0, PCR without RT reaction; 00, PCR without template. The structure of the amplified products is indicated on the right. M, 100 bp ladder. Evaluation of the RT–PCR result is shown.

of the IL-4R minigene in HEK239 resulted in a variant that predominantly skips the alternative exon 8 (Figure 40A, lane 1). The cotransfection of c-Abl does not change the splicing pattern (Figure 40A, lane 5) when no exogenous YT521-B is present. However, an increase in the concentration of YT521-B by cotransfection promotes retention of intron D in a dose-dependent manner (Figure 40A, lane 2-4). The presence of c-Abl significantly reduces this influence of YT521-B on alternative splice site selection (Figure 40A, lane 6-8). Western blots of immunoprecipitations made in parallel experiments confirmed the dose-dependent increase of YT521-B (Figure 40B) and verified that YT521-B was phosphorylated by c-Abl (Figure 40C). Quantification of the RT-PCR experiments (Figure 40D) revealed that the c-Abl evoked differences were statistically significant (t-test: $p < 0.0001$, < 0.006) for 2 and 3 µg transfected EGFP-YT521-B, respectively (Figure 40D). The observed 2-3 fold effect is comparable with the 1.5 to four fold effects of phosphorylation on splice site selection observed *in vivo* (Meshorer et al., 2002; Stamm, 2002; Abdennebi et al., 2002; Holdiman et al., 2002). To rule out that this is an unspecific tyrosine phosphorylation effect, the same experiment was performed in the presence of an expression clone for Sik, which

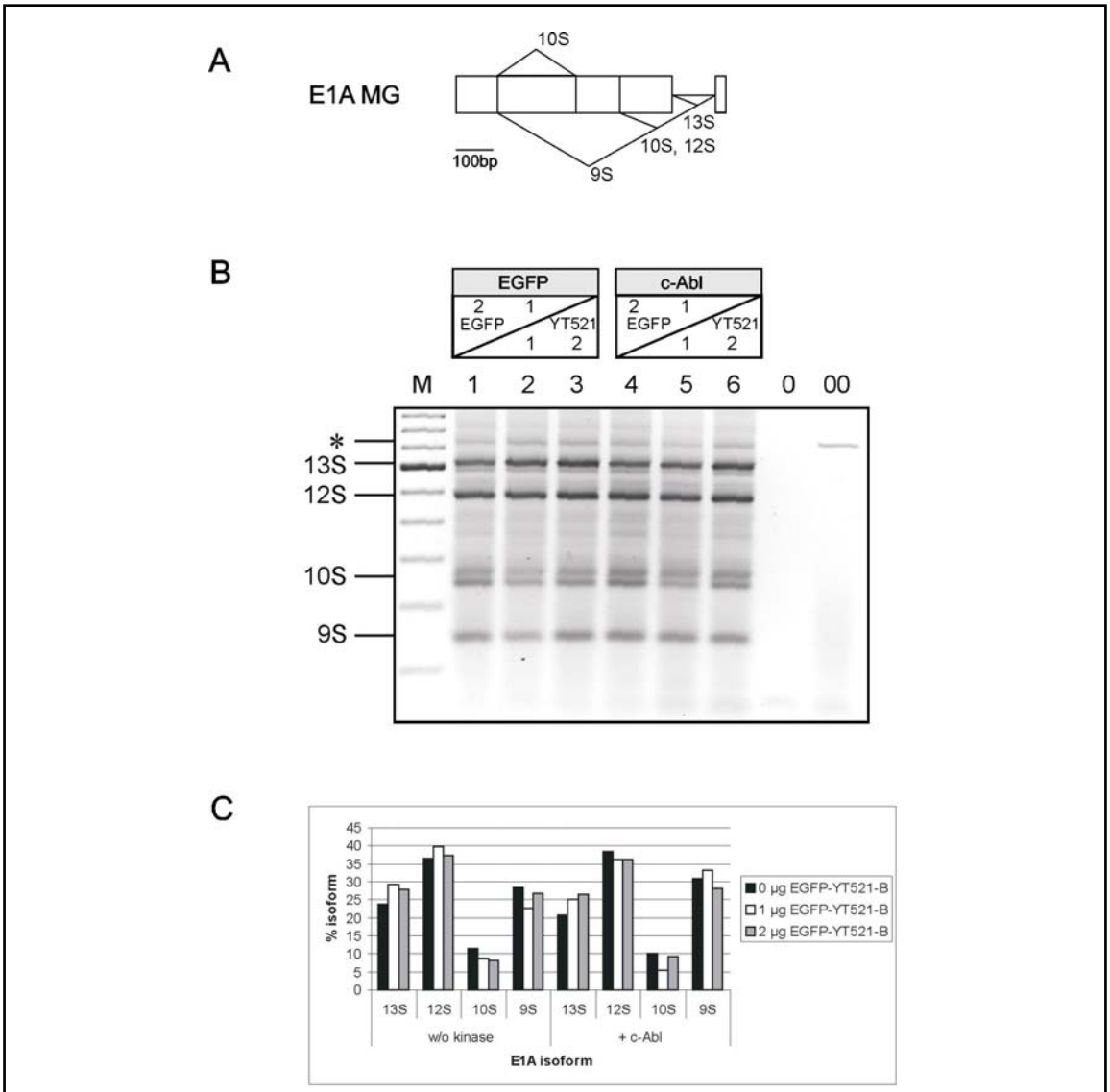


Figure 42. Phosphorylation of YT521-B does not change splice site selection of adenovirus E1A pre-mRNA. (**A**) Structure of the E1A minigene used (pMTE1A (Caceres et al., 1994)). Exons are shown as boxes, introns as lines. The E1A transcript isoforms are shown schematically. (**B**) HEK293 cells were transiently transfected with increasing amounts of EGFP-YT521-B in the presence of the E1A minigene. The amount of transfected EGFP-YT521-B is indicated and normalized using pEGFP-C2. The RNA was analyzed by RT–PCR. 0, PCR without RT reaction; 00, PCR without template. The isoforms are marked on the left; asterisk indicates unspliced pre-mRNA. M, 100 bp ladder. (**C**) Evaluation of the RT–PCR result.

does not phosphorylate YT521-B (Figure 28A). As shown in Figure 40E, Sik does not influence the YT521-B mediated changes in alternative splicing. These results indicate that tyrosine phosphorylation of YT521-B changes its influence on IL-4R splice site selection. The phosphorylation dependency of the localization of YT521-B (Figures 32,

34) suggests that changes in local concentration and solubility of YT521-B are responsible for these effects.

4.18. Phosphorylation changes the ability of YT521-B to influence CD44 alternative pre-mRNA splicing, but does not affect the splicing pattern of tra2-beta and E1A pre-mRNAs

To confirm that the dependency of YT521-B function on its phosphorylation status is present in other systems, several more minigenes were tested. Firstly, hTra2-beta and E1A minigenes (Zerler et al., 1986; Caceres et al., 1994) were employed in the experiment. As shown in Figure 41, formation of tra2-beta1 isoform is promoted by both

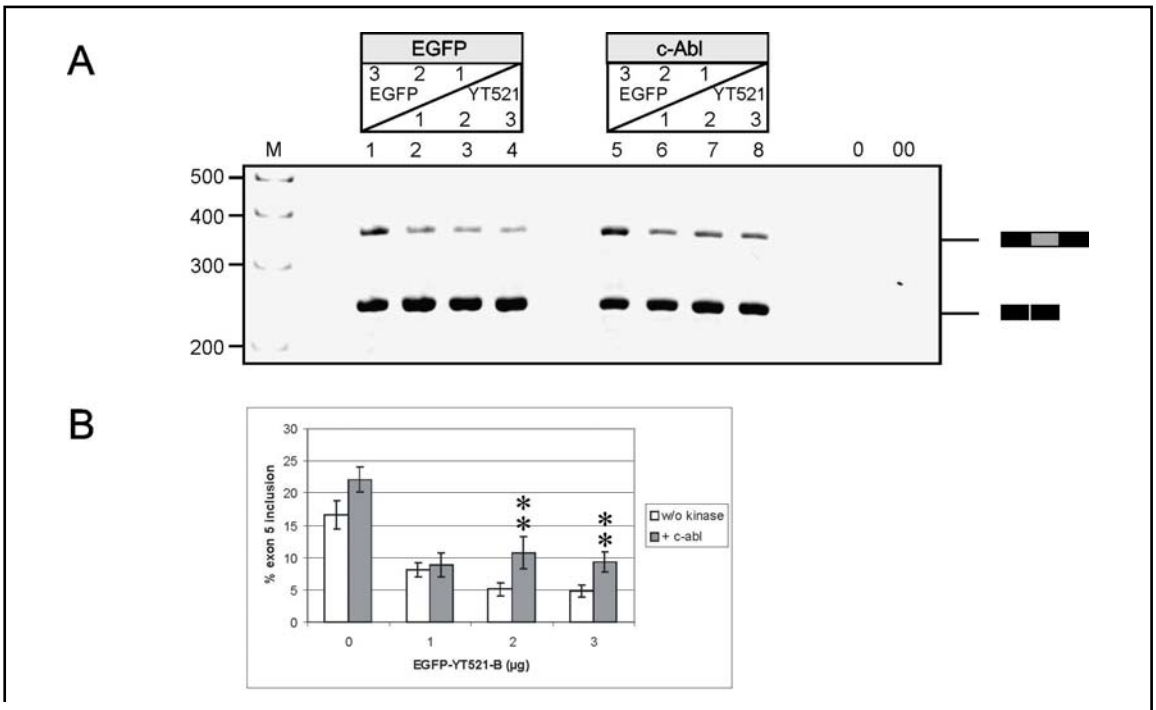


Figure 43. Phosphorylation of YT521-B changes splice site selection of CD44, exon v5. (**A**) Influence of c-Abl on YT521-B mediated changes in CD44 exon v5 splice site selection. HEK293 cells were transiently transfected with increasing amounts of EGFP-YT521-B in the presence of the CD44 v5 minigene (lanes 1–4). The amount of transfected EGFP-YT521-B is indicated and normalized using pEGFP-C2. An additional 0.5 µg EGFP-C2 was present in lanes 1–4 to allow comparison with the c-Abl co-transfections in lanes 5–8. In lanes 5–8, 0.5 µg of c-Abl expression plasmid was co-transfected with an increasing amount of EGFP-YT521-B. The RNA was analyzed by RT–PCR. 0, PCR without RT reaction; 00, PCR without template. The structure of the amplified products is indicated on the right; the alternative exon is shown in gray. M, 100 bp ladder. (**B**) Statistical evaluation of the RT–PCR results. The ratio between the signal corresponding to exon inclusion and all products was determined from at least three different independent experiments. Stars indicate statistical significant differences (P=0.005 and P=0.001, for 2 and 3 µg, respectively).

non-phosphorylated and phosphorylated YT521-B. The splicing of adenovirus E1A pre-mRNA is not altered by YT521-B and phosphorylation of protein does not change the E1A splicing pattern (Figure 42). Next, CD44 exon v5 minigene was tested. It is known that this minigene is dependent on the YT521-B interacting proteins rSLM-2 and SAM68 and that its splicing is regulated by phosphorylation (Stoss et al., 2001; Weg-Remers et al., 2001; Matter et al., 2002). As shown in Figure 43, an increase of the concentration of YT521-B reduces inclusion of the alternative exon CD44 v5 in a concentration-dependent manner. This dependency is partially abolished by c-Abl induced phosphorylation, as c-Abl does not influence the effect of small concentrations of YT521-B (Figure 43B, lanes 2 and 6), but blocks the further increase of exon skipping promoted by YT521-B (Figure 43B, lanes 3-4, and 7-8; Figure 43C). In further studies, similar results were achieved also for SRp20 minigene where the effect of YT521-B is modulated by its phosphorylation status (Rafalska et al., 2004). Together, these data show that tyrosine phosphorylation modulates the influence of YT521-B on splice site selection of several, but not all, pre-mRNAs.

5. DISCUSSION

5.1. YT521-B orthologs and YTH-containing proteins

Computational analysis presented in this work shows the existence of two groups of proteins related to rat splicing factor YT521-B. The first is a group of protein orthologs of a very high (60%–96%) sequence identity and the same modular structure. It is striking that in the performed searches no YT521-B orthologs were found in invertebrates (Figure 13). Additional searches performed in *C.elegans* and *D.melanogaster* databases confirm this observation and suggest that YT521-B may be a vertebrate-specific protein. The first described rat YT521-B protein was shown to be a component of the large protein complex acting on pre-mRNA and regulated by extracellular stimuli (Hartmann et al., 1999; Stoss et al., 2001; Rafalska et al., 2004) (Figure 44). It was also shown to influence splicing of several reporter genes *in vivo* (Hartmann et al., 1999; Rafalska et al., 2004). A high degree of evolutionary conservation among YT521-B vertebrate orthologs suggests that they fulfill similar roles and indicates the importance of the proteins in the process of alternative splice site selection.

The second group includes proteins containing YTH domain, but without any other structural similarities. These proteins belong to the YTH domain family (Pfam04146), but their function(s) remain unknown. One of the human YTH-containing proteins, high-glucose-regulated protein 8, was identified to be a tumor-associated antigen in renal-cell carcinoma (NY-REN-2) (Scanlan et al., 1999) and in chronic lymphocytic leukemia (CLL-associated antigen KW-14) (Krackhardt et al., 2002). Besides the putative RNA-binding YTH domain, it contains a proline-rich region and several potential serine and tyrosine phosphorylation sites. This may indicate that the protein plays a role or is an endpoint in the signal transduction pathway. Another YTH domain containing protein, FLJ21940, seems to belong to the DEAD helicases family and may be involved in nucleic acid metabolism. However, the structural and functional characterization of the YTH domain and proteins belonging to the YTH domain family remains to be performed. It is possible that YT521-B-like proteins form a protein family conserved during eukaryotic evolution, similar to the SR proteins family. The very high degree of phylogenetic conservation of the SR proteins (Zahler et al., 1992; Birney et al.,

1993), and in particular of their RS and RRM domains, reflects the unique properties of individual SR proteins in subcellular targeting (Caceres et al., 1997; Misteli et al., 1998; Caceres et al., 1998), recruitment of the splicing apparatus (Wu and Maniatis, 1993), and RNA binding specificity (Tacke and Manley, 1999) (reviewed in Graveley, 2000; Sanford et al., 2003). Similarly, as it is likely for YTH domain containing proteins, individual SR proteins arose by gene duplications and individual family members are more closely related to the corresponding members from different species than to other family members from the same species. For example, the RS domains of human and mouse SF2/ASF, human and mouse SRp20, and human and chicken SC35 are 98%–100% conserved at the amino acid level, whereas the sequence identity for human RS domains present in SF2/ASF, SRp20, and SC35 is 40%–56%. In comparison, YT521-B orthologs share 60%–96% sequence identity, whereas human YTH domains from distant proteins only 31%–49%.

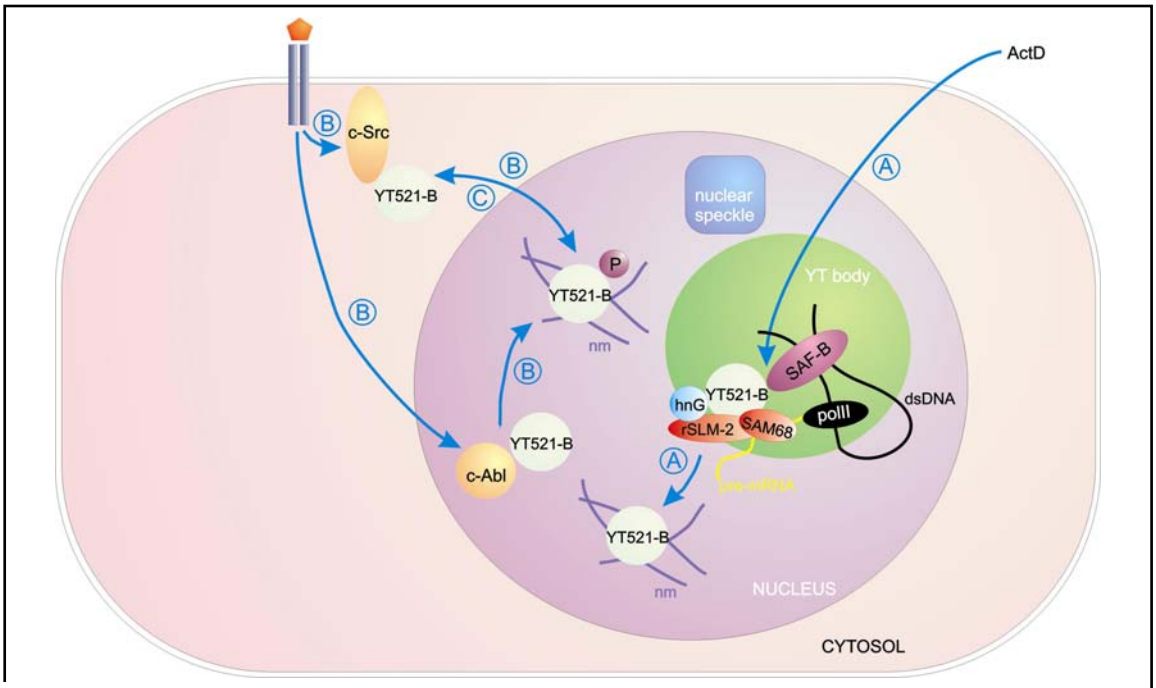


Figure 44. YT521-B interacts with splicing associated proteins and can be regulated by extracellular stimuli. YT521-B is a part of a protein complex assembled around DNA-binding SAF-B (Nayler et al., 1998c) and it is predicted to bind RNA. This complex is thought to help recognizing exons by binding to exon enhancer sequences on the pre-mRNA (*yellow line*) generated by RNA pol II. YT521-B defines a novel compartment in the nucleus, the YT bodies (*green circle*) (Nayler et al., 2000). After actinomycin D treatment (**A**) YT bodies disperse throughout the nucleoplasm and YT521-B associates with insoluble nuclear structures (*nm*). YT521-B is a target for tyrosine kinases (**B**). Phosphorylation causes dispersion of YT bodies and moves the protein into the insoluble nuclear fraction. Due to its shuttling ability (**C**), YT521-B can be also a target for membrane-bound SRC family kinases.

5.2. YTH is a possible novel nucleic acid binding domain

Regulation of gene expression at the post-transcriptional level is mainly achieved by proteins containing domains involved in RNA binding. Most of the proteins characterized as splicing factors include in their structures at least one motif that allows them RNA recognition. YT521-B is known to be involved in pre-mRNA splicing, but in contrast to the other known splicing factors, it lacks a distinguishable domain conferring RNA binding. The YTH domain, resembling RNA recognition motif with its mixed αβ-fold and aromatic residues conservation seems to fill this gap (Stoilov et al., 2002b). Experimental results presented in this work confirm earlier computational prediction (section 4.5.). For the first time, the direct interaction between YT521-B and purine-rich RNA is shown here. Moreover, mutagenesis analysis reveals the importance of the conserved residues in the YTH domain. The nuclear staining of seven mutants (Figure 25) is reminiscent of that observed after actinomycin D treatment (Nayler et al., 2000) or after phosphorylation (Rafalska et al., 2004). It is possible that some of the conserved residues are important for YTH domain function. Their change may lead to association of the protein with insoluble nuclear matrix similarly as it happens after tyrosine phosphorylation of YT521-B (Figure 45). Results of *in vivo* splicing assay on CD44 v5 minigene (Figure 27) confirm this hypothesis. Mutant K364L, present in YT bodies, influences CD44 v5 pre-mRNA splicing in the same way as the wild-type YT521-B. On the other hand, mutant F412D, showing the dispersed staining of the nucleus, is not capable to act on pre-mRNA at all. This could happen due to alterations in the structure of YTH domain or substitution of the residue crucial for interaction with RNA. It is known for RRM proteins, like U1A, Pab1p, and Sex-lethal, that a single amino acid substitution can significantly reduce their RNA binding ability (Nagai et al., 1990; Deardorff and Sachs, 1997; Lee et al., 1997). Therefore, further experiments, such as SELEX (Systematic Evolution of Ligands by Exponential enrichment) and EMSA, are needed to identify RNA sequence targets of YT521-B and to explain the dynamics of RNA:protein interaction.

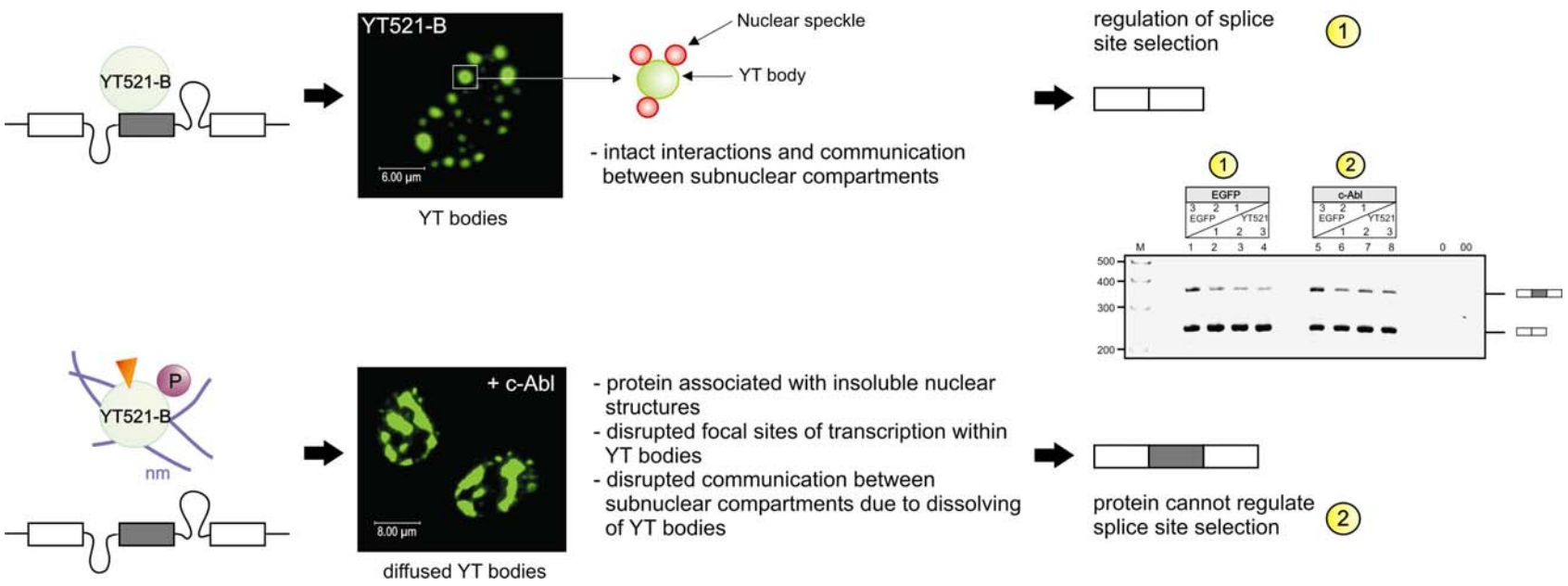


Figure 45. Change of YT521-B intranuclear distribution results in alteration of YT521-B influence on splice site selection. YT521-B is present in YT bodies, which contain sites of active transcription and are in close proximity to nuclear speckles (Nayler et al., 2000). YT521-B can regulate splice site selection in concentration-dependent manner (Hartmann et al., 1999). Disruption of YTH domain (*orange triangle*) or phosphorylation of the protein (*purple circle*) causes disruption of focal sites of transcription and protein association with insoluble nuclear structures (*violet lines*). The protein is no longer present in YT bodies and shows dispersed staining of the nucleus. Most probably due to its altered binding properties, protein cannot regulate splice site selection. Picture of the gel in this figure is part of Figure 43.

5.3. Tyrosine-phosphorylation regulates the nuclear localization of YT521-B

An increasing number of specialized domains have been described for the nucleus. In several cases, these domains have been shown to be highly dynamic structures that respond to stimulation, changes in cell cycle, malignant transformation or DNA damage (reviewed in Spector, 2001; Lamond and Spector, 2003; Matera, 1999; Lamond and Earnshaw, 1998; Spector, 2003). The experiments described here show that the intranuclear distribution of YT521-B is dependent on its tyrosine-phosphorylation status. Non-phosphorylated YT521-B is predominantly located in YT bodies, nuclear structures that contain focal sites of transcription and partially overlap with SC35 nuclear speckles (Nayler et al., 2000). YT bodies are soluble in non-denaturing buffers. Tyrosine phosphorylation of YT521-B changes that distribution (Figure 34) and solubility (Figure 35). It causes diffuse nuclear appearance of YT521-B and its tight association with an insoluble nuclear fraction (Figure 46). The release of YT521-B from YT bodies after tyrosine phosphorylation is reminiscent of the dispersion of nuclear speckles after serine/threonine phosphorylation of their proteins through cdc2 like kinases or SR-protein kinases (Colwill et al., 1996a; Nayler et al., 1998a; Stojdl and Bell, 1999; Colwill et al., 1996b). Functionally, serine/threonine phosphorylation of splicing factors can change splice site selection (Stoilov et al., 2004; Hartmann et al., 2001) by changing phosphorylation-dependent protein:protein interactions (Shin et al., 2004). Since c-Abl causes YT521-B phosphorylation (Figure 28), colocalizes with it (Figure 32) and binds tightly to it in immunoprecipitations (Figure 30), YT521-B may be a novel nuclear target for c-Abl. It is possible that the first association between c-Abl and YT521-B is mediated by the proline-rich region of YT521-B and the SH3 domain of c-Abl. Tyrosine phosphorylation would then stabilize the binding by phosphotyrosine-SH2 domain interaction. Several other non-receptor tyrosine kinases do not phosphorylate YT521-B, indicating a specific interaction between c-Abl and YT521-B. It is well established that RNA polII CTD is phosphorylated by c-Abl (Baskaran et al., 1993). Since YT bodies contain transcriptional start sites, it is possible that c-Abl phosphorylates YT521-B and RNA polII in the same nuclear compartment. The activity of nuclear c-Abl is controlled

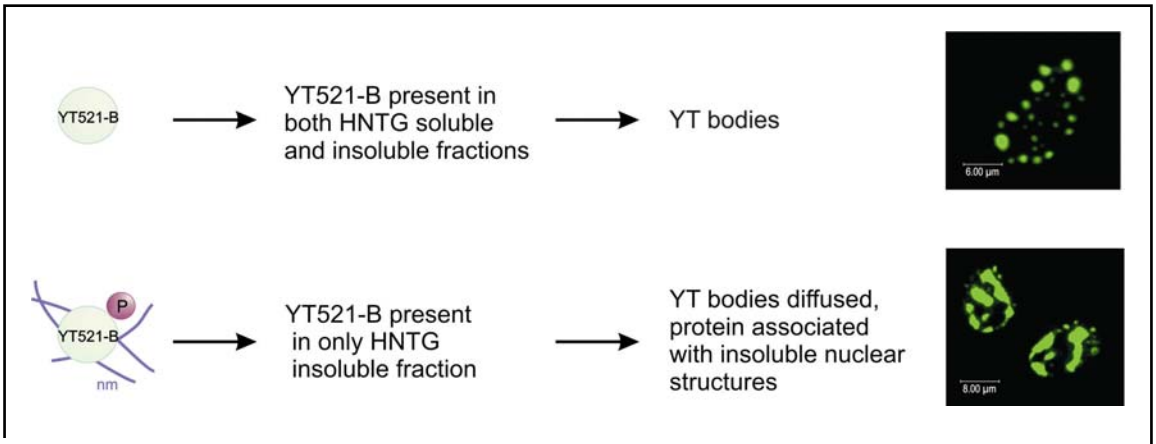


Figure 46. Phosphorylation of YT521-B results in protein association with insoluble nuclear structures.

during the cell cycle via interaction with pRB (Welch and Wang, 1993). The cyclin-dependent hyperphosphorylation of pRB releases binding to c-Abl and activates the kinase at the S-phase. This correlates with the finding that YT bodies cannot be detected during S-phase and form at the entrance of G1 (Van Etten, 1999; Welch and Wang, 1993). It also explains why YT bodies are markers for differentiated cells and are less pronounced in rapidly dividing transformed cells (Nayler et al., 2000). In contrast to the CTD (Baskaran et al., 1993), YT521-B is phoshorylated by src family kinases. YT521-B shuttles between nucleus and the cytosol (Figure 33), which shows that the protein can contact the membrane bound src-family kinases c-Src and $p59^{fyn}$ that phosphorylate YT521-B and bind to it tightly. It is therefore likely that these kinases can directly phosphorylate YT521-B and influence its properties. The Scansite Motif Scan program predicts 22 possible tyrosine phosphorylation sites in YT521-B. It remains to be determined which sites are responsible for the localization effects. The exact mechanism of actinomycin D influence on YT body localization remains to be determined as well. A general increase in tyrosine phosphorylation after actinomycin D treatment is not observed, but it cannot rule out the phosphorylation of specific proteins. Since YT521-B is insoluble after this treatment, it is not possible to determine whether it is phosphorylated. It is possible that actinomycin D causes a disruption of the nuclear matrix due to changes in DNA structure and RNA polymerase activity.

5.4. Similarities between effects of tyrosine phosphorylation in the cell membrane and the nucleus

After tyrosine phosphorylation, YT521-B changes its binding properties and associates with an insoluble nuclear structure. There are several parallels between the behavior of YT521-B and membrane bound receptor tyrosine kinases that form large signaling complexes after phosphorylation using scaffolding adaptors. First, in both cases tyrosine phosphorylation induces binding to other proteins. After phosphorylation YT521-B binds tightly to c-Abl, most likely to its SH2 domain. Secondly, the complex formation induced by phosphorylation changes the function of the molecules. Membrane bound receptor tyrosine kinases gain signaling function in the activated complexes, whereas YT521-B seems to lose its function in splice site regulation. It is not clear however, whether there are nuclear scaffolding adaptor proteins in non-transformed cells. One possibility is that DNA or RNA present in the nuclear matrix serves as such an adaptor. All interacting proteins of YT521-B can bind to RNA, or to both RNA and matrix attachment region DNA (Hartmann et al., 1999). Biochemically, benzonase is needed to effectively solubilize phosphorylated YT521-B, which suggests an involvement of nucleic acids. Finally, there is no YT521-B ortholog in yeast or bacteria, suggesting that similar to receptor tyrosine kinases (Darnell, 1997), YT521-B is characteristic for metazoans.

5.5. Splice site selection is regulated by tyrosine phosphorylation

A functional consequence of YT521-B phosphorylation is to modulate the influence of YT521-B on splice site selection. YT521-B binds to several proteins implicated in splice site selection, which it can change in a concentration-dependent manner (Hartmann et al., 1999). In the IL-4 receptor system, YT521-B causes intron retention (Figure 40), which is a frequent form of alternative splicing occurring in an estimated 6% of all alternative splicing events (Stamm et al., 2000). The retention of the intron between exon 9 and 10 can be detected in lymphocytes (Figure 39), indicating that it is a physiological event. However, the form is clearly not very abundant in cells, especially when compared with the results from minigene transfection. This low abundance can be explained by in frame stop codons present in the intron that destine the

mRNA to nonsense-mediated decay. RNA derived from minigenes most likely escapes nonsense-mediated decay because it is not translated, which explains the difference between minigene transfections and endogenous RNA. Phosphorylation changes the influence of YT521-B on splice site selection about 2-3 fold (Figure 40), which is comparable with the effects of phosphorylation on splice site selection observed *in vivo* (Meshorer et al., 2002; Stamm, 2002; Abdennebi et al., 2002; Holdiman et al., 2002). A similar influence of phosphorylation on the ability of YT521-B to regulate the alternative cassette exons in CD44 (Figure 43) and SRp20 (Rafalska et al., 2004) is observed. This shows that this mechanism regulates not only intron retention, but also other modes of alternative splicing.

Non-phosphorylated YT521-B colocalizes with transcriptional start sites and shows overlapping localization with nuclear speckles (Nayler et al., 2000). There, its most likely function is to bind nascent mRNA via its YTH domain (Stoilov et al., 2002b) or to form specific mRNA-protein complexes that bridge splice sites (Maniatis and Reed, 2002; Wu and Maniatis, 1993). As with other splicing factors, interactions between YT521-B and other proteins functioning in pre-RNA processing are intrinsically weak (Hartmann et al., 1999). This allows for dynamic association and reassociation to the nascent pre-mRNA-protein complexes. Such dynamic behavior is only possible when the protein is soluble. As a consequence, non-phosphorylated YT521-B can regulate splice sites in a concentration-dependent manner. Phosphorylation results in a strong association between YT521-B and nuclear structures (Figure 35). Furthermore, phosphorylated YT521-B is dispersed throughout the nucleus and removed from the vicinity of actively transcribed genes (Figure 45). Due to its insolubility and spatial distance, phosphorylated YT521-B is effectively removed from pre-mRNA processing events in the cell and can no longer influence splice site selection. A similar phosphorylation-dependant redistribution of nuclear splicing factors has been demonstrated for SF2/ASF (Caceres et al., 1998), where serine/threonine phosphorylation induced by the Clk/Sty kinases causes accumulation of SF2/ASF in the cytosol. Interestingly, the phosphorylation affects mainly the localization of proteins, but not their ability to change splice sites *in vitro* (Cazalla et al., 2002). Serine/threonine phosphorylation under stress conditions (van der Houven van Oordt et al., 2000) or hypoxia (Daoud et al., 2002) causes

hyperphosphorylation and accumulation of SR-proteins or hnRNPs in the cytosol. This could serve the similar purpose of removing factors from pre-mRNA processing events. In fact, *in vivo* splice site selection is altered after application of osmotic stress or hypoxia (van der Houven van Oordt et al., 2000; Daoud et al., 2002). Using cotransfection assays, it was previously demonstrated that sequestration of splicing factors can regulate splice site selection (Stoilov et al., 2004). Although this mechanism can explain differences in splice site selection between tissues with different concentration of regulatory proteins, it cannot account for rapid changes in splice site selection after cellular stimulation (Stamm, 2002). It is possible that phosphorylation-dependent sequestration of YT521-B in an insoluble state is a mechanism to regulate splice site selection in response to a kinase signal. After phosphorylation, YT521-B is localized in nuclear regions distant from the areas where transcription and pre-mRNA processing occurs. The phosphorylated protein is insoluble and most likely not able to participate in the dynamic rearrangement of protein complexes necessary for splice site recognition. This mechanism allows the cell to temporarily lower the active concentration of YT521-B without destroying the protein. The mechanism partially explains the frequently observed changes of splice site selection in malignant transformation (Xu and Lee, 2003) and during development (Black, 2003), as these processes are triggered or are concomitant with an increase in receptor tyrosine phosphorylation. It implies that in addition to the combination of splicing factors (Smith and Valcarcel, 2000) the tyrosine phosphorylation status of splicing factors can contribute to cell-specific alternative splicing.

5.6. Working model of YT521-B function

To summarize data presented in this work, the working model of YT521-B function consistent with the current state of the art is being proposed (Figures 44–46). As shown in Figure 44, YT521-B is a part of a protein complex assembled around DNA-binding SAF-B (Nayler et al., 1998c; Hartmann et al., 1999). In the nucleus, YT521-B is localized in YT bodies, which mostly overlap with focal sites of transcription (Nayler et al., 2000). Speckles, which are often found in close contact with YT bodies, serve as accessory domain supplying the transcriptionally active site with the necessary helper factors. The YT bodies can be dynamically regulated by signal transduction pathways

emanting from non-receptor tyrosine kinases. A non-phosphorylated YT521-B can form transient complexes with the processed RNA, most probably through its YTH domain, and with other splicing factors. These interactions help recognizing exons and regulate splice site selection (Figure 45). The phosphorylation changes the interaction between YT521-B and proteins involved in splice site selection (Hartmann et al., 1999). Phosphorylated YT521-B associates with nuclear components, most probably with the nuclear matrix (Figure 44) and is present in the insoluble nuclear fraction (Figure 46). Due to this sequestration its influence on splice site selection is abolished (Figure 45).

REFERENCES

Abdennebi, L., Lesport, A.S., Remy, J.J., Grebert, D., Pisselet, C., Monniaux, D. and Salesse, R. (2002) Differences in splicing of mRNA encoding LH receptor in theca cells according to breeding season in ewes. *Reproduction*, **123**, 819-826.

Altschul, S.F., Gish, W., Miller, W., Meyers, E.W. and Lipman, D.J. (1990) Basic Local Alignment Search Tool. *Journal of Molecular Biology*, **215**, 403-410.

Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W. and Lipman, D.J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucl. Acids Res.*, **25**, 3389-3402.

Ando, S., Sarlis, N.J., Krishnan, J., Feng, X., Refetoff, S., Zhang, M.Q., Oldfield, E.H. and Yen, P.M. (2001) Aberrant Alternative Splicing of Thyroid Hormone Receptor in a TSH-Secreting Pituitary Tumor Is A Mechanism for Hormone Resistance. *Mol. Endocrinol.*, **15**, 1529-1538.

Aravind, L. and Koonin, E.V. (1999a) G-patch: a new conserved domain in eukaryotic RNA-processing proteins and type D retroviral polyproteins. *Trends Biochem. Sci.*, **24**, 342-344.

Aravind, L. and Koonin, E.V. (1999b) Novel predicted RNA-binding domains associated with the translation machinery. *J. Mol. Evol.*, **48**, 291-302.

Aravind, L. and Landsman, D. (1998) AT-hook motifs identified in a wide variety of DNA-binding proteins. *Nucl. Acids Res.*, **26**, 4413-4421.

Assier, E., Bouzinba-Segard, H., Stolzenberg, M.-C., Stephens, R., Bardos, J., Freemont, P., Charron, D., Trowsdale, J. and Rich, T. (1999) Isolation, sequencing and expression of RED, a novel human gene encoding an acidic-basic dipeptide repeat. *Gene*, **230**, 145-154.

Banerjee, H., Rahn, A., Gawande, B., Guth, S., Valcarcel, J. and Singh, R. (2004) The conserved RNA recognition motif 3 of U2 snRNA auxiliary factor (U2AF65) is essential in vivo but dispensable for activity in vitro. *RNA*, **10**, 240-253.

Baskaran, R., Dahmus, M.E. and Wang, J.Y.J. (1993) Tyrosine Phosphorylation of Mammalian RNA Polymerase II Carboxyl-Terminal Domain. *Proc. Natl Acad. Sci. U. S. A.*, **90**, 11167-11171.

Ben-Neriah, Y., Daley, G.Q., Mes-Masson, A.M., Witte, O.N. and Baltimore, D. (1986) The chronic myelogenous leukemia-specific P210 protein is the product of the bcr/abl hybrid gene. *Science*, **233**, 212-214.

Birnboim, H.C. and Doly, J. (1979) A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic Acids Res.*, **7**, 1513-1523.

Birney, E., Kumar, S. and Krainer, A.R. (1993) Analysis of the RNA-recognition motif and RS and RGG domains: conservation in metazoan pre-mRNA splicing factors. *Nucleic Acids Res.*, **21**, 5803-5816.

Black, D.L. (2003) Mechanisms of alternative pre-messenger RNA splicing. *Annu. Rev. Biochem.*, **72**, 291-336.

Black, D.L., Chan, R., Min, H., Wang, J. and Bell, L. (1998) The electrophoretic mobility shift assay for RNA binding proteins. In Smith, C.W.J. (ed.), *RNA:Protein Interactions A Practical Approach*. Oxford University Press, Vol. 192, pp. 109-136.

Blum, H., Wolf, M., Enssle, K., Rollinghoff, M. and Gessner, A. (1996) Two distinct stimulus-dependent pathways lead to production of soluble murine interleukin-4 receptor. *J. Immunol.*, **157**, 1846-1853.

Bogden, C.E., Fass, D., Bergman, N., Nichols, M.D. and Berger, J.M. (1999) The structural basis for terminator recognition by the Rho transcription termination factor. *Mol. Cell*, **3**, 487-493.

Borsi, L., Balza, E., Gaggero, B., Allemanni, G. and Zardi, L. (1995) The Alternative Splicing Pattern of the Tenascin-C Pre-mRNA Is Controlled by the Extracellular pH. *J. Biol. Chem.*, **270**, 6243-6245.

Boukis, L.A., Liu, N., Furuyama, S. and Bruzik, J.P. (2004) Ser/Arg-rich Protein-mediated Communication between U1 and U2 Small Nuclear Ribonucleoprotein Particles. *J. Biol. Chem.*, **279**, 29647-29653.

Brett, D., Hanke, J., Lehmann, G., Haase, S., Delbruck, S., Krueger, S., Reich, J. and Bork, P. (2000) EST comparison indicates 38% of human mRNAs contain possible alternative splice forms. *FEBS Letters*, **474**, 83-86.

Brett, D., Pospisil, H., Valcarcel, J., Reich, J. and Bork, P. (2002) Alternative splicing and genome complexity. *Nat. Genet.*, **30**, 29-30.

Buckanovich, R.J. and Darnell, R.B. (1997) The neuronal RNA binding protein Nova-1 recognizes specific RNA targets in vitro and in vivo. *Mol. Cell. Biol.*, **17**, 3194-3201.

Buckanovich, R.J., Posner, J.B. and Darnell, R.B. (1993) Nova, the paraneoplastic Ri antigen, is homologous to an RNA-binding protein and is specifically expressed in the developing motor system. *Neuron*, **11**, 657-672.

Bullock, W.O., Fernandez, J.M. and Short, J.M. (1987) XL1-Blue: A high efficiency plasmid transforming recA Escherichia coli strain with beta-galactosidase selection. *Biotechniques*, **5**, 376-378.

Bunnell, S.C., Henry, P.A., Kolluri, R., Kirchhausen, T., Rickles, R.J. and Berg, L.J. (1996) Identification of Itk/Tsk Src Homology 3Domain Ligands. *J. Biol. Chem.*, **271**, 25646-25656.

Burge, C.B., Tuschl, T. and Sharp, P.A. (1999) Splicing of precursors to mRNAs by the spliceosomes. In Gesteland, R.F., Cech, T.R. and Atkins, J.F. (eds.), *The RNA World Second Edition*. CSHL Press, pp. 525-560.

Burset, M., Seledtsov, I.A. and Solovyev, V.V. (2000) Analysis of canonical and non-canonical splice sites in mammalian genomes. *Nucl. Acids Res.*, **28**, 4364–4375.

Burset, M., Seledtsov, I.A. and Solovyev, V.V. (2001) SpliceDB: database of canonical and non-canonical mammalian splice sites. *Nucl. Acids Res.*, **29**, 255-259.

Bycroft, M., Grunert, S., Murzin, A.G., Proctor, M. and St Johnston, D. (1995) NMR solution structure of a dsRNA binding domain from Drosophila staufen protein reveals homology to the N-terminal domain of ribosomal protein S5. *EMBO J.*, **14**, 3563-3571.

Bycroft, M., Hubbard, T.J., Proctor, M., Freund, S.M. and Murzin, A.G. (1997) The solution structure of the S1 RNA binding domain: a member of an ancient nucleic acid-binding fold. *Cell*, **88**, 235-242.

Caceres, J.F., Misteli, T., Screaton, G.R., Spector, D.L. and Krainer, A.R. (1997) Role of the Modular Domains of SR Proteins in Subnuclear Localization and Alternative Splicing Specificity. *J. Cell Biol.*, **138**, 225-238.

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

Caceres, J.F., Stamm, S., Helfman, D.M. and Krainer, A.R. (1994) Regulation of alternative splicing in vivo by overexpression of antagonistic splicing factors. *Science*, **265**, 1706-1709.

Cao, W., Jamison, S.F. and Garcia-Blanco, M.A. (1997) Both phosphorylation and dephosphorylation of ASF/SF2 are required for pre-mRNA splicing in vitro. *RNA*, **3**, 1456-1467.

Caputi, M. and Zahler, A.M. (2001) Determination of the RNA Binding Specificity of the Heterogeneous Nuclear Ribonucleoprotein (hnRNP) H/H'/F/2H9 Family. *J. Biol. Chem.*, **276**, 43850-43859.

Carmichael, G.G. (1975) Isolation of bacterial and phage proteins by homopolymer RNA-cellulose chromatography. *J. Biol. Chem.*, **250**, 6160-6167.

Cartegni, L., Maconi, M., Morandi, E., Cobianchi, F., Riva, S. and Biamonti, G. (1996) hnRNP A1 Selectively Interacts Through its Gly-rich Domain with Different RNA-binding Proteins. *J. Mol. Biol.*, **259**, 337-348.

Cazalla, D., Zhu, J., Manche, L., Huber, E., Krainer, A.R. and Caceres, J.F. (2002) Nuclear Export and Retention Signals in the RS Domain of SR Proteins. *Mol. Cell. Biol.*, **22**, 6871-6882.

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

Chalfant, C.E., Watson, J.E., Bisnauth, L.D., Kang, J.B., Patel, N., Obeid, L.M., Eichler, D.C. and Cooper, D.R. (1998) Insulin Regulates Protein Kinase Cbeta II Expression through Enhanced Exon Inclusion in L6 Skeletal Muscle Cells. A novel mechanism of insulin- and insulin-like growth factor-I-induced 5' splice site selection. *J. Biol. Chem.*, **273**, 910-916.

Chan, W.Y., Soloviev, M.M., Ciruela, F. and McIlhinney, R.A.J. (2001) Molecular Determinants of Metabotropic Glutamate Receptor 1B Trafficking. *Mol. Cell. Neurosci.*, **17**, 577-588.

Chen, C. and Okayama, H. (1987) High-efficiency transformation of mammalian cells by plasmid DNA. *Mol. Cell Biol.*, **7**, 2745-2752.

Chen, J.D. and Evans, R.M. (1995) A transcriptional co-repressor that interacts with nuclear hormone receptors. *Nature*, **377**, 454-457.

Chen, T., Boisvert, F.-M., Bazett-Jones, D.P. and Richard, S. (1999) A Role for the GSG Domain in Localizing Sam68 to Novel Nuclear Structures in Cancer Cell Lines. *Mol. Biol. Cell*, **10**, 3015-3033.

Chen, T., Cote, J., Carvajal, H.V. and Richard, S. (2001) Identification of Sam68 Arginine Glycine-rich Sequences Capable of Conferring Nonspecific RNA Binding to the GSG Domain. *J. Biol. Chem.*, **276**, 30803-30811.

Chiquet-Ehrismann, R., Matsuoka, Y., Hofer, U., Spring, J., Bernasconi, C. and Chiquet, M. (1991) Tenascin variants: differential binding to fibronectin and distinct distribution in cell cultures and tissues. *Cell Regul.*, **2**, 927–938.

Clark, F. and Thanaraj, T.A. (2002) Categorization and characterization of transcript-confirmed constitutively and alternatively spliced introns and exons from human. *Hum. Mol. Genet.*, **11**, 451-464.

Colwill, K., Feng, L.L., Yeakley, J.M., Gish, G.D., Caceres, J.F., Pawson, T. and Fu, X.-D. (1996b) SRPK1 and Clk/Sty Protein Kinases Show Distinct Substrate Specificities for Serine/Arginine-rich Splicing Factors. *J. Biol. Chem.*, **271**, 24569-24575.

Colwill, K., Pawson, T., Andrews, B., Prasad, J., Manley, J.L., Bell, J.C. and Duncan, P. (1996a) The Clk/Sty protein kinase phosphorylates SR splicing factors and regulates their intranuclear distribution. *EMBO J.*, **15**, 265-275.

Coulter, L.R., Landree, M.A. and Cooper, T.A. (1997) Identification of a new class of exonic splicing enhancers by in vivo selection. *Mol. Cell. Biol.*, **17**, 2143–2150.

Cremer, T. and Cremer, C. (2001) Chromosome territories, nuclear architecture and gene regulation in mammalian cells. *Nat. Rev. Genet.*, **2**, 292-301.

Cremer, T., Kreth, G., Koester, H., Fink, R.H., Heintzmann, R., Cremer, M., Solovei, I., Zink, D. and Cremer, C. (2000) Chromosome territories, interchromatin domain compartment, and nuclear matrix: an integrated view of the functional nuclear architecture. *Crit. Rev. Eukaryot. Gene Expr.*, **10**, 179-212.

Crispino, J.D., Blencowe, B.J. and Sharp, P.A. (1994) Complementation by SR proteins of pre-mRNA splicing reactions depleted of U1 snRNP. *Science*, **265**, 1866-1869.

Daley, G.Q., Van Etten, R.A. and Baltimore, D. (1990) Induction of chronic myelogenous leukemia in mice by the P210bcr/abl gene of the Philadelphia chromosome. *Science*, **247**, 824-830.

Daoud, R., Mies, G., Smialowska, A., Olah, L., Hossmann, K.-A. and Stamm, S. (2002) Ischemia Induces a Translocation of the Splicing Factor tra2-beta 1 and Changes Alternative Splicing Patterns in the Brain. *J. Neurosci.*, **22**, 5889-5899.

Darnell, J.E., Jr. (1997) Phosphotyrosine signaling and the single cell:metazoan boundary. *Proc. Natl Acad. Sci. U. S. A.*, **94**, 11767-11769.

Deardorff, J.A. and Sachs, A.B. (1997) Differential effects of aromatic and charged residue substitutions in the RNA binding domains of the yeast Poly(A)-binding protein. *Journal of Molecular Biology*, **269**, 67-81.

Debnath, J., Chamorro, M., Czar, M.J., Schaeffer, E.M., Lenardo, M.J., Varmus, H.E. and Schwartzberg, P.L. (1999) rlk/TXK Encodes Two Forms of a Novel Cysteine String Tyrosine Kinase Activated by Src Family Kinases. *Mol. Cell. Biol.*, **19**, 1498-1507.

Declerck, N., Vincent, F., Hoh, F., Aymerich, S. and van Tilbeurgh, H. (1999) RNA recognition by transcriptional antiterminators of the BglG/SacY family: functional and structural comparison of the CAT domain from SacY and LicT1. *J. Mol. Biol.*, **294**, 389-402.

Denegri, M., Chiodi, I., Corioni, M., Cobianchi, F., Riva, S. and Biamonti, G. (2001) Stress-induced Nuclear Bodies Are Sites of Accumulation of Pre-mRNA Processing Factors. *Mol. Biol. Cell*, **12**, 3502-3514.

Denhez, F. and Lafyatis, R. (1994) Conservation of regulated alternative splicing and identification of functional domains in vertebrate homologs to the Drosophila splicing regulator, suppressor-of-white-apricot. *J. Biol. Chem.*, **269**, 16170-16179.

Derry, J.J., Richard, S., Valderrama Carvajal, H., Ye, X., Vasioukhin, V., Cochrane, A.W., Chen, T. and Tyner, A.L. (2000) Sik (BRK) Phosphorylates Sam68 in the

Nucleus and Negatively Regulates Its RNA Binding Ability. *Mol. Cell. Biol.*, **20**, 6114-6126.

Di Fruscio, M., Chen, T. and Richard, S. (1999) Characterization of Sam68-like mammalian proteins SLM-1 and SLM-2: SLM-1 is a Src substrate during mitosis. *Proc. Natl Acad. Sci. U. S. A.*, **96**, 2710-2715.

Dominski, Z. and Kole, R. (1994) Identification of exon sequences involved in splice site selection. *J. Biol. Chem.*, **269**, 23590-23596.

Domon, C., Lorkovic, Z.J., Valcarcel, J. and Filipowicz, W. (1998) Multiple Forms of the U2 Small Nuclear Ribonucleoprotein Auxiliary Factor U2AF Subunits Expressed in Higher Plants. *J. Biol. Chem.*, **273**, 34603-34610.

D'Souza, I., Poorkaj, P., Hong, M., Nochlin, D., Lee, V.M.-Y., Bird, T.D. and Schellenberg, G.D. (1999) Missense and silent tau gene mutations cause frontotemporal dementia with parkinsonism-chromosome 17 type, by affecting multiple alternative RNA splicing regulatory elements. *Proc. Natl Acad. Sci. U. S. A.*, **96**, 5598-5603.

Dundr, M. and Misteli, T. (2001) Functional architecture in the cell nucleus. *Biochem. J.*, **356**, 297-310.

Eissa, N.T., Strauss, A.J., Haggerty, C.M., Choo, E.K., Chu, S.C. and Moss, J. (1996) Alternative Splicing of Human Inducible Nitric-Oxide Synthase mRNA. Tissue-specific regulation and induction by cytokines. *J. Biol. Chem.*, **271**, 27184-27187.

Elliott, D.J., Ma, K., Kerr, S.M., Thakrar, R., Speed, R., Chandley, A.C. and Cooke, H. (1996) An RBM homologue maps to the mouse Y chromosome and is expressed in germ cells. *Hum. Mol. Genet.*, **5**, 869-874.

Elrick, L.L., Humphrey, M.B., Cooper, T.A. and Berget, S.M. (1998) A Short Sequence within Two Purine-Rich Enhancers Determines 5' Splice Site Specificity. *Mol. Cell. Biol.*, **18**, 343-352.

English, J.M., Vanderbilt, C.A., Xu, S., Marcus, S. and Cobb, M.H. (1995) Isolation of MEK5 and differential expression of alternatively spliced forms. *Journal of Biological Chemistry*, **270**, 28897-28902.

Eperon, I.C., Ireland, D.C., Smith, R.A., Mayeda, A. and Krainer, A.R. (1993) Pathways for selection of 5' splice sites by U1 snRNPs and SF2/ASF. *EMBO J.*, **12**, 3607-3617.

Eperon, I.C., Makarova, O.V., Mayeda, A., Munroe, S.H., Caceres, J.F., Hayward, D.G. and Krainer, A.R. (2000) Selection of Alternative 5' Splice Sites: Role of U1 snRNP and Models for the Antagonistic Effects of SF2/ASF and hnRNP A1. *Mol. Cell. Biol.*, **20**, 8303-8318.

Fahrenkrog, B., Koser, J. and Aebi, U. (2004) The nuclear pore complex: a jack of all trades? *Trends in Biochemical Sciences*, **29**, 175-182.

Faustino, N.A. and Cooper, T.A. (2003) Pre-mRNA splicing and human disease. *Genes Dev.*, **17**, 419-437.

Filipowicz, W. and Pogacic, V. (2002) Biogenesis of small nucleolar ribonucleoproteins. *Curr. Opin. Cell Biol.*, **14**, 319-327.

Fischer, U., Liu, Q. and Dreyfuss, G. (1997) The SMN-SIP1 complex has an essential role in spliceosomal snRNP biogenesis. *Cell*, **90**, 1023-1029.

Gall, J.G. (2000) Cajal bodies: the first 100 years. *Annu. Rev. Cell Dev. Biol.*, **16**, 273-300.

Gao, Q.-S., Memmott, J., Lafyatis, R., Stamm, S., Screaton, G. and Andreadis, A. (2000) Complex Regulation of Tau Exon 10, Whose Missplicing Causes Frontotemporal Dementia. *J. Neurochem.*, **74**, 490-500.

Garrison, S., Hojgaard, A., Patillo, D., Weis, J.J. and Weis, J.H. (2001) Functional Characterization of Pactolus, a beta -Integrin-like Protein Preferentially Expressed by Neutrophils. *J. Biol. Chem.*, **276**, 35500-35511.

Ghetti, A., Piñol-Roma, S., Michael, W.M., Morandi, C. and Dreyfuss, G. (1992) hnRNP I, the polypyrimidine tract-binding protein: distinct nuclear localization and association with hnRNAs. *Nucl. Acids Res.*, **20**, 3671–3678.

Gorlach, M., Wittekind, M., Beckman, R.A., Mueller, L. and Dreyfuss, G. (1992) Interaction of the RNA-binding domain of the hnRNP C proteins with RNA. *EMBO J.*, **11**, 3289-3295.

Grabowski, P.J. (1998) Splicing regulation in neurons: tinkering with cell-specific control. *Cell*, **92**, 709-712.

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

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

Guinamard, R., Fougereau, M. and Seckinger, P. (1997) The SH3 domain of Bruton's tyrosine kinase interacts with Vav, Sam68 and EWS. *Scand. J. Immunol.*, **45**, 587-595.

Haegebarth, A., Heap, D., Bie, W., Derry, J.J., Richard, S. and Tyner, A.L. (2004) The Nuclear Tyrosine Kinase BRK/Sik Phosphorylates and Inhibits the RNA-binding Activities of the Sam68-like Mammalian Proteins SLM-1 and SLM-2. *J. Biol. Chem.*, **279**, 54398-54404.

Hall, S.L. and Padgett, R.A. (1994) Conserved Sequences in a Class of Rare Eukaryotic Nuclear Introns with Non-consensus Splice Sites. *J. Mol. Biol.*, **239**, 357-365.

Hao, Q.L., Ferris, D.K., White, G., Heisterkamp, N. and Groffen, J. (1991) Nuclear and cytoplasmic location of the FER tyrosine kinase. *Mol. Cell Biol.*, **11**, 1180-1183.

Harris, H. (1999) *The Birth of the Cell*. Yale University Press.

Hartmann, A.M., Nayler, O., Schwaiger, F.W., Obermeier, A. and Stamm, S. (1999) The Interaction and Colocalization of Sam68 with the Splicing-associated Factor YT521-B in Nuclear Dots Is Regulated by the Src Family Kinase p59fyn. *Mol. Biol. Cell*, **10**, 3909-3926.

Hartmann, A.M., Rujescu, D., Giannakouros, T., Nikolakaki, E., Goedert, M., Mandelkow, E.-M., Gao, Q.S., Andreadis, A. and Stamm, S. (2001) Regulation of Alternative Splicing of Human Tau Exon 10 by Phosphorylation of Splicing Factors. *Molecular and Cellular Neuroscience*, **18**, 80-90.

Hasegawa, M., Smith, M.J., Iijima, M., Tabira, T. and Goedert, M. (1999) FTDP-17 mutations N279K and S305N in tau produce increased splicing of exon 10. *FEBS Lett.*, **443**, 93-96.

Hastings, M.L. and Krainer, A.R. (2001) Pre-mRNA splicing in the new millennium. *Curr. Opin. Cell Biol.*, **13**, 302-309.

Henras, A., Henry, Y., Bousquet-Antonelli, C., Noaillac-Depeyre, J., Gelugne, J.P. and Caizergues-Ferrer, M. (1998) Nhp2p and Nop10p are essential for the function of H/ACA snoRNPs. *EMBO J.*, **17**, 7078-7090.

Hermann, H., Fabrizio, P., Raker, V.A., Foulaki, K., Hornig, H., Brahms, H. and Luhrmann, R. (1995) snRNP Sm proteins share two evolutionarily conserved sequence motifs which are involved in Sm protein-protein interactions. *EMBO J.*, **14**, 2076-2088.

Hertel, K.J., Lynch, K.W. and Maniatis, T. (1997) Common themes in the function of transcription and splicing enhancers. *Curr. Opin. Cell Biol.*, **9**, 350-357.

Hoffman, B.E. and Grabowski, P.J. (1992) U1 snRNP targets an essential splicing factor, U2AF65, to the 3' splice site by a network of interactions spanning the exon. *Genes Dev.*, **6**, 2554-2568.

Hoffman, D.W., Query, C.C., Golden, B.L., White, S.W. and Keene, J.D. (1991) RNA-Binding Domain of the A Protein Component of the U1 Small Nuclear

Ribonucleoprotein Analyzed by NMR Spectroscopy is Structurally Similar to Ribosomal Proteins. *Proc. Natl Acad. Sci. U.S.A.*, **88**, 2495-2499.

Holdiman, A.J., Fergus, D.J. and England, S.K. (2002) 17[beta]-Estradiol upregulates distinct maxi-K channel transcripts in mouse uterus. *Molecular and Cellular Endocrinology*, **192**, 1-6.

Huang, S. (2000) Review: Perinucleolar Structures. *J. Struct. Biol.*, **129**, 233-240.

Hutton, M., Lendon, C.L., Rizzu, P., Baker, M., Froelich, S., Houlden, H., Pickering-Brown, S., Chakraverty, S., Isaacs, A., Grover, A., Hackett, J., Adamson, J., Lincoln, S., Dickson, D., Davies, P., Petersen, R.C., Stevens, M., de Graaff, E., Wauters, E., van Baren, J., Hillebrand, M., Joosse, M., Kwon, J.M., Nowotny, P., Che, L.K., Norton, J., Morris, J.C., Reed, L.A., Trojanowski, J., Basun, H., Lannfelt, L., Neystat, M., Fahn, S., Dark, F., Tannenberg, T., Dodd, P.R., Hayward, N., Kwok, J.B.J., Schofield, P.R., Andreadis, A., Snowden, J., Craufurd, D., Neary, D., Owen, F., Oostra, B.A., Hardy, J., Goate, A., van Swieten, J., Mann, D., Lynch, T. and Heutink, P. (1998) Association of missense and 5'-splice-site mutations in tau with the inherited dementia FTDP-17. *Nature*, **393**, 702-705.

Imai, Y., Matsuo, N., Ogawa, S., Tohyama, M. and Takagi, T. (1998) Cloning of a gene, YT521, for a novel RNA splicing-related protein induced by hypoxia/reoxygenation. *Molecular Brain Research*, **53**, 33-40.

International Human Genome Sequencing Consortium. (2004) Finishing the euchromatic sequence of the human genome. *Nature*, **431**, 931-945.

Jensen, L.E. and Whitehead, A.S. (2001) IRAK1b, a Novel Alternative Splice Variant of Interleukin-1 Receptor-associated Kinase (IRAK), Mediates Interleukin-1 Signaling and Has Prolonged Stability. *J. Biol. Chem.*, **276**, 29037-29044.

Jessen, T.-H., Oubridge, C., Teo, C.H., Pritchard, C. and Nagai, K. (1991) Identification of molecular contacts between the U1 A small nuclear ribonucleoprotein and U1 RNA. *EMBO J.*, **10**, 3447-3456.

Jiang, Z., Cote, J., Kwon, J.M., Goate, A.M. and Wu, J.Y. (2000) Aberrant Splicing of tau Pre-mRNA Caused by Intronic Mutations Associated with the Inherited Dementia Frontotemporal Dementia with Parkinsonism Linked to Chromosome 17. *Mol. Cell. Biol.*, **20**, 4036-4048.

Johnson, J.M., Castle, J., Garrett-Engele, P., Kan, Z., Loerch, P.M., Armour, C.D., Santos, R., Schadt, E.E., Stoughton, R. and Shoemaker, D.D. (2003) Genome-Wide Survey of Human Alternative Pre-mRNA Splicing with Exon Junction Microarrays. *Science*, **302**, 2141-2144.

Johnson, K.R., Merrick, W.C., Zoll, W.L. and Zhu, Y. (1997) Identification of cDNA Clones for the Large Subunit of Eukaryotic Translation Initiation Factor 3. Comparison of homologues from human, Nicotiana tabacum, Caenorhabditis elegans, and Saccharomyces cerevisiae. *J. Biol. Chem.*, **272**, 7106-7113.

Jumaa, H. and Nielsen, P.J. (1997) The splicing factor SRp20 modifies splicing of its own mRNA and ASF/SF2 antagonizes this regulation. *EMBO J.*, **16**, 5077-5085.

Kan, Z., Rouchka, E.C., Gish, W.R. and States, D.J. (2001) Gene Structure Prediction and Alternative Splicing Analysis Using Genomically Aligned ESTs. *Genome Res.*, **11**, 889-900.

Kay, B.K., Williamson, M.P. and Sudol, M. (2000) The importance of being proline: the interaction of proline-rich motifs in signaling proteins with their cognate domains. *FASEB J.*, **14**, 231-241.

Kent, O.A., Ritchie, D.B. and MacMillan, A.M. (2005) Characterization of a U2AF-Independent Commitment Complex (E') in the Mammalian Spliceosome Assembly Pathway. *Mol. Cell. Biol.*, **25**, 233-240.

Kent, W.J. (2002) BLAT - The BLAST-Like Alignment Tool. *Genome Res.*, **12**, 656-664.

Kent, W.J., Sugnet, C.W., Furey, T.S., Roskin, K.M., Pringle, T.H., Zahler, A.M. and Haussler, D. (2002) The Human Genome Browser at UCSC. *Genome Res.*, **12**, 996-1006.

Kestler, D.P., Agarwal, S., Cobb, J., Goldstein, K.M. and Hall, R.E. (1995) Detection and analysis of an alternatively spliced isoform of interleukin-6 mRNA in peripheral blood mononuclear cells. *Blood*, **86**, 4559-4567.

Kiledjian, M. and Dreyfuss, G. (1992) Primary structure and binding activity of the hnRNP U protein: binding RNA through RGG box. *EMBO J.*, **11**, 2655-2664.

Kim, N., Lim, D., Lee, S. and Kim, H. (2005) ASePCR: Alternative splicing electronic RT-PCR in multiple tissues and organs. *Nucl. Acids Res.*, **33 (Web Server Issue)**, Accepted.

Kimura, M. (1983) *The neutral theory of molecular evolution*. Cambridge University Press, Cambridge.

Koch, T., Schulz, S., Pfeiffer, M., Klutzny, M., Schroder, H., Kahl, E. and Hollt, V. (2001) C-terminal Splice Variants of the Mouse {micro}-Opioid Receptor Differ in Morphine-induced Internalization and Receptor Resensitization. *J. Biol. Chem.*, **276**, 31408-31414.

Konig, H., Ponta, H. and Herrlich, P. (1998) Coupling of signal transduction to alternative pre-mRNA splicing by a composite splice regulator. *EMBO J.*, **17**, 2904-2913.

Kornblihtt, A.R., de la Mata, M., Fededa, J.P., Munoz, M.J. and Nogues, G. (2004) Multiple links between transcription and splicing. *RNA*, **10**, 1489-1498.

Kowalska, A., Hasegawa, M., Miyamoto, K., Akiguchi, I., Ikemoto, A., Takahashi, K., Araki, W. and Tabira, T. (2002) A novel mutation at position +11 in the intron following exon 10 of the tau gene in FTDP-17. *J. Appl. Genet.*, **43**, 535-543.

Kozmik, Z., Kurzbauer, R., Dorfler, P. and Busslinger, M. (1993) Alternative splicing of Pax-8 gene transcripts is developmentally regulated and generates isoforms with different transactivation properties. *Mol. Cell Biol.*, **13**, 6024-6035.

Krackhardt, A.M., Witzens, M., Harig, S., Hodi, F.S., Zauls, A.J., Chessia, M., Barrett, P. and Gribben, J.G. (2002) Identification of tumor-associated antigens in chronic lymphocytic leukemia by SEREX. *Blood*, **100**, 2123-2131.

Kramer, A. (1996) The Structure and Function of Proteins Involved in Mammalian Pre-mRNA Splicing. *Ann. Rev. Biochem.*, **65**, 367-409.

Kramer, A. and Utans, U. (1991) Three protein factors (SF1, SF3 and U2AF) function in pre-splicing complex formation in addition to snRNPs. *EMBO J.*, **10**, 1503-1509.

Krecic, A.M. and Swanson, M.S. (1999) hnRNP complexes: composition, structure, and function. *Curr. Opin. Cell Biol.*, **11**, 363-371.

Krummheuer, J., Lenz, C., Kammler, S., Scheid, A. and Schaal, H. (2001) Influence of the Small Leader Exons 2 and 3 on Human Immunodeficiency Virus Type 1 Gene Expression. *Virology*, **286**, 276-289.

Kunkel, T.A. (1985) Rapid and efficient site-specific mutagenesis without phenotypic selection. *Proc. Natl Acad. Sci. U.S.A.*, **82**, 488-492.

Kunkel, T.A., Roberts, J.D. and Zakour, R.A. (1987) Rapid and efficient site-specific mutagenesis without phenotypic selection. *Methods Enzymol.*, **154**, 367-382.

Lamond, A.I. and Earnshaw, W.C. (1998) Structure and Function in the Nucleus. *Science*, **280**, 547-553.

Lamond, A.I. and Spector, D.L. (2003) Nuclear speckles: a model for nuclear organelles. *Nat. Rev. Mol. Cell Biol.*, **4**, 605-612.

Lander, E.S., Linton, L.M., Birren, B., Nusbaum, C., Zody, M.C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W., Funke, R., Gage, D., Harris, K., Heaford, A., Howland, J., Kann, L., Lehoczky, J., LeVine, R., McEwan, P., McKernan, K., Meldrim, J., Mesirov, J., Miranda, C., Morris, W., Naylor, J., Raymond, C., Rosetti, M., Santos, R., Sheridan, A., Sougnez, C., Stange-Thomann, N., Stojanovic, N., Subramanian, A., Wyman, D., Rogers, J., Sulston, J., Ainscough, R., Beck, S., Bentley, D., Burton, J., Clee, C., Carter, N., Coulson, A., Deadman, R., Deloukas, P., Dunham, A., Dunham, I., Durbin, R., French, L., Grafham, D., Gregory, S., Hubbard, T., Humphray, S., Hunt, A., Jones, M., Lloyd, C., McMurray, A., Matthews, L., Mercer, S., Milne, S., Mullikin, J.C., Mungall, A., Plumb, R., Ross, M., Shownkeen, R., Sims, S., Waterston, R.H., Wilson, R.K., Hillier, L.W., McPherson, J.D., Marra, M.A., Mardis, E.R., Fulton, L.A., Chinwalla, A.T., Pepin, K.H., Gish, W.R., Chissoe, S.L., Wendl, M.C., Delehaunty, K.D., Miner, T.L., Delehaunty, A., Kramer, J.B., Cook, L.L., Fulton, R.S., Johnson, D.L., Minx, P.J., Clifton, S.W., Hawkins, T., Branscomb, E., Predki, P., Richardson, P., Wenning, S., Slezak, T., Doggett, N., Cheng, J.F., Olsen, A., Lucas, S., Elkin, C., Uberbacher, E., Frazier, M., Gibbs, R.A., Muzny, D.M., Scherer, S.E., Bouck, J.B., Sodergren, E.J., Worley, K.C., Rives, C.M., Gorrell, J.H., Metzker, M.L., Naylor, S.L., Kucherlapati, R.S., Nelson, D.L., Weinstock, G.M., Sakaki, Y., Fujiyama, A., Hattori, M., Yada, T., Toyoda, A., Itoh, T., Kawagoe, C., Watanabe, H., Totoki, Y., Taylor, T., Weissenbach, J., Heilig, R., Saurin, W., Artiguenave, F., Brottier, P., Bruls, T., Pelletier, E., Robert, C., Wincker, P., Smith, D., Doucette-Stamm, L., Rubenfield, M., Weinstock, K., Lee, H.M., Dubois, J., Rosenthal, A., Platzer, M., Nyakatura, G., Taudien, S., Rump, A., Yang, H., Yu, J., Wang, J., Huang, G., Gu, J., Hood, L., Rowen, L., Madan, A., Qin, S., Davis, R.W., Federspiel, N.A., Abola, A.P., Proctor, M.J., Myers, R.M., Schmutz, J., Dickson, M., Grimwood, J., Cox, D.R., Olson, M.V., Kaul, R., Raymond, C., Shimizu, N., Kawasaki, K., Minoshima, S., Evans, G.A., Athanasiou, M., Schultz, R., Roe, B.A., Chen, F., Pan, H., Ramser,

J., Lehrach, H., Reinhardt, R., McCombie, W.R., de la Bastide, M., Dedhia, N., Blocker, H., Hornischer, K., Nordsiek, G., Agarwala, R., Aravind, L., Bailey, J., Bateman, A., Batzoglou, S., Birney, E., Bork, P., Brown, D.G., Burge, C.B., Cerutti, L., Chen, H., Church, D., Clamp, M., Copley, R.R., Doerks, T., Eddy, S., Eichler, E.E., Furey, T.S., Galagan, J., Gilbert, J., Harmon, C., Hayashizaki, Y., Haussler, D., Hermjakob, H., Hokamp, K., Jang, W., Johnson, L., Jones, T., Kasif, S., Kasprzyk, A., Kennedy, S., Kent, W., Kitts, P., Koonin, E.V., Korf, I., Kulp, D., Lancet, D., Lowe, T.M., McLysaght, A., Mikkelsen, T., Moran, J., Mulder, N., Pollara, V., Ponting, C.P., Schuler, G., Schultz, J., Slater, G., Smit, A., Stupka, E., Szustakowski, J., Thierry-Mieg, D., Thierry-Mieg, J., Wagner, L., Wallis, J., Wheeler, R., Williams, A., Wolf, Y., Wolfe, K., Yang, S., Yeh, R., Collins, F., Guyer, M.S., Peterson, J., Felsenfeld, A., Wetterstrand, K.A., Patrinos, A. and Morgan, M.J. (2001) Initial sequencing and analysis of the human genome. *Nature*, **409**, 860-921.

Lee, A.L., Volkman, B.F., Robertson, S.A., Rudner, D.Z., Barbash, D.A., Cline, T.W., Kanaar, R., Rio, D.C. and Wemmer, D.E. (1997) Chemical Shift Mapping of the RNA-Binding Interface of the Multiple-RBD Protein Sex-Lethal. *Biochemistry*, **36**, 14306-14317.

Lee, S., Neumann, M., Stearman, R., Stauber, R., Pause, A., Pavlakis, G.N. and Klausner, R.D. (1999) Transcription-Dependent Nuclear-Cytoplasmic Trafficking Is Required for the Function of the von Hippel-Lindau Tumor Suppressor Protein. *Mol. Cell. Biol.*, **19**, 1486-1497.

Lee, Y., Sultana, R., Pertea, G., Cho, J., Karamycheva, S., Tsai, J., Parvizi, B., Cheung, F., Antonescu, V., White, J., Holt, I., Liang, F. and Quackenbush, J. (2002) Cross-Referencing Eukaryotic Genomes: TIGR Orthologous Gene Alignments (TOGA). *Genome Res.*, **12**, 493-502.

Letunic, I., Goodstadt, L., Dickens, N.J., Doerks, T., Schultz, J., Mott, R., Ciccarelli, F., Copley, R.R., Ponting, C.P. and Bork, P. (2002) Recent improvements to the

SMART domain-based sequence annotation resource. *Nucl. Acids Res.*, **30**, 242-244.

Levi-Strauss, M., Carroll, M.C., Steinmetz, M. and Meo, T. (1988) A previously undetected MHC gene with an unusual periodic structure. *Science*, **240**, 201-204.

Lewis, B.P., Green, R.E. and Brenner, S.E. (2003) Evidence for the widespread coupling of alternative splicing and nonsense-mediated mRNA decay in humans. *Proc. Natl Acad. Sci. U. S. A.*, **100**, 189-192.

Lewis, H.A., Chen, H., Edo, C., Buckanovich, R.J., Yang, Y.Y., Musunuru, K., Zhong, R., Darnell, R.B. and Burley, S.K. (1999) Crystal structures of Nova-1 and Nova-2 K-homology RNA-binding domains. *Structure*, **7**, 191-203.

Lewis, H.A., Musunuru, K., Jensen, K.B., Edo, C., Chen, H., Darnell, R.B. and Burley, S.K. (2000) Sequence-Specific RNA Binding by a Nova KH Domain: Implications for Paraneoplastic Disease and the Fragile X Syndrome. *Cell*, **100**, 323-332.

Li, Y., Chen, Z.Y., Wang, W., Baker, C.C. and Krug, R.M. (2001) The 3'-end-processing factor CPSF is required for the splicing of single-intron pre-mRNAs in vivo. *RNA*, **7**, 920-931.

Lopez, A.J. (1995) Developmental Role of Transcription Factor Isoforms Generated by Alternative Splicing. *Dev. Biol.*, **172**, 396-411.

Lyman, S.D., James, L., Escobar, S., Downey, H., de Vries, P., Brasel, K., Stocking, K., Beckmann, M.P., Copeland, N.G., Cleveland, L.S. and al., e. (1995) Identification of soluble and membrane-bound isoforms of the murine flt3 ligand generated by alternative splicing of mRNAs. *Oncogene*, **10**, 149-157.

Lynch, K.W. and Weiss, A. (2000) A Model System for Activation-Induced Alternative Splicing of CD45 Pre-mRNA in T Cells Implicates Protein Kinase C and Ras. *Mol. Cell. Biol.*, **20**, 70-80.

MacMillan, A.M., McCaw, P.S., Crispino, J.D. and Sharp, P.A. (1997) SC35-mediated reconstitution of splicing in U2AF-depleted nuclear extract. *PNAS*, **94**, 133-136.

Maniatis, T. and Reed, R. (2002) An extensive network of coupling among gene expression machines. *Nature*, **416**, 499-506.

Manley, J.L. and Tacke, R. (1996) SR proteins and splicing control. *Genes Dev.*, **10**, 1569-1579.

Maquat, L.E. (2004) Nonsense-mediated mRNA decay: splicing, translation and mRNP dynamics. *Nat. Rev. Mol. Cell Biol.*, **5**, 89-99.

Marfatia, K.A., Crafton, E.B., Green, D.M. and Corbett, A.H. (2003) Domain Analysis of the Saccharomyces cerevisiae Heterogeneous Nuclear Ribonucleoprotein, Nab2p. Dissecting the requirements for Nab2p-facilitated poly(A) RNA export. *J. Biol. Chem.*, **278**, 6731-6740.

Matera, A.G. (1999) Nuclear bodies: multifaceted subdomains of the interchromatin space. *Trends Cell Biol.*, **9**, 302-309.

Matera, A.G., Frey, M.R., Margelot, K. and Wolin, S.L. (1995) A perinucleolar compartment contains several RNA polymerase III transcripts as well as the polypyrimidine tract-binding protein, hnRNP I [published erratum appears in J Cell Biol 1995 Jul;130(2):497-500]. *J. Cell Biol.*, **129**, 1181-1193.

Matter, N., Herrlich, P. and Konig, H. (2002) Signal-dependent regulation of splicing via phosphorylation of Sam68. *Nature*, **420**, 691-695.

Maul, G.G., Negorev, D., Bell, P. and Ishov, A.M. (2000) Review: Properties and Assembly Mechanisms of ND10, PML Bodies, or PODs. *J. Struct. Biol.*, **129**, 278-287.

McKay, N.G., Hunter, D.J., Haites, N.E. and Power, D.A. (1994) Regulation of Alternative Splicing of the Fibronectin IIICS Domain by Cytokines. *Biochem. Biophys. Res. Commun.*, **199**, 1005-1011.

Mermoud, J.E., Cohen, P. and Lamond, A.I. (1992) Ser/Thr-specific protein phosphatases are required for both catalytic steps of pre-mRNA splicing. *Nucleic Acids Res.*, **20**, 5263-5269.

Meshorer, E., Erb, C., Gazit, R., Pavlovsky, L., Kaufer, D., Friedman, A., Glick, D., Ben-Arie, N. and Soreq, H. (2002) Alternative Splicing and Neuritic mRNA Translocation Under Long-Term Neuronal Hypersensitivity. *Science*, **295**, 508-512.

Miki, T., Bottaro, D.P., Fleming, T.P., Smith, C.L., Burgess, W.H., Chan, A.M.-L. and Aaronson, S.A. (1992) Determination of ligand-binding specificity by alternative splicing: two distinct growth factor receptors encoded by a single gene. *Proc. Natl Acad. Sci. U. S. A.*, **89**, 246-250.

Mironov, A.A., Fickett, J.W. and Gelfand, M.S. (1999) Frequent Alternative Splicing of Human Genes. *Genome Res.*, **9**, 1288-1293.

Missler, M. and Sudhof, T.C. (1998) Neurexins: three genes and 1001 products. *Trend. Genet.*, **14**, 20-26.

Misteli, T., Caceres, J.F., Clement, J.Q., Krainer, A.R., Wilkinson, M.F. and Spector, D.L. (1998) Serine Phosphorylation of SR Proteins Is Required for Their Recruitment to Sites of Transcription In Vivo. *J. Cell Biol.*, **143**, 297-307.

Misteli, T., Caceres, J.F. and Spector, D.L. (1997) The dynamics of a pre-mRNA splicing factor in living cells. *Nature*, **387**, 523-527.

Modrek, B. and Lee, C. (2002) A genomic view of alternative splicing. *Nat. Genet.*, **30**, 13-19.

Modrek, B., Resch, A., Grasso, C. and Lee, C. (2001) Genome-wide detection of alternative splicing in expressed sequences of human genes. *Nucl. Acids Res.*, **29**, 2850-2859.

Morisawa, G., Han-yama, A., Moda, I., Tamai, A., Iwabuchi, M. and Meshi, T. (2000) AHM1, a Novel Type of Nuclear Matrix-Localized, MAR Binding Protein with a Single AT Hook and a J Domain-Homologous Region. *Plant Cell*, **12**, 1903-1916.

Morrogh, L.M., Hinshelwood, S., Costello, P., Cory, G.O. and Kinnon, C. (1999) The SH3 domain of Bruton's tyrosine kinase displays altered ligand binding properties when auto-phosphorylated in vitro. *Eur. J. Immunol.*, **29**, 2269-2279.

Musco, G., Stier, G., Joseph, C., Castiglione Morelli, M.A., Nilges, M., Gibson, T.J. and Pastore, A. (1996) Three-dimensional structure and stability of the KH domain: molecular insights into the fragile X syndrome. *Cell*, **85**, 237-245.

Nagai, K., Oubridge, C., Jessen, T.H., Li, J. and Evans, P.R. (1990) Crystal structure of the RNA-binding domains of the U1 small nuclear ribonucleoprotein A. *Nature*, **348**, 515-520.

Nakai, K. and Kanehisa, M. (1992) A knowledge base for predicting protein localization sites in eukaryotic cells. *Genomics*, **14**, 897-911.

Nasim, M.T., Chernova, T.K., Chowdhury, H.M., Yue, B.-G. and Eperon, I.C. (2003) HnRNP G and Tra2beta: opposite effects on splicing matched by antagonism in RNA binding. *Hum. Mol. Genet.*, **12**, 1337-1348.

Nayler, O., Cap, C. and Stamm, S. (1998b) Human Transformer-2-beta Gene (SFRS10): Complete Nucleotide Sequence, Chromosomal Localization, and Generation of a Tissue-Specific Isoform. *Genomics*, **53**, 191-202.

Nayler, O., Hartmann, A.M. and Stamm, S. (2000) The ER Repeat Protein YT521-B Localizes to a Novel Subnuclear Compartment. *J. Cell Biol.*, **150**, 949-962.

Nayler, O., Schnorrer, F., Stamm, S. and Ullrich, A. (1998a) The Cellular Localization of the Murine Serine/Arginine-rich Protein Kinase CLK2 Is Regulated by Serine 141 Autophosphorylation. *J. Biol. Chem.*, **273**, 34341-34348.

Nayler, O., Stratling, W., Bourquin, J., Stagljar, I., Lindemann, L., Jasper, H., Hartmann, A., Fackelmayer, F., Ullrich, A. and Stamm, S. (1998c) SAF-B protein couples transcription and pre-mRNA splicing to SAR/MAR elements. *Nucl. Acids Res.*, **26**, 3542-3549.

Nestler, E.J., Kelz, M.B. and Chen, J. (1999) DeltaFosB: a molecular mediator of long-term neural and behavioral plasticity. *Brain Res.*, **835**, 10-17.

Neugebauer, K.M. (2002) On the importance of being co-transcriptional. *J. Cell Sci.*, **115**, 3865-3871.

Neumann, M., Afonina, E., Ceccherini-Silberstein, F., Schlicht, S., Erfle, V., Pavlakis, G.N. and Brack-Werner, R. (2001) Nucleocytoplasmic transport in human astrocytes: decreased nuclear uptake of the HIV Rev shuttle protein. *J Cell Sci*, **114**, 1717-1729.

Nichols, R.C., Wang, X.W., Tang, J., Hamilton, B.J., High, F.A., Herschman, H.R. and Rigby, W.F.C. (2000) The RGG Domain in hnRNP A2 Affects Subcellular Localization. *Exp.Cell Res.*, **256**, 522-532.

Nikolakaki, E., Kohen, R., Hartmann, A.M., Stamm, S., Georgatsou, E. and Giannakouros, T. (2001) Cloning and Characterization of an Alternatively Spliced Form of SR Protein Kinase 1 That Interacts Specifically with Scaffold Attachment Factor-B. *J. Biol. Chem.*, **276**, 40175-40182.

Nishizawa, Y., Usukura, J., Singh, D.P., Chylack, L.T. and Shinohara, T. (2001) Spatial and temporal dynamics of two alternatively spliced regulatory factors, lens epithelium-derived growth factor (ledgf/p75) and p52, in the nucleus. *Cell Tissue Res.*, **305**.

Nissim-Rafinia, M., Chiba-Falek, O., Sharon, G., Boss, A. and Kerem, B. (2000) Cellular and viral splicing factors can modify the splicing pattern of CFTR transcripts carrying splicing mutations. *Hum. Mol. Genet.*, **9**, 1771-1778.

Obenauer, J.C., Cantley, L.C. and Yaffe, M.B. (2003) Scansite 2.0: proteome-wide prediction of cell signaling interactions using short sequence motifs. *Nucl. Acids Res.*, **31**, 3635-3641.

Ogino, S. and Wilson, R.B. (2002) Genetic testing and risk assessment for spinal muscular atrophy (SMA). *Hum. Genet.*, **111**, 477 - 500.

Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S., Nikaido, I., Osato, N., Saito, R., Suzuki, H., Yamanaka, I., Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schonbach, C., Gojobori, T., Baldarelli, R., Hill, D.P., Bult, C., Hume, D.A., Quackenbush, J., Schriml, L.M., Kanapin, A., Matsuda, H., Batalov, S., Beisel, K.W., Blake, J.A., Bradt, D., Brusic, V., Chothia, C., Corbani, L.E., Cousins, S., Dalla, E., Dragani, T.A., Fletcher, C.F., Forrest, A., Frazer, K.S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S., Gustincich, S., Hirokawa, N., Jackson, I.J., Jarvis, E.D., Kanai, A., Kawaji, H., Kawasawa, Y., Kedzierski, R.M., King, B.L., Konagaya, A., Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R., Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T., Numata, K., Okido, T., Pavan, W.J., Pertea, G., Pesole, G., Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramachandran, S., Ravasi, T., Reed, J.C., Reed, D.J., Reid, J., Ring, B.Z., Ringwald, M., Sandelin, A., Schneider, C., Semple, C.A., Setou, M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale, R.D., M, T., Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y., Wells, C., Wilming, L.G., Wynshaw-Boris, A., Yanagisawa, M., Yang, I., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P., Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura, M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K., Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E.S., Rogers, J., Birney, E. and Hayashizaki, Y. (2002) Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs. *Nature*, **420**, 563-573.

Ono, Y., Ohno, M. and Shimura, Y. (1994) Identification of a putative RNA helicase (HRH1), a human homolog of yeast Prp22. *Mol. Cell Biol.*, **14**, 7611-7620.

Ormondroyd, E., de la Luna, S. and La Thangue, N.B. (1995) A new member of the DP family, DP-3, with distinct protein products suggests a regulatory role for alternative splicing in the cell cycle transcription factor DRTF1/E2F. *Oncogene*, **11**, 1437-1446.

Oubridge, C., Ito, N., Evans, P.R., Teo, C.H. and Nagai, K. (1994) Crystal structure at 1.92 A resolution of the RNA-binding domain of the U1A spliceosomal protein complexed with an RNA hairpin. *Nature*, **372**, 432-438.

Pagni, M., Ioannidis, V., Cerutti, L., Zahn-Zabal, M., Jongeneel, C.V. and Falquet, L. (2004) MyHits: a new interactive resource for protein annotation and domain identification. *Nucl. Acids Res.*, **32**, W332-335.

Pan, Z., Selyanko, A.A., Hadley, J.K., Brown, D.A., Dixon, J.E. and McKinnon, D. (2001) Alternative splicing of KCNQ2 potassium channel transcripts contributes to the functional diversity of M-currents. *J. Physiology*, **531**, 347-358.

Patel, N.A., Chalfant, C.E., Watson, J.E., Wyatt, J.R., Dean, N.M., Eichler, D.C. and Cooper, D.R. (2001) Insulin Regulates Alternative Splicing of Protein Kinase C beta II through a Phosphatidylinositol 3-Kinase-dependent Pathway Involving the Nuclear Serine/Arginine-rich Splicing Factor, SRp40, in Skeletal Muscle Cells. *J. Biol. Chem.*, **276**, 22648-22654.

Pendergast, A.M. (1996) Nuclear tyrosine kinases: from Abl to WEE1. *Current Opinion in Cell Biology*, **8**, 174-181.

Perez-Canadillas, J.-M. and Varani, G. (2001) Recent advances in RNA-protein recognition. *Curr. Opin. Struct. Biol.*, **11**, 53-58.

Platani, M. and Lamond, A.I. (2004) Nuclear organisation and subnuclear bodies. In Jeanteur, P. (ed.), *RNA Trafficking and Nuclear Structure Dynamics*. Springer, Vol. 35, pp. 1-22.

Po, S.S., Wu, R.C., Juang, G.J., Kong, W. and Tomaselli, G.F. (2001) Mechanism of alpha -adrenergic regulation of expressed hKv4.3 currents. *Am. J. Physiol. Heart Circ. Physiol.*, **281**, H2518-2527.

Ponting, C.P. (1997) Tudor domains in proteins that interact with RNA. *Trends Biochem. Sci.*, **22**, 51-52.

Poorkaj, P., Bird, T.D., Wijsman, E., Nemens, E., Garruto, R.M., Anderson, L., Andreadis, A., Wiederholt, W.C., Raskind, M. and Schellenberg, G.D. (1998) Tau is a candidate gene for chromosome 17 frontotemporal dementia. *Ann. Neurol.*, **43**, 815-825.

Popielarz, M., Cavaloc, Y., Mattei, M.-G., Gattoni, R. and Stvenin, J. (1995) The Gene Encoding Human Splicing Factor 9G8. *J. Biol. Chem.*, **270**, 17830-17835.

Prasad, J., Colwill, K., Pawson, T. and Manley, J.L. (1999) The Protein Kinase Clk/Sty Directly Modulates SR Protein Activity: Both Hyper- and Hypophosphorylation Inhibit Splicing. *Mol. Cell. Biol.*, **19**, 6991-7000.

Pret, A.M., Balvay, L. and Fiszman, M.Y. (1999) Regulated splicing of an alternative exon of beta-tropomyosin pre-mRNAs in myogenic cells depends on the strength of pyrimidine-rich intronic enhancer elements. *DNA Cell Biol.*, **18**, 671-683.

Rafalska, I., Zhang, Z., Benderska, N., Wolff, H., Hartmann, A.M., Brack-Werner, R. and Stamm, S. (2004) The intranuclear localization and function of YT521-B is regulated by tyrosine phosphorylation. *Hum. Mol. Genet.*, **13**, 1535-1549.

Reed, R. (2000) Mechanisms of fidelity in pre-mRNA splicing. *Curr. Opin. Cell Biol.*, **12**, 340-345.

Reed, R. and Palandjian, L. (1997) Spliceosome assembly. In Krainer, A.R. (ed.), *Eukaryotic mRNA Processing*. IRL Press at Oxford University Press, pp. 103-129.

Reeves, R. and Nissen, M.S. (1990) The A-T-DNA-binding domain of mammalian high mobility group I chromosomal proteins. A novel peptide motif for recognizing DNA structure. *J. Biol. Chem.*, **265**, 8573-8582.

Rehm, T., Huber, R. and Holak, T.A. (2002) Application of NMR in Structural Proteomics: Screening for Proteins Amenable to Structural Analysis. *Structure*, **10**, 1613-1618.

Riteau, B., Rouas-Freiss, N., Menier, C., Paul, P., Dausset, J. and Carosella, E.D. (2001) HLA-G2, -G3, and -G4 Isoforms Expressed as Nonmature Cell Surface Glycoproteins Inhibit NK and Antigen-Specific CTL Cytolysis. *J. Immunol.*, **166**, 5018-5026.

Robinson, D.R., Wu, Y.-M. and Lin, S.-F. (2000) The protein tyrosine kinase family of the human genome. *Oncogene*, **19**, 5548-5557.

Rosenfeld, R.G. (1994) Circulating growth hormone binding proteins. *Horm. Res.*, **42**, 129-132.

Rost, B. (1996) PHD: predicting one-dimensional protein structure by profile-based neural networks. *Methods Enzymol.*, **266**, 525-539.

Rubin, G.M., Yandell, M.D., Wortman, J.R., Gabor Miklos, G.L., Nelson, C.R., Hariharan, I.K., Fortini, M.E., Li, P.W., Apweiler, R., Fleischmann, W., Cherry, J.M., Henikoff, S., Skupski, M.P., Misra, S., Ashburner, M., Birney, E., Boguski, M.S., Brody, T., Brokstein, P., Celniker, S.E., Chervitz, S.A., Coates, D., Cravchik, A., Gabrielian, A., Galle, R.F., Gelbart, W.M., George, R.A., Goldstein, L.S.n.B., Gong, F., Guan, P., Harris, N.L., Hay, B.A., Hoskins, R.A., Li, J., Li, Z., Hynes, R.O., Jones, S.J.n.M., Kuehl, P.M., Lemaitre, B., Littleton, J.T., Morrison, D.K., Mungall, C., O'Farrell, P.H., Pickeral, O.K., Shue, C., Vosshall, L.B., Zhang, J., Zhao, Q., Zheng, X.H., Zhong, F., Zhong, W., Gibbs, R., Venter, J.C., Adams, M.D. and Lewis, S. (2000) Comparative Genomics of the Eukaryotes. *Science*, **287**, 2204-2215.

Ryter, J.M. and Schultz, S.C. (1998) Molecular basis of double-stranded RNA-protein interactions: structure of a dsRNA-binding domain complexed with dsRNA. *EMBO J.*, **17**, 7505-7513.

Ryther, R.C.C., McGuinness, L.M., Phillips III, J.A., Moseley, C.T., Magoulas, C.B., Robinson, I.C.A.F. and Patton, J.G. (2003) Disruption of exon definition produces a dominant-negative growth hormone isoform that causes somatotroph death and IGHD II. *Hum Genet*, **113**, 140-148.

Saitou, N. and Nei, M. (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.*, **4**, 406-425.

Sampson, N.D. and Hewitt, J.E. (2003) SF4 and SFRS14, two related putative splicing factors on human chromosome 19p13.11. *Gene*, **305**, 91-100.

Sanford, J.R., Longman, D. and Caceres, J.F. (2003) Multiple roles of the SR protein family in splicing regulation. In Jeanteur, P. (ed.), *Regulation of alternative splicing*. Springer, Vol. 31, pp. 33-58.

Scanlan, M.J., Gordan, J.D., Williamson, B., Stockert, E., Bander, N.H., Jongeneel, V., Gure, A.O., Jager, D., Jager, E., Knuth, A., Chen, Y.T. and Old, L.J. (1999) Antigens recognized by autologous antibody in patients with renal-cell carcinoma. *Int. J. Cancer*, **83**, 456-464.

Schaeffer, E.M. and Schwartzberg, P.L. (2000) Tec family kinases in lymphocyte signaling and function. *Curr. Opin. Immunol.*, **12**, 282-288.

Scheer, U. and Hock, R. (1999) Structure and function of the nucleolus. *Curr. Opin. Cell Biol.*, **11**, 385-390.

Schnare, M., Blum, H., Juttner, S., Rollinghoff, M. and Gessner, A. (1998) Specific Antagonism of Type I IL-4 Receptor with a Mutated Form of Murine IL-4. *J Immunol*, **161**, 3484-3492.

Schultz, J., Milpetz, F., Bork, P. and Ponting, C.P. (1998) SMART, a simple modular architecture research tool: Identification of signaling domains. *Proc. Natl Acad. Sci. U.S.A.*, **95**, 5857-5864.

Screaton, G.R., Bell, M.V., Jackson, D.G., Cornelis, F.B., Gerth, U. and Bell, J.I. (1992) Genomic structure of DNA encoding the lymphocyte homing receptor CD44 reveals at least 12 alternatively spliced exons. *Proc. Natl Acad. Sci. U. S. A.*, **89**, 12160-12164.

Sharp, P.A. (1994) Split genes and RNA splicing. *Cell*, **77**, 805-815.

Shatkin, A.J. and Manley, J.L. (2000) The ends of the affair: capping and polyadenylation. *Nat. Struct. Biol.*, **7**, 838-842.

Shifrin, V.I. and Neel, B.G. (1993) Growth factor-inducible alternative splicing of nontransmembrane phosphotyrosine phosphatase PTP-1B pre-mRNA. *J. Biol. Chem.*, **268**, 25376-25384.

Shin, C., Feng, Y. and Manley, J.L. (2004) Dephosphorylated SRp38 acts as a splicing repressor in response to heat shock. *Nature*, **427**, 553-558.

Shin, C. and Manley, J.L. (2004) Cell signalling and the control of pre-mRNA splicing. *Nat. Rev. Mol. Cell Biol.*, **5**, 727-738.

Shu, C.J. and Zhulin, I.B. (2002) ANTAR: an RNA-binding domain in transcription antitermination regulatory proteins. *Trends Biochem. Sci.*, **27**, 3-5.

Siddiqa, A., Sims-Mourtada, J.C., Guzman-Rojas, L., Rangel, R., Guret, C., Madrid-Marina, V., Sun, Y. and Martinez-Valdez, H. (2001) Regulation of CD40 and CD40 ligand by the AT-hook transcription factor AKNA. *Nature*, **410**, 383-387.

Siomi, H., Matunis, M.J., Michael, W.M. and Dreyfuss, G. (1993) The pre-mRNA binding K protein contains a novel evolutionarily conserved motif. *Nucleic Acids Res.*, **21**, 1193-1198.

Smith, C.W.J. and Valcarcel, J. (2000) Alternative pre-mRNA splicing: the logic of combinatorial control. *Trends Biochem. Sci.*, **25**, 381-388.

Smith, M.A., Fanger, G.R., O'Connor, L.T., Bridle, P. and Maue, R.A. (1997) Selective Regulation of Agrin mRNA Induction and Alternative Splicing in PC12 Cells by Ras-dependent Actions of Nerve Growth Factor. *J. Biol. Chem.*, **272**, 15675-15681.

Soret, J. and Tazi, J. (2003) Phosphorylation-dependent control of the pre-mRNA splicing machinery. In Jeanteur, P. (ed.), *Regulation of alternative splicing*. Springer, Vol. 31, pp. 89-126.

Spector, D.L. (1990) Higher Order Nuclear Organization: Three-Dimensional Distribution of Small Nuclear Ribonucleoprotein Particles. *Proc. Natl Acad. Sci. U. S. A.*, **87**, 147-151.

Spector, D.L. (2001) Nuclear domains. *J Cell Sci*, **114**, 2891-2893.

Spector, D.L. (2003) The dynamics of chromosome organization and gene regulation. *Annual Review of Biochemistry*, **72**, 573-608.

Speiser, P.W. and White, P.C. (1989) Structure of the human RD gene: a highly conserved gene in the class III region of the major histocompatibility complex. *DNA*, **8**, 745-751.

Spillantini, M.G. and Goedert, M. (1998) Tau protein pathology in neurodegenerative diseases. *Trends Neurosci.*, **21**, 428-433.

Spritz, R.A., Strunk, K., Surowy, C.S. and Mohrenweiser, H.W. (1990) Human U1-70K ribonucleoprotein antigen gene: organization, nucleotide sequence, and mapping to locus 19q13.3. *Genomics*, **8**, 371-379.

St Johnston, D., Brown, N., Gall, J. and Jantsch, M. (1992) A Conserved Double-Stranded RNA-Binding Domain. *Proc. Natl. Acad. Sci. U. S. A.*, **89**, 10979-10983.

Stamm, S. (2002) Signals and their transduction pathways regulating alternative splicing: a new dimension of the human genome. *Hum. Mol. Genet.*, **11**, 2409-2416.

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

Stamm, S., Zhang, M.Q., Marr, T.G. and Helfman, D.M. (1994) A sequence compilation and comparison of exons that are alternatively spliced in neurons. *Nucleic Acids Res.*, **22**, 1515–1526.

Stamm, S., Zhu, J., Nakai, K., Stoilov, P., Stoss, O. and Zhang, M.Q. (2000) An Alternative-Exon Database and Its Statistical Analysis. *DNA Cell Biol.*, **19**, 739-756.

Stasiv, Y., Regulski, M., Kuzin, B., Tully, T. and Enikolopov, G. (2001) The Drosophila Nitric-oxide Synthase Gene (dNOS) Encodes a Family of Proteins That Can Modulate NOS Activity by Acting as Dominant Negative Regulators. *J. Biol. Chem.*, **276**, 42241-42251.

Stoffler, D., Fahrenkrog, B. and Aebi, U. (1999) The nuclear pore complex: from molecular architecture to functional dynamics. *Curr. Opin. Cell Biol.*, **11**, 391-401.

Stoilov, P., Daoud, R., Nayler, O. and Stamm, S. (2004) Human tra2-beta1 autoregulates its protein concentration by influencing alternative splicing of its pre-mRNA. *Hum. Mol. Genet.*, **13**, 509-524.

Stoilov, P., Meshorer, E., Gencheva, M., Glick, D., Soreq, H. and Stamm, S. (2002a) Defects in pre-mRNA processing as causes of and predisposition to diseases. *DNA Cell Biol.*, **21**, 803-818.

Stoilov, P., Rafalska, I. and Stamm, S. (2002b) YTH: a new domain in nuclear proteins. *Trends in Biochemical Sciences*, **27**, 495-497.

Stojdl, D.F. and Bell, J.C. (1999) SR protein kinases: the splice of life. *Biochem. Cell Biol.*, **77**, 293-298.

Stoss, O., Novoyatleva, T., Gencheva, M., Olbrich, M., Benderska, N. and Stamm, S. (2004) p59fyn-mediated phosphorylation regulates the activity of the tissue-specific splicing factor rSLM-1. *Mol. Cell. Neurosci.*, **27**, 8-21.

Stoss, O., Olbrich, M., Hartmann, A.M., Konig, H., Memmott, J., Andreadis, A. and Stamm, S. (2001) The STAR/GSG Family Protein rSLM-2 Regulates the Selection of Alternative Splice Sites. *J. Biol. Chem.*, **276**, 8665-8673.

Stoss, O., Stoilov, P., Daoud, R., Hartmann, A.M., Olbrich, M. and Stamm, S. (2000) Misregulation of pre-mRNA splicing that causes human diseases. *Gene Ther. Mol. Biol.*, **5**, 9-28.

Stoss, O., Stoilov, P., Hartmann, A.M., Nayler, O. and Stamm, S. (1999) The in vivo minigene approach to analyze tissue-specific splicing. *Brain Research Protocols*, **4**, 383-394.

Studier, F.W., Rosenberg, A.H., Dunn, J.J. and Dubendorff, J.W. (1990) Use of T7 RNA polymerase to direct expression of cloned genes. *Methods Enzymol.*, **185**, 60-89.

Sureau, A. and Perbal, B. (1994) Several mRNAs with variable 3' untranslated regions and different stability encode the human PR264/SC35 splicing factor. *Proc. Natl Acad. Sci. U. S. A.*, **91**, 932-936.

Surowy, C.S., Hoganson, G., Gosink, J., Strunk, K. and Spritz, R.A. (1990) The human RD protein is closely related to nuclear RNA-binding proteins and has been highly conserved. *Gene*, **90**, 299-302.

Taagepera, S., McDonald, D., Loeb, J.E., Whitaker, L.L., McElroy, A.K., Wang, J.Y.J. and Hope, T.J. (1998) Nuclear-cytoplasmic shuttling of C-ABL tyrosine kinase. *Proc. Natl Acad. Sci. U. S. A.*, **95**, 7457–7462.

Tacke, R. and Manley, J.L. (1999) Determinants of SR protein specificity. *Curr. Opin. Cell Biol.*, **11**, 358-362.

Tang, Y., Novoyatleva, T., Benderska, N., Kishore, S., Thanaraj, A. and Stamm, S. (2005) Analysis of alternative splicing in vivo using minigenes. In Hartmann, R.K., Bindereif, A., Schön, A. and Westhof, E. (eds.), *Handbook of RNA Biochemistry*. Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, pp. 755-782.

Tarn, W. and Steitz, J. (1995) Modulation of 5' Splice Site Choice in Pre-Messenger RNA by Two Distinct Steps. *Proc. Natl Acad. Sci. U. S. A.*, **92**, 2504-2508.

Tarn, W.Y. and Steitz, J.A. (1996) Highly diverged U4 and U6 small nuclear RNAs required for splicing rare AT-AC introns. *Science*, **273**, 1824-1832.

Thanaraj, T.A. (2000) Positional characterisation of false positives from computational prediction of human splice sites. *Nucl. Acids Res.*, **28**, 744-754.

Thanaraj, T.A. and Stamm, S. (2003) Prediction and statistical analysis of alternatively spliced exons. In Janteur, P. (ed.), *Regulation of Alternative Splicing*. Springer, Vol. 31, pp. 1-31.

Thanaraj, T.A., Stamm, S., Clark, F., Riethoven, J.-J., Le Texier, V. and Muilu, J. (2004) ASD: the Alternative Splicing Database. *Nucl. Acids Res.*, **32**, D64-69.

Thompson, J.D., Higgins, D.G. and Gibson, T.J. (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.*, **22**, 4673–4680.

Thomson, A.M., Rogers, J.T., Walker, C.E., Staton, J.M. and Leedman, P.J. (1999) Optimized RNA gel-shift and UV cross-linking assays for characterization of cytoplasmic RNA-protein interactions. *Biotechniques*, **27**, 1032-1042.

Tone, M., Tone, Y., Fairchild, P.J., Wykes, M. and Waldmann, H. (2001) Regulation of CD40 function by its isoforms generated through alternative splicing. *Proc. Natl Acad. Sci. U. S. A.*, **98**, 1751-1756.

van der Houven van Oordt, W., Diaz-Meco, M.T., Lozano, J., Krainer, A.R., Moscat, J. and Caceres, J.F. (2000) The MKK3/6-p38-signaling Cascade Alters the Subcellular Distribution of hnRNP A1 and Modulates Alternative Splicing Regulation. *J. Cell Biol.*, **149**, 307-316.

Van Etten, R.A. (1999) Cycling, stressed-out and nervous: cellular functions of c-Abl. *Trends in Cell Biology*, **9**, 179-186.

Varani, G. and Nagai, K. (1998) RNA recognition by RNP proteins during RNA processing. *Annu. Rev. Biophys. Biomol. Struct.*, **27**, 407-445.

Varani, L., Hasegawa, M., Spillantini, M.G., Smith, M.J., Murrell, J.R., Ghetti, B., Klug, A., Goedert, M. and Varani, G. (1999) Structure of tau exon 10 splicing regulatory element RNA and destabilization by mutations of frontotemporal dementia and parkinsonism linked to chromosome 17. *Proc. Natl Acad. Sci. U. S. A.*, **96**, 8229-8234.

Venables, J.P., Elliott, D.J., Makarova, O.V., Makarov, E.M., Cooke, H.J. and Eperon, I.C. (2000) RBMY, a probable human spermatogenesis factor, and other hnRNP G proteins interact with Tra2beta and affect splicing. *Hum. Mol. Genet.*, **9**, 685-694.

Venables, J.P., Vernet, C., Chew, S.L., Elliott, D.J., Cowmeadow, R.B., Wu, J., Cooke, H.J., Artzt, K. and Eperon, I.C. (1999) T-STAR/ETOILE: a novel relative of SAM68 that interacts with an RNA- binding protein implicated in spermatogenesis. *Hum. Mol. Genet.*, **8**, 959-969.

Wang, C., Politz, J.C., Pederson, T. and Huang, S. (2003) RNA Polymerase III Transcripts and the PTB Protein Are Essential for the Integrity of the Perinucleolar Compartment. *Mol. Biol. Cell*, **14**, 2425-2435.

Wang, H.-Y., Lin, W., Dyck, J.A., Yeakley, J.M., Songyang, Z., Cantley, L.C. and Fu, X.-D. (1998) SRPK2: A Differentially Expressed SR Protein-specific Kinase Involved in Mediating the Interaction and Localization of Pre-mRNA Splicing Factors in Mammalian Cells. *J. Cell Biol.*, **140**, 737-750.

Wang, L., Oh, D.Y., Bogerd, J., Choi, H.S., Ahn, R.S., Seong, J.Y. and Kwon, H.B. (2001) Inhibitory Activity of Alternative Splice Variants of the Bullfrog GnRH Receptor-3 on Wild-Type Receptor Signaling. *Endocrinology*, **142**, 4015-4025.

Waragai, M., Lammers, C., Takeuchi, S., Imafuku, I., Udagawa, Y., Kanazawa, I., Kawabata, M., Mouradian, M. and Okazawa, H. (1999) PQBP-1, a novel polyglutamine tract-binding protein, inhibits transcription activation by Brn-2 and affects cell survival. *Hum. Mol. Genet.*, **8**, 977-987.

Weber, T., Baumgartner, R., Renner, C., Marahiel, M.A. and Holak, T.A. (2000) Solution structure of PCP, a prototype for the peptidyl carrier domains of modular peptide synthetases. *Structure*, **8**, 407-418.

Wedekind, J.E., Dance, G.S.C., Sowden, M.P. and Smith, H.C. (2003) Messenger RNA editing in mammals: new members of the APOBEC family seeking roles in the family business. *Trends Genet.*, **19**, 207-216.

Weg-Remers, S., Ponta, H., Herrlich, P. and Konig, H. (2001) Regulation of alternative pre-mRNA splicing by the ERK MAP-kinase pathway. *EMBO J.*, **20**, 4194-4203.

Weighardt, F., Biamonti, G. and Riva, S. (1996) The roles of heterogeneous nuclear ribonucleoproteins (hnRNP) in RNA metabolism. *Bioessays*, **18**, 747-756.

Welch, P. and Wang, J. (1993) A C-terminal protein-binding domain in the retinoblastoma protein regulates nuclear c-Abl tyrosine kinase in the cell cycle. *Cell*, **75**, 779-790.

Wheeler, D.L., Barrett, T., Benson, D.A., Bryant, S.H., Canese, K., Church, D.M., DiCuccio, M., Edgar, R., Federhen, S., Helmberg, W., Kenton, D.L., Khovayko, O., Lipman, D.J., Madden, T.L., Maglott, D.R., Ostell, J., Pontius, J.U., Pruitt,

K.D., Schuler, G.D., Schriml, L.M., Sequeira, E., Sherry, S.T., Sirotkin, K., Starchenko, G., Suzek, T.O., Tatusov, R., Tatusova, T.A., Wagner, L. and Yaschenko, E. (2005) Database resources of the National Center for Biotechnology Information. *Nucl. Acids Res.*, **33**, D39-45.

Wilkinson, F.L., Holaska, J.M., Zhang, Z., Sharma, A., Manilal, S., Holt, I., Stamm, S., Wilson, K.L. and Morris, G.E. (2003) Emerin interacts in vitro with the splicing-associated factor, YT521-B. *Eur. J. Biochem.*, **270**, 2459-2466.

Wong, B.R., Besser, D., Kim, N., Arron, J.R., Vologodskaia, M., Hanafusa, H. and Choi, Y. (1999) TRANCE, a TNF family member, activates Akt/PKB through a signaling complex involving TRAF6 and c-Src. *Mol. Cell*, **4**, 1041-1049.

Wrighton, N., Campbell, L.A., Harada, N., Miyajima, A. and Lee, F. (1992) The murine interleukin-4 receptor gene: genomic structure, expression and potential for alternative splicing. *Growth Factors*, **6**, 103-118.

Wu, J.Y. and Maniatis, T. (1993) Specific interactions between proteins implicated in splice site selection and regulated alternative splicing. *Cell*, **75**, 1061-1070.

Wu, S., Romfo, C.M., Nilsen, T.W. and Green, M.R. (1999) Functional recognition of the 3' splice site AG by the splicing factor U2AF35. *Nature*, **402**, 832-835.

Xiao, S.H. and Manley, J.L. (1997) Phosphorylation of the ASF/SF2 RS domain affects both protein-protein and protein-RNA interactions and is necessary for splicing. *Genes Dev.*, **11**, 334-344.

Xie, J., Beickman, K., Otte, E. and Rymond, B.C. (1998) Progression through the spliceosome cycle requires Prp38p function for U4/U6 snRNA dissociation. *EMBO J.*, **17**, 2938-2946.

Xie, J. and Black, D.L. (2001) A CaMK IV responsive RNA element mediates depolarization-induced alternative splicing of ion channels. *Nature*, **410**, 936-939.

Xie, J. and McCobb, D.P. (1998) Control of Alternative Splicing of Potassium Channels by Stress Hormones. *Science*, **280**, 443-446.

Xu, Q. and Lee, C. (2003) Discovery of novel splice forms and functional analysis of cancer-specific alternative splicing in human expressed sequences. *Nucl. Acids Res.*, **31**, 5635-5643.

Xu, Q., Modrek, B. and Lee, C. (2002) Genome-wide detection of tissue-specific alternative splicing in the human transcriptome. *Nucl. Acids Res.*, **30**, 3754-3766.

Yamaguchi, Y., Takagi, T., Wada, T., Yano, K., Furuya, A., Sugimoto, S., Hasegawa, J. and Handa, H. (1999) NELF, a multisubunit complex containing RD, cooperates with DSIF to repress RNA polymerase II elongation. *Cell*, **97**, 41-51.

Yanagisawa, H., Bundo, M., Miyashita, T., Okamura-Oho, Y., Tadokoro, K., Tokunaga, K. and Yamada, M. (2000) Protein binding of a DRPLA family through arginine-glutamic acid dipeptide repeats is enhanced by extended polyglutamine. *Hum. Mol. Genet.*, **9**, 1433-1442.

Yao, K.S., Godwin, A.K., Johnson, C. and O'Dwyer, P.J. (1996) Alternative splicing and differential expression of DT-diaphorase transcripts in human colon tumors and in peripheral mononuclear cells in response to mitomycin C treatment. *Cancer Res*, **56**, 1731-1736.

Zacharias, D.A. and Strehler, E.E. (1996) Change in plasma membrane Ca2(+)-ATPase splice-variant expression in response to a rise in intracellular Ca2+. *Curr. Biol.*, **6**, 1642-1652.

Zahler, A.M., Lane, W.S., Stolk, J.A. and Roth, M.B. (1992) SR proteins: a conserved family of pre-mRNA splicing factors. *Genes Dev.*, **6**, 837-847.

Zavolan, M., Kondo, S., Schonbach, C., Adachi, J., Hume, D.A., Hayashizaki, Y. and Gaasterland, T. (2003) Impact of Alternative Initiation, Splicing, and Termination on the Diversity of the mRNA Transcripts Encoded by the Mouse Transcriptome. *Genome Res.*, **13**, 1290-1300.

Zerler, B., Moran, B., Maruyama, K., Moomaw, J., Grodzicker, T. and Ruley, H.E. (1986) Adenovirus E1A coding sequences that enable ras and pmt oncogenes to transform cultured primary cells. *Mol. Cell Biol.*, **6**, 887-899.

Zhang, J., Berenstein, E.H., Evans, R.L. and Siraganian, R.P. (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**, 71-79.

Zhang, M., Zamore, P.D., Carmo-Fonseca, M., Lamond, A.I. and Green, M.R. (1992) Cloning and intracellular localization of the U2 small nuclear ribonucleoprotein auxiliary factor small subunit. *Proc. Natl Acad. Sci. U. S. A.*, **89**, 8769-8773.

Zhuang, Y. and Weiner, A.M. (1986) A compensatory base change in U1 snRNA suppresses a 5' splice site mutation. *Cell*, **46**, 827-835.

Zorio, D.A.R. and Bentley, D.L. (2004) The link between mRNA processing and transcription: communication works both ways. *Exp. Cell Res.*, **296**, 91-97.

CURRICULUM VITAE

Name:	Ilona Rafalska-Metcalf
Date of Birth:	September 2nd, 1976
Place of Birth:	Kraków, Poland
Nationality:	Polish
Marital Status:	Married

Education and Research Experience

2001 – 2005	Friedrich-Alexander-University, Erlangen, Germany Graduate Student
1999	Sanquin Research at CLB, Amsterdam, The Netherlands Visiting Scientist
1996 – 2001	Jagiellonian University, Kraków, Poland Master of Science, Molecular Biology M.Sc. Thesis: "Obtaining the interspecies hybridomas producing monoclonal antibodies against Tumor Necrosis Factor α-Convertase (TACE) domains and establishing the cell line stably transfected with cDNA for TACE"
1991 – 1995	Adam Mickiewicz High School, Kraków, Poland High School Diploma

Languages

Polish (Mother Tongue), English (fluent, both spoken and written), German and French (good practical skills), Russian (basic knowledge)

Awards received

5-month Tempus Mobility Grant Contract SJEP-11113, February – July 1999

Publications

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

Rafalska, I., Zhang, Z., Benderska, N., Wolff, H., Hartmann, A.M., Brack-Werner, R., Stamm, S. (2004) The intranuclear localization and function of YT521-B is regulated by tyrosine phosphorylation. *Hum. Mol. Genet.* 13(15), 1535-1549

Rafalska, I., Stoilov, P., Heinrich, B., Stamm, S. YTH: a new RNA binding domain? RNA 2003 8th Annual Meeting of the RNA Society, July 1-6, 2003, p. 617 (conference materials)

Stoilov, P., Rafalska, I., Stamm, S. (2002) YTH: a new domain in nuclear proteins. *Trends Biochem. Sci.* 27(10), 495-497

Bzowska, M., Rafalska, I., Lassak, A., Mezyk, R., Jura, N., Bereta, M., Bereta J. Monoclonal antibodies as tools for studying TNF-a release from cell membrane. "Progress in Molecular Biology" 29th Annual Conference of the Jagiellonian University Institute of Molecular Biology and Biotechnology, June 7-9, 2002, p.10 (conference materials)

Wicher, K.B., Jura, N., Kowanetz, M., Lorenowicz, M., Rafalska, I., Szybalski, W., Bereta, M. (1999) Eukaryotic expression system operating on the basis of Flp recombinase activity. *Biotechnologia*, 1(44), 144-152 (Polish)