Structural basis for dual regulation of spliceosomal helicase Brr2 by the Prp8 Jab1 domain and links to retinal disease

Inaugural-Dissertation
to obtain the academic degree
Doctor rerum naturalium (Dr. rer. nat.)
submitted to the Department of Biology, Chemistry and Pharmacy
of Freie Universität Berlin

by

Traudy Wandersleben

from Temuco, Chile

Berlin, 2013

The present work was carried out from March 2010 until May 2013 at the Institute of Chemistry and Biochemistry – Freie Universität Berlin, under the supervision of Prof. Dr. Markus Wahl.

1^{st} Reviewer: Prof. Dr. Markus Wahl

2^{nd} Reviewer: Prof. Dr. Christian Freund

Date of Defence: 25^{th} of October 2013

Acknowledgements

I would like to thank my PhD advisor, Prof. Markus Wahl, for giving me the great opportunity to work in his laboratory. Markus provided me with guidance, support and an unlimited flow of ideas that allowed me to finally finish this work. I am also very grateful to Karine Santos for all the help and advice during my PhD and for her precious friendship. She is a great scientist and I wish her all the best in her future career.

I will be always thankful to my friends and colleges Gert Weber, Christian Becke and Eva Absmeier for the insightful discussions about the research and their good will to read and correct the manuscript of my thesis.

Special thanks to Haydar Bulut for his friendship during this long journey. I would like to thank my colleges in the AG Wahl, whose I considered my friends, for the great working environment and the good time spent together inside and outside the laboratory. Oleg, Katja, Jan D., Jan W., Ronja, Nicole, Tonio, Nelly, Sunbin, Bernhard, Martin, Alex and Holger. Thank you all.

To Jia Hui Li, who worked with me during her Bachelor thesis. I am grateful for all her help, enthusiasm and motivation.

I am also very grateful to Karin Hesse, Claudia Alings, Clemens Langner and Carsten Jakob for their assistance during all these years. Dr. Uwe Mueller, Dr. Manfred Weiss, Dr. Sandra Puehringer and all MX-group at BESSY for the technical support. Dr. Gleb Bourenkov for his help and technical support at DESY.

I would like to thank my family for their valuable support. Specially to my parents, whose determination and great effort to give my siblings and me a good education were fundamental to set the basis that allowed me to be here.

Last, but certainly not least, I would like to thank my beloved husband, Ricardo. He has been at my side since I began this PhD giving me support, love and unwavering faith in me. There are no words to convey how much I love him.

Contents

List of Figures

List of Tables

Abbreviations

β-ME 2-Mercaptoethanol

ADP adenosine diphosphate

APS ammonium peroxodisulfate

ASCC activating signal cointegrator complex

ATP adenosine triphosphate

BAC bacterial artificial chromosome

BBP branch point binding protein

BP branch point

BPS branch point sequence

CNS crystallography and NMR system

Coot crystallographic object-oriented toolkit

DEN deformable elastic network

DMSO dimethylsulfoxide

DNA deoxyribonucleic acid

DSF differential scanning fluorimetry

DTT 1,4-Dithiothreitol

EDTA ethylenediaminetetraacetic acid

EJC exon junction complex

EM electron microscopy

ESE exonic splicing enhancer

ESI exonic splicing inhibitor

exoG exogenous guanosine

FRET Förster resonance energy transfer

GST Glutathione S-Transferase

HB domain helix bundle domain

HEPES 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid

HGNC Human Nomenclature Committee

HLH helix-loop-helix domain

IEP intron encoded protein

IG immunoglobulin-like domain

IPTG isopropyl-β-D-1-thiogalactopyranoside

ISE intronic splicing enhancer

ISI intronic splicing inhibitor

ISL intramolecular stem-loop

lncRNA long non-coding RNA

LSm Sm-like protein

m7G 7-methylguanosine

miRNA micro RNA

mRNA messenger RNA

mRNP messenger ribonucleoprotein complex

ncRNA non-coding RNA

NLS nuclear localization signal

NMR nuclear magnetic resonance

nt nucleotide

NTC NineTeen Complex

NTP nucleoside triphosphate

NTR complex NineTeen related complex

OD optical density

ORF open reading frame

PBS phosphate-buffered saline

PCR polymerase chain reaction

PIPES piperazine-N,N-bis-(2-ethanesulfonic acid)

PMSF phenylmethylsulfonyl fluoride

PPIases peptidyl-prolyl cis/trans isomerases

PPT polypyrimidine tract

pre-mRNA precursor mRNA

Prp pre-mRNA processing factor

RBD RNA binding domain

RES complex pre-mRNA retention and splicing complex

RNA ribonucleic acid

RP retinitis pigmentosa

RRM RNA recognition motif

rRNA ribosomal RNA

RS domain arginine/serine rich domain

SDS sodium dodecylsulfate

SDS-PAGE SDS-polyacrylamide gel electrophoresis

SF superfamily

SF1 splicing factor 1

SL stem loop

snRNA small nuclear RNA

snRNP small nuclear ribonucleoprotein complex

SR protein serine-arginine protein

SS splice site

TBE buffer TRIS-borate-EDTA buffer

TBS TRIS-buffered saline

TEMED N,N,N',N'-tetraethylenediamide

TEV tobacco etch virus

TRIS Tris-(hydroxymethyl)-aminomethane

tRNA transfer RNA

U2AF U2 auxiliary factor

UTR untranslated region

UV ultra-violet

WH domain winged helix domain

XDS X-ray Detector Software

Summary

The spliceosome is a very dynamic cellular ribonucleoprotein (RNP) machinery responsible for the removal of non-coding regions (introns) and ligation of coding regions (exons) during pre-mRNA maturation in eukaryotes. Throughout the splicing cycle the spliceosome undergoes major structural rearrangements, especially in the RNA-RNA interaction network. The spliceosome does not have a preformed catalytic core, therefore, these rearrangements are determinant for its formation. RNA helicases are considered the driving forces in the dynamic remodeling of RNA-RNA, RNA-protein and protein-protein interactions. Eight members of the superfamily 2 (SF2) of helicases are present in the spliceosome, the Ski2-like RNA helicase Brr2 being one of them.

Brr2 carries out the unwinding of the U4/U6 RNA duplex that is a crucial step for the spliceosome activation. The action of Brr2 ends with the release of the U4 snRNP and thus the U6 snRNA is able to extensively base pair with U2 snRNA, whose interaction is essential for catalysis. This enzyme has also been involved in spliceosome disassembly but, must probably, not as a helicase. Brr2 encounters the U4/U6 di-snRNA early, when U4/U6.U5 tri-snRNP is formed. Hence, Brr2 must be tightly regulated to ensure correct timing of spliceosome activation and disassembly. However, Brr2 does not only require inhibition to avoid premature unwinding, but also activation, because this helicase is very inefficient and has to unwind the most stable RNA duplex of the spliceosome. Two protein factors have been observed to modulate Brr2's activity, Prp8 and Snu114. Both proteins, like Brr2, are constitutive members of the U5 snRNP, but only Prp8 showed a direct effect on Brr2's activity. Nevertheless, the detailed features of this interaction are unknown. The main goals of this thesis were to learn more about the structural basis that make this interaction possible and how Prp8 influences Brr2's activity.

During my PhD thesis I produced recombinantely for the first time two fragments of human Prp8, hPrp8$^{\mathrm{Jab1/MPN}}$ and hPrp8$^{\mathrm{CTF}}$ (the latter was successfully co-expressed in complex with hBrr2$^{\mathrm{HR}}$). I was also able to produce most of the *Retinitis pigmentosa* (RP) related hPrp8$^{\mathrm{Jab1/MPN}}$ mutants described in literature. For each of the produced

proteins I established and optimized purification protocols. The so produced pure proteins were used for interaction studies, complex reconstitution and crystallization trials. In parallel, I reconstituted the yeast Brr2-Prp8 complexes using yPrp8 fragments already described in literature (yPrp8$^{Jab1/MPN}$ and yPrp8^{CTF}) and a novel construct of yBrr2, which includes the helicase region and an extended N-terminus (yBrr2^{enHR}). All the complexes generated in this work have been crystallized successfully. Nevertheless, only hBrr2^{HR}-hPrp8$^{Jab1/MPN}$ complex allowed me to solve the structure with good resolution (3.4 Å).

Based on the results of the structural and functional studies, I demonstrated how the Jab1/MPN domain of Prp8 binds to Brr2 and can inhibit RNA loading and Brr2-mediated U4/U6 snRNA unwinding by transiently inserting its C-terminal tail into Brr2's RNA binding channel. The same domain acts as a coactivator under conditions favoring RNA binding, enhancing the coupling of ATP hydrolysis to duplex unwinding. Thus, my data uncovered a unique dual-mode regulation of a SF2 helicase by a protein cofactor and revealed that its disruption of Brr2-Prp8 interaction constitutes a disease principle underlying certain forms of RP.

Zusammenfassung

Das Spleißosom ist eine sehr dynamische Ribonukleoprotein (RNP)-Maschine, die für das Entfernen von nicht-kodierenden Regionen (Introns) und das Ligieren von kodierenden Regionen (Exons) während der Reifung von prä-mRNA in Eukaryoten verantwortlich ist. Während des Spleißzyklus durchläuft das Spleißosom weitreichende strukturelle Reorganisationen, besonders im RNA-RNA-Interaktionsnetzwerk. Das Spleißosom besitzt kein vorgefertigtes katalytisches Zentrum, es wird erst durch umfangreiche Reorganisationen ausgebildet. Als treibende Kraft hinter diesen dynamischen Umgestaltungen von RNA-RNA, RNA-Protein und Protein-Protein-Interaktionen werden RNA-Helikasen angenommen. Acht Mitglieder der Superfamilie-2 Helikasen (SF2) sind im Spleißosom vertreten, eine von ihnen ist die Ski2-ähnliche Helikase Brr2.

Ein für die Aktivierung des Spleißosoms essentieller Schritt ist die Entwindung der U4/U6 RNA-Duplex, die von Brr2 katalysiert wird. Nach der Entwindung durch Brr2 wird das U4 snRNP vom Spleißosom losgelöst und ermöglicht die Ausbildung extensiver Basenpaarung zwischen den U2 und U6 snRNAs; diese Interaktion ist essentiell für die katalytische Aktivierung des Spleißosoms. Brr2 wird außerdem eine Funktion bei der Disassemblierung des Spleißosoms zugesprochen, wirkt dort mit großer Wahrscheinlichkeit jedoch nicht als Helikase. Sie trifft bereits früh im Spleißzyklus auf ihr Substrat, nämlich beim Aufbau des U4/U6.U5 tri-snRNPs. Daher muss Brr2 strikt reguliert werden, um einen korrekten zeitlichen Ablauf von Aktivierung und Disassemblierung zu gewährleisten. Diese Regulation beinhaltet nicht nur die Inhibierung von Brr2, um eine verfrühte Entwindung zu verhindern, sondern auch eine Aktivierung, da das Enzym nur über schwache Helikase-Aktivität verfügt, jedoch den stabilsten RNA-Duplex des Spleißosoms auflösen muss. Es konnte gezeigt werden, dass die Proteine Prp8 und Snu114 die Brr2-Aktivität beeinflussen können. Beide Proteine sind, wie Brr2, konstitutive Bestandteile des U5 snRNPs, aber nur für Prp8 konnte ein direkter Effekt auf die Brr2-Aktivität nachgewiesen werden. Dennoch sind die Details dieser Interaktion unbekannt. Das Ziel dieser Arbeit war es, mehr über die strukturellen Details dieser Interaktion zu lernen, und

wie Prp8 Brr2 beeinflusst.

Während meiner Doktorarbeit gelang es mir erstmals zwei Prp8-Fragmente, hPrp8$^{\mathrm{Jab1/MPN}}$ und hPrp8$^{\mathrm{CTF}}$, rekombinant herzustellen (letzteres Fragment durch Koexpression mit hBrr2$^{\mathrm{HR}}$). Des Weiteren gelang es mir, die meisten der mit Retinitis pigmentosa in Verbindung gebrachten Mutanten von hPrp8$^{\mathrm{Jab1/MPN}}$ zu produzieren. Für jedes dieser Proteine etablierte und optimierte ich Reinigungsprotokolle. Die so erhaltenen Proteine benutzte ich für Interaktionsstudien, Rekonstitution von Proteinkomplexen und Kristallisationsexperimente. Zeitgleich gelang es mir, mit bereits in der Literatur beschriebenen Prp8-Fragmenten (yPrp8$^{\mathrm{Jab1/MPN}}$ und yPrp8$^{\mathrm{CTF}}$) und einem neuen Brr2-Konstrukt, das zusätzlich zur Helikase-Region weitere N-terminale Reste beinhaltet (yBrr2$^{\mathrm{enHR}}$), Brr2-Prp8-Komplexe aus Hefe zu rekonstituieren. Alle in dieser Arbeit hergestellten Komplexe konnten erfolgreich kristallisiert werden, ich konnte jedoch nur die Struktur des hBrr2$^{\mathrm{HR}}$-hPrp8$^{\mathrm{Jab1/MPN}}$-Komplexes mit guter Auflösung (3.4 Å) bestimmen.

Basierend auf den Ergebnissen der strukturellen und funktionellen Untersuchungen konnte ich zeigen, dass die Jab1/MPN-Domäne von Prp8 Brr2 bindet und RNA-Bindung und Entwindung inhibiert, in dem der C-terminale Schwanz vorrübergehend in dem RNA-Bindungstunnel von Brr2 inseriert wird. Dieselbe Domäne wirkt auch als Koaktivator unter Bedingungen, die die RNA-Bindung begünstigen. Dabei wird die Kopplung von ATP-Hydrolyse und Duplex-Entwindung verstärkt. Meine Ergebnisse zeigen einen einzigartigen Regulationsmechanismus einer SF2-Helikase durch einen Protein-Kofaktor, der eine Doppelrolle als Aktivator und Inhibitor ausübt, und dass bestimmte Formen der Krankheit Retinitis pigmentosa in einer Störung der Interaktion zwischen Brr2 und Prp8 begründet sind.

Chapter 1

Introduction

In the wide molecular repertoire of the cell, three types of macromolecules stand out due to their functions, deoxyribonucleic acid (DNA), ribonucleic acid (RNA) and proteins. The flow of information between these macromolecules is determinant for life and in the last century it was described as the central dogma of molecular biology [Crick, 1970]. Originally, this flow was thought to be unidirectional, where the DNA contains the information required to build, maintain and propagate the organism, the RNA acts as a coding intermediate and the proteins are the functional final product [Crick, 1958]. However, this simple picture continuously acquires new layers of complexity, starting with the discovery of retroviruses that use RNA as template for DNA synthesis [Temin and Mizutani, 1970] and continuing with the multiple functions of the non-coding RNAs (ncRNAs) that have been unveiled in the last years.

RNAs are central players in gene expression and three major classes of RNA are main actors in the production of proteins: ribosomal RNA (rRNA), transfer RNA (tRNA) and messenger RNA (mRNA). Furthermore, RNA molecules are implicated in the processing and modification of the three mentioned RNA types, in virus replication and gene expression regulation acting as ncRNAs, such as ribozymes (catalytic RNA molecules), riboswitches (RNA sensors), long non-coding RNAs (lncRNAs) and micro RNAs (miRNAs) [Caprara and Nilsen, 2000; Brantl, 2002; Serganov and Patel, 2007; Wan et al., 2011]. Due to the versatility of RNA, the existence of an ancient world centred on this polymer has been postulated [Spirin, 2002].

Along with RNA's functional assortment, it was observed that they hardly function alone in a cellular environment. An RNA may interact with itself, other RNAs, ligands or proteins that would modulate its function and give rise to ribonucleoprotein

(RNP) machineries like the ribosome or the spliceosome.

1.1 Gene expression

The complete hereditary information for an individual organism is called genome and in the case of cellular life forms it exists as DNA. The functional unit of the genome is the gene. Gene is defined by The Human Nomenclature Committee (HGNC) as: "a DNA segment that contributes to phenotype/function. In the absence of demonstrated function, a gene may be characterized by sequence, transcription or homology" [Wain et al., 2002]. However, this definition seems to present several problems integrating some discoveries, such as overlapping genes (a DNA region may code for two different protein products in different reading frames [Contreras et al., 1977; Normark et al., 1983]) and trans-splicing phenomenon (ligation of portions of two separate pre-mRNA molecules to form a mature species [Konarska et al., 1985; Solnick, 1985; Lasda and Blumenthal, 2011]). Therefore, an updated definition has been proposed: "The gene is a union of genomic sequences encoding a coherent set of potentially overlapping functional products" [Gerstein et al., 2007]. Nevertheless, based on both definitions, the term gene may enclose sequences coding for diverse products like proteins and ncRNAs [Baetu, 2012]. In this thesis, gene expression will mainly refer to protein synthesis.

When the cell synthesizes a protein, the information contained in the DNA has to be transformed from nucleotides (nts) to amino acids. However, a gene is not directly translated into protein and has to undergo two main steps, transcription and translation. Transcription, for protein-encoding genes, is the process in which an mRNA copy is generated of one strand of the DNA that codes for the protein and it is carried out by the enzyme RNA polymerase (RNA polymerase II in eukaryotes).

The next step is translation, when the nucleotide sequence of the mRNA is converted into the amino acid sequence of the protein. This process is performed by the ribosomes, which are ribonucleoprotein particles consisting of two subunits, small and large, both with a high content of catalytic rRNA [Moore and Steitz, 2011]. During translation, nascent polypeptides are elongated from the amino (N) to the carboxy (C) terminus by the addition of one amino acid at a time. The small ribosomal subunit is responsible for the decoding process mediating the correct interactions between the anti-codons of the aminoacyl-tRNAs and the codons of the mRNA. The large ribosomal subunit builds the peptide bond [Melnikov et al., 2012]. Finally, when the stop codon in the mRNA is reached, the protein synthesis

is terminated and the ribosomal particles are recycled [Steitz, 2008].

Although the mRNA has the same function in all cells, there are some significant differences between prokaryotic and eukaryotic mRNA generation and processing. Since prokaryotes do not have a nucleus, the processes of transcription and translation are carried out in the same compartment and in a concerted manner. The mRNA generated is unstable and has a short half-life. Furthermore, the bacterial genes are mainly uninterrupted, except for few non-coding regions called introns (section 1.2) that are able to self-catalyze their excision. In contrast, transcription in eukaryotes occurs in the nucleus and the RNA generated is a precursor mRNA (pre-mRNA) which has to undergo maturation before it can be transported to the cytoplasm and translated. Maturation includes the addition of a 5' 7-methylguanosine (m7G) cap, the removal of introns with the subsequent ligation of the coding regions (exons) – in a process called splicing – and the final addition of the 3' polyadenine (poly A) tail. Not all eukaryotic genes are interrupted by introns and, in the case of simple organisms, such as yeast, most of the genes are uninterrupted. In comparison, most of the higher eukaryotic genes contain introns which are usually much longer than the exons [Lewin, 2003]. Another important feature of higher eukaryotes is that a vast majority of genes undergo alternative splicing. Recent studies using high-throughput sequencing indicate that more than 90% of human multi-exon genes are spliced in this manner [Wang et al., 2008; Luco et al., 2011]. During this process exons are differentially spliced to generate diverse mRNA products starting from a single pre-mRNA and producing multiple protein isoforms that are determinant for cell and tissue identity. Alternative splicing is considered the main mechanism by which a small number of protein-coding genes in the human genome (around 20.000) can give rise to more than 100.000 proteins of a cell's proteome. Thus, alternative splicing appears as a clear explanation for mammalian proteomic complexity [Luco et al., 2011].

1.2 Introns and their splicing mechanisms

Introns can be separated into four major classes depending on their splicing mechanism: tRNA/archaeal introns, autocatalytic group I and group II introns and spliceosomal introns [Haugen et al., 2005]. tRNA/archaeal introns are generally small (14-160 nt) with a prevalence of at least 5% in the tRNA genes of Archaea and Eukarya. These introns do not share any sequence homology, but their locations are very conserved, being found predominantly in the anticodon stem. tRNA splicing depends exclusively on protein enzymes: an endonuclease excises the intron, a ligase

joins the exons and, in the case of eukaryotes, a 2'-phosphotransferase is required to remove a 2'-phosphate left at the ligation junction (Fig.1.1 A)[Calvin and Li, 2008].

Group I introns are small (between 250-500 nt) and able to self-splice from RNA transcripts via a series of two transesterification reactions using a specific mechanism that requires the binding of an exogenous guanosine (exoG) cofactor to the catalytic core of the intron, called G-binding site. The first step starts with the nucleophilic attack at the 5' splice site (SS) by the 3'OH of the exoG, which attaches to the intron and releases the 5' exon as a consequence. The second transesterification step is initiated by the attack on the 3'SS by the 3'OH of the previously released exon, resulting in exon ligation and also intron loss (Fig.1.1 B). Group I introns can be encountered in bacteria and eukaryotes. Nevertheless, most of such introns identified to date are found in eukaryotic genomes. Among them, plants, algae and fungi are particularly enriched taxa, with a prevalence of 90% of all group I introns. From the eukaryotic group I introns, 58% are in the nucleus (forming part of rDNA), 16% are mitochondrial and 26% are in plastid DNA. In contrast, the prokaryotic introns are rare (less than 4% of all known group I introns) and none of them have been found in Archaea [Haugen et al., 2005].

Group II introns are mobile genetic elements with sizes varying from 400 to 800 nt and a highly conserved secondary structure consisting of six domains radiating from a central wheel (DI-VI) (Fig.1.2 A). These domains interact to form an equally conserved tertiary structure (Fig.1.2 B) that brings distant sequences together, some of which form the active site [Toor et al., 2008]. Apart from the catalytic RNA, the majority of group II introns contain an open reading frame coding for a multifunctional intron-encoded protein (IEP) that assists in splicing. Group II introns can be found in bacteria, archaea (rare) and the organelles of various eukaryotes (mostly fungi and plants) [Fedorova and Zingler, 2007; Lambowitz and Zimmerly, 2010].

Like group I introns, group II introns catalyze their own excision via transesterification reactions, but use a mechanism similar to the spliceosomal intron removal. In the first step, the 2'OH of a bulged adenosine in DVI attacks the phosphodiester bond on the 5'SS, generating an intron lariat bound to the 3' exon and a free 5' exon. In the second step, the 3'OH of the cleaved 5' exon is the nucleophile and attacks the phosphodiester bond on the 3'SS, resulting in exon ligation and intron lariat RNA excision (Fig.1.3). Other less common mechanisms, such as reverse splicing, hydrolytic splicing and circle formation have also been described for this kind of introns, however, they will not be addressed in this thesis (for more information refer to [Fedorova and Zingler, 2007]).

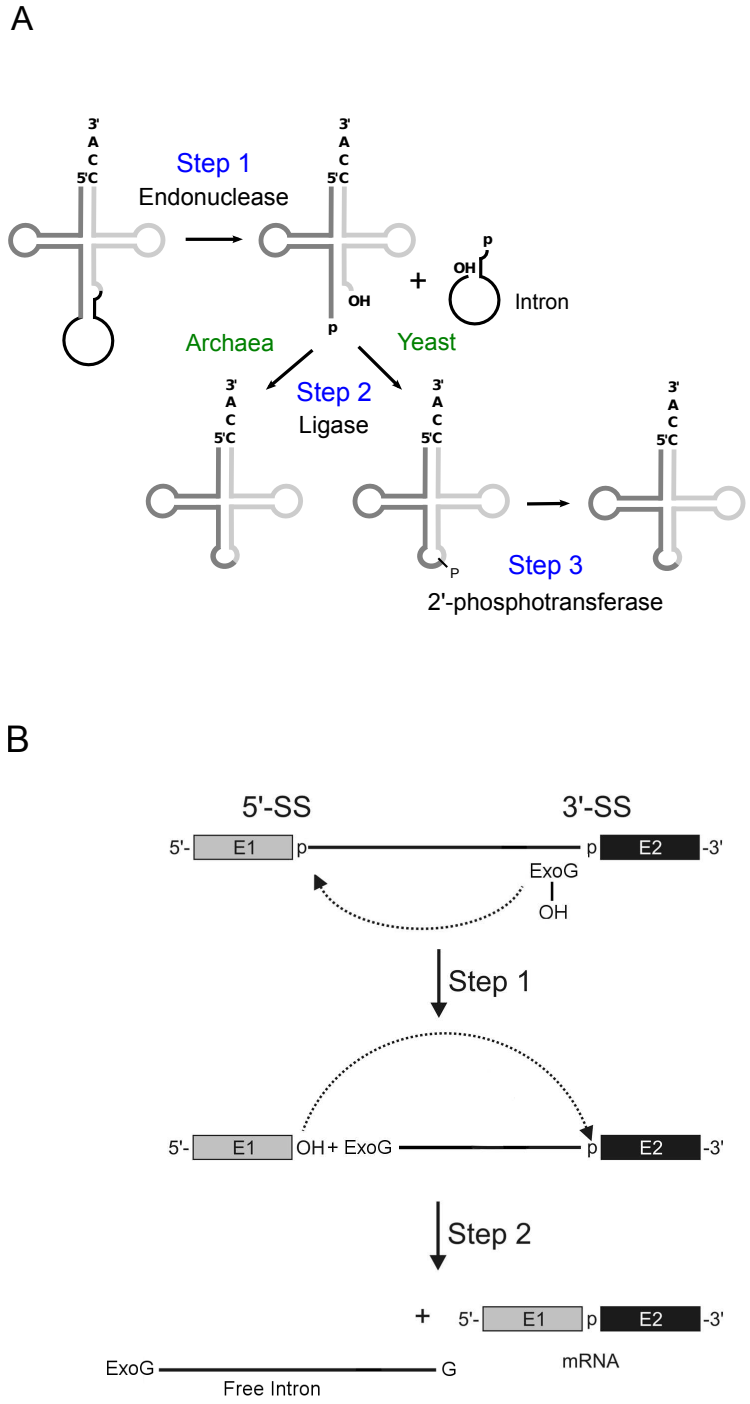

Figure 1.1: Intron splicing mechanisms. **(A)** The eukaryotic and archaeal tRNA splicing pathway. The pathways split at the ligation step. Required co-factors are highlighted at each step [Calvin and Li, 2008]; **(B)** Schematic representation of the two-step splicing mechanism of group I introns. Boxes and solid lines represent the exons and the intron, respectively. Step 1, nucleophilic attack at the 5'splice site by the 3'-OH of guanosine. Step 2, the second transesterification reaction leading to exon ligation and release of the intron RNA [Haugen et al., 2005].

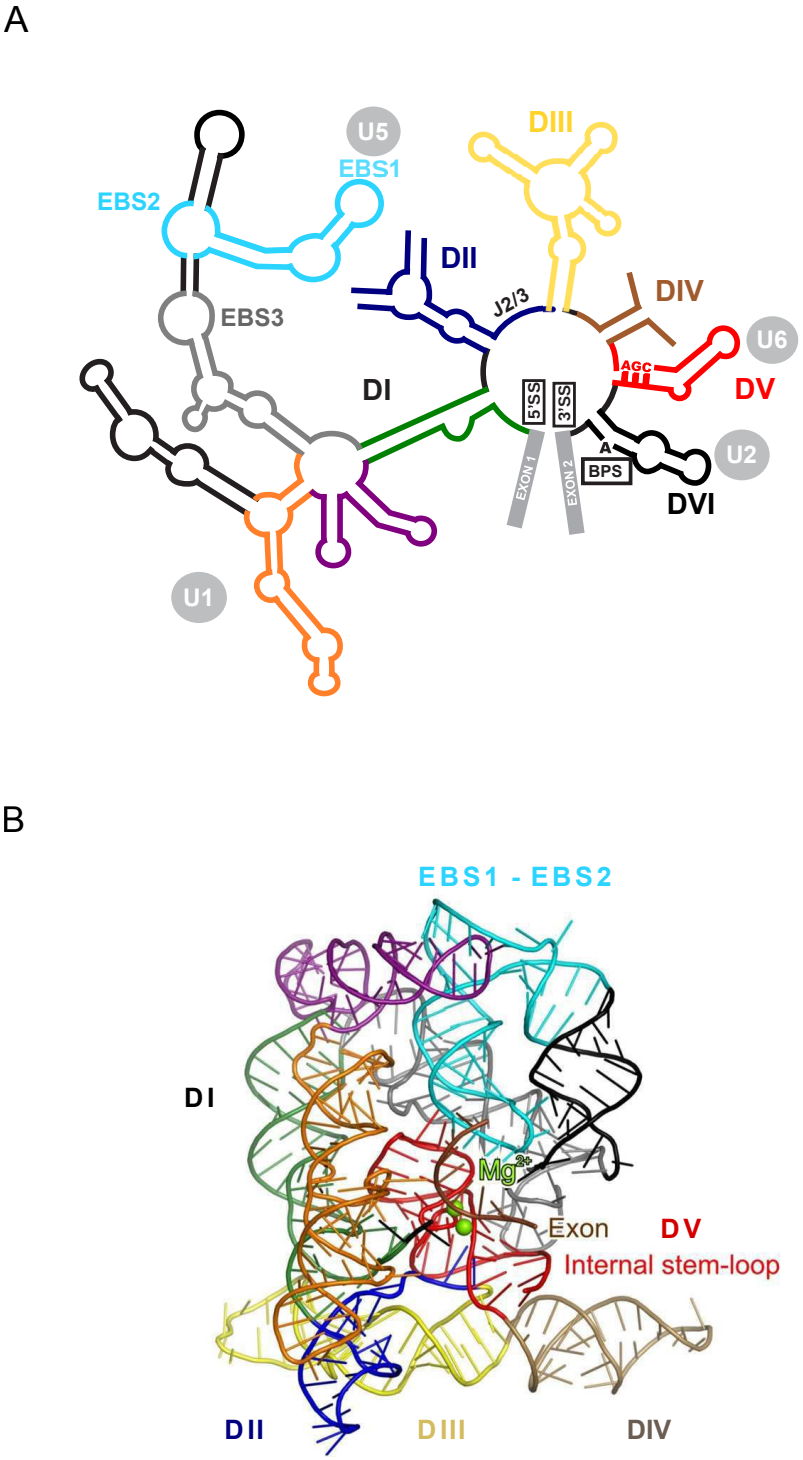

Figure 1.2: Group II introns. **(A)** Schematic representation of the secondary structure of a group II intron. The exons are shown as gray boxes. Each domain (DI-VI) is coloured as in (B) and the identity of each RNA domain is indicated close to each domain. The sites of 5' and 3' splice sites and branch sequence are marked as 5' SS, 3' SS and BS, respectively. The functional and/or structural homologue of each RNA element in the major spliceosome is shown in gray circles; **(B)** Tertiary structure of a group II intron [Toor et al., 2008].

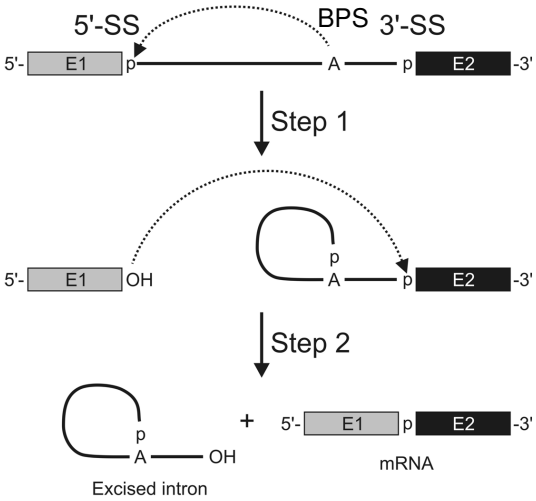

Figure 1.3: Schematic representation of the two-step mechanism of group II introns and pre-mRNA splicing. Boxes and solid lines represent the exons and the intron, respectively. The branch site adenosine is indicated by the letter A and the phosphate groups by the letter p. The arrows indicate nucleophilic attacks.

Spliceosomal or pre-mRNA introns are excised through a similar mechanism as group II introns, but they lack the catalytic domains of the latter and require a very dynamic RNP machinery, called spliceosome, to carry out the two transesterification reactions. Since pre-mRNA splicing is the central topic of this thesis, it will be described in more detail in the following sections. Introns in nuclear pre-mRNAs are very variable in size and contain little conserved secondary and tertiary structural information. These introns are mainly defined by very short consensus sequences, the 5'SS, the 3'SS and the branch point sequence (BPS). These elements are active players in two transesterification reactions (Fig.1.3) which resemble the 1^{st} and 2^{nd} step of group II introns splicing. The consensus sequences exhibit a different degree of conservation depending on the organism (Fig.1.4) [Wahl et al., 2009; Will and Lührmann, 2010]. Introns in *Saccharomyces cerevisiae* show a high sequence conservation in the three *cis*-elements. The 5'SS sequence entails 5'-**GUAUGU**-3' (the sequences typeset in bold represent nts with at least 90% of conservation), the consensus BPS is 5'-**CUA<u>A</u>C**-3' and contains the conserved adenosine (underlined) implicated in the first step of catalysis, and, finally, the sequence that represents the vast majority of the 3' splice sites, 5'-**YAG**-3' (Y = pyrimidine) [Spingola et al., 1999]. In the case of higher eukaryotes, these sequences are less conserved. However, it is still possible to identify nucleotide patterns. The 5'SS consensus sequence is 5'-**GUR**AGU-3' (where the sequences typeset in red bold face, from now on, are the invariable nt, in bold black are the nt with more than 90% of conservation, the non bold parts of the sequence have over 50% of conservation; R = purine), the consensus BPS is generally located 18-40 nt upstream of the 3'SS and comprises

the sequence 5'-C**U**R**A**C-3' (bulged adenosine underlined). The last exon/intron recognition signal is the 3'SS and includes the sequence 5'-**Y**AG-3' [Sharp and Burge, 1997]. Furthermore, just upstream of the 3'SS element, another essential and highly conserved segment (> 70% of conservation) of 10-12 nt has been identified. This stretch is mainly composed of pyrimidines and, therefore, it was called the polypyrimidine tract (PPT).

Interestingly, another class of introns has been recognized in the genomes of many metazoans, the U12-type introns. The latter are distinguished by an unusual combination of dinuclotides at the 5' and 3' termini (AU and AC, respectively), by a very conserved sequence at their 5'SS and BPS, and by the lack of PPT. These introns are generally rare, comprising 0.15%-0.35% of all human introns, being concentrated in so-called information processing genes, which include functions such as DNA replication, DNA repair and gene expression [Burge et al., 1998; Pessa, 2006].

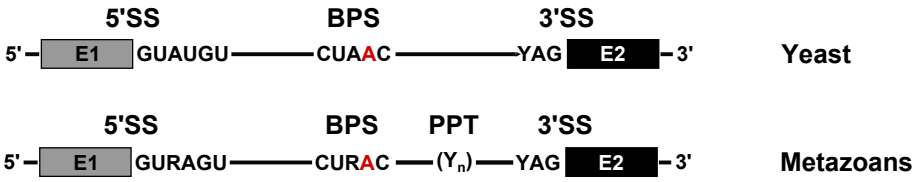

Figure 1.4: Conserved sequence elements found in introns from metazoans and yeast *S. cerevisiae*. The 5' and 3' exons are shown as boxes. The branch point adenosine is highlighted in red typeset. Y = pyrimidine and R = purine. The polypyrimidine tract is represented as Y(n).

Due to the scarce information contained in the short consensus sequences and their general poor conservation, additional *cis*-elements that could ensure intron recognition are necessary. Among them are the exonic and intronic splicing enhancers (ESEs and ISEs) or silencers (ESSs and ISSs). These elements are short and diverse sequences which regulate constitutive and alternative splicing by binding specific RNA-binding proteins that include heterogeneous nuclear ribonucleoproteins (hnRNPs) and serine/arginine rich (SR) proteins [Will and Lührmann, 2010; Montes et al., 2012].

1.3 The spliceosome

The splicing of pre-mRNA introns is catalyzed by the spliceosome, a multi-megadalton, highly dynamic RNP machine. In many eukaryotes and most metazoans, two splicing machineries coexist: the major or U2-dependent spliceosome, responsible for the removal of the most abundant class of introns (GU/AG introns),

and the minor or U12-dependent spliceosome, which catalyzes the splicing of the U12-type (AU/AC) of introns. The main building blocks of both spliceosomes are the uridine-rich small nuclear RNPs (snRNPs). These particles are generally composed of a post-transcriptionally modified small nuclear RNA (snRNA), a common set of seven Sm proteins (B/B', D3, D2, D1, E, F and G in humans), arranged around the snRNA forming a heteroheptameric ring, and a variable number of particle-specific proteins (Fig.1.5) [Patel and Bellini, 2008]. There are two exceptions to this general composition, U6 and U6atac snRNPs. Both snRNPs do not contain Sm proteins and seven Sm like proteins (LSm 2-8) have been found associated with these particles instead [Schneider et al., 2002; Will and Lührmann, 2005; Patel and Bellini, 2008].

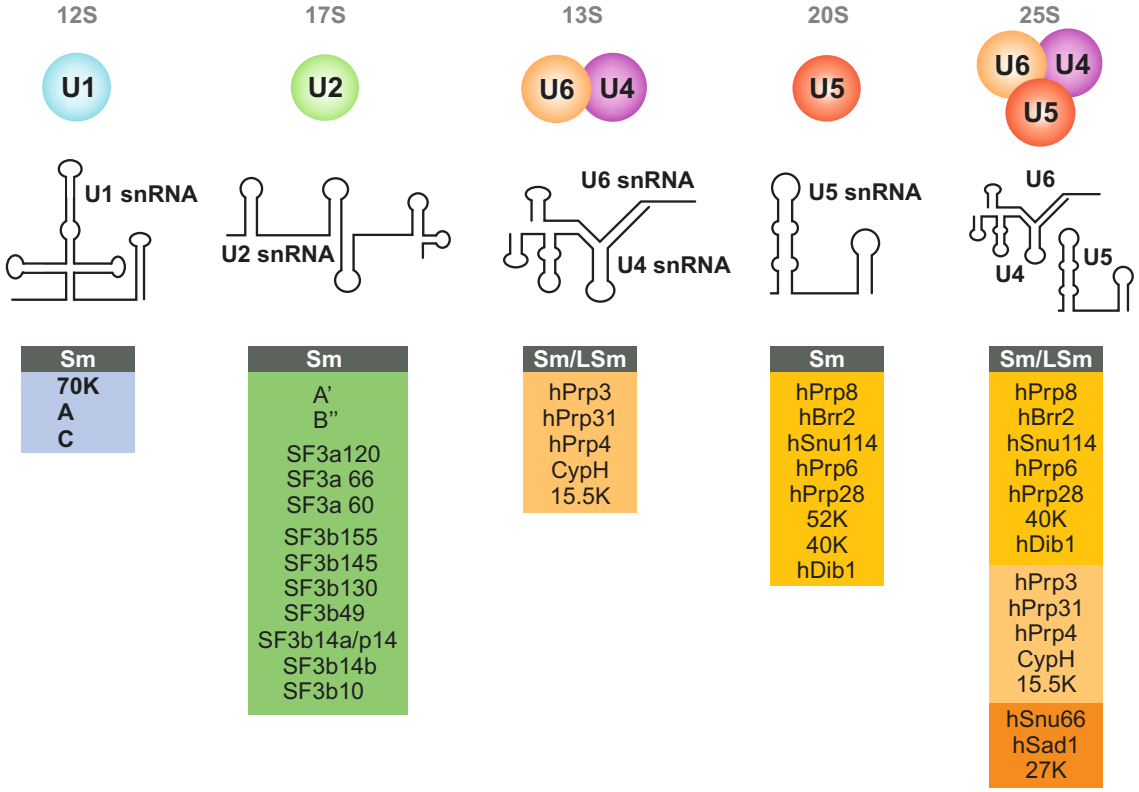

Figure 1.5: Protein composition and snRNA secondary structures of the human major spliceosomal snRNPs. The seven Sm and LSm proteins are indicated by grey boxes labeled Sm and LSm, respectively. The colored boxes list the specific proteins associated with each snRNP. The tri-snRNP contains two sets of Sm proteins and one set of LSm proteins [Will and Lührmann, 2010].

1.4 The major spliceosome

The major spliceosome requires 5 snRNPs U1, U2, U4, U5 and U6 (named according to their discovery order). U4 and U6 form a single particle through extensive base-pairing of their snRNAs, the U4/U6 di-snRNP. U5 snRNP is also not

recruited alone during spliceosome assembly and forms an early complex together with the U4/U6 di-snRNP, the 25S tri-snRNP (Fig.1.5). Since most of the snRNP-associated proteins identified so far are conserved in metazoans and yeast, the following sections will mainly refer to the human nomenclature of the proteins [Jurica and Moore, 2003].

The spliceosome assembles on the pre-mRNA in a stepwise manner and none of the mentioned particles contain a preformed catalytic center to carry out the transesterification reactions. Thus, the catalytic center has to be formed anew in each splicing cycle (Fig.1.6). The canonical assembly observed in all eukaryotic cells is based on intron recognition. However, in metazoans, when introns exceed more than 250nt, an alternative assembly pathway has been observed. This pathway differs from the cross-intron assembly mainly in the early steps, where the splicing complexes first gather across the exons in a process called exon definition [Fox-Walsh et al., 2005].

1.4.1 Stepwise assembly of snRNP particles

The first step of the cross-intron assembly is the recruitment of U1 to the 5'SS, then two non-snRNP factors, the the splicing factor 1 (SF1)/ branch point binding protein (BBP) and the U2 auxiliary factor (U2AF), bind to the BPS and PPT respectively. These initial interactions formed the center of the spliceosomal complex E.

The next step involves the ATP-dependent U2 snRNP incorporation. This snRNP base pairs with the BPS, in a process which comprehends the displacement of SF1/BBP and leads to complex A formation (also called pre-spliceosome). After U2 engagement, the pre-assembled U4/U6.U5 tri-snRNP is recruited, generating the pre-catalytic complex B. At this point, all the components required for catalysis are present, but it is necessary that they undergo major rearrangements in RNA-RNA and RNA-protein interactions that will lead to the release of U1 and U4 snRNPs and the subsequent complex B activation (B^{act} complex). However, for complete catalytic activation, the DEAH-box RNA helicase Prp2 is additionally required, yielding the complex B^{*}. In this way, the first transesterification reaction can be performed. The first step of catalysis generates the complex C and after additional rearrangements the second catalytic step is carried out. Finally, the mature mRNP is released and the spliceosome dissociates, freeing the U2, U5 and U6 snRNPs to be recycled for further splicing cycles [Wahl et al., 2009; Will and Lührmann, 2010].

In the case of spliceosome assembly across the exon, U1 snRNP binds to the

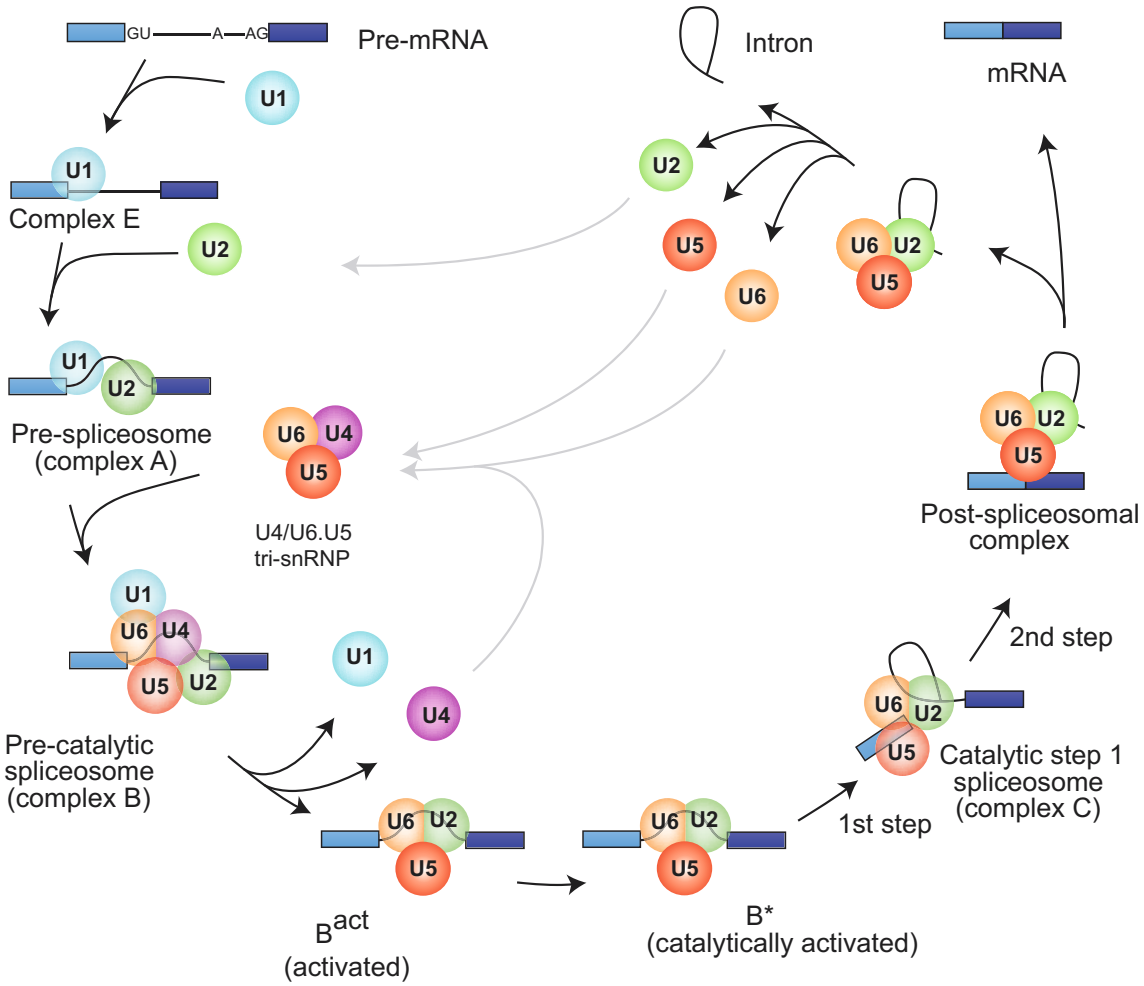

Figure 1.6: Schematic representation of the assembly and disassembly of the U2-dependent spliceosome. For simplicity, the ordered interactions of the snRNPs (indicated by circles), but not those of non-snRNP proteins, are shown. The various spliceosomal complexes are named according to the metazoan nomenclature. Exon and intron sequences are indicated by boxes and lines, respectively [Will and Lührmann, 2010].

5'SS downstream of an exon and assists U2AF recruitment on the PPT upstream of the exon. Then U2 snRNP is incorporated by binding to the BPS upstream of the exon and proteins of the SR protein family are also recruited by the ESEs within the exon. Together, these proteins form a stabilizing network across the exon that holds the exon-defined complex up. Nevertheless, the catalytic steps of splicing occur across an intron and, after splice site recognition, the system should switch from an exon-defined to intron-defined splicing complex. Currently, not much is known about this latter event, but it has been postulated that cross-exon interactions are first disrupted and transformed into cross-intron interactions being an essential step for exon inclusion during alternative splicing [Reed, 2000; Smith and Valcárcel, 2000; Will and Lührmann, 2010].

1.4.2 RNA rearrangements

The spliceosome undergoes major structural rearrangements during the splicing cycle, especially in the RNA-RNA interaction network. Most of the information gathered to date is related to changes observed in the RNA secondary structure interactions and little is known about the dynamics of the tertiary structure interactions, hence the current rearrangement models have to be considered as simplified representations of the real process.

At the initial stages, U1 snRNA base pairs with the 5'SS and afterwards U2 snRNA base pairs with the BPS, forming a short U2/BPS duplex in which the branch point adenosine is bulged out. Then, upon recruitment of the U4/U6.U5 tri-snRNP, the pre-catalytic complex B is formed. In this complex, the U6 and U4 snRNAs are extensively base paired with each other, U5 snRNA contacts nucleotides of the 5' and 3' exons, and the 3' end of U6 snRNA base pairs with the 5' end of U2 [Nilsen, 1994; Staley and Guthrie, 1998a] (Fig.1.7).

During catalytic activation the U4/U6 RNA duplex is disrupted and the 5' end of U6 snRNA displaces the U1 snRNA, establishing an interaction with the 5'SS. In addition, U6 and U2 snRNAs interact with each other generating an extensively base paired duplex and a central region of the U6 snRNA forms an intramolecular stem-loop (U6-ISL) which has been shown to be crucial for catalysis [Staley and Guthrie, 1998a] (Fig.1.7).

After the first step of splicing, the RNA network is rearranged again and the splicing intermediates are relocated to allow the nucleophilic attack of the 5'exon at the 3'SS. There is not much information about the changes that occur at this stage. Nevertheless, recent studies have shown that the U6/5'SS interaction must be disrupted and the U2/BPS interaction is also not required for the second step [Will and Lührmann, 2010]. The U5 snRNA, on the other hand, is responsible for 5' exon tethering and the alignment of the exons for the second splicing step [Turner et al., 2004].

1.4.3 Major spliceosome snRNP particles

1.4.3.1 U1

The human 12S U1 snRNP consists of a 165 nucleotide long snRNA molecule and ten proteins, including the seven Sm proteins common to all snRNPs. The other three proteins are specific for U1 and are named U1-70K, U1-A and U1-C

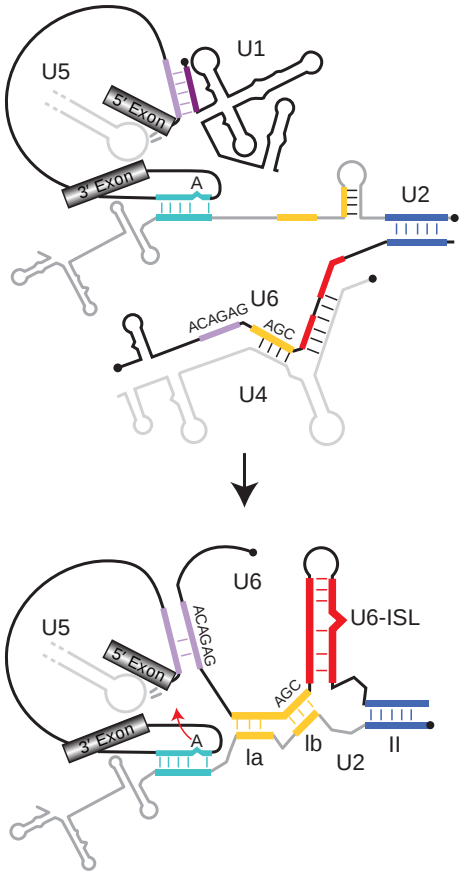

Figure 1.7: RNA rearrangement during activation. Exon sequences are indicated by grey boxes and intron sequences by a thin black line. snRNAs are shown schematically in grey or black, the regions engaging in base pairing interactions (indicated by short lines) highlighted in color (not drawn to scale). The 5' end of the snRNAs is indicated by a black dot. Solely stem loop 1 of the U5 snRNA is shown. During the transition from a precatalytic spliceosome (upper diagram) to a catalytically activated spliceosome (lower diagram) U1 and U4 are displaced, and U6 and U2 engage in novel base pairing interactions [Will and Lührmann, 2010].

[Patel and Bellini, 2008; Pomeranz Krummel et al., 2009; Buratti and Baralle, 2010]. A couple of years ago, the crystal structures of the core of human U1 snRNP at 5.5 Å and 4.4 Å resolution have been published [Pomeranz Krummel et al., 2009; Weber et al., 2010]. The first U1 crystal structure contained the U1 snRNA, the Sm proteins and portions of U1-C and U1-70K, but the U1-A protein was not present [Pomeranz Krummel et al., 2009]. From this map, it was determined that the U1 snRNA forms four stem-loops, named SL1 to SL4, and a short helix (H) confirming the model suggested by Krol et al. [Krol et al., 1990]. Following the short helix, and in between SL3 and SL4, the single-stranded uridine rich Sm site is located, where the seven Sm proteins assemble forming a heteroheptameric ring. U1-70K binds to SL1 through its RNA binding domain (RBD), while U1-C binds the N-terminal residues of U1-70K and, by means of this binding, the protein is able to contact

the U1 core domain. The contact of U1-C to the RNA was observed in the crystal structure where the 5' end of U1 snRNA base-pairs with its counterpart from an adjacent complex, mimicking the binding of the pre-mRNA 5'SS to U1 snRNA. Thus, a helix from the zinc finger of U1-C binds across the minor groove of this RNA duplex, being consistent with its role as a stabilizer of the U1/5'SS base-pairing [Pomeranz Krummel et al., 2009].

1.4.3.2 U2

The U2 snRNP, analogous to U1, is also formed by a modified snRNA and seven Sm proteins. The human U2 snRNA is 186 nt long and the Sm proteins are arranged in the corresponding ring structure around the Sm site. Apart from this basic core, the U2 snRNP is stably associated to the U2-A' and U2-B" polypeptides and the heteromeric protein complexes SF3a and SF3b. All together, these molecules form a stable particle, the so called 17S U2 snRNP (Fig.1.5). The SF3a complex consists of at least three subunits –SF3a120, SF3a66, and SF3a60– and the SF3b complex contains at least 7 subunits –SF3b155, SF3b145, SF3b130, SF3b49, SF3b14a/p14, SF3b14b, and SF3b10. All mentioned subunits are named according to their apparent molecular weight. Additionally, a series of other proteins have been linked to the U2 snRNP (e.g. CHERP, hPrp5, hPrp43, SPF31 and SPF45), the majority of which are present in substoichiometric amounts [Will et al., 2002; Dybkov et al., 2006; Patel and Bellini, 2008]. These protein factors were isolated by immunoaffinity chromatography in complex with the 17S particle and subsequently analyzed through mass spectrometry. They were found mostly in low amounts and their identification was only possible due to mild buffer conditions during the purification procedure [Will et al., 2002].

Functionally, U2 snRNP plays a crucial role in splicing. In the initial steps of assembly, U2 participates in the recognition of the pre-mRNA through base pairing with the branch point site. Subsequently, it positions the branch point adenosine within the catalytic core of the spliceosome and, together with U6 snRNA, U2 snRNA forms part of the RNA network that brings into close proximity the reactive sites of the pre-mRNA [Will et al., 2002; Dybkov et al., 2006].

1.4.3.3 U5

The 20S U5 snRNP is the largest RNP in the spliceosome and the only particle shared by major and minor spliceosomes. It consists of a 116 nt long snRNA,

the Sm proteins and eight specific protein factors: hPrp8 (220K), hPrp6 (102K), hBrr2 (200K), hPrp28 (100K), hSnu114 (116K), hLin1 (52K), 40K and hDib1 (15K) [Liu et al., 2006; Patel and Bellini, 2008; Valadkhan and Jaladat, 2010; Will and Lührmann, 2010]. Two of these proteins, hBrr2 and hPrp8, are the central topic of this thesis and will be described in more detail in sections 1.8 and 1.9, respectively. The overall distribution of the different components of this particle was determined by cryo-EM and affinity labelling, suggesting that the snRNA occupies the center of the U5 snRNP and the associated proteins are located in the periphery [Sander et al., 2006; Häcker et al., 2008].

U5 snRNA bears a highly conserved stem-loop that interacts with exon sequences at the 5'SS and 3'SS, thus being implicated in the tethering of the 5' exon splicing intermediate produced by the first transesterification and in the alignment of the exons for the second catalytic step [Newman, 1997; Turner et al., 2004]. The proteins, on the other hand, are associated with the activation and regulation of the spliceosome prior to and during catalysis. hPrp6 is necessary for the formation of a stable tri-snRNP U4/U6.U5 [Makarov et al., 2000; Lutzelberger et al., 2009]. hPrp28 and hBrr2 are ATP dependent helicases belonging to the SF2 superfamily and it has been shown that both are important for the RNA rearrangements required for spliceosome activation. hPrp28 has a role in destabilizing the U1 snRNA interaction with the 5'SS and hBrr2 unwinds the U4/U6 RNA duplex leaving U6 snRNA free for new base pairing with U2, hence leading to the catalytic core formation [Staley et al., 1999; Turner et al., 2004; Liu et al., 2006]. hPrp8 and hSnu114 are two of the constitutive members of U5 that have been linked to hBrr2 activity modulation [Small et al., 2006; Maeder et al., 2008; Wahl et al., 2009]. Besides, hPrp8 can be crosslinked to the 5'SS, BPS and the 3'SS during B complex formation indicating a potential role in the spliceosomal active site [Grainger and Beggs, 2005].

1.4.3.4 U4/U6 di-snRNP

U6 is an unusual snRNP, since its highly conserved snRNA is transcribed by RNA polymerase III instead of RNA polymerase II, which transcribes all the others snRNAs present in the spliceosome. This 106 nt long RNA forms, after activation and remodelling, a stem-loop which is almost identical to the the catalytic domain V in group II introns (Fig.1.2 A). As in the group II introns, the U6 snRNA stem-loop binds a divalent cation. Furthermore, the U6 snRNP is thought to be assembled entirely in the nucleus and, as previously mentioned, it is associated with a set of seven LSm proteins that form a ring structure around the RNA at the 3'end [Will and Lührmann, 2001; Patel and Bellini, 2008; Wahl et al., 2009]. After activation

of the spliceosome, U6 snRNA interacts with the 5'SS and extensively base pairs with its counterpart in the U2 snRNP. Together, both snRNAs actively participate in the intron excision [Wahl et al., 2009]. The presence of the U4 snRNA in the duplex has a regulatory purpose, preventing premature catalytic core formation. U4 snRNP contains the U4 snRNA (of 142 nt in humans) and the heptameric ring of Sm proteins. Additionally, human 13S U4/U6 particle contains five specific proteins: hPrp3 (90K), hPrp31 (61K), hPrp4 (60K), CypH (20K) and 15.5K (hSnu13) [Will and Lührmann, 2001; Patel and Bellini, 2008]. 15.5K has a special function, it binds directly to the U4 snRNA 5' stem-loop and facilitates binding of other U4 associated proteins [Nottrott et al., 2002]. hPrp31 binds to the U4 core/15.5K complex and it is crucial for tri-snRNP formation [Liu et al., 2011]. The molecule forms a bridge that stabilizes the trimeric particle by means of interaction with hPrp6 [Makarov et al., 2000; Liu et al., 2007]. The rest of the proteins, hPrp3, hPrp4 and CypH interact with each other forming a stable complex that is recruited by the U4/U6 di-snRNP and only hPrp3 directly contacts the U6 snRNA. This heteromeric complex has been related to stabilization of the stem II of the U4/U6 snRNA duplex [Nottrott et al., 2002].

1.4.3.5 U4/U6.U5 tri-snRNP

In the 25S trimeric U4/U6.U5 particle, apart from the specific proteins present in each of the individual snRNPs, three new proteins are recruited: hSnu66 (110K), hSad1 (65K), and 27K. All proteins of the 25S tri-snRNP have been highly conserved during evolution and almost all orthologs are present in the yeast *Saccharomyces cerevisiae*, with the exception of the 27K, 40K and CypH proteins. Also, one of the members of the U5 snRNP, hLin1 (52K), leaves the particle upon tri-snRNP assembly. Yeast two-hybrid assays were employed to identify the binding partners of the three newly recruited proteins. Full-length hSnu66 was found to interact with the U4/U6 specific protein hPrp3 and also with the full length U5 specific proteins, hPrp6 and hBrr2 [Liu et al., 2007].

1.5 The minor spliceosome

The minor or U12 spliceosome, as the U2 spliceosome, is composed of five snRNPs U11, U12, U4atac, U6atac and U5. The latter is the only snRNP shared by both types of spliceosomes.

1.5.1 Minor spliceosome assembly

The minor spliceosome, as the major counterpart, assembles stepwise and in a highly ordered manner on the pre-mRNA. In general terms, the assembly of the U12-dependent spliceosome is similar to that of the U2-dependent spliceosome and presents the main differences at the earliest stage.

The first particle bound to the pre-mRNA is the preformed and highly stable U11/U12 di-snRNP, which binds cooperatively to the 5'SS and BPS. Therefore, the earliest step of assembly resembles complex A, whereas the complex E does not exist. The next step is the binding of U4atac/U6atac and U5 snRNPs as a preformed tri-snRNP complex; this interaction generates the complex B. Afterwards, during activation of the minor spliceosome, U11 and U4atac snRNPs are destabilized and released, mimicking the sequence of events observed for the major spliceosome activation. The remaining steps appear to mirror those of the major spliceosome, with the generation of a C-like complex after the first transesterification reaction and subsequent dissociation of the minor spliceosome once the second transesterification comes to an end [Will and Lührmann, 2005].

1.5.2 Minor spliceosome snRNP particles

1.5.2.1 U11/U12 di-snRNP

The U11 and U12 particles are involved in the recognition of the consensus sequences of the U12 type of introns in the first step of splicing. Furthermore, the U12 snRNA, after spliceosome activation, plays the same catalytic role observed for U2 snRNA. In general, the snRNAs of the minor U11 and U12 particles are completely unrelated to their analogues U1 and U2, but they present similar structural folds [Tarn and Steitz, 1996; Will and Lührmann, 2005]. Moreover, analysis of the protein composition of affinity purified human 18S U11/U12 di-snRNPs and 12S U11 mono particle by matrix-assisted laser desorption/ionization (MALDI) mass spectrometry (MS) showed that the protein content of the particles is similar to the one of the major counterparts, containing the conserved Sm heptameric ring and, in the case of U12, the SF3b complex. However, some proteins are missing, like all the U1 related proteins and the SF3a complex of U2. Additionally, seven new proteins were identified for the di-snRNP with apparent molecular masses of 20, 25, 31, 35, 48, 59 and 65 kDa. From these seven proteins, four were also present in the U11 mono particle (25K, 35K, 48K and 59K) [Will, 2004]. Nevertheless, the precise function of these proteins has not yet been established.

Although the U11 and U12 snRNPs can be found in the cell as individual particles (12S U11 and 15S U12), they are recruited as a di-snRNP during spliceosome assembly.

1.5.2.2 U4atac/U6atac di-snRNP

The U4atac/U6atac di-snRNP shares many similarities with the equivalent particle in the major spliceosome, U4/U6 di-snRNP. The U4atac and U6atac snRNAs undergo base pairing with each other like the major snRNAs U4 and U6. Although the overall sequence conservation is not very high between these minor and major snRNAs (ca. 40%), the structures of the U4atac/U6atac and U4/U6 snRNA duplexes are very similar [Tarn and Steitz, 1996]. The protein composition of the di-snRNP and the tri-snRNP (after U5 recruitment) is also largely conserved as shown by immunoprecipitation of the snRNPs with antibodies specific for the U4/U6.U5 proteins [Schneider et al., 2002], but a comprehensive analysis has still not been performed [Will, 2004]. In terms of their function, U4atac and U6atac particles play the same role as the one described for U4 and U6.

1.6 Non-snRNP factors and spliceosomal protein dynamics

Not only snRNPs are recruited in this catalytic process, multiple non-snRNP protein factors join the spliceosome during the splicing cycle. Mass spectrometric analyses conducted on *S. cerevisiae* spliceosomes showed that the spliceosome undergoes dramatic compositional changes upon assembly and activation, recruiting and releasing a broad number of proteins in each step (Fig.1.8) [Fabrizio et al., 2009].

The inventory of non-snRNP factors includes the already mentioned hnRNPs and SR proteins that can act as inhibitors and activators of the spliceosome activity, respectively, upon binding on specific sequences [Wahl et al., 2009]. Other examples are:
- The Prp19/CDC5 complex (NTC in yeast), which is required to stabilize binding of U5 and U6 snRNPs to the spliceosome after U4 dissociation [Chan et al., 2003; Makarova et al., 2004; Chan and Cheng, 2005].
- The RES (retention and splicing) complex is important for splicing efficiency and the retention of unspliced pre-mRNA in the nucleus [Dziembowski et al., 2004].
- The members of the exon junction complex (EJC) which are involved in the export and processing of mature mRNA [Tange et al., 2005].

- The SF2 RNA helicases (for more details see section 1.7) and peptidyl-prolyl cis/trans isomerases (PPIases) are responsible for RNP conformational changes.

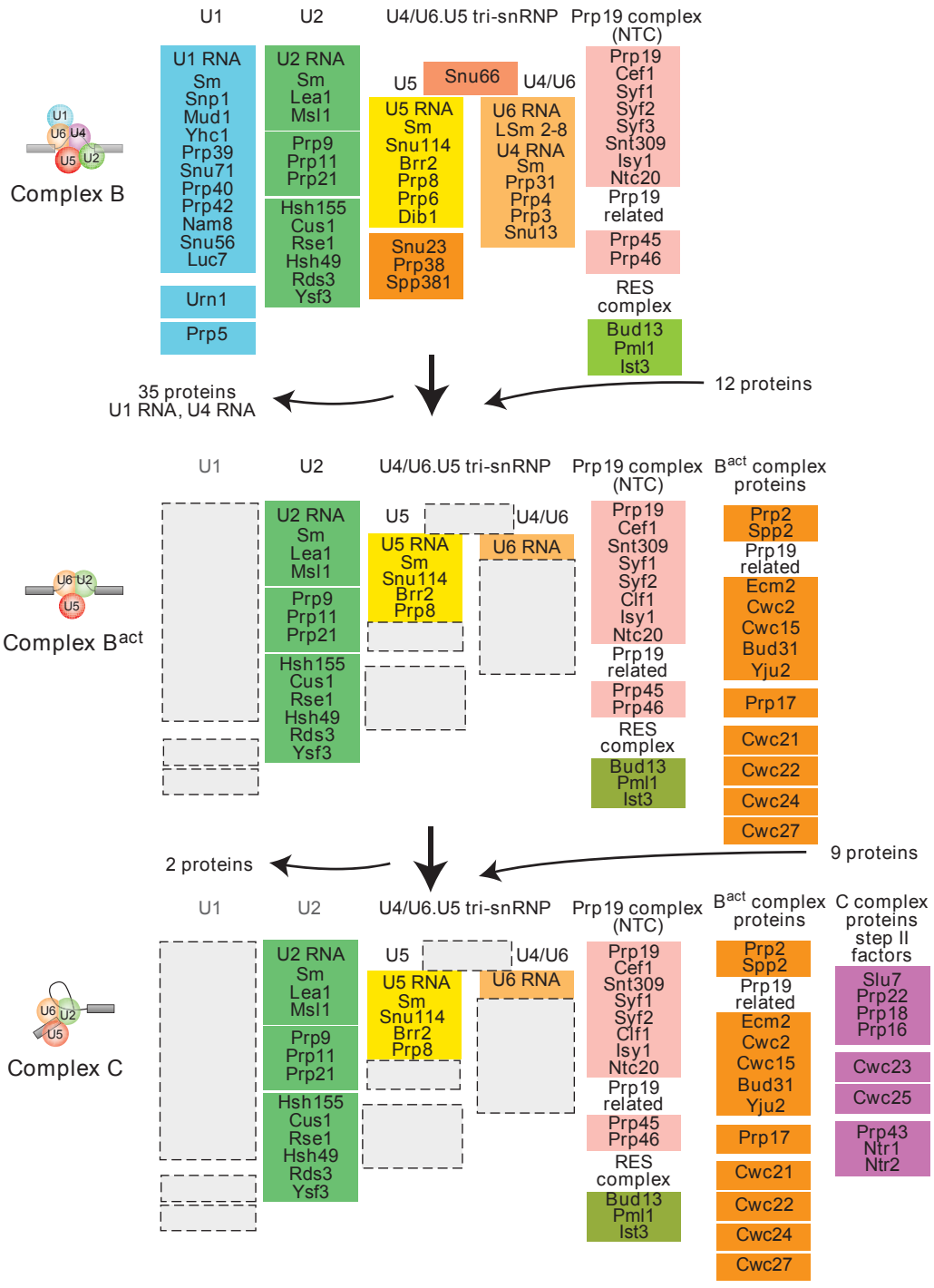

Figure 1.8: Protein dynamics during spliceosome activation and catalytic steps of the *S. cerevisiae* spliceosome. Proteins identified by mass spectrometry in yeast B, B^{act}, and C spliceosomal complexes are shown. Proteins are grouped according to their function or association with an snRNP, protein complex or spliceosomal complex [Fabrizio et al., 2009].

1.7 Spliceosomal helicases, their activity and regulation

RNA helicases are central players in splicing being considered the main driving forces in the dynamic remodeling of RNA-RNA, RNA-protein and protein-protein interactions [Staley and Guthrie, 1998b]. Eight conserved members of the SF2 superfamily of helicases were identified at various steps of the splicing cycle: Sub2/UAP56, Prp5, U5-100K/Prp28, U5-200K/Brr2, Prp2, Prp16, Prp22, and Prp43. All these enzymes present an RNA dependent ATPase activity (Fig.1.9).

The spliceosomal RNA helicases can be divided into two groups, RNA helicases transiently associated with the spliceosome (majority) and the constitutive members of the snRNPs.

1.7.1 Transiently associated helicases

There are six transiently associated helicases. Sub2/UAP56 and Prp5 are members of the DEAD-box family. Both are recruited after complex E formation and are involved in the rearrangements required for the transition to complex A. Sub2/UAP56 seems to facilitate U2 association by displacing U2AF and/or BBP from the BPS. It is also an essential factor in mRNA export [Strässer and Hurt, 2001; Luo et al., 2001] and it has been recently postulated that it contributes to pre-mRNA retention in the nucleus [Takemura et al., 2011]. In terms of its activity, UAP56 has an RNA-stimulated ATPase activity and it can only hydrolyze ATP but not other nucleoside-triphosphate (NTP). UAP56 is a weak ATPase (kcat = 0.25/min) and it does not bind RNA in a sequence dependent manner. Nevertheless, UAP56 is able to unwind short double stranded (ds) RNA with blunt ends, 5' overhangs or 3' overhangs with similar efficiency. The weak ATPase and helicase activity of UAP56 raises the possibility that other protein factors may modulate its activity. However, two of its known interaction partners (U2AF65 and Aly) do not have any influence on UAP56's ATPase and helicase activity [Shen et al., 2007].

Prp5 has been implicated in the U2 snRNP remodeling and U2/BPS stability check [Kistler and Guthrie, 2001; Xu and Query, 2007]. The ATPase activity of Prp5 is stimulated by RNA and it shows a strong binding preference for the U2 snRNA compared to the other spliceosomal snRNAs [O'Day et al., 1996]. No helicase activity has been observed so far. [O'Day et al., 1996; Kosowski et al., 2009]. Prp5 is required for the conformational switch between two mutually exclusive forms of U2

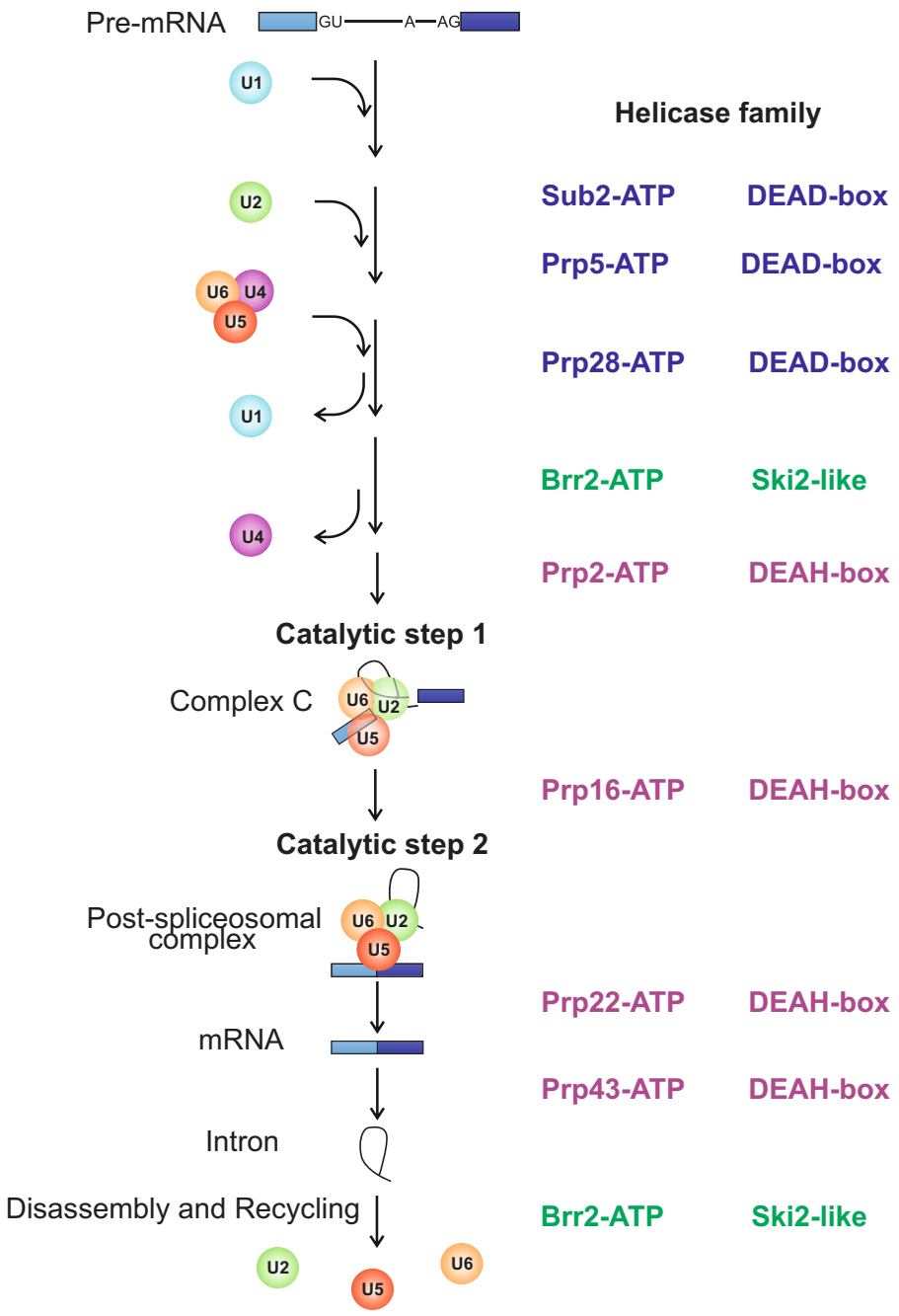

Figure 1.9: Helicases participating in spliceosomal RNA rearrangement. The five snRNPs are represented as coloured circles. RNA helicases facilitate conformational transitions that allow assembly and activation of the spliceosome, catalysis, product release and snRNP recycling. Spliceosomal helicases can be assigned to the DEAD-box, DEAH-box or Ski2-like helicase families (typeset in blue, purple and green, respectively). The DEAH-box helicases are not specific for ATP and are able to bind any nucleoside-triphosphate (NTP) [Fairman-Williams et al., 2010], but in the case of spliceosomal helicases, they have only been associated to ATP hydrolysis.

snRNA, stem-loop IIC and stem IIA, but the exact mechanism and regulation is not clear. However, it has been postulated that Prp5 may promote the displacement of a protein called Cus2 which, when bound to U2 snRNA, stabilizes the stem-loop IIC

conformation. To trigger this conformational exchange the helicase requires ATP [Perriman et al., 2003; Perriman and Ares, 2007]. Moreover, Prp5 seems to be present along the whole splicing cycle and several additional roles have been attributed to the molecule (for more details see [Perriman et al., 2003; Kosowski et al., 2009]).

The DEAH-box protein Prp2 plays a role before the first transesterification reaction by promoting a poorly understood rearrangement that converts B^{act} into B^{*} complex [Kim and Lin, 1996]. Prp2 is an RNA-dependent ATPase but does not exhibit measurable helicase activity on a test duplex [Kim et al., 1992]. It has been suggested that the activity of Prp2 may be regulated by the action of protein cofactors, as has been described for other helicases of the SF2 superfamily (i.e. eukaryotic translation initiation factor 4A, eIF4A, and the hepatitis C virus helicase NS3 [Silverman et al., 2003]). The Spp2 protein was shown to interact with Prp2 and this interaction is important for splicing and cell viability. The specific way of action of Spp2 is unknown, but a domain of Spp2 called G-patch is required for Prp2 binding [Silverman et al., 2004].

Prp16 is also a DEAH-box helicase and takes part in the conformational remodeling necessary for the second catalytic step, but the mechanistic details remain obscure. Nevertheless, genetic studies in yeast showed that Prp16 interacts with Prp8, U6 snRNA [Query and Konarska, 2004] and Isy1 [Villa and Guthrie, 2005] during this stage of the splicing cycle. Moreover, its activity is linked to the release of two factors, Yju2 and Cwc25 [Tseng et al., 2011]. This molecule also promotes the elimination of aberrant lariat intermediates, contributing to the fidelity of pre-mRNA splicing [Burgess and Guthrie, 1993]. Prp16 is an RNA-dependent NTPase [Schwer and Guthrie, 1991], which is able to unwind RNA duplexes *in vitro* and its unwinding activity seems to be non-sequence-specific [Wang et al., 1998]. However, Prp16 helicase activity is weak and not very efficient compared to other DNA unwindases [Lohman and Bjornson, 1996] and may not be relevant for its function.

The Prp22 helicase is recruited during the second catalytic step, where it functions consecutive to Prp16 in an ATP-independent manner. Afterwards, Prp22 is required again for the release of the mRNA product from the spliceosome, most likely, because it disrupts mRNA/U5 snRNP contacts. This reaction requires ATP [Schwer and Gross, 1998; Schwer, 2008]. Prp22 has also been associated with a fidelity check of exon ligation during the second step of splicing [Mayas et al., 2006]. Prp22 can unwind duplex RNA that have 3' overhangs in an ATP-dependent fashion and RNA stimulates the ATPase activity of the helicase [Schwer and Gross, 1998; Tanaka and Schwer, 2005]. Additionally, the RNA-stimulated ATP hydrolysis is affected by chain length, and optimal activity requires RNA oligomers of $\geq$20 nt [Tanaka and Schwer,

2005]. It has been found that more than 13 nucleotides downstream of the splice junction are necessary for mRNA release and Prp22 interacts with this segment in the mRNA, suggesting that the region downstream of the exon-exon junction is a cofactor for Prp22 action within the spliceosome [Tanaka and Schwer, 2005; Schwer, 2008].

The DEAH-box RNA helicase that participates last in the splicing cycle is Prp43, which has also been involved in ribosome biogenesis [Leeds et al., 2006; Combs et al., 2006]. Prp43, together with Ntr1 and Ntr2, forms the NTR complex that is implicated in spliceosome disassembly [Tsai et al., 2005]. The recruitment of the complex is probably mediated by the Ntr2 protein through its interaction with Brr2 and U5 snRNP [Tsai et al., 2007]. Ntr1, on the other hand, enhances Prp43 activity and, therefore, stimulates the disruption of RNA-RNA or RNA-protein contacts leading to intron lariat dissociation [Tanaka et al., 2007]. In general, Prp43 has a rather poor helicase activity and preferentially unwinds substrates with single-stranded tails in 5' to 3' direction *in vitro* [Tanaka and Schwer, 2006]. However, addition of Ntr1 increases Prp43 activity and allows it to unwind model substrates both in 5' to 3' and in 3' to 5' direction [Tanaka et al., 2007]. During ribosome biogenesis, Prp43 activity is also regulated by another protein, Pfa1 [Lebaron et al., 2005, 2009]. Interestingly, the two proteins have a G-patch domain in common which has been also observed in the Prp2 co-factor Spp2 [Silverman et al., 2004].

1.7.2 Constitutive helicases of snRNP particles

Prp28 and Brr2 are the only helicases in the spliceosome constitutively associated with one of the snRNPs (U5). Prp28 is a member of the DEAD-box family of proteins and it displays RNA-dependent ATPase activity [Strauss and Guthrie, 1994]. Prp28 promotes the exchange of U1 for U6 at the 5'SS in an ATP-dependent manner during the transition from B to B^{act} complexes [Staley et al., 1999]. However, Prp28 showed no detectable helicase activity *in vitro* so far and it may require a cofactor to activate it. Alternatively, it may use another strategy, like displacement of U1C, to destabilize the U1-5'SS interaction or to strengthen the U6-5'SS interaction [Staley et al., 1999]. In humans, Prp28 association with the tri-snRNP and the spliceosome is regulated by phosphorylation [Mathew et al., 2008]. This post-translational modification in Prp28 is mediated by the kinase SRPK2 and its target is the RS domain (domain rich in alternating residues of mixed charge, Arg-Asp, Arg-Glu and Arg-Ser dipeptides) of the helicase. RS domains are known to have a role in stabilizing the binding of snRNPs to their pre-mRNA binding sites by engaging in protein-protein interactions [Kohtz et al., 1994; Pasman et al., 1995]. The phosphorylation of hPrp28's RS

domain has been shown to be crucial to ensure its interaction with the spliceosome [Mathew et al., 2008].

Brr2 is the central focus of this study, thus it will be discussed in more detail in the next section.

1.8 Brr2

Brr2 stands for Bad Response to Refrigeration factor 2 and it was discovered in 1996 by analysis of genes that, upon mutation, generated cold sensitive strains in *S. cerevisiae* and were associated with defects in pre-mRNA splicing [Noble and Guthrie, 1996]. Shortly after the discovery of the yeast protein, the human ortholog was also identified in HeLa cells as an intrinsic component of the U5 snRNP with an apparent molecular weight of 200 kDa and therefore termed U5-200K [Laggerbauer et al., 1996].

1.8.1 Brr2 in the Ski2-like family context

Brr2 belongs to the Ski2-like family of helicases, another division of SF2, named after the founding member Ski2 [Fairman-Williams et al., 2010]. Formerly, the Ski2-like and NS3/NPH-II families were grouped together in the DExH box family (based on the composition of the functional motif II, see below) [Jankowsky and Bowers, 2006], but in the last years additional structural information was acquired allowing a more specific classification [Bleichert and Baserga, 2007].

To date, four Ski2-like RNA helicases have been identified in *S. cerevisiae*: Ski2 [Widner and Wickner, 1993], Mtr4 [Searfoss and Wickner, 2000], Brr2 [Raghunathan et al., 1998] and Slh1 [Martegani et al., 1997]. Human homologs have been observed for each of these proteins. In addition, this family includes several DNA helicases like HEL308 [Büttner et al., 2007], Mer3/HFM1 [Nakagawa and Kolodner, 2002] and the ASCC3 subunit of the activating signal cointegrator complex (ASCC) [Dango et al., 2011]. These helicases are involved in diverse processes in the cell, including pre-mRNA splicing, nuclear and cytoplasmic RNA decay, DNA repair and genome maintenance.

The members of the Ski2-like family present a conserved domain composition that includes a variable N-terminus, two RecA domains followed by a winged helix domain that connects to a helical bundle/ratchet domain (HB), a helix-loop-helix domain

(HLH) and, in several of the helicases of this family, an Immunoglobulin/fibronectin-like domain (IG) at the C terminus of the protein. The last three domains, when present, form the Sec63 homology unit, named by its resemblance to a portion of the Sec63p subunit of the protein translocation apparatus of the endoplasmic reticulum [Ponting, 2000]. In all known eukaryotic helicases of this family, the two RecA, the WH and HB domains are present, but just a few have a complete Sec63 homolgy unit.

Brr2 contains all the domains already mentioned, including the full Sec63 unit, but it exhibits an uncommon architecture with two consecutive helicase units (herein referred to as "cassettes") that have been only observed in two other members of this family, Slh1 [Martegani et al., 1997] and ASCC3 [Dango et al., 2011] (Fig.1.10 A).

1.8.1.1 N-terminus

The N-terminus of Brr2 is expected to be mainly unstructured. However, recent systematic structural bioinformatics analysis predicted a PWI-like fold in the region between residues 258-338 of the human protein. A similar domain has been also predicted in the N-terminus of other spliceosomal helicases, like Prp2 and Prp22. This motif seems to be related to nucleic acid binding [Korneta et al., 2012].

1.8.1.2 RecA domains

The RecA domains form the helicase core and are common to all helicases of the SF1 and SF2 superfamilies, containing the NTP and RNA binding motifs. The ATP in the Ski2-like RNA unwindases is bound by an adenine recognition motif, the Q-motif, that interacts specifically with the adenine base. Additionally, the ribose of the ATP is bound by residues from motif V and motif VI. Residues from motifs I, II, V and VI also contact the phosphate chain of the ATP and a coordinated magnesium ion. The RNA substrate, on the other hand, is recognized by residues located in the conserved motifs Ia, Ib and Ic in the RecA-1 domain, and in motifs IV and Vb in the RecA-2 domain. These residues interact directly with the phosphate backbone of single-stranded oligonucleotides. Additional structural features may also contribute to the regulation of RNA binding and assist unwinding, like a long β-hairpin (5'HP) located between motifs Vb and VI. The 5'HP is proposed to control the access to the RNA binding groove in a nucleotide dependent manner in DEAH-box helicases [He et al., 2010]. However, the 5'HP in Ski2-like helicases is shorter and, therefore, another function was proposed, the separation of the duplex strands with the 5'HP

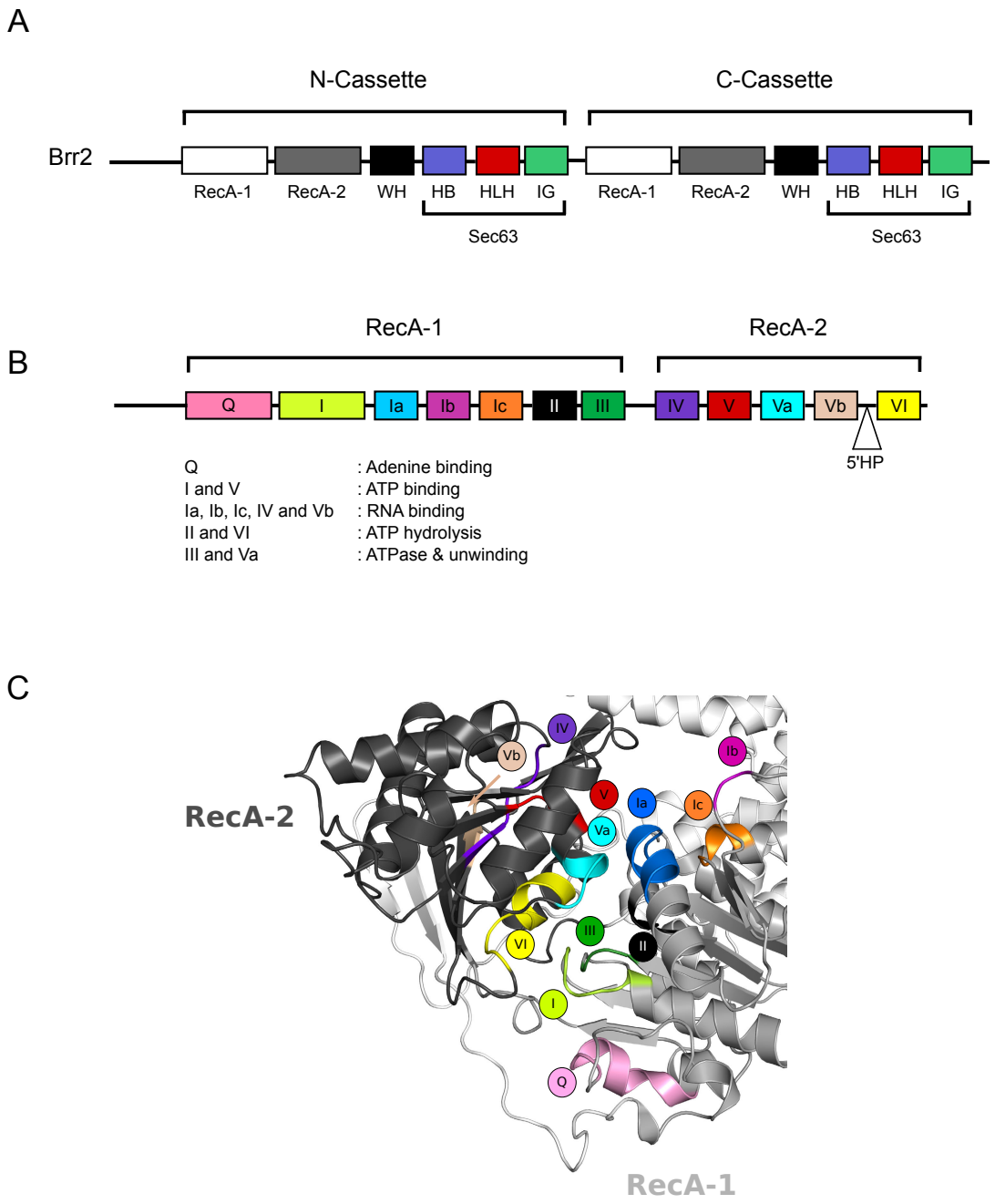

Figure 1.10: **(A)** Domain organization of Brr2. Domains are represented as boxes and not to scale. RecA-1 – light gray, RecA-2 – dark gray, WH – black, HB – blue, HLH – red, IG – green ; **(B)** Motif organization of RecA domains in Ski2-like helicases [Fairman-Williams et al., 2010]; **(C)** Close-up view of the RecA domains in the N-terminal cassette of Brr2. Conserved sequence motifs are colored as in panel (B).

acting as a wedge when one strand of the RNA substrate is pulled by the HB [Büttner et al., 2007; Woodman et al., 2007; Zhang et al., 2009; Pena et al., 2009]. The motifs III and Va are implicated in the coordination of the ATPase and unwinding activities [Büttner et al., 2007; Fairman-Williams et al., 2010; Cordin et al., 2012] (Fig.1.10 B).

Brr2 encompasses a duplicate of the helicase unit, *ergo* four RecA domains with the potential to bind the corresponding substrates, RNA and ATP. The C-terminal cassette exhibits significant deviations in the putative helicase motifs in comparison to the canonical motifs. In motif II, typically represented by DExD/H, the catalytic glutamate is replaced by an aspartate (DDAH) and thus the ability to hydrolyze ATP seems to be hindered in this cassette [Santos et al., 2012]. However, the binding of the nucleotide is still possible [Santos et al., 2012]. Motif III also presents differences, instead of the triad Ser-Ala-Thr (SAT), which has been associated with interdomain interactions of the RecA domains upon ATP and RNA binding, the two last residues have been exchanged giving rise to the array SNC or SSS, in yeast and human sequences, respectively. Furthermore, the C-terminal cassette lacks obvious motifs IV and V, which are involved in RNA and ATP binding [Santos, 2012]. Together, these non-canonical residues render the C-terminal cassette catalytically inactive. The N-terminal cassette, on the other hand, entails all the canonical motifs and it is catalytically active (section 1.8.2) [Santos et al., 2012].

1.8.1.3 Sec63 unit

In the last years, several new structures of members of the Ski2-like family have been unveiled, contributing to the knowledge of domain packing, intramolecular interactions and substrate binding. For example, the crystal structure of the DNA helicase Hel308 from *Archaeoglobus fulgidus* in complex with a partially unwound DNA-duplex at 3.1 Å resolution showed that the HLH domain accommodates the 3' single-stranded overhang of the substrate and it seems to play an important role in the coupling of ATPase and helicase activities [Büttner et al., 2007; Woodman et al., 2007]. It was also suggested that the helical bundle (HB) may act as a ratchet for 3'→ 5' directional transport of the single-stranded product in an ATP dependent manner [Büttner et al., 2007]. The Sec63 unit of Brr2 exhibits a similar fold compared to Hel308 HB-HLH domains, as shown in the crystal structures of the C-terminal Sec63 unit alone [Pena et al., 2009; Zhang et al., 2009] and the helical region ($Brr2^{HR}$, containing both cassettes) [Santos et al., 2012], therefore it may behave likewise in presence of substrate. Consistent with this observation, mutations of positively charged residues (that are not involved in intramolecular interactions) in the HLH domain of Brr2 reduce unwinding activity and enhance ATPase activity [Pena et al., 2009; Santos, 2012]. Similar effects were observed in mutational analysis conducted on the HLH of the Hel308 helicase [Woodman et al., 2007]. The last IG domain is not present in all helicases and little is known about its function, but it is presumed to work as a scaffold for protein-protein interactions.

1.8.2 Brr2 activity and regulation

Brr2 is an essential RNA helicase of the spliceosome. It has been observed that this enzyme stays bound to the catalytic core of the spliceosome throughout splicing and is required at least twice during the process. First, it enters the spliceosome as an intrinsic component of the U4/U6.U5 tri-snRNP and carries out the U4/U6 duplex unwinding, which will lead to the activation of the complex B [Laggerbauer et al., 1998; Raghunathan et al., 1998]. Once the U4 snRNA leaves the spliceosome, Brr2 should be inhibited to avoid unspecific unwinding of other RNA duplexes. In addition, it has been proposed that this helicase may act during spliceosome disassembly [Small et al., 2006]. However, Brr2 most likely does not function as an helicase during this process, because Brr2 hydrolyzes exclusively ATP upon RNA binding and it has been shown that spliceosome disassembly can be carried out using any NTP as energy source [Fourmann et al., 2013].

The helicase and ATPase activities of Brr2 lie in the N-terminal cassette, as shown in recent experiments conducted on the individual cassettes [Santos et al., 2012]. The C-terminal cassette proved to be inactive, as suggested in previous studies [Kim and Rossi, 1999]; it also seems to be unable to bind the RNA substrate, observation supported by modeling and mutational analysis [Pena et al., 2009; Santos et al., 2012]. Nevertheless, the C-terminal cassette is still important for Brr2 activity. The N-terminal cassette alone exhibits a ten fold lower helicase activity than the enzyme containing both helicase units [Santos et al., 2012]. Mutations on residues involved in inter-cassette interactions (not related to the ATP and RNA binding motifs) also led to altered ATPase activity and reduction or abrogation of the helicase activity, corroborating the importance of the communication between both cassettes [Santos et al., 2012]. Moreover, mutations in the linker and the ATP binding site at the C-terminal cassette also affect Brr2 catalytic activities. The information gathered in all these studies converge to present the C-terminal cassette as an intramolecular cofactor which can directly affect the N-terminal cassette and, by means of interaction with regulatory factors, modulate the enzyme activities [Santos et al., 2012]. Along with these results, previous yeast two-hybrid and co-immunoprecipitation assays showed that the C-terminal cassette interacts with several splicing factors (Prp2, Prp8, Prp16, Snp1, Slu7, Snu66, Snu114 and Ntr2) and suggested a function for this cassette as protein-protein interaction platform [van Nues and Beggs, 2001; Liu et al., 2006].

In terms of RNA unwinding, several members of the Ski2-like family exhibit confirmed 3'$\rightarrow$ 5' unwinding polarity, therefore, the same is expected for Brr2. It has also been observed that these helicases require duplex substrates with a single-

stranded overhang at the 3' terminus. In the spliceosome context, the U4/U6 duplex is the first substrate of Brr2 and it has been postulated that the enzyme binds onto the single-stranded region of U4 snRNA located downstream of stem 1 (Fig. 1.11). Then it continues the translocation on U4, unwinding primarily the U4/U6 stem 1 of the RNA duplex [Hahn et al., 2012; Mozaffari-Jovin et al., 2012]. However, the enzyme is a poor helicase [Laggerbauer et al., 1998; Raghunathan et al., 1998; Maeder et al., 2008] and U4/U6 is the most stable duplex in the spliceosome [Guthrie and Patterson, 1988]. Thus, Brr2 most probably requires an activator to stimulate its unwinding activity. On the other hand, despite the low efficiency of the enzyme, Brr2 encounters its substrate very early at the tri-snRNP and to avoid premature unwinding it has to be inhibited. A recent study showed that this inhibition might be partially accomplished by the RNaseH-like domain of Prp8 which competes with Brr2 for the same binding site on the U4/U6 RNA duplex [Mozaffari-Jovin et al., 2012].

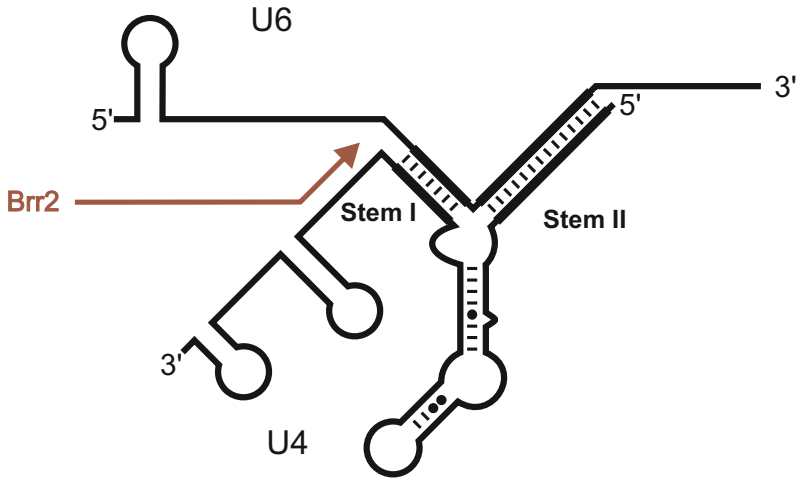

Figure 1.11: Schematic representation of U4/U6 snRNA duplex, brown arrow indicates the binding site of Brr2

The factors Snu114 and Prp8 have been shown as clear modulators of Brr2 activity. Both molecules are constitutive members of U5 snRNP and, in human, they were isolated forming an stable complex with hBrr2 and 40K [Achsel et al., 1998]. Additionally, Snu114 and Prp8 showed a strong physical interaction [van Nues and Beggs, 2001] and the assembly of Snu114 into U5 snRNP requires the presence of Prp8 and an intact GTP binding domain [Brenner and Guthrie, 2006]. Snu114 is the only GTPase encountered in the spliceosome. It was shown that this enzyme influences U4/U6 unwinding [Bartels et al., 2002]. The function of the protein seems to require GTP binding, but not its hydrolysis and it was suggested that it modulates U4/U6 unwinding through Brr2 activity regulation [Small et al., 2006]. This regulation might be carried out by direct interaction with Brr2 or indirectly

through Prp8 [Small et al., 2006].

Prp8 directly interacts with Brr2 and has been shown to influence its activity [Kuhn et al., 2002; Maeder et al., 2008]. Genetic studies suggested that Prp8 may act as a negative regulator of Brr2 activity [Kuhn et al., 2002]. On the other hand, some *in vitro* analysis, conducted with purified full lenght Brr2 and a C-terminal fragment of Prp8 (residues 1806-2413), showed that the latter stimulates Brr2 helicase activity, but inhibits the RNA-dependent ATPase activity [Maeder et al., 2008]. Based on these results, it is possible to speculate on a dual regulatory effect of Prp8 on Brr2's activity and that the interaction between these factors might also be further regulated. In line with this hypothesis, it has been observed that Prp8's very C-terminus is able to bind conjugated ubiquitin and that ubiquitin represses U4/U6 unwinding suggesting that this interaction may influence Brr2 activity [Bellare et al., 2005]. Prp8 is the interaction partner of interest in this thesis and, therefore, will be addressed in more detail in section 1.9.

1.8.3 Brr2 structure

Recently, the structure of a 200 kDa fragment of human Brr2 was solved at 2.65 Å resolution (Protein Data Bank (PDB) ID code 2Q0Z). This is the first structure of a SF2 helicase encompassing both tandem repeats of helicase cassettes and a N-terminal extension of 60 conserved residues, henceforth referred as helical region (HR) (borders of the deposited structure 403-2136) [Santos et al., 2012]. The structure gives important insights into the cassette interaction and their relative assembly. Regardless of only 27% sequence identity, both cassettes display similar overall architecture and domain distribution forming a ring like structure around a central tunnel, which resembles the one observed in the Hel308 structure (Fig.1.12 B). The interaction between the cassettes is very tight presenting a large contact area of 1200 $Å^2$, where the C-terminal cassette is rotated by 115° and translated by 63 Å relative to the N-terminal cassette within $hBrr2^{HR}$ (Fig.1.12 A). Hence, the N-terminal IG domain is placed between the C-terminal RecA-2 and WH domains. Additionally, the N-terminal RecA-1 domain and WH domains contact the C-terminal RecA-2 domain. This extensive communication surface is in line with the fact that the two cassettes form a functional unit.

The cassettes are covalently connected by a conserved 20 residue linker (residues 1289-1309), which is deeply buried in the interaction surface and lacks exposed loops. Besides, the Brr2 structure contains a N-terminal extension that precedes the RecA-1 domain and surrounds the N-terminal cassette adopting an irregular but well defined

fold. Deletion of this extension leads to helicase activity reduction. Consequently, it has been postulated that this N-terminal extension may be involved in productive domain organization or substrate binding [Santos et al., 2012].

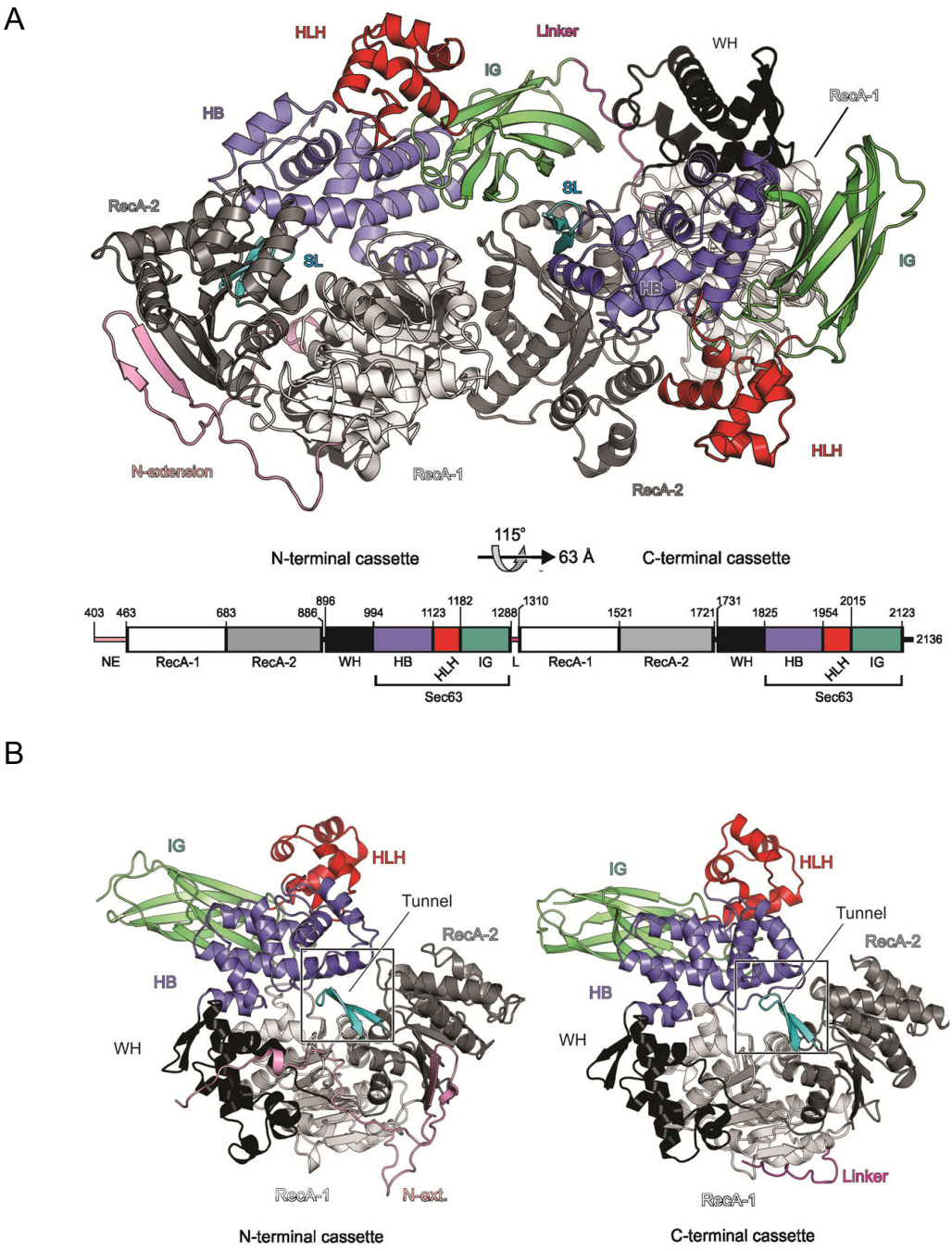

Figure 1.12: Helicase region (HR) of human Brr2. **(A)** Ribbon plot of hBrr2^{HR}. Coloring by domains and functional elements: N-terminal extension – pink, RecA1 – light gray, RecA2 – dark gray, WH – black, HB – blue, HLH – red, IG – green, linker – magenta, separator loop (SL) – cyan. Rotation/translation symbols below the panel indicate the relationship between the cassettes within hBrr2^{HR}; **(B)** Ribbon plots of N-terminal (left) and C-terminal (right) cassettes [Santos et al., 2012].

1.9 Prp8

Prp8 stands for pre-mRNA processing factor 8 and it was first identified as a splicing factor on heat-sensitive *S. cerevisiae* strains generated upon mutation of the corresponding gene (initially RNA8 and renamed afterwards PRP8 gene) [Lustig et al., 1986]. Later, the factor was found to be a stable component of the U5 snRNP and the U4/U6.U5 tri-snRNP [Lossky et al., 1987; Stevens and Abelson, 1999]. After its identification, Prp8 was cloned, sequenced and characterized by immunoprecipitation and *in vitro* splicing reactions, which showed a large protein product of 280 kDa determinant for splicing activity *in vitro* [Jackson et al., 1988].

The human ortholog of Prp8 was first detected in HeLa cells by immunoblotting with the anti-yeast Prp8 antibodies and it was associated with the U5 and U4/U6.U5 particles as well as affinity purified spliceosomes [Anderson et al., 1989; Pinto and Steitz, 1989]. The human protein exhibits a smaller size, of apparent 220 kDa on a denaturating polyacrylamide gel electrophoresis (PAGE) and it was initially named U5-220K, but the molecular weight calculated from its protein sequence is 273 kDa. Prp8 is highly conserved among all eukaryotes. The comparison of the amino acid sequences between yeast and human displayed an overall identity of 61% [Hodges et al., 1995] and the alignment with all confirmed mammalian Prp8 proteins (*Homo sapiens, Mus musculus, Bos taurus* and *Rattus norvegicus*) showed almost no difference, having a mismatch of only two amino acids from human to mouse.

It is noteworthy that Prp8 does not present any enzymatic activity and it has been suggested that it would mainly act as a scaffold for RNA and protein interaction. Therefore its interaction with Brr2 and implications in spliceosome activation are fundamental for a better understanding of this important process.

1.9.1 Prp8 structure and domain distribution

For a long time little was known about the folding and domain distribution of Prp8. Database analysis of the protein sequence were able to reveal some information about conserved functional regions or resemblance to known protein domains, but only two small domains of the C-terminal region had their structures solved, the RNase H-like and Jab1/MPN domains. This year, a structure of a longer fragment of Prp8 (residues 885-2413) has been elucidated, giving new insights into the molecule's composition and offering structural bases to explain previously observed behaviour upon mutation [Galej et al., 2013]. The Prp8 fragment was co-crystallized with full length Aar2 and the crystal diffracted to 1.9Å (PDB ID: 4I43) (Fig.1.13 A) [Galej

et al., 2013]. The structure revealed a new large domain (residues 885-1824) that can be subdivided into a large polymerase/reverse transcriptase (RT)-like domain, a linker domain and a small type II restriction endonuclease-like domain. The structure also shows that the C-terminal domains, RNase H-like and Jab1/MPN, are connected by a disordered linker. Another disordered linker connects the RNase H-like domain to the type II restriction endonuclease-like domain. Nevertheless, the N-terminal region of Prp8 is still missing, thus previous predictions are the only information available so far (Fig.1.13 B).

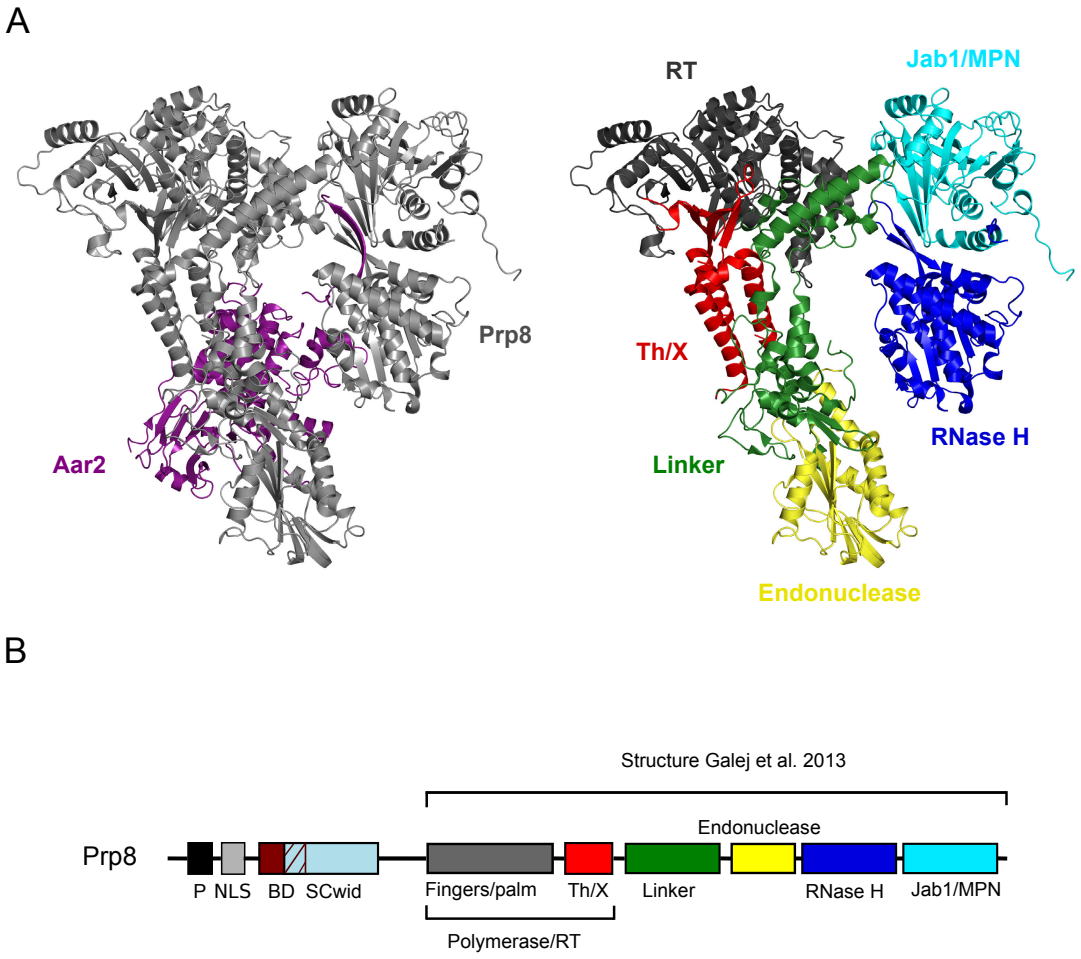

Figure 1.13: **(A)** Three dimensional structure of Prp8 fragment-Aar2 complex (left) of *S. cerevisiae* [Galej et al., 2013] and the same structure without Aar2 and coloured by domain (right); **(B)** Domain organization of Prp8. Domains are represented as boxes and not to scale. P = proline-rich – black, NLS – light gray, BD – brown, SCwid – light blue, RT fingers/palm – dark gray, Thumb/maturase X – red, Linker – green, Endonuclease – yellow, RNase H-like domain – blue and Jab1/MPN domain – cyan [Grainger and Beggs, 2005; Turner, 2006; Dlakic and Mushegian, 2011; Galej et al., 2013]

1.9.1.1 Proline-rich region

A proline-rich region has been observed at the N-terminus of all fungal Prp8 proteins and in some plants, but its presence is not common in most organisms [Grainger and Beggs, 2005].

1.9.1.2 Nuclear localization signal (NLS)

A NLS can be observed within the first 300 residues of the N-terminal region of Prp8 in all the organisms. This motif is generally found as a pair of bipartite sequences separated by a variable stretch of 10-12 amino acids and it was identified using the algorithm of Nakai [Nakai and Horton, 1999]. Studies conducted on yeast revealed that this sequence alone was enough to place the Prp8 molecule into the nucleus [Boon, 2005].

1.9.1.3 Putative bromodomain

A putative bromodomain was identified by means of the bioinformatics tool HHPred, in the region between residues 200 and 315 in yeast [Dlakic and Mushegian, 2011]. Bromodomains are related to acetyl-lysine recognition. Since several spliceosomal proteins are acetylated [Choudhary et al., 2009], it has been postulated that Prp8 assists in the interaction with the so modified proteins.

1.9.1.4 Snu114/Cwc21 interacting domain (SCwid)

The SCwid domain was identified in *S. cerevisiae* by means of yeast two-hybrid, immunoprecipitation and genetic analyses. This region encompasses residues 253-543 and it was shown to interact with fragments of the N-terminal half of Snu114 and with the Cwc21 factor (yeast ortholog of human SRm300), a member of the Niteen complex (NTC) in yeast (Prp19 complex in human) [Grainger et al., 2009].

1.9.1.5 Polymerase/RT-like domain

Residues 950-1220 of yeast Prp8 were previously reported to exhibit a high similarity to the reverse transcriptase (RT) domain encoded by several prokaryotic mobile elements, such as group II introns (P-value = 1.9×10^{-10}) [Dlakic and Mushegian, 2011]. The new structure reveals a polymerase-like fold in this region

(residues 885-1375), but also it shows a high similarity to the RTs of group II introns, therefore referred as polymerase/RT-like domain in this thesis [Galej et al., 2013]. The polymerase/RT-like domain is composed of three canonical subdomains: palm, fingers and thumb. They form a deep cleft similar to the one that accommodates the nucleic acid template and primer in polymerases. However, the residues responsible for the metal coordination required for the polymerase activity are only partially conserved in Prp8 (1 out of three aspartate residues). Therefore, the Prp8 polymerase-like domain is most probably a pseudo enzyme. Residues 1059-1115 were predicted to form an RNA recognition motif (RRM), but this region does not resemble the RRM fold and is embedded in the finger-like subdomain. The region encompassing residues 1257-1375 showed a significant sequence similarity with the thumb/maturase X (Th/X) domain of fungal group II intron RT and it was correctly predicted to form a helical bundle [Dlakic and Mushegian, 2011].

1.9.1.6 Linker domain

The region between residues 1376-1649 was identified in the latest structure as a linker domain that allows an intimate interaction between the polymerase-like domain and the endonuclease-like domain [Galej et al., 2013]. Moreover, this region overlaps with the U5 snRNA interaction region determined by cross-linking and proteolytic mapping within Prp8 [Turner, 2006]. Based on the results of the mentioned study, two discrete regions (770-871 and 1281-1413) of Prp8 were found as contact point for the invariant uridine at position 97 in U5 loop 1 and one of them (residues 1281-1413) lies in the observed linker domain. Interestingly, the 3'SS fidelity region, named after studies on yeast Prp8 mutations that supressed 3'SS splicing defects, also falls into the linker domain. The 3'SS fidelity region was previously identified between residues 1372-1660 in yeast (1300-1588 in human) and was split in two fractions based on sequence conservation. The 3.1 region (residues y1372-1546) is less conserved (39%-68% identity) and contains only two of the known 3'SS fidelity mutations. The 3.2 region (residues y1547-1660), on the other hand, is highly conserved (55%-87% identity) containing almost all the 3'SS fidelity mutations and a consensus site for phosphorylation [Grainger and Beggs, 2005]. Furthermore, a U6 snRNA interaction domain (1503-1673) broadly overlaps with the 3.2 region [Turner, 2006].

1.9.1.7 RNase H-like domain

The RNase H-like domain encloses the region just before the very C-terminal Jab1/MPN domain (residues 1827-2092 in yeast and 1755-2016 in human). Initial

studies based on cross-linking analyses indicated an association between this region and the 5'SS [Reyes et al., 1999]. The ortholog domains of yeast and human were cloned, purified and crystallized at 2 Å and 1.9 Å resolution, respectively [Pena et al., 2008; Ritchie et al., 2008; Yang et al., 2008]. The structures showed a RNase H like fold that resembles a left-hand mitten, where the central six-stranded β-sheet surrounded by N-terminal α-helices forms the palm, an extended β-hairpin corresponds to the thumb and the C-terminal α-helical array represents the fingers. Albeit, the protein is a pseudo enzyme lacking the catalytic motifs and it is unlikely that it binds metal ions. Recent studies showed the interaction of the RNase H domain with the U4/U6 duplex snRNA and it was suggested that this domain competes with Brr2 for the same binding region, the single-stranded stretch of U4 snRNA downstream of stem I [Mozaffari-Jovin et al., 2012].

1.9.1.8 Jab1/MPN domain

Sequence comparison allowed the identification of a MPN (Mpr1-1, Pad-1, N-terminal) domain at the very C-terminus of Prp8 [Grainger and Beggs, 2005]. Afterwards, using the approach of transposon insertion, yeast Prp8 was dissected into four putative indivisible functional units and the one encompassing the C-terminus showed boundaries consistent with the predicted Jab1/MPN borders (2173-2310) [Boon, 2005]. However, the final confirmation was only obtained with the crystal structures of the Jab1/MPN domains of *S. cerevisiae* (2.0 Å) [Pena et al., 2007] and *C. elegans* (2.3 Å) [Zhang et al., 2007]. Both structures include N-terminal and C-terminal extensions which are determinant for domain stability and correct folding. The borders of the yeast fragment are amino acids 2147-2397, lacking the last 16 residues on the C-terminus, because they formed an unstructured acidic tail that hindered crystallization. The *C. elegans* fragment extends from residue 2057 to 2329 and includes the last residues of the C-terminal tail (22 residues), but they are disordered and not visible in the crystal structure. The overall structure exhibits the MPN core (a β barrel flanked by helices on opposite sides), the N-terminal extension, which entails a long helix, and the C-terminal appendix that contains a significant portion of loops. Some MPN domain containing proteins have been associated with de-ubiquitination activity [Maytal-Kivity et al., 2002]. This subset of MPN proteins contains the JAMM motif which is thought to coordinate a Zn^{2+} ion and function as Zn^{2+} dependent isopeptidases to remove ubiquitin or ubiquitin-like molecules from target proteins [Maytal-Kivity et al., 2002]. The MPN domain of Prp8 contains a partial JAMM motif unable to bind Zn^{2+} [Zhang et al., 2009], thus this domain is inactive and most probably acts as a protein-protein interaction platform. Yeast

two-hybrid analyses showed that the C-terminal region of yPrp8 interacts with yBrr2 [van Nues and Beggs, 2001] and hPrp8 (1986-2335) interacts with hBrr2 and hSnu114 [Liu et al., 2006]. It is thought that this region is responsible for the modulation of Brr2 activity [Maeder et al., 2008].

Furthermore, the Jab1/MPN domain harbors all the Prp8 related mutations that cause Retinitis pigmentosa (section 1.10). Nineteen different mutations have been identified to date, including missense, premature stop codon and deletions [McKie et al., 2001; De Erkenez et al., 2002; Kondo et al., 2003; Martinez-Gimeno et al., 2003; Ziviello et al., 2005; Sullivan et al., 2006; Towns et al., 2010]. Almost all mutations are localized in a highly conserved region within the last exon of Prp8 gene (exon 42), with the exception of two mutations which were found in exon 38 and in the exon 41-42 junction, respectively (Table 1.1).

1.10 Retinitis pigmentosa and splicing

Retinitis pigmentosa (RP) is a progressive rod-cone dystrophy that leads to night blindness, gradual reduction of the peripheral visual field with eventual development of tunnel vision and, in some severe cases, to total blindness. The name of the condition is based on the perivascular deposits of melanin with a characteristic bone spicule pattern, which are the result of the retinal pigmented epithelium (RPE) decay [Mordes et al., 2006]. The worldwide incidence of RP is approximately 1 in 4000, leading to more than 1.5 million affected individuals. The disease is genetically diverse and can be inherited as an autosomal-dominant (adRP) (30-40% of the cases), autosomal-recessive (arRP) (50-60%), X-linked trait (xlRP) (5-15%) and, in rare cases, as a non-Mendelian trait [Hartong et al., 2006]. To date, 62 genes have been associated with nonsyndromic RP (RetNet database) [Benaglio et al., 2011] and 23 of them are the source for the adRP forms. As it could be expected, most of the genes are predominantly or specifically expressed in the retina and are determinant for survival and function of the retinal cells [Swaroop and Zack, 2002]. Interestingly, five genes related to adRP code for ubiquitously expressed pre-mRNA splicing factors: hPrp3 (RP18) [Chakarova et al., 2002], hPrp31 (RP11) [Vithana et al., 2001], hPrp8 (RP13) [McKie et al., 2001], hBrr2 (RP33, SNRNP200) [Zhao et al., 2009] and PAP-1 (RP9) [Keen et al., 2002]. The first two are essential for the di-snRNP U4/U6 and tri-snRNP U4/U6.U5 assembly [Nottrott et al., 2002; Weidenhammer et al., 1996, 1997]. Prp8 and Brr2 have important roles in spliceosome activation, RNA rearrangements and splicing regulation. The function of PAP-1 is unknown, but it has been observed that it interacts with splicing factors including Prp3 and U2AF35

Table 1.1: Table summarising all known PRP8 (RP13) RP causing mutations

Mutation	Mutation type	Amino acid change	Reference
c.6353C>T	missense	S2118F	[Towns et al., 2010]
c.6901C>A	missense	P2301T	[McKie et al., 2001]
c.6901C>T	missense	P2301S	[Ziviello et al., 2005]
c.6912C>G	missense	F2304L	[McKie et al., 2001]
c.6926A>C	missense	H2309P	[McKie et al., 2001]
c.6926A>G	missense	H2309R	[McKie et al., 2001]
c.6928A>G	missense	R2310G	[McKie et al., 2001]
c.6929G>A	missense	R2310K	[McKie et al., 2001]
c.6930G>C	missense	R2310S	[Towns et al., 2010]
c.6942C>A	missense	F2314L	[McKie et al., 2001]
c.6943-6944 del C	frame-shift	L2315fsX2358	[Martinez-Gimeno et al., 2003]
c.6974-6994 del 21 bp	frame-shift	V2325fsX2329	[Martinez-Gimeno et al., 2003]
c.6893-6896 del ins 7 bp	frame-shift	L2298fsX2336	[Martinez-Gimeno et al., 2003]
c.6972-6977 del 6 bp ins 11 bp	frame-shift	2325fsX2359	[Kondo et al., 2003]
c.6991 del G	frame-shift	E2331fsX2358	[Sullivan et al., 2006]
Exon junction 41-42G>A	splice site change		[Sullivan et al., 2006]
c.6961CAG>TAG	stop	Q2321X	[De Erkenez et al., 2002]
c.7000T>A	missense	Y2334N	[McKie et al., 2001]
c.7006T>C	frame-shift	Stop2336fsX2377	[Martinez-Gimeno et al., 2003]

[Maita et al., 2004, 2005]. RP9 is a rare disease and the patients carrying these mutations can be asymptomatic [Kim et al., 1995].

In view of the ubiquitous expression of spliceosomal factors, it is intriguing that the mentioned mutations only affect the retinal cells. The explanation which seems to have most support rests on the fact that the retina contains the highest amount of processed pre-mRNA in the whole body and, consistently, the retinal cells express 7-fold more major snRNAs and twice as much of the minor counterparts compared to other human tissues. Thus, the retina may be more sensitive to spliceosomal component deficiency than other tissues [Tanackovic et al., 2011]. Additionally, since

the patients carry heterozygous mutations, the wild type orthologs of the proteins are always co-expressed. Therefore, the effect of the defective molecules is compensated in other cell types [McKie et al., 2001].

The RP mutations related to the Prp8 factor are of great interest for this thesis, because most of them are localized in the C-terminal tail of the Jab1/MPN domain [McKie et al., 2001] that has been implicated in Brr2 interaction [Weber et al., 2011] and may be responsible for Brr2 modulation.

1.11 Structural studies of the spliceosome

Different approaches have been used to gain insights into the mechanism of splicing and the spliceosome's structural organization. Mass spectrometry allowed the characterization of the protein content of intermediate spliceosomal assemblies [Herold et al., 2008; Fabrizio et al., 2009] showing that more than 170 protein factors are associated to the human spliceosome [Will and Lührmann, 2010]. Single-particle cryo-electron microscopy (Cryo-EM) was also an important tool for the elucidation of the overall architecture of the spliceosomal snRNPs. Low to moderate resolution (10-30 Å) 3D structures of the U1 snRNP, the heteromeric protein complex SF3b, the U11/U12 di-snRNP, the U5 snRNP, the U4/U6 di-snRNP and the U4/U6.U5 tri-snRNP have been obtained by EM [Stark et al., 2001; Golas et al., 2003, 2005; Sander et al., 2006; Häcker et al., 2008; Herold et al., 2008; Fabrizio et al., 2009]. However, to have a complete picture of the splicing process, precise information at an atomic level is required. To this end, high-resolution structures were obtained by means of X-ray crystallography or NMR. Nevertheless, due to technical limitations and the dynamic nature of the spliceosome, the existing structures have been restricted to individual snRNPs (i.e. U1 snRNP [Pomeranz Krummel et al., 2009; Weber et al., 2010] and U4 snRNPs [Leung et al., 2011]), spliceosomal proteins (i.e. human 15.5K [Reuter et al., 1999]) and protein fragments (i.e. $Brr2^{HR}$ [Santos et al., 2012] and $Prp8^{Jab1/MPN}$ [Pena et al., 2007]). Many of the solved structures correspond to complexes of spliceosomal proteins with RNA or a protein binding partner (i.e U1A domain RNA complex [Oubridge et al., 1994], U2B"-U2A'-RNA complex [Price et al., 1998] and human p14 in a complex with a fragment of SF3b155 [Schellenberg et al., 2006]). Although the structures of proteins and complexes were not obtained in the spliceosomal context, each model contributed with important information that, complemented with other assays, help to assemble some pieces of the spliceosomal puzzle.

1.12 Aims of the project

Its high complexity and dynamics make it difficult to analyzed the spliceosome as a whole. This thesis deals with events involved in the transition from complex B to B^{act}, which entails dramatic RNA rearrangements that lead to the release of U4 and U1 snRNPs. At this stage, two SF2 helicases are required, Brr2 and Prp28, but only Brr2 and its interaction partner Prp8 will be addressed in this work. Little is known about the molecular basis of this interaction and how Prp8 mediates Brr2 regulation. Thus, the overall aim of this work is to gather information, by means of structural and biochemical analyses, that will allow us to better understand the nature of the Brr2-Prp8 interaction and, consequently, of spliceosome activation. To this end, the following specific goals were proposed:

1. Establishment of efficient recombinant production for human and yeast Prp8 C-terminal fragment (encompassing RNase H and Jab1/MPN domains), $Prp8^{Jab1/MPN}$ and their complexes with Brr2 or fragments of Brr2.

2. Delimitation of the precise interaction regions of Brr2 and Prp8.

3. Establishment of efficient purification protocols for complexes to obtain sufficient amounts of highly homogeneous material for further structural and functional studies.

4. Crystallization and crystal structure analysis of the complexes that can be produced.

5. Investigation of effects of the RP mutants in Prp8 and Brr2 interaction an regulation.

Chapter 2

Materials and methods

2.1 Materials

2.1.1 Chemicals

Standard chemicals that are not listed here were purchased from Sigma-Aldrich, Merck or Fluka.

Table 2.1: List of chemicals

Chemical	Supplier
1 kb DNA ladder, 2-log DNA ladder	New England Biolabs, Germany
1,4-Dithiothreitol (DTT)	Roth, Germany
2-Mercaptoethanol (β-ME)	Roth, Germany
2-Propanol	Merck, Germany
4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES)	Roth, Germany
Acetic acid	Merck, Germany
Acetone	Merck, Germany
Acrylamide solutions:	Roth, Germany
Rotiphorese Gel 40 (38% acrylamide, 2% B N,N'-methylene-bis-acrylamide)	
Rotiphorese Gel 30 (29.2% acrylamide, 0.8% N,N'-methylene-bis-acrylamide)	
Adenosine diphosphate (ADP)	Sigma-Aldrich, Germany
Agarose (electrophoresis grade)	Invitrogen, Germany
Ammonium peroxodisulfate (APS)	Merck, Germany

Continued on next page

Table 2.1 – *Continued from previous page*

Chemical	**Supplier**
Ammonium sulphate	Fluka, Switzerland
Ampicillin, sodium salt	Fluka, Switzerland
Bluo-Gal	Invitrogen, Germany
Boric acid	Merck, Germany
Bovine serum albumin	Sigma-Aldrich, Germany
Bradford assay reagent	Bio-Rad, Germany
Bromophenol blue, sodium salt	Merck, Germany
Butanediol	Sigma-Aldrich, Germany
Calcium chloride dihydrate	Merck, Germany
Casyton (buffer for Casy TT counter)	Roche, Germany
Chloramphenicol	Roche, Germany
Complete-EDTA free protease inhibitor	Roche, Germany
Coomassie brilliant blue G-250	Serva, Germany
Coomassie brilliant blue R-250	Serva, Germany
Dimethylsulfoxide (DMSO)	Roth, Germany
Ethanol	Merck, Germany
Ethidium bromide solution (10mg/ml)	Roth, Germany
Ethylene glycol	Serva, Germany
Ethylenediaminetetraacetic acid, disodium salt dihydrate (EDTA)	Roth, Germany
Express Five SFM	Invitrogen, Germany
Formaldehyde	Merck, Germany
Glycerol	Merck, Germany
Guanidine hydrochloride	Roth, Germany
Hydrochloric acid (HCl)	Merck, Germany
IgepalR CA-630 (NP-40)	Sigma-Aldrich, Germany
Imidazole	Merck, Germany
Isopropyl-β-D-1-thiogalactopyranoside (IPTG)	Roth, Germany
Izit protein crystal stain	Hampton research, USA
Kanamycine sulfate	Roth, Germany
LB medium	Q-Bio-gene, USA
LB-agar medium	Q-Bio-gene, USA
L-glutamine	Invitrogen, Germany
Lithium chloride	Fluka, Switzerland
Magnesium chloride hexahydrate	Fluka, Switzerland
Methanol	Merck, Germany

Continued on next page

Table 2.1 – *Continued from previous page*

Chemical	**Supplier**
Milk powder, dry, instant	Heirler, Germany
N,N,N',N'-Tetraethylenediamide (TEMED)	Sigma-Aldrich, Germany
Phenylmethylsulfonyl fluoride (PMSF)	Roche, Germany
Piperazine-N,N-bis-(2-ethanesulfonic acid) (PIPES)	Roth, Germany
Polyethylene glycol 200-8000	Sigma-Aldrich, Germany
Polyethylene glycol 3350	Sigma-Aldrich, Germany
Ponceau S	Serva, Germany
Potassium chloride	Merck, Germany
Potassium hydroxide	Merck, Germany
Protein molecular weight marker (unstained, pre-stained)	Bio-rad, Germany
Sf-900^{TM} III SFM	Invitrogen, Germany
Silver nitrate	Merck, Germany
Sodium carbonate	Merck, Germany
Sodium chloride	Merck, Germany
Sodium dodecylsulfate (SDS)	Serva, Germany
Sodium hydroxide	Merck, Germany
SYPRO Orange Protein Stain	Invitrogen, USA
Tetracyclin	Fluka, Switzerland
Trifluoroacetic acid	Fluka, Switzerland
Tris-(hydroxymethyl)-aminomethane (TRIS)	VWR International, Germany
Trimethylamine N-oxide dehydrate (TMAO)	Sigma-Aldrich, Germany
Triton X-100	Merck, Germany
Tween-20	Roth, Germany
Urea	Merck, Germany
X-tremeGene 9 Transfection Reagent	Roche, Germany
Xylene cyanol FF	Sigma-Aldrich, Germany
Yeast extract powder	Roth, Germany

2.1.2 Buffer solutions and media components

All buffers and media were prepared with Milli-Q water and filtered (0.22 μm) or autoclaved. The pH was adjusted with 37% or 1 M HCl and 10 M or 1 M NaOH if not stated otherwise. The buffers and solutions used in this work are listed in table 2.2.

Table 2.2: List of buffers

Buffer	Composition
10×TAE	0.4 M Tris pH 8.0, 0.2 M M Acetic Acid, 10 mM EDTA
5×SDS loading buffer	250 mM Tris pH 6.8, 8% (w/v) SDS, 10% (v/v) β-ME, 30% (v/v) Glycerol, 0.02% (w/v), Bromophenol blue
4×agarose gel loading buffer	1×TAE, 30% (v/v) Glycerol, 0.05% (w/v) Xylene cyanol, 0.05% (w/v) Bromophenol blue
10×PBS	1.35 M NaCl, 25 mM KCl, 100 mM NaH_2PO_4 17 mM KH_2PO_4 pH 7.4
10×SDS running buffer	250 mM Tris pH 6.8, 2 mM Glycine, 1% (w/v) SDS
Binding buffer for GST-tagged proteins	50 mM Tris pH 8.0, 300 mM NaCl, 5% glycerol (v/v), 1 mM DTT (added before use)
Binding buffer for His-tagged proteins	50 mM HEPES pH 7.5, 600 mM NaCl, 10% glycerol (v/v), 10 mM Imidazole, 2 mM β-ME (added before use)
Binding buffer heparin low salt	25 mM Tris ph 8.0, 50 mM NaCl, 5% glycerol (v/v); 2 mM β-ME (added before use)
Coomassie staining solution	0.025% (w/v) Coomassie (R250); 0.025% (w/v) Coomassie (G250); 30% (v/v) isopropanol; 7.5% (v/v) acetic acid
Destaining solution I	40% (v/v) ethanol, 20% (v/v) acetic acid
Destaining solution II	10% (v/v) acetic acid
Elution buffer for GST-tagged proteins	50 mM Tris pH 8.0, 300 mM NaCl, 5% glycerol (v/v), 1 mM DTT, 10 mM L-Gluthathione (added before use)
Elution buffer for His-tagged proteins	50 mM HEPES pH 7.5, 600 mM NaCl, 10% glycerol (v/v), 250 mM Imidazole, 2 mM β-ME (added before use)
Elution buffer heparin high salt	25 mM Tris ph 8.0, 1.5 M NaCl, 5% glycerol (v/v); 2 mM β-ME (added before use)
LB medium	1% Peptone, 0.5% yeast extract, 0.5% NaCl
LiCl solution for protein purification	2 M LiCl, 50 mM HEPES pH 7.5
Lysis buffer for *E.coli* cells	50 ml Binding buffer for affinity purification; 0.05% NP-40, 1 tablet protease inhibitor (added before use)
Lysis buffer for insect cells	15 ml Binding buffer for affinity purification per 400 ml culture pellet; 0.05% NP-40, 1 tablet protease inhibitor (added before use)
Resolving gel buffer	1.5 M Tris pH 8.8, 0.4% (w/v) SDS
Stacking gel buffer	0.5 M Tris-HCl pH 6.8, 0.4% (w/v) SDS
TB buffer	10 mM PIPES pH 6.7, 250 mM KCl, 55 mM $MnCl_2$, 15 mM $CaCl_2$
SOB medium	2% (w/v) Tryptone, 0.55% (w/v) yeast extract, 10 mM NaCl, 10 mM KCl, 10 mM $MgCl_2$, 10 mM $MgSO_4$

Continued on next page

Table 2.2 – *Continued from previous page*

Buffer	Composition
SOC medium	2% (w/v) Tryptone, 0.55% (w/v) yeast extract, 10 mM NaCl, 10 mM KCl, 10 mM $MgCl_2$, 10 mM $MgSO_4$, 20 mM Glucose

2.1.3 Consumables

Table 2.3 lists the items used routinely in this thesis.

Table 2.3: Consumables

Item	Supplier
Acupuncture needle	Moxom Medical, Germany
Concentrators (Amicon Ultra)	Millipore, USA
Cover slides (glass, ϕ 22 mm)	Hampton Research, USA
Crystallization plates MRC, 96 well, sitting drop	Molecular Dimensions, UK
Dialysis membranes	Spectra/Por, USA
Disposable plastic cuvettes	Brand, Germany
Durapore filters	Millipore, USA
Electroporation cuvettes	Bio-Rad, Germany
Eppendorf safe-lock micro test tubes	Eppendorf, Germany
Fluotrac-600, 96-well plates	Greiner-Bio-One, Germany
Gloves	Kimberley-Clark, USA
Linbro plates, 24 well, hanging drop	Jena Biosciences, Germany
Linbro plates, 24 well, sitting drop	Jena Biosciences, Germany
Needles	Henke Sass Wolf, Germany
Parafilm	Pechiney Plastic Packaging, USA
Slide-A-lyzer	Pierce, USA
Sterile filters 0.22 μm, 0.45 μm	Sarstedt, Germany
Surgical blades	Martin, Germany
Syringes	Braun, Germany
Tubes (5 ml, 10 ml, 15 ml and 50 ml)	Greiner-Bio-One, Germany
Weighting dishes	Roth, Germany

2.1.4 Chromatographic resins and columns

Table 2.4 lists the resins and columns used routinely in this thesis.

Table 2.4: Chromatographic resins and columns

Matrix	Supplier
Glutathione Sepharose 4 Fast Flow	GE Healthcare, Germany
HiLoad 16/60 Superdex 200 prep grade	GE Healthcare, Germany
HiLoad 16/60 Superdex 75 prep grade	GE Healthcare, Germany
HiPrep Heparin FF 16/10	GE Healthcare, Germany
HiTrap Crude Ni^{+2}-NTA column	GE Healthcare, Germany
MonoQ 5/50 GL, MonoQ HR 10/10	GE Healthcare, Germany
Ni^{+2}-NTA agarose beads	Qiagen, Germany
PD-10 column	GE Healthcare, Germany
Superdex 200 10/300 GL	GE Healthcare, Germany
Superdex 200 26/60,	GE Healthcare, Germany
Superdex 200 PC 3.2/30	GE Healthcare, Germany
Superdex 75 PC 3.2/30	GE Healthcare, Germany

2.1.5 Molecular biology kits

All molecular biology kits used in this thesis are listed in table 2.5.

Table 2.5: Commercial molecular biology kits

Kit	Supplier
QIAquick PCR Purification Kit	Qiagen, Germany
QIAprep spin miniprep Kit	Qiagen, Germany
QIAquick Gel Extraction Kit	Qiagen, Germany
QuikChange II XL Site-Directed Mutagenesis Kit	Agilent Technologies, Germany
Maxiprep Kit	Invitrogen, Germany

2.1.6 Crystallization screens

All kits for crystallization screening used in this work are listed in table 2.6. The crystallization screen solutions were transferred to 96-well MRC crystallization plates for automated pipetting using a Cartesian robot. The Additive, Silver Bullet

and Silver Bullet Bio solutions were also transferred into a 96-well plate format for automated pipetting using a Cartesian robot.

Table 2.6: List of crystallization screens

Screen	Supplier
Additives, Index, SaltRX, Silver Bullets, Silver Bullets Bio	Hampton Research, USA
Classics Suite, Classics Lite Suite, Classics II Suite, PEGs Suite, MPD Suite, Anions Suite, Cations Suite, Protein Complex Suite, PEGs II Suite, PACT Suite, Nucleix Suite, JCSG+ Suite	Qiagen, Germany

2.1.7 Instrumentation

All devices used in this work are listed in table 2.7.

Table 2.7: Devices

Device	Manufacturer
Äkta Explorer, Purifier, Prime, Micro	GE Healthcare, Germany
Allegra X-15R	Beckman Coulter, Germany
Autoclaves H+P	Labortechnik, Germany
Avanti J-26 XP	Beckman Coulter, Germany
Axiovert100	Zeiss, Germany
Beamline 14.2	HZB, Berlin, Germany
Beamline 14.3	HZB, Berlin, Germany
Beamline P14, Petra III	DESY, Hamburg, Germany
Biofuge fresco, Biofuge pico	Heraeus, Germany
Cartesian crystallization robot, 4 channels	Zinsser Analytic, Germany
Cartesian crystallization robot, 8 channels	Digilab, USA
CASY TT Counter	Innovatis, Germany
Cryofuge 6000i	Heraeus, Germany
Digital camera Camedia C-5050 Zoom	Olympus, Japan
Electroporator EasyjecT	Prima Equibio, England
Elexsys 500 EPR spectrometer	Bruker ASX B.V., Netherlands
Filter KV 408	Schott, Germany
Gel electrophoresis equipment	Bio-Rad, Germany
Glass-ware	VWR International, Germany
Heating blocks	Eppendorf, Germany

Continued on next page

Table 2.7 – *Continued from previous page*

Device	**Manufacturer**
HT multitron culture shaker	Infors, Switzerland
Ice machine	Ziegra, UK
Jasco J-810 CD Spectropolarimeter	Jasco, Inc., USA
Milli-Q synthesis A10	Millipore, USA
Magnetic stirrer	IKA, Germany
Megafuge 1.0R, swing out rotor type 2704	Heraeus, Germany
Micro fluidizer M110S	Micro fluidics, USA
Microliter syringes	Hamilton, Switzerland
Microscope SZ-PT	Olympus, Japan
Microwave oven	Bosch
Mounted CryoLoop	Hampton Research, USA
12-way multichannel pipette	Eppendorf, Germany
Mx3000P QPCR Systems	Agilent Technologies, Germany
NanoDrop 2000 Spectrometer	Thermo Fisher Scientific, USA
PCR cycler T-Professional	Biometra, Germany
PCR cycler Peqstar 2x gradient	PeqLab, Germany
pH-meter, Professional Meter PP-20	Sartorius, Germany
Photometer DU 530	Beckmann, Germany
Pipettes	Gilson, USA
Rocking platform	Biometra, Germany
Power supplies	Bio-Rad, Germany
Quartz Cuvette 1 mm, 110-QS	Hellma, Germany
Scales BP4100	Sartorius, Germany
Scales XS4002S DeltaRangeR Mettler	Toledo, Germany
Scales XS205 DualRangeR Mettler	Toledo, Germany
Scanner (gel documentation)	Epson, Germany
Sonopuls Ultrasonic Homogenizer HD 3100	Bandelin, Germany
Spectrophotometer Ultrospec 3000 pro	GE Healthcare, Germany
Speed vac concentrator	5301 Eppendorf, Germany
Table centrifuge 5415R	Eppendorf, Germany
Thermomixer comfort	Eppendorf, Germany
Tunair flasks	Sigma-Aldrich, Germany
Vortex Genie 2	Scientific Industries, USA

2.1.8 Enzymes and proteins

All enzymes used in this work are listed in table 2.8. Proteases were dissolved in Milli-Q water if not stated otherwise.

Table 2.8: List of enzymes

Enzyme	Supplier
Cloned Pfu DNA polymerase (10 U/μl)	Agilent Technologies, Germany
DNase I	Roche, Germany
Lysozyme	Merck, Germany
Phusion DNA polymerase	Finnzymes OY, Finland
PreScission protease	GE Healthcare, Germany
Restriction endonucleases	New England Biolabs, Germany
RNase A, T1	Ambion, Germany
TEV Protease	Home-made, recombinant
T4 DNA ligase	New England Biolabs, Germany
T4 DNA polymerase	New England Biolabs, Germany

2.1.9 Plasmids

Table 2.9 summarizes all plasmids and constructs for the recombinant proteins used in this work.

Table 2.9: List of plasmids

Name	Description	Reference
pETM11	Vector for expression of genes with a TEV protease-cleavable N-terminal His_6 tag in *E. coli*; KanR	EMBL, Heidelberg
pGEX6P1	Vector for expression of genes with a PreScission protease-cleavable N-terminal GST tag in E. coli; AmpR	GE Healthcare
pFL	MultiBac acceptor vector for bacmid preparation; AmpR and GenR	Dr. I. Berger, EMBL, Grenoble
pFL-10His	MultiBac acceptor vector for bacmid preparation; MCSII was modified to code for an N-terminal His_{10} and a TEV protease cleavage site; AmpR and GenR	[Santos, 2012]

Continued on next page

Table 2.9 – *Continued from previous page*

Name	Description	Reference
pFL-10His hBrr2^{HR} (Truncation 1)	N-His$_{10}$; TEV protease cleavage site; hBrr$^{395-2129}$, cloned via EcoRI and HindIII	[Santos, 2012]
pFL-10His HsN (N-cassette)	N-His10; TEV protease cleavage site; hBrr$^{395-1324}$, cloned via EcoRI and HindIII	[Santos, 2012]
pFL-10His HsC (C-cassette)	N-His10; TEV protease cleavage site; hBrr2$^{1282-2136}$, cloned via EcoRI and HindIII	[Santos, 2012]
pFL yBrr2	N-His$_6$; full length yBrr2 cloned via SacI and HindIII	[Santos, 2012]
pFL-10His enHR	N-His$_{10}$; TEV protease cleavage site; yBrr2$^{271-2163}$	Christian Becke, FU-Berlin
pETM11 Prp8CTF$^{1836-2413}$	N-His$_6$; TEV protease cleavage site; cloned via BsaI and NotI	[Santos, 2012]
pGEX6P1 Prp8CTF$^{1836-2413}$	N-GST; PreScission protease cleavage site; cloned via BamHI and NotI	[Santos, 2012]
pETM11 Prp8CTF$^{1836-2398}$	N-His$_6$; TEV protease cleavage site; cloned via BsaI and NotI	Dr. Gert Weber, FU-Berlin
pFL GST-hJab1$^{2064-2335}$	N-GST; PreScission protease cleavage site; cloned via EcoRI and HindIII	This study
pFL GST-hJab1$^{2064-2320}$	N-GST; PreScission protease cleavage site; cloned via EcoRI and HindIII	This study
pFL GST-hJab1^{P2301T}	N-GST; PreScission protease cleavage site; cloned via EcoRI and HindIII	This study
pFL GST-hJab1^{F2304L}	N-GST; PreScission protease cleavage site; cloned via EcoRI and HindIII	This study
pFL GST-hJab1^{H2309P}	wN-GST; PreScission protease cleavage site; cloned via EcoRI and HindIII	This study
pFL GST-hJab1^{H2309R}	N-GST; PreScission protease cleavage site; cloned via EcoRI and HindIII	This study
pFL GST-hJab1^{R2310G}	N-GST; PreScission protease cleavage site; cloned via EcoRI and HindIII	This study

Continued on next page

Table 2.9 – *Continued from previous page*

Name	Description	Reference
pFL GST-hJab1^{R2310K}	wN-GST; PreScission protease cleavage site; cloned via EcoRI and HindIII	This study
pFL GST-hJab1^{R2310G}	N-GST; PreScission protease cleavage site; cloned via EcoRI and HindIII	This study
pFL GST-hJab1^{F2314L}	N-GST; PreScission protease cleavage site; cloned via EcoRI and HindIII	This study
pFL GST-hJab1^{Y2334N}	N-GST; PreScission protease cleavage site; cloned via EcoRI and HindIII	This study
pIDK hPrp8CTF$^{1760-2335}$	cloned via XhoI and NheI	This study
pFL-10His hBrr2$^{\mathrm{HR}}$ fused to pIDK hPrp8CTF$^{1760-2335}$	N-His$_{10}$; TEV protease cleavage site	This study
pFL-10His hPrp8CTF$^{1760-2335}$	N-His$_{10}$; cloned via EcoRI and HindIII	This study
pFL-10His hPrp8CTF$^{1760-2320}$	N-His$_{10}$; cloned via EcoRI and HindIII	This study

2.1.10 Bacterial strains

E. coli strains XL1 blue and DH5α were used for routine cloning applications of plasmid DNA. BL21 (DE3) "Rosetta2" strain was utilized for production of protein from target genes cloned in T7-driven expression vectors. DH10MultiBacY strain was used for recombinant bacmid preparation.

2.1.11 Insect cell lines

Sf9 and Sf21 cells were used to isolate and propagate recombinant baculovirus stocks. Additionally, Sf9 cells were used in the co-expression of some of the complexes. High Five^{TM} cells were used for large scale expression of recombinant proteins. All insect cell lines used are adapted to serum-free suspension culture. The table 2.11

Table 2.10: Bacterial strains

Strain	Description	Supplier
XL1 blue	*recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac [F' proAB lacIqZΔM15 Tn10 (Tet^R)]*	Agilent, Germany
DH5α	*F^- endA1 glnV44 thi-1 recA1 relA1 gyrA96 deoR nupG Φ80dlacZΔM15 Δ(lacZYA-argF)U169, hsdR17($r_K{}^-$ $m_K{}^+$), $^-$*	
BL21 (DE3) "Rosetta2"	*F^- ompT hsdSc$_B$($r_B{}^-$ $m_B{}^-$) gal dcm pRARE2 (Cam^R)*	Novagen, USA
DH10 MultiBacY	-	Dr. I. Berger, EMBL, Grenoble

lists the insect cell lines used in this work.

Table 2.11: Insect cell lines

Cell line	Description	Supplier
Sf9 cells	clonal isolate derived from the parental *Spodoptera frugiperda* (Fall Armyworm) cell line IPLB-Sf21-AE. Originated at the USDA Insect Pathology Laboratory (Vaughn et al., 1977)	Invitrogen, Germany
Sf21 cells	isolated from *Spodoptera frugiperda* (Fall Armyworm) pupal ovarian tissue (Vaughn et al., 1977)	Invitrogen, Germany
High Five^{TM} cells	BTI-TN-5B1-4 was developed by the Boyce Thompson Institute for Plant Research, Ithaca, NY and originated from a clonal isolate derived from the ovarian cells of the cabbage looper, *Trichoplusia ni* (Wickham et al., 1992)	Invitrogen, Germany

2.1.12 Software

Table 2.12 lists the software used in this work.

Table 2.12: List of software

Software	Reference
ApE 2.0.45	Davis, USA, 2012
CNS	[Brünger et al., 1998]
Collaborative Computational Project Number 4 (CCP4i) program suite	[Potterton et al., 2004]
Coot	[Emsley and Cowtan, 2004]
Corel Draw	Corel Corporation, USA
iMosflm 1.0.7.	[Battye et al., 2011]
Phaser	[McCoy et al., 2007]
Phenix suite	[Adams et al., 2002]
Pymol	Schrödinger LLC, USA
XDS	[Kabsch, 2010]

2.2 Methods

Unless otherwise stated, all molecular biological methods were performed as described in Sambrook, 1989. Commercial kits were used according to the instructions of the manufacturer.

2.2.1 Nucleic acid methods

2.2.1.1 Determination of nucleic acid concentration

To determine the concentration of nucleic acids, the light absorption of an aqueous solution was measured at the wavelength of 260 nm using a Nanodrop spectrophotometer. The concentration was then calculated using theoretical absorption values at 260 nm [Sambrook and Russell, 2001].

double-stranded DNA $\quad 1\ OD_{260} = 50\ \mu g/ml$

single-stranded DNA $\quad 1\ OD_{260} = 33\ \mu g/ml$

2.2.1.2 Agarose gel electrophoresis for DNA

To separate nucleic acids, agarose gel electrophoresis was used for both analytical visualization and purification of preparative amounts of DNA. Depending on the size

of the DNA fragments, the gels were prepared with varying agarose concentrations of 0.8–1.5% in 1×TAE buffer. Commercial DNA ladders were used as marker on each gel. Before casting, ethidium bromide was added to the final concentration of 0.7 μg/ml in gel solution. Samples were mixed with DNA loading dye. Gels were run at a constant voltage of 120 V in 1×TAE buffer. DNA bands were visualized by UV illumination at 254 nm.

2.2.1.3 DNA purification using agarose gel electrophoresis

DNA bands were illuminated at a wavelength of 254 nm and excised from the gel with a sterile razor blade. DNA was extracted from the gel using QIAquick gel extraction kit following the instruction of the manufacturer.

2.2.1.4 Polymerase chain reaction (PCR)

Polymerase chain reaction (PCR) was used for amplification of genes or gene fragments. Depending on the target gene, cloned Phusion polymerase or *Pfu* polymerase were used according to the instructions of manufacturers. Typical PCR conditions are shown below (Table 2.13).

Table 2.13: Conditions for PCR

Compound	**Amount**
10× polymerase reaction buffer	1×
dNTP mixture (10 mM of each dNTP)	0.2 mM
DNA template (plasmid DNA)	10-50 ng
Primers	0.3 μM each
DNA polymerase	0.1-0.2 U/μl
DMSO	0-4%

Cycle step	**Temperature (°C)**	**Time**	**Cycles**
Initial denaturation	95-98*	30 sec	1
Denaturation	95-98*	5-10 sec	
Annealing	T_m - 5°C	10-30 sec	25-30
Extension	68-72*	15 sec/kb*	
Final Extension	68-72*	7 min	1

* Depending on the polymerase used.

2.2.1.5 Site-directed mutagenesis

Desired mutations were introduced using the QuikChange II XL Site-Directed Mutagenesis Kit according to the manufacturer's instructions. The resulting clones were verified by DNA sequencing.

2.2.1.6 Restriction digestion and ligation of DNA

Restriction digestions were carried out to generate compatible ends in vectors and PCR products for subsequent ligation reactions. Buffers and temperatures were chosen according to the manufacturer's instruction. The PCR amplified fragments were purified before digestion using a QIAquick PCR purification kit. A typical digestion reaction contained 2-4 μg DNA and 3 units of restriction enzyme per μg of DNA in a total volume of 50 μl. For ligation, the digested DNA was resolved on a preparative agarose gel and the band containing the desired product was excised and extracted as explained previously. Ligation reactions typically contained 100 ng of plasmid DNA, two to five fold molar excess of the insert DNA, 800 units of T4 DNA ligase in 10-20 μl reaction volume. Reaction mixtures were incubated at 4°C overnight or at 25°C for 1 h.

2.2.1.7 Plasmid isolation from *Escherichia coli* cells

A single colony from an overnight grown LB-agar plate was used to inoculate LB-medium. Cells were grown in LB medium overnight at 37°C. Plasmid purification was carried out using Mini- or Maxiprep kits, according to the manufacturer's instructions.

2.2.1.8 Plasmid verification

All cloned plasmids were verified for the presence of the correct insert by PCR or analytical restriction digestion. Sequences of the inserts which showed the correct size in agarose gel analysis were verified by DNA sequencing (Seqlab, Göttingen).

2.2.2 Cells and cell culture methods

2.2.2.1 *Escherichia coli* strains cultivation

The *Escherichia coli (E. coli)* cell strain DH5α was used for plasmid propagation and the strain Rosetta 2 (Novagen, Madison, WI, USA) for protein production by T7 RNA polymerase induced expression with IPTG. Optical densities of cultures were measured in plastic cuvettes with 1 cm path-length in an Ultrospec 3000 pro spectral photometer at 600 nm wavelength using the respective plain media as reference. *E. coli* cells were grown in liquid medium or on agar plates supplemented with adequate antibiotics in the following concentrations:

- ampicillin 100 μg/ml,
- chloramphenicol 34 μg/ml,
- tetracycline 15 μg/ml,
- gentamicin 7 μg/ml,
- kanamycin 50 μg/ml.

2.2.2.2 Transformation of *E. coli* cells

For electroporation, 50-100 ng DNA were mixed with 50 μl electro-competent *E. coli* cells on ice. The mixture was transferred to an ice-cold electroporation cuvette (1-2 mm electrode gap) and subjected to a 4.8 ms pulse of 1.8-2.5 kV. Cells were collected in 1 ml of LB or SOC medium and incubated for 1 h at 37°C in a shaker. Subsequently, the corresponding volume of recovered culture (50-200 μl) was streaked out on an agar-plate containing the selective antibiotics.

For chemical transformation, 100-200 ng of plasmid DNA was mixed with 100 μl of chemically competent *E. coli* cells and incubated for 30 min on ice. Ice-cold cells were then heat-shocked for 45 sec at 42°C and cooled on ice for 3 min. Cells were mixed with 1 ml of LB or SOC medium and incubated at 37°C for 1 h in a shaker. The cells were collected and selected on LB-agar plates supplemented with appropriate antibiotics.

2.2.2.3 Protein expression in *E. coli*

For protein expression, cells were grown in LB or TB medium supplemented with the corresponding antibiotic until OD_{600} values between 0.6-0.8 and then induced with 1mM IPTG. After induction, the expression was carried out at 18°C overnight. The decreased temperature arrests cell growth and slows down protein synthesis rate, which helps to keep the over-expressed protein soluble.

2.2.2.4 Baculovirus expression vector system

Baculovirus gene expression takes advantage of the viral life cycle by use of the late and very late genes promoters. In tissue cultures, these late and very late proteins (such as p10 and polyhedrin (polh)) are non-essential for viral replication and their genes can be exchanged with other genes to create a recombinant baculovirus. This circumstance was exploited for producing the first recombinant baculoviruses in 1983 [Smith et al., 1983] by standard homologous recombination using transfer plasmids carrying the foreign genes.

2.2.2.5 The MultiBac system

The MultiBac system [Trowitzsch et al., 2010] is based on an *Ac*NPV baculovirus genome derived from the Tn7-based BAC variant [Luckow and Summers, 1988]. The MultiBac baculoviral genome is propagated as a bacterial artificial chromosome (BAC) in *E. coli* cells and utilizes a Tn7 attachment site embedded in a *LacZα* gene for integrating foreign genes via transfer plasmids (Acceptor plasmids) into the baculoviral genome. Successful integration of the expression cassette leads to disruption of the *LacZα* gene and enables the selection of positive clones by blue/white screening.

Baculoviral genes v-cath and chiA, coding for a cathepsin protease and a chitinase, were substituted by a LoxP imperfect inverted repeat introduced together with an enhanced YFP protein coding gene. The resulting BAC is called EMBacY. Two families of modular transfer plasmids denominated Acceptors (pFL and pKL) and Donors (pUCDM, pIDC, pIDS, pIDK and pSPL) are currently used to deliver the heterologous genes, under control of p10 or polh promoters, to the BAC. Both Acceptor and Donor plasmids harbor a LoxP site that can be used to combine one Acceptor with one or more Donor plasmids before integration into the EMBacY. This strategy is widely used in case of multiprotein complex expression. The transposition involving the Acceptor plasmid or Acceptor/Donor fusions and the EMBacY occurs

in vivo in *E. coli* cells, called DH10MultiBacY, tailored for this purpose. These cells provide a helper plasmid encoding the Tn7 transposon complex for accessing the Tn7 site on the EMBacY. For virus production, the composite EMBacY is isolated from DH10MultiBacY and used to transfect insect cells.

The presence of YFP gene in the EMBacY serves the purpose of directly observing virus performance by using a fluorescence spectrophotometer [Trowitzsch et al., 2010]. YFP is under the control of a very late promoter (polh) as are, typically, the heterologous genes. Consequently, it is possible to follow protein production levels by monitoring YFP expression.

2.2.2.6 CreLox Recombination

For the co-expression of hPrp8$^{\mathrm{CTF}}$-hBrr2$^{\mathrm{HR}}$ complex, the genes of each protein were primarily cloned separated. hBrr2$^{\mathrm{HR}}$ was cloned into the Acceptor vector pFL (the pFL10His variant that includes a His$_{10}$-tag) as described in previous publication [Santos et al., 2012]. hPrp8$^{\mathrm{CTF}}$, in the other hand, was cloned untagged into the Donor vector pIDK. The vectors containing the proper insert were fused by CreLox recombination [Fitzgerald et al., 2006]. Subsequent to the Cre-mediated fusion of Acceptor and Donor vectors, *in vitro* Tn7 transposition into EMBacY was performed. The recombinant bacmids were isolated and used to transfect SF9 cells.

2.2.2.7 Insect cells strains culture

All handling of insect cell lines was carried out under sterile conditions in a laminar flow hood. Sf9, Sf21 and High Five^{TM} cell stocks (25 ml) were maintained in 250 ml shaker flasks at 27°C while shaking at 80 rpm. Sf9 and Sf21 cells were cultured in SF900 III SFM and High Five^{TM} cells in Express Five^{R} SFM medium supplemented with L-glutamine prior to use.

- Sf9 cells were used for virus production of the hPrp8$^{\mathrm{Jab1/MPN}}$, hBrr2$^{\mathrm{HR}}$, yBrr2$^{\mathrm{enHR}}$ constructs and co-expression of hPrp8$^{\mathrm{CTF}}$-hBrr2$^{\mathrm{HR}}$ complex.

- Sf21 cells were used for virus amplification of hPrp8$^{\mathrm{Jab1/MPN}}$, hBrr2$^{\mathrm{HR}}$ and yBrr2$^{\mathrm{enHR}}$.

-High Five^{TM} were used for the expression of hPrp8$^{\mathrm{Jab1/MPN}}$, hBrr2$^{\mathrm{HR}}$ and yBrr2$^{\mathrm{enHR}}$.

2.2.2.8 Bacmid preparation, transfection and virus amplification

Recombinant bacmid was prepared using the MultiBac system as described by Fitzgerald et al., 2006 [Fitzgerald et al., 2006]. The genes of interest were cloned in a pFL vector or in a modified version called pFL10His and transformed into electrocompetent DH10MultiBacY cells. Following electroporation, cells were incubated at 37°C for 4 h in SOC medium and plated on LB-agar media containing 50 μg/ml kanamycin, 7 μg/ml gentamicin, 10 μg/ml tetracycline, 100 μg/ml ampicillin, 100 μg/ml Bluo-Gal, and 1 mM IPTG.

The expression constructs were individually integrated via Tn7 transposition into the EMBacY. The Tn7 transposition site is embedded in a *LacZ*α gene allowing the selection of positive EMBacY recombinants by blue/white screening. White positive clones were re-streaked on a new LB-agar plate. After the second round of blue-white screening, single white colonies were inoculated into LB media containing 50 μg/ml kanamycin, 7 μg/ml gentamicin, 10 μg/ml tetracycline and 100 μg/ml ampicillin, and cultivated overnight with vigorous shaking. Recombinant EMBacYs were isolated from the bacterial hosts using QIAprep miniprep kit, omitting the column purification step. The recombinant bacmids were precipitated by addition of isopropanol and pelleted by centrifugation in a table-top centrifuge at 17000×g for 30 min. The pellets were washed with 70% ethanol and the dried bacmid pellets were dissolved in 40 μl ddH_2O and mixed with 200 μl of SF-900 III medium in a sterile laminar flow hood. The transfection reagent, X-tremeGene 9, was diluted in the same medium (1:10, 10 μl X-tremeGene 9 in 100 μl medium) and mixed with the bacmid solution. The mixture was incubated for 1 h and distributed equally in two wells of a 6-well plate, each containing 3 ml of Sf9 cells (0.3×10^6 cells/ml). To obtain initial virus (V_0 generation), Sf9 cells grown as adhesive culture in 6-well plates were transfected with composite EMBacY BACs. The efficiency of transfection was monitored by eYFP fluorescence using an Axiovert 40 CFL microscope (Zeiss) equipped with an HBO illuminator and a proper filter set for eYFP fluorescence visualization. The V_0 generation was harvested 60 h post-transfection and immediately used to infect a 25 ml suspension culture of Sf21 cells in an Erlenmeyer shaker flask for further virus amplification (V_1 generation). Infected Sf21 were diluted to a cell count below 10^6 cells/ml every 24 h until cell proliferation arrest (typically 2-4 days). After proliferation arrest, 10^6 cells were sampled from the infected culture every 12 h for cell counting and diameter determination using a Casy TT cell counter. For YFP fluorescence signal measurements (performed in a FLUOTRAC-600 black 96 well polystyrene plates using a Victor X3 multilabel plate reader), the cells were resuspended in 500 μl PBS and lysed by sonication. The amplified virus (V_1

generation) was harvested by centrifugation and storage of the medium containing the virus 60 h after cell proliferation arrest and the same volume of fresh media was replenished to the culture. YFP signal from the 10^6 cells sampled every 12 h continues to be recorded until it reaches a plateau. At this point, cells were harvested and protein production was analyzed by SDS-PAGE using a fraction of the samples used for YFP fluorescence measurements.

2.2.2.9 Protein expression in insect cells

To determine the appropriate amount of virus (V_1) to be used for large scale expression, 3 flasks each containing 25 ml of High Five^{TM} or Sf9 (depending on the expression) at 0.5×10^6 cells/ml were infected with varying volumes of V_1, such as 25 μl (1:1000), 50 μl (1:500) and 100 μl (1:250). Cell counting and diameter determination was performed every 24 h. The amount of V_1 that allowed one round of cell division and yielded high protein production was chosen for scalling up the expression. For large scale expression, 400 ml of High Five^{TM} (or Sf9) cells kept in suspension in an Erlenmeyer shaker flasks at 0.5×10^6 cells/ml were infected with appropriate volume of V_1 virus. Samples of 10^6 cells were taken from the infected culture every 12 h for cell counting and diameter determination. The samples were further used for YFP fluorescence signal measurements as described above and for protein production analysis by SDS-PAGE. The infected cells were harvested when the YFP fluorescence signal reached a plateau (typically 2-3 days) or before the cell viability dropped below 90%.

2.2.3 Protein methods

2.2.3.1 Determination of protein concentration

Protein concentrations were determined using 280 nm absorbance on a NanoDrop and the theoretical extinction coefficients of the proteins were used to calculate the concentration.

2.2.3.2 Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE)

Protein samples were analysed by SDS-PAGE as described by Laemmli, 1970 [Laemmli, 1970]. By this method, denaturated proteins are resolved according to

their molecular weight. According to the protein sizes to be separated, 10-12% (final acrylamide concentration) gels were prepared and run vertically in SDS-PAGE running buffer. The gels typically consist of acrylamide, bisacrylamide, SDS, and a buffer with an adjusted pH, the acrylamide mixture used in this thesis is Rotiphorese Gel 30 solution. A common setup consists of a resolving gel and a stacking gel. The resolving gel is prepared with a pH of 8.8 and variable acrylamide percentage, while the stacking gel has pH 6.8 and 4% final acrylamide concentration. The acrylamide solution was polymerized using 0.3% (w/v) APS and 0.03% (v/v) TEMED. Protein samples were denatured in protein loading buffer and heated to 95°C for 5 min. Electrophoresis was typically stopped when the bromophenol blue border reached the bottom of the gel. The gels were incubated in Coomassie staining solution for 15 minutes and destained by the sequential addition of destaining solution I until bands became visible. The gel was destained further with destaining solution II [Fairbanks et al., 1971].

2.2.3.3 Purification of human and yeast Brr2 full length and truncations

If not mentioned otherwise, the same purification protocol was used for all human and yeast Brr2 constructs. The High Five^{TM} cell pellet was resuspended in 50 mM HEPES, pH 7.5, 600 mM NaCl, 2 mM β-mercaptoethanol, 0.05% NP40, 10% glycerol, 10 mM imidazole, supplemented with EDTA-free protease inhibitors and lysed by sonication using a Sonopuls Ultrasonic Homogenizer HD 3100. Cell debris was removed by centrifugation and the soluble extract was filtered using a 0.44 μm filter. The target was captured on a 5 ml HisTrap FF column and eluted with a linear gradient from 10 to 250 mM imidazole. The His-tag was cleaved with TEV protease during overnight dialysis at 4°C against 50 mM HEPES, pH 7.5, 600 mM NaCl, 2 mM β-mercaptoethanol, 10% (v/v) glycerol and 15 mM imidazole. The cleaved protein was again loaded on a 5 ml HisTrap FF column to remove the His-tagged protease, uncut protein and cleaved His-tag. The flow-through containing the protein of interest was diluted to a final concentration of 80 mM sodium chloride and loaded on a MonoQ 10/100 GL column equilibrated with 25 mM Tris-HCl, pH 8.0, 50 mM NaCl, 2 mM β-mercaptoethanol. The protein was eluted with a linear 50 to 600 mM sodium chloride gradient and further purified by gel filtration on a 26/60 Superdex 200 gel filtration column in 10 mM Tris-HCl, pH 7.5, 200 mM NaCl, 2 mM DTT. In case of yeast Brr2, after the recycling step, the protein was diluted to a final concentration of 80 mM sodium chloride and loaded on a HiPrep Heparin FF column equilibrated with 25 mM Tris-HCl, pH 8.0, 50 mM NaCl, 2 mM β-mercaptoethanol. The protein was eluted with a linear 50 to 600 mM sodium chloride gradient and

further purified by gel filtration on a 26/60 Superdex 200 gel filtration column in 10 mM Tris-HCl, pH 7.5, 200 mM NaCl, 2 mM DTT. For the purification of the N-terminal cassette construct, all solutions used were buffered at pH 8.0.

2.2.3.4 Expression and purification of C-terminal domain of Prp8 from *S. cerevisiae (y)*

Two different constructs of yPrp8^{CTF} were used, a short construct (residues 1836-2398) His_6-tagged and a longer construct (residues 1836-2413) glutathione S-transferase (GST)-tagged. Both constructs were expressed in *E. coli* Rosetta 2(DE3). Cells were grown in TB medium to an OD_{600} of 0.6 at 37°C, induced with IPTG , shifted to 18°C for 60 h and harvested by centrifugation.

Cell pellets of the His_6-tagged protein were resuspended in 50 mM HEPES, pH 7.5, 500 mM NaCl, 2 mM β-mercaptoethanol, 5% (v/v) glycerol, supplemented with EDTA-free protease inhibitors and DNase I. The cells were lysed by three passes through a microfluidizer. The cell lysate was clarified by centrifugation and then loaded onto a 5 ml HisTrap FF column equilibrated in resuspension buffer without detergent. Protein was eluted with a linear gradient from 10 mM to 300 mM imidazole. Fractions containing the protein of interest were pooled, mixed with TEV protease and dialyzed over night at 4°C against 25 mM HEPES, pH 7.5, 200 mM NaCl, 5% (v/v) glycerol, 2 mM -mercaptoethanol. The digested sample was again loaded on a 5 ml HisTrap FF column equilibrated in resuspension buffer (without detergent) in order to remove TEV protease, uncut protein and the cleaved His_6-tag. The flow-through containing Prp8$^{CTF1836-2398}$ was concentrated and injected directly onto a Superdex 200 gel filtration column equilibrated in 10 mM Tris-HCl, pH 7.5, 150 mM NaCl, 1mM DTT.

Cell pellets of the GST-tagged protein were resuspended in 50 mM Tris, pH 8.0, 300mM NaCl, 1 mM DTT, 5% (v/v) glycerol, supplemented with EDTA-free protease inhibitors and DNase I. The cells were lysed by three passes through a microfluidizer. The cell lysate was clarified by centrifugation and then loaded onto GSH beads previously equilibrated with binding buffer (resuspension buffer without protease inhibitors and DNase I). The clear lysate and beads were incubated together at 4°C for 1 hour with gentle agitation. Afterwards, the flow through was collected and the beads were washed twice with 10 column volumes (CV) of binding buffer. Elution was performed through fractionation (2 ml fractions) with elution buffer (binding buffer supplemented with 10 mM Glutathione). The samples were run on an SDS-PAGE and the fractions containing the protein were collected and buffer exchanged using

a desalting column and the same binding buffer. The so buffer-exchanged sample was incubated overnight with PreScission protease at 4°C. The next day the sample was recycled over the GSH beads. The flow through was collected, concentrated and loaded onto a gel filtration column S200 (1-2 ml injection volume and 3 ml fractions) equilibrated in, HEPES 20mM pH 7.5, 200mM NaCl and 1mM DTT. In case of co-crystallization with Brr2^{enHR}, the complex was assembled in the last concentration step, incubated for 20 min and run over a S200 column.

2.2.3.5 Expression and purification of human (h) Prp8$^{\mathbf{Jab1/MPN}}$ and RP mutants

The GST-tagged hJab1 domain was expressed in High Five^{TM} cells. Cell pellets were resuspended in 50 mM Tris, pH 8.0, 300mM NaCl, 5% glycerol, 1mM DTT, 0.05% NP40 supplemented with EDTA-free protease inhibitors and DNase I. The suspension was then lysed by sonication using a Sonopuls Ultrasonic Homogenizer HD 3100, cell debris was removed by centrifugation and the soluble extract was filtered using a 0.44 μm filter. The clear lysate was loaded onto GSH beads previously equilibrated with binding buffer and incubated 1 h at 4°C with gentle agitation. Afterwards, the flow through was collected and the beads were washed twice with 10 column volumes (CV) of binding buffer. Elution was performed through fractionation (2 ml fractions) with elution buffer (binding buffer supplemented with 10 mM Glutathione). The samples were run on an SDS-PAGE and the fractions containing the protein were collected and buffer exchanged using a desalting column and the same binding buffer. The so buffer-exchanged sample was incubated overnight with Precission protease at 4°C. The next day the sample was recycled over the GSH beads. The flow through was collected, concentrated and loaded onto a S200 equilibrated in Tris 10mM pH 8.0, 150mM NaCl and 1mM DTT. In case of co-crystallization with hBrr2$^{\mathrm{HR}}$, the complex is made in the last concentration step, incubated for 20 min and run over a S200 gel filtration column.

2.2.3.6 Co-expression and purification of hBrr2$^{\mathbf{HR}}$- hPrp8$^{\mathbf{Jab1/MPN}}$ complex

The complex was expressed in Sf9 cells. To co-purify the complex of hBrr2 and Prp8 fragments, we performed Ni^{2+}-NTA chromatography on beads, using the tag present in the Brr2 construct. The hPrp8$^{Jab1/MPN}$ fragment was untagged. Insect cells pellet were resuspended in 50mM Tris, pH 8.0, 300mM NaCl, 5% glycerol, 1mM DTT, 0.05% NP40 supplemented with EDTA-free protease inhibitors and DNase I

and lysed *via* sonification as previously described. Centrifuged and filtered lysate was loaded on equilibrated 1 ml Ni^{2+}-NTA beads for gravity flow. The flow through was collected, followed by two washing steps with binding buffer and three elution steps with high imidazole elution buffer (5 CV each step). The elution fraction was treated with 0.5 mg TEV protease and dialysed to allow recycling. During recycling over the same Ni^{2+}-NTA beads, the target proteins were again collected in the flow through and wash fractions, which was then followed by elution. The flow through was concentrated and loaded onto a gel filtration column (S200) for the last purification step. The gel filtration buffer used was HEPES 20mM pH 7.5, 200mM NaCl and 1mM DTT.

2.2.3.7 Analytical gel filtration analysis

The Brr2 constructs and complexes with Prp8 fragments were analyzed by analytical size exclusion (gel filtration) chromatography on a Superdex 200 PC3.2 column in 10 mM Tris-HCl, pH 7.5, 200 mM NaCl, 2 mM DTT at a flow rate of 70 μl/min. For analysis of complex formation, proteins were mixed in defined ratios in a total volume of 60 μl and incubated for 20 min on ice prior to the chromatography. Fractions were concentrated and analyzed by SDS-PAGE.

2.2.3.8 Differential scanning fluorimetry (DSF)

DSF experiments were done in a 96-well plate in a plate reader combined with a thermocycler (Stratagene Mx3005P). In order to determine a suitable protein concentration, three different final concentrations were chosen varying from 1 μM to 10 μM, the dilution was made in the gel filtration buffer supplemented with 10×SYPRO orange (1:500 dilution of the stock) in a total volume of 20 μl and pipetted into a 96-well plate. The temperature was increased from 25°C to 95°C and the fluorescence emission was monitored in steps of 1°C/min with hold steps of 30 sec between reads. The fluorescence intensity was then plotted as a function of temperature. The protein concentration that showed a clear sigmoidal curve was chosen for buffer optimization and compound screens. The sigmoidal curve from each condition was normalized and corrected for the background signal of the fluorophore in the buffer. The inflection points of the normalized curves, representing the thermal melting temperature of the protein in the respective conditions, were compared. In case of $hPrp8^{CTF}$, different buffer compositions and compounds were tested for their stabilizing effect on the protein at the defined protein concentration. This method was also used to compare the difference in stability between the $hPrp8^{Jab1/MPN}$ and

the RP mutants in the defined buffer.

2.2.4 Crystallographic methods

2.2.4.1 General crystallography setup

All the *S. cerevisie* complexes were concentrated until 3.5 mg/ml, since higher concentrations lead to precipitation, while the human complexes were concentrated between 10-12mg/ml. The samples were subjected to an initial round of crystallization trials using different commercially available crystallization reagents in 96-well MRC plates by sitting drop vapor diffusion technique. Drops of 200 nl (100 nl protein solution + 100 nl reservoir) were dispensed using a Cartesian liquid dispensing robot with 4 or 8 channels. Initial hits were usually refined by manual setups in 96-well MRC plates and 24-well format by sitting and hanging drop. Commercial additive screens were routinely tested to improve crystallization conditions.

2.2.4.2 Crystallization and diffraction data collection

Crystallization of the yeast complexes was carried out at 18°C. In case of yBrr2$^{\mathrm{enHR}}$-yPrp8$^{\mathrm{CTF}}$ complex, needle-like crystals were obtained for both constructs of Prp8 in the same three conditions (Table 2.14). We failed to reproduce the crystals on a grid screen using sitting drop and 96-well MRC plates, which covered a pH range from 6.2 to 7.0 (0.1 M MES or HEPES) and PEG concentrations from 7% to 14%. Therefore, no improvement was possible for these crystals.

Table 2.14: Crystallization conditions yBrr2$^{\mathrm{enHR}}$-yPrp8$^{\mathrm{CTF}}$ complexes

Screen	Condition
PEGs II suite	C4: 0.2 M $MgCl_2$, 0.1 M MES pH 6.5, 10% PEG 4000
PEGs II suite	H10: 0.2 M MgAc, 10% PEG 8000
JCSG Core I Suite	C4: 0.2 M $MgCl_2$, 0.1 M Tris pH 7.0, 10% PEG 8000

The complex yBrr2$^{\mathrm{enHR}}$-yPrp8$^{\mathrm{Jab1/MPN}}$ was also crystallized in similar conditions as the CTF complexes (Table 2.15). The crystals were reproduced on a 96-well MRC plates grid screen using 0.5 μl sample and 0.5μl reservoir. One condition could be also reproduced in 24-well format and bigger drops were used, 1 μl sample plus 1 μl reservoir. The crystals were cryo-protected by transfer into a solution containing the reservoir solution supplemented with different cryoprotectants like glycerol, PEG

400, PEG 300, propylene glycol, 2,3-butanediol or oils and then flash-cooled in liquid nitrogen. A cryoprotectant consisting of 0.1M MES, pH 6.2, 10% PEG 4000, 0.2 M $MgCl_2$, 25% PEG 400 offered the best cryoprotection for the best diffracting crystals that grew in the condition 0.1M MES, pH 6.2, 10% PEG 8000, 0.2 M $MgCl_2$. Diffraction data were collected at beamline 14.2 of BESSY II (HZB, Berlin, Germany) and processed with XDS [Kabsch, 2010].

Table 2.15: Crystallization conditions yBrr2$^{\mathbf{enHR}}$-yPrp8$^{\mathbf{Jab1/MPN}}$ complex

Screen	Condition
Classic suite I	F4: 0.1 M HEPES pH 7.5, 10% PEG 8000
PEGs II suite	C4: 0.2 M $MgCl_2$, 0.1 M MES pH 6.5, 10% PEG 4000
PEGs II suite	C6: 0.1 M Tris pH 8.5, 12% PEG 4000
JCSG Core I Suite	C4: 0.2 M $MgCl_2$, 0.1 M Tris pH 7.0, 10% PEG 8000
JSCG Core II Suite	A12: 0.1 M Tris pH 8.5, 10% PEG 8000

The hBrr2$^{\mathrm{HR}}$-hPrp8$^{\mathrm{CTF}}$ complex crystallized in four different conditions and at two temperatures, 4°C and 18°C (Table 2.16). The conditions were easily reproduced on a 24-well grid screen using hanging drop vapour diffusion technique. The drops were also bigger, with a final volume of 2 μl (1 μl sample plus 1 μl reservoir). As for the yeast crystals, several cryorotectants were tested, glycerol, PEG 400, ethylene glycol (EG), 2,3-butanediol, TMAO and perfluoropolyether (PFO). The crystals were then flash-cooled in liquid nitrogen. The best diffracting crystals were obtained at 4°C in 1 M LiCl, 0.1 M MES pH 6.0, 8% PEG 6000 and using 25% EG as cryoprotectant. The crystals were submitted to further treatments to improve like resolution cross-linking with glutaraldehyde, dehydration and seeding (dilution used 1:5000) [Heras and Martin, 2005]. Diffraction data were collected at beamline P14 of PETRA III (DESY, Hamburg, Germany) and processed with XDS [Kabsch, 2010].

Table 2.16: Crystallization conditions hBrr2$^{\mathbf{HR}}$-hPrp8$^{\mathbf{CTF}}$ complex

Screen	Temperature	Condition
JCSG Core II Suite	4°C	C9: 1 M LiCl, 0.1 M MES pH 6.0, 10% PEG 6000
JCSG Core II Suite	4°C	E4: 1 M LiCl, 0.1 M HEPES pH 7.0, 10% PEG 6000
JCSG Core II Suite	18°C	G1: 2 M ammonium sulphate, 5% isopropanol
JCSG Core II Suite	18°C	D8: 0.2 M $MgCl_2$, 0.1 M sodium cacodylate pH 6.5, 10% PEG 3000

The complex hBrr2$^{\mathrm{HR}}$-hPrp8$^{\mathrm{Jab1/MPN}}$ yielded several hits in different screens and at different temperatures. The best diffraction was obtained after optimization of

crystals grown in 24 well formant and hanging drop in 0.2M $MgCl_2$, 0.1M HEPES pH 7.9, 12% Ethanol at 4°C. Crystals were obtained by mixing 1 μl of protein solution at 12 mg/ml with 0.5 μl of reservoir solution. Several cryoprotectans were tested and 30% EG proved to be the best solution preserving the integrity of the crystal during flash-cooling in liquid nitrogen. The crystals were further optimized using ADP as a cofactor, which was added as an additive to the crystallization drop during setup (0.5 μl of 5 mM stock dissolved in gel filtration buffer). The last improvement was achieved using a dehydration method. The latter entails the addition to the drop of 8× the drop volume of a dehydration solution (22% Ethanol,12%EG, 0.2M $MgCl_2$, 0.1M HEPES pH 7.9) and to let the drop equilibrate against air for 30 minutes. Diffraction data were collected at beamline P14 of PETRA III (DESY, Hamburg, Germany) and at beamline 14.2 of BESSY II (HZB, Berlin, Germany) and processed with XDS [Kabsch, 2010].

2.2.4.3 Structure solution, model building and refinement

The structures were solved by molecular replacement using the coordinates of the $hBrr2^{HR}$ and the $yPrp8^{Jab1/MPN}$ structures as search models and the software Molrep [Vagin and Teplyakov, 1997] and Phaser [McCoy et al., 2007] of the CCP4i suite [Potterton et al., 2004]. The refinement of molecular replacement solutions was carried out by Phenix software suite [Adams et al., 2002], including TLS refinement, and manual model building in COOT [Emsley and Cowtan, 2004]. All figures displaying coordinate files and electron density maps were generated with Pymol.

Chapter 3

Results

3.1 Production of proteins

3.1.1 Production of human proteins

3.1.1.1 hBrr2

All the fragments of hBrr2 used in this thesis were produced in our laboratory using the constructs and purification protocols described previously [Santos, 2012]. The helical region that includes the two helicase cassettes and an N-terminal extension was named hBrr2^{HR} (residues 395-2136). The fragment consisting of the N-terminal cassette and the inter-cassette linker was called hBrr2^{NC} (residues 395-1324). The fragment formed by the C-terminal cassette and the inter-cassette linker is referred to as hBrr2^{CC} (residues 1282-2136) [Santos, 2012].

3.1.1.2 hPrp8$^{Jab1/MPN}$

No previous information was available in terms of hPrp8^{CTF} and the hPrp8$^{Jab1/MPN}$ domains expression, albeit post-translational modifications were expected like phosphorylation [Dephoure et al., 2008; Olsen et al., 2010] and ubiquitylation [Bellare et al., 2005, 2008] (dbPTM: data base of post-translational modifications). Initial attempts to express these domains in *E. coli* by members of our laboratory failed. Therefore, a baculovirus-based expression system using insect cells was chosen to produce the human protein fragments. Additionally, in order to facilitate the cloning and expression of the constructs, we worked with expression-optimized synthetic genes.

I cloned the region of the PRP8 gene encoding for hPrp8$^{Jab1/MPN}$ domain (residues 2064-2335) fused to a GST (Glutathione S-Transferase) tag into a pFL vector and the protein was expressed in High Five cells. The final yields were in the order of 10 mg of the fusion protein per litre of culture (Fig. 3.1). The protein fragment was purified through its GST-tag, based on standard glutathione sepharose purification protocol. After removal of the tag, the protein was ready for the last purification step using gel filtration chromatography. During the purification a degradation product of the fragment was observed and it was confirmed as part of the Prp8$^{Jab1/MPN}$ domain by mass spectrometry. Most probably, the shorter fragment has a C-terminal deletion of the last 15 residues which conform to the flexible tail of the protein and, in case of the human domain, seem to be prone to degradation. The purified protein was then used for crystallization trials, interaction studies and complex assembly.

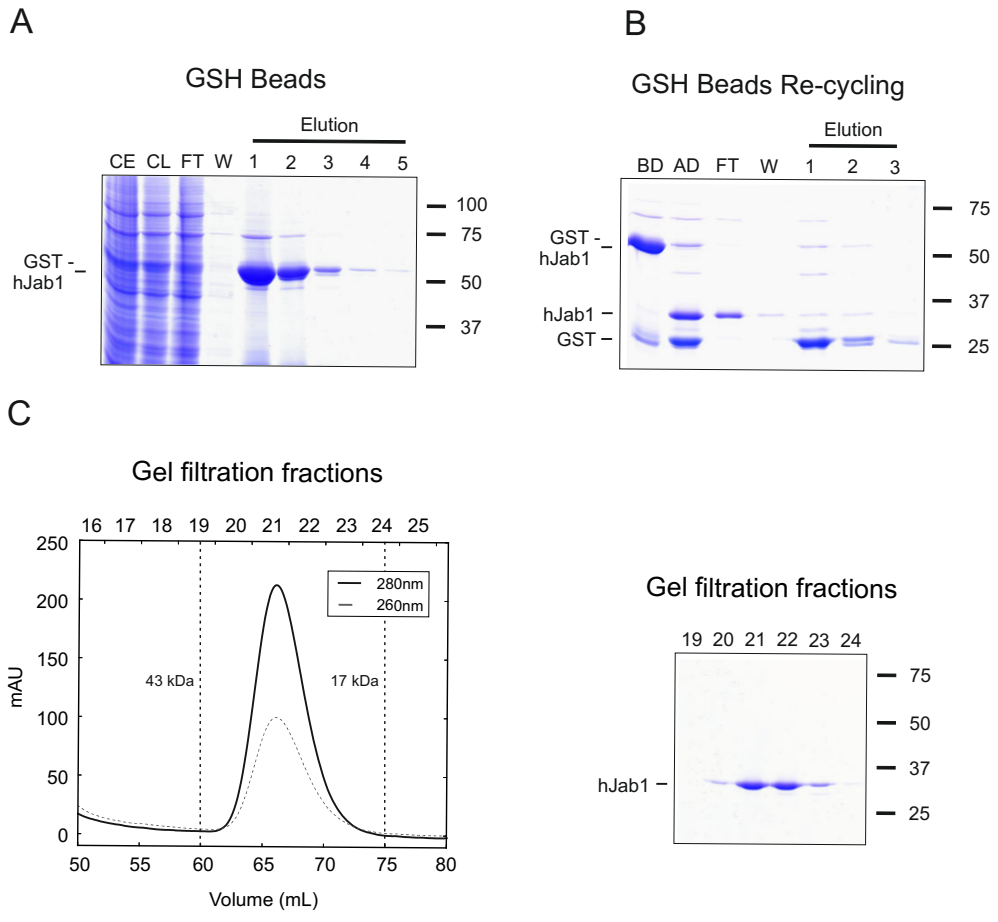

Figure 3.1: Purification of hPrp8$^{Jab1/MPN}$. Coomassie-stained SDS gels showing GSH purification (**A** and **B**) and gel filtration on a Superdex 75 16/60 gel filtration column **(C)** of hPrp8$^{Jab1/MPN}$. The molecular weight marker is shown on the right (sizes in kDa). The protein migrated as a monomer. CE – whole cell extract, CL – cleared lysate, FT – flow through, W – wash, BD – before digestion and AD – after digestion.

3.1.1.3 hPrp8^{CTF}

I also cloned the hPrp8^{CTF} (residues 1755-2335) using a modified pFL vector that includes a cleavable N-terminal His_{10}-tag. I was able to express the protein fragment using insect cells. The produced protein was soluble, but it had a high tendency to aggregate, running in the void volume of a Superdex 200 gel filtration column (Fig. 3.2). Since it is known that the stability and solubility of a protein is greatly influenced by the buffer conditions, I screened a broad range of buffers using differential scanning fluorimetry (DSF). The results of the assay were not conclusive, with little increase in the melting temperatures and most of the obtained curves showed more than one thermal transition. The latter phenomenon has been associated to the presence of aggregates, oligomers or complexes [Kopec and Schneider, 2011] and, in the case of this protein, I presumed it was related to the tendency of the molecule to aggregate. Nevertheless, I selected few conditions for further analysis (Fig. 3.3). After testing the putative hits using gel filtration chromatography I could not observe any improvement in the behaviour of the protein, therefore I decided to try a new approach.

3.1.1.4 Co-production of hBrr2^{HR}-hPrp8^{CTF} complex

Since the purification of hPrp8^{CTF} failed, I tried to co-express the molecule together with hBrr2^{HR} in insect cells. The gene for hBrr2^{HR} was cloned in an acceptor vector (pFL) as previously described [Santos, 2012] and the gene for hPrp8^{CTF} was cloned in a donor vector (pIDK). Both plasmids were then fused by Cre-lox recombination. The expression was only possible in Sf9 cells. This strategy proved to be successful and the purification of the complex was initially carried out through the N-terminal His_{10}-tag of the Brr2 fragment. After the removal of the tag, the complex was loaded on a size exclusion column for the last purification step. The chromatogram showed a single peak with the retention volume expected for a 1:1 complex. Additionally, the apparent amount of each complex component appeared to be equimolar on the SDS-PAGE (Fig. 3.4). The purified complex was used for crystallization trials (section 3.3.5).

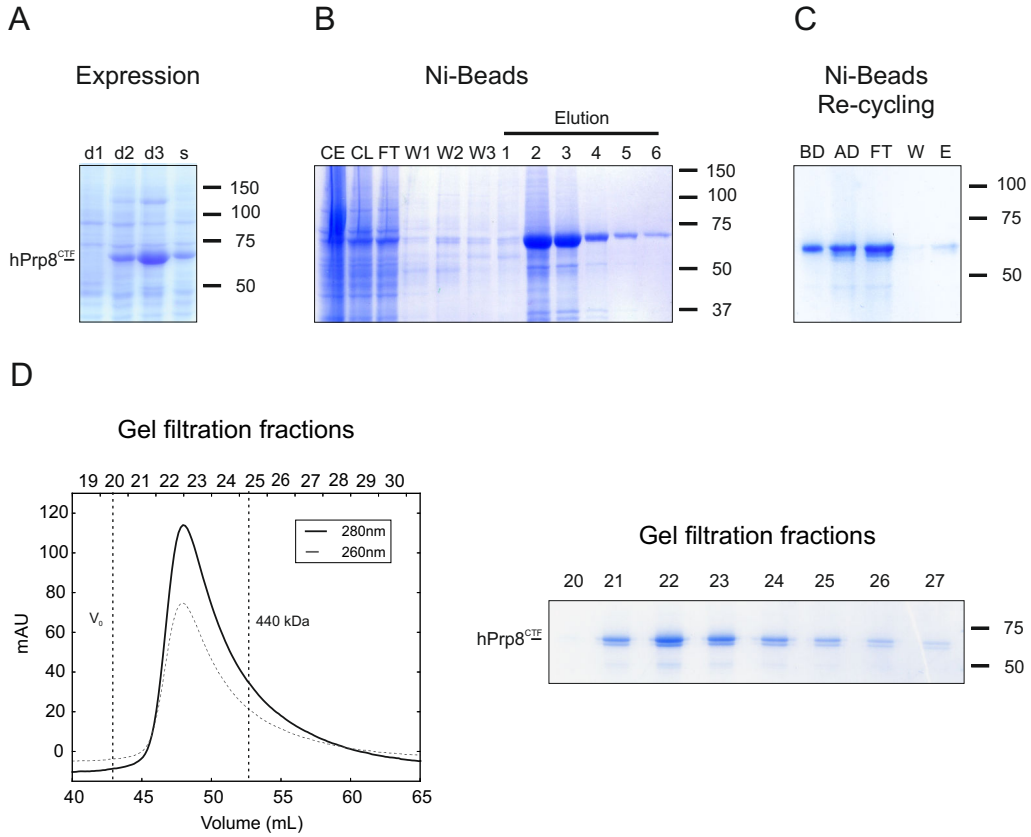

Figure 3.2: Production and purification of hPrp8^{CTF}. Coomassie-stained SDS gels showing expression **(A)** Ni^{+2}-NTA affinity purification (**B** and **C**) and gel filtration chromatography **(D)**. The molecular weight marker is shown on the right (sizes in kDa). The protein was solubly expressed, but it migrated in the exclusion volume of an S200 16/60 column. d1 – whole cell extract day 1 (24h), d2 – whole cell extract day 2 (48h), d3 – whole cell extract day 3 (72h), s – soluble supernatant, CE – whole cell extract, CL – cleared lysate, FT – flow through, W – wash, BD – before digestion, AD – after digestion and E – Elution.

3.1.1.5 Production of hPrp8$^{Jab1/MPN}$ *Retinitis pigmentosa* (RP) mutants

Nearly all Prp8 mutations linked to RP13 (P2301, F2304, H2309, R2310, F2314, Y2334 in human and mutations at other positions that lead to frameshifts or a stop codon) (Fig. 3.5) cluster in the C-terminal 35 amino acids of hPrp8. Additionally, S2118 is an RP13-linked residue located in the globular core of hPrp8$^{Jab1/MPN}$ but close to the C-terminal tail.

In order to test the effect of the mutations on hPrp8$^{Jab1/MPN}$ tail, I modified this domain using almost all the single point mutations previously mapped in the genes of patients with RP. The chosen hPrp8$^{Jab1/MPN}$ point mutations were: P2301S, F2304L, H2309P, H2309R, R2310G, R2310K, F2314L, Q2321stop and Y2334N. All the mutants were produced in insect cells. The first four mutants (P2301S, F2304L,

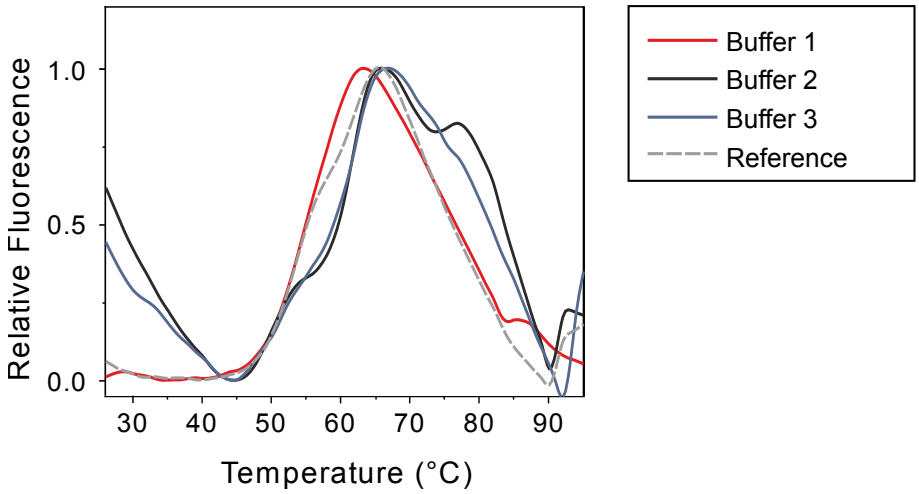

Figure 3.3: DSF analysis plot of hPrp8^{CTF} in different buffer conditions. Normalized melting curves of three different buffer conditions which increased the melting temperature of hPrp8^{CTF} or displayed a single thermal transition. Buffer 1 = 50 mM NaCitrate pH 5.4. Buffer 2 = 50 mM Borax pH 9.0, 100 mM NaCl. Buffer 3 = 50 mM Borax pH 9.0, 200 mM NaCl. Reference = 20 mM Tris pH 8.0, 200mM NaCl.

H2309P, H2309R) were produced, but the proteins were insoluble or precipitated after tag removal. The rest of the mutants were successfully expressed, purified and characterized. Additionally, I performed interaction studies together with hBrr2^{HR}.

3.1.1.6 Differential scanning fluorimetry (DSF) assay of hPrp8$^{Jab1/MPN}$ soluble mutants

Aiming to further characterize the hPrp8$^{Jab1/MPN}$ soluble mutants, the samples were analyzed using DSF. I tested hPrp8$^{Jab1/MPN}$ variants alone and in complex hBrr2^{HR}. All of the obtained spectra of the hPrp8$^{Jab1/MPN}$ variants including wild-type exhibited cooperative transitions with comparable melting temperatures (Tms) (Fig. 3.6A), meaning that the point mutations do not affect the stability of the molecule. The complexes showed a moderate increase in the Tms compare to hBrr2^{HR} alone (2°C higher), although no clear differences were observed when the hPrp8 mutants-hBrr2^{HR} complexes were compared (Fig. 3.6B). Based on these results I can conclude that complex formation stabilizes hBrr2^{HR} but no appreciable effects were observed when wild-type hPrp8$^{Jab1/MPN}$ was exchanged by the mutants.

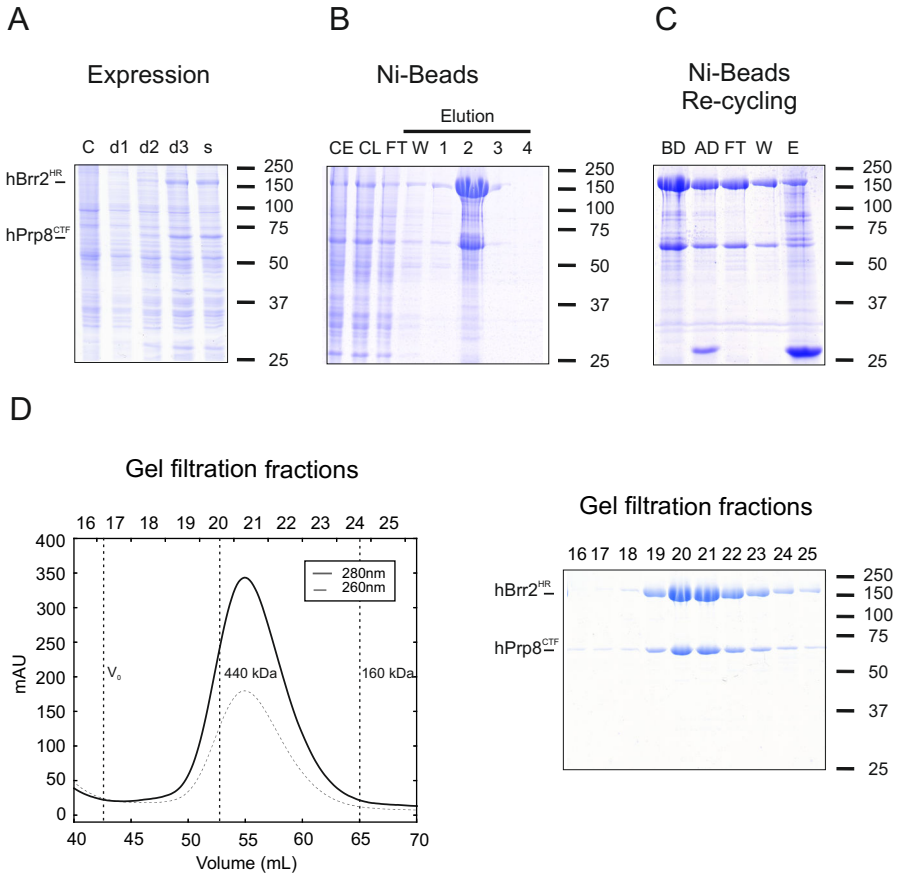

Figure 3.4: Production and purification of the hBrr2^{HR}-hPrp8^{CTF} complex in Sf9 insect cells. **(A)** Coomassie-stained SDS gel showing a time course after proliferation arrest of the production of hBrr2^{HR}-hPrp8^{CTF} complex. (**B** and **C**) Coomassie-stained SDS gel showing Ni^{+2}-NTA affinity purification. **(D)** Gel filtration chromatography on a S200 16/60 column. The molecular weight marker is displayed on the right (sizes are in kDa). C – control, d1 – whole cell extract day 1 (24h), d2 – whole cell extract day 2 (48h), d3 – whole cell extract day 3 (72h), s – soluble supernatant, CE – whole cell extract, CL – cleared lysate, FT – flow through, W – wash, BD – before digestion, AD – after digestion and E – Elution.

3.1.2 Production of *Saccharomyces cerevisiae* (yeast) proteins

3.1.2.1 yBrr2

The gene of yBrr2 full length and a fragment that includes the two helicase cassettes and an extended N-terminus which covers a region predicted as a PWI domain (yBrr2^{enHR} residues 271- 2163) were cloned into a pFL vector with a cleavable N-terminal His_{10}-tag. In this case we also worked with expression-optimized synthetic genes. The proteins were expressed in High Five insect cells [Santos, 2012]. The purification was optimized in our laboratory yielding up to 4 mg of highly pure protein per 800 ml of insect cell culture (Fig. 3.7).

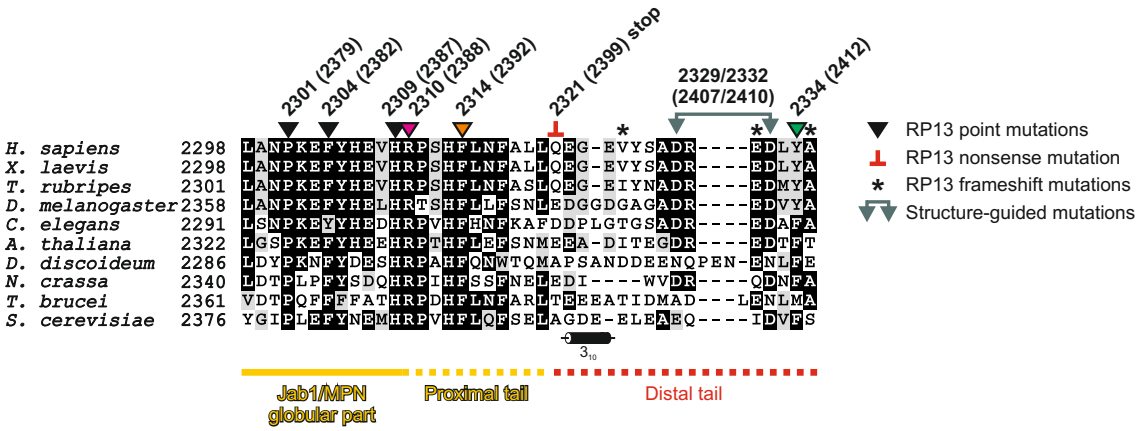

Figure 3.5: Sequence alignment of the C-terminal residues of the Prp8$^{\text{Jab1/MPN}}$ domain from various organisms. Identical residues are boxed in black and similar residues are shaded grey. Secondary structure symbol below the alignment, 3_{10} helix. Solid and dashed golden and red lines below the alignment, borders of the globular part of the Prp8$^{\text{Jab1/MPN}}$ domain, of the proximal part and of the distal part of the C-terminal tail. Triangles/T-bar/asterisks above the alignment, positions of RP13 point/nonsense/frameshift mutations. Dark green double arrow, position of conserved acidic residues mutated additionally in Mozaffari-Jovin et al. [Mozaffari-Jovin et al., 2013]. Symbols for mutations are coloured like the curves and bars in the following panels. Numbers above the alignment, human (yeast) residues.

3.1.2.2 yPrp8

The fragments of yeast Prp8, yPrp8$^{\text{CTF}}$ long (residues 1836-2413), yPrp8$^{\text{CTF}}$ short (residues 1836-2398) and yPrp8$^{\text{Jab1/MPN}}$ (residues 2147-2413) used in this thesis were expressed in *E. coli* and purified following the protocols described in previous publications [Santos, 2012; Santos et al., 2012; Weber et al., 2011]. The purified proteins were used for complex reconstitution with the yBrr2 fragments and functional studies.

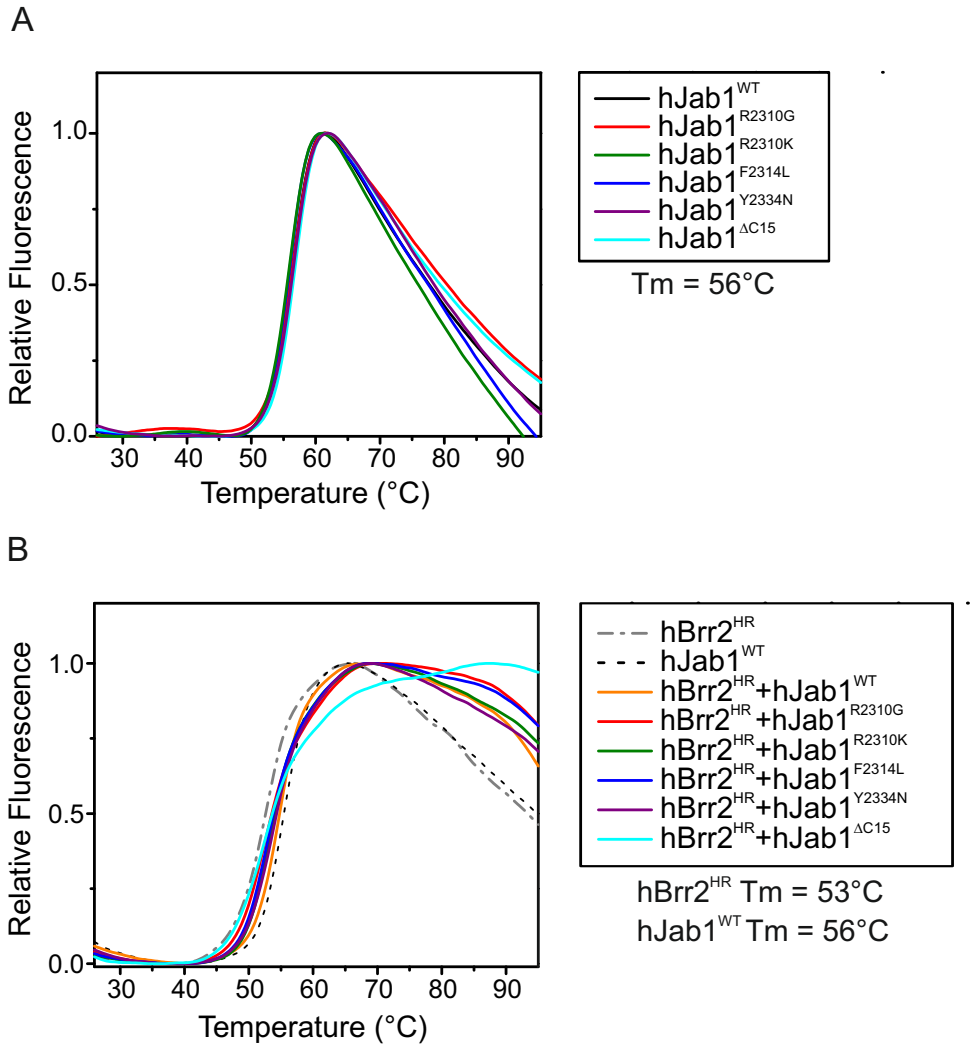

Figure 3.6: DSF analysis plot. **(A)** Normalized melting curves of all the hPrp8$^{Jab1/MPN}$ mutants used in this thesis. **(B)** Normalized melting curves of the hPrp8$^{Jab1/MPN}$ mutants in complex with hBrr2^{HR}. The wild-type hPrp8$^{Jab1/MPN}$ domain and hBrr2^{HR} alone were used as controls.

3.2 Brr2-Prp8 complex assembly and interaction studies

3.2.1 Human complexes

3.2.1.1 hBrr2^{HR}-hPrp8$^{Jab1/MPN}$ complex assembly

I tested the ability of the hPrp8$^{Jab1/MPN}$ to form a complex with hBrr2^{HR} using gel filtration chromatography. The complex hBrr2^{HR}-hPrp8$^{Jab1/MPN}$ was assembled at least 20 minutes prior to loading the sample onto the column. First, I tried an

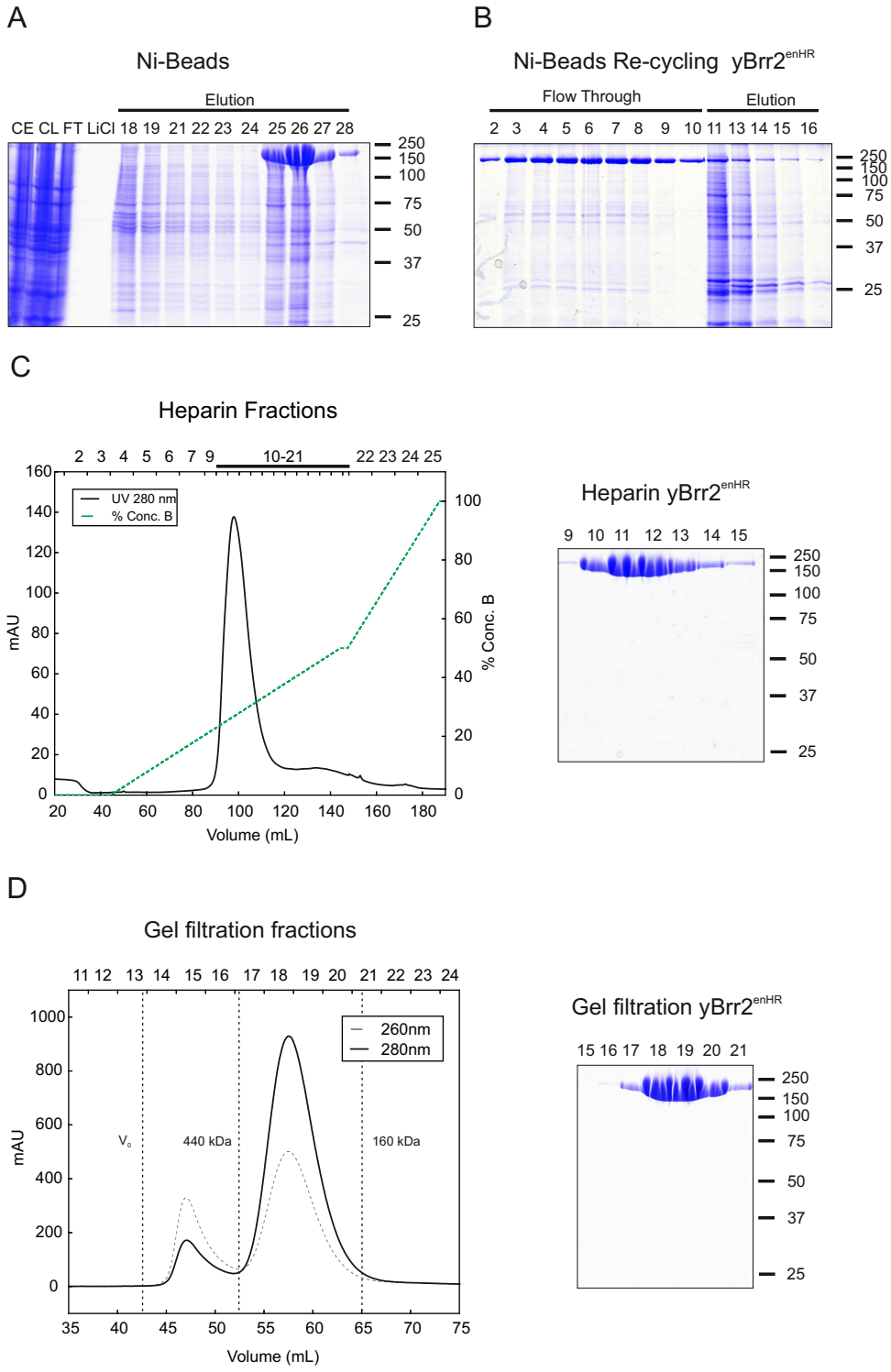

Figure 3.7: yBrr2^{enHR} purification. (**A** and **B**) Coomassie-stained SDS gel showing Ni^{+2}-NTA affinity purification. **(C)** Chromatogram and corresponding SDS gel of heparin column purification, the gel only shows the fractions covering the elution peak. **(D)** Gel filtration chromatography on a S200 16/60 column and SDS-PAGE of the peak fractions. The molecular weight marker is displayed on the right (sizes are in kDa). CE – whole cell extract, CL – cleared lysate, FT – flow through, LiCl – Lithium chloride wash, W – wash, BD – before digestion and AD – after digestion.

excess of the hPrp8$^{Jab1/MPN}$, choosing a 1:2 molar ratio (hBrr2^{HR}:hPrp8$^{Jab1/MPN}$). I could observe that the two molecules interacted and formed a stable complex which migrated as a single peak in the chromatogram, as shown in the SDS-PAGE of the run (Fig. 3.8 lower panels left and right). Another important finding was the higher affinity of the larger fragment of hPrp8$^{Jab1/MPN}$ over the degradation product, yielding a homogeneous complex containing only the longer hPrp8$^{Jab1/MPN}$ domain (Fig. 3.8 lower panel right). However, the shorter degradation fragment was still able to bind Brr2 as observed when equimolar amounts of the two proteins were used (Fig. 3.8 lower panel left), although with lower affinity than the longer protein.

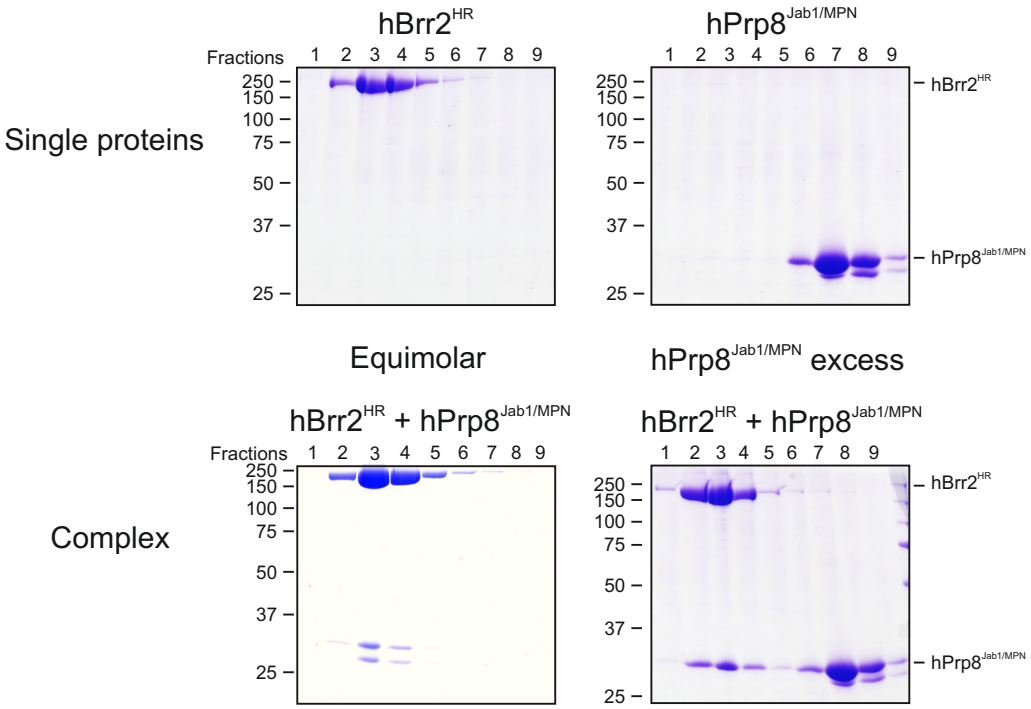

Figure 3.8: The hPrp8$^{Jab1/MPN}$ domain interacts with hBrr2 in solution and the longer fragment shows higher affinity than the shorter degradation product. Gel filtration analyses were performed with the indicated individual proteins or protein mixtures. Proteins were mixed at 1:1 and 1:2 (hBrr2^{HR}:hPrp8$^{Jab1/MPN}$) molar ratios, prior to S200 size exclusion chromatography, and proteins in the eluted fractions 1-9 (identical in each panel) were analyzed by SDS-PAGE. Numbers on the left indicate molecular weight standards in kDa.

3.2.1.2 hPrp8$^{Jab1/MPN}$ interaction studies with Brr2 cassettes

Since it was not known whether both cassettes of Brr2 were required for the interaction with the Jab1/MPN domain of Prp8, I conducted binding studies using gel filtration chromatography and the individual cassettes. These analyses showed that the hPrp8$^{Jab1/MPN}$ co-migrated exclusively with the isolated hBrr2^{NC} and no interaction was observed with hBrr2^{CC} (Fig. 3.9).

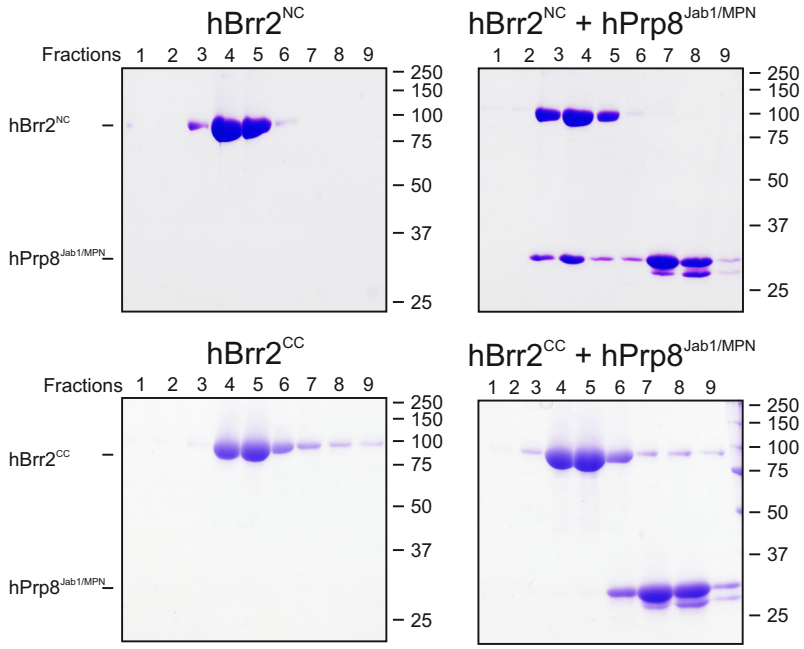

Figure 3.9: The hPrp8$^{Jab1/MPN}$ domain interacts with the N-terminal helicase cassette of hBrr2 in solution. Gel filtration analyses were performed with the indicated individual proteins or protein mixtures. Proteins were mixed at 1:2 (hBrr2 variant:hPrp8$^{Jab1/MPN}$) molar ratio prior to S200 size exclusion chromatography, and proteins in the eluted fractions 1-9 (identical in each panel) were analyzed by SDS-PAGE. Numbers on the right indicate molecular weight standards in kDa.

After the successful optimization of the purification protocols and subsequent complex assembly, I was able to use the purified complexes for crystallization trials as described in the following sections. In case of hBrr2^{NC} -hPrp8$^{Jab1/MPN}$ complex all the attempts for crystallization failed, however, favourable outcomes were obtained with the complexes containing hBrr2^{HR} (hBrr2^{HR}-hPrp8$^{Jab1/MPN}$ and hBrr2^{HR}-hPrp8^{CTF}).

3.2.2 Yeast complexes reconstitution

The yeast complexes were reconstituted using a molar excess of the Prp8 fragments. A typical ratio was 1:2 of yBrr2^{enHR}:Prp8 fragments. All formed complexes were very stable and I could observe two main peaks in the gel filtration chromatogram (not counting the aggregates peak that comes at the column's void volume, V_0), one at the expected retention volume for the chosen column for every complex with 1:1 ratio and another peak of the corresponding Prp8 fragment excess (Fig. 3.10). The SDS-PAGE of the runs showed an apparent equimolarity of the components of each complex and high purity which is suitable for crystallization screening. After gel filtration chromatography, I concentrated the complexes to 3.5 mg/ml. Any attempt to increase the concentration of the samples led to precipitation.

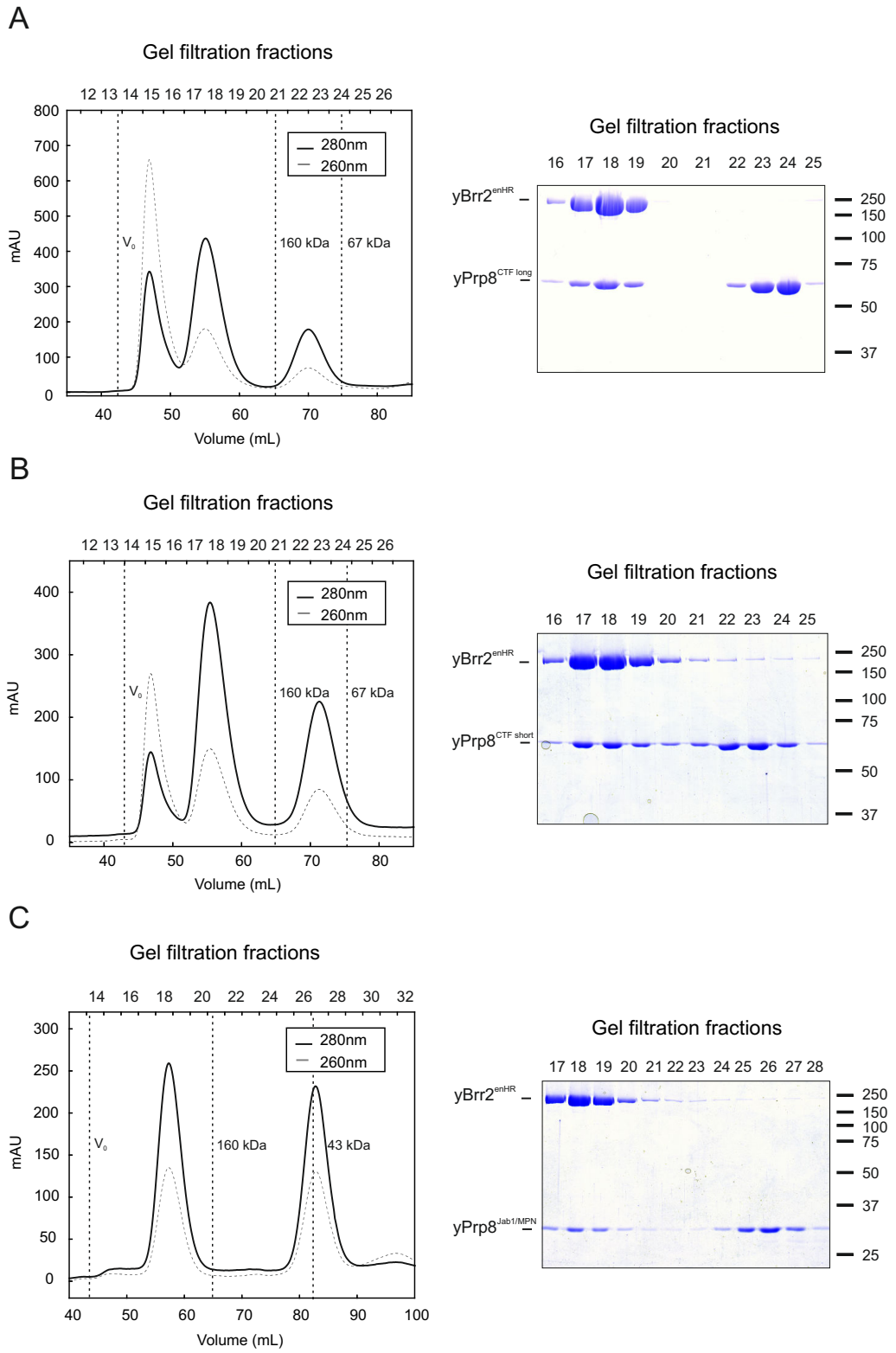

Figure 3.10: Reconstitution of yeast complexes on a S200 16/60 gel filtration column and corresponding SDS-PAGEs of the runs. **(A)** yBrr2$^{\text{enHR}}$-yPrp8$^{\text{CTF 1836-2413}}$. **(B)** yBrr2$^{\text{enHR}}$-yPrp8$^{\text{CTF 1836-2398}}$. **(C)** yBrr2$^{\text{enHR}}$-yPrp8$^{\text{Jab1/MPN}}$. Proteins were mixed at 2:1 (yPrp8 variant:yBrr2) molar ratio respectively. The molecular weight marker is shown on the right (sizes in kDa) in the SDS-PAGEs.

3.3 Crystallization and structural analysis

3.3.1 hBrr2^{HR}-hPrp8$^{Jab1/MPN}$ complex crystallization

Aiming for the elucidation of the molecular basis for hBrr2^{HR}-hPrp8$^{Jab1/MPN}$ interaction, the purified complex was subjected to crystallization. Initial screens were prepared using a crystallization robot which facilitates testing a broader spectrum of conditions. Several initial hits were obtained and I was able to successfully reproduce most of the crystallization conditions (Fig. 3.11A). The crystals were then tested on a synchrotron beamline (P14, Petra III DESY, Hamburg). The resolution of the best diffracting crystal in this initial screening was 6.0 Å which allowed us to position hBrr2^{HR} (PDB ID: 4F91 [Santos et al., 2012]) by molecular replacement and also the globular part of hPrp8$^{Jab1/MPN}$ using the yeast Prp8$^{Jab1/MPN}$ structure (PDB ID: 2OG4 [Pena et al., 2007]) as search model. The crystallization conditions were further optimized using additives as described in Material and Methods (see section 2.2.4). Other optimization techniques were also tested, like glutaraldehyde cross-linking and dehydration methods. One of these dehydration methods, which consists in the direct addition of a solution with higher precipitant concentration than the original crystallization condition to the crystallization drop and air dehydration for 30 min at 4°C (see Materials and Methods for more details), proved to be successful. Dehydration together with ADP as additive yielded the best diffracting crystal with a resolution of 3.4 Å. The images of the obtained dataset were processed using XDS (X-ray Detector Software) which allowed me to extract the reflection data (h,k,l,I,σI) from the images. The phase problem was again addressed by molecular replacement using hBrr2^{HR} and yPrp8$^{Jab1/MPN}$ structural coordinates as search models.

To establish the number of complex molecules in the asymmetric unit I calculated the Matthews coefficient for the given space group and unit cell (Table 3.1). The software offered four solutions and estimated, in terms of probability, that the best option should contain two complexes per asymmetric unit and a solvent content of 56%. However, the molecular replacement failed to place two complexes and presented a unique solution with one complex per asymmetric unit and 78% solvent content. The next step was to generate a model of hPrp8$^{Jab1/MPN}$ using the known Prp8$^{Jab1/MPN}$ structures as references and the online tool HHpred (toolkit.tuebingen.mpg.de/hhpred). The generated hPrp8$^{Jab1/MPN}$ was superimposed on the molecular replacement model using the program Coot (Crystallographic Object-Oriented Toolkit) which displays and allows the manipulation of atomic models of macromolecules using 3D computer graphics. Thus, yPrp8$^{Jab1/MPN}$ domain was replaced for the human model. The first model that came out from molecular

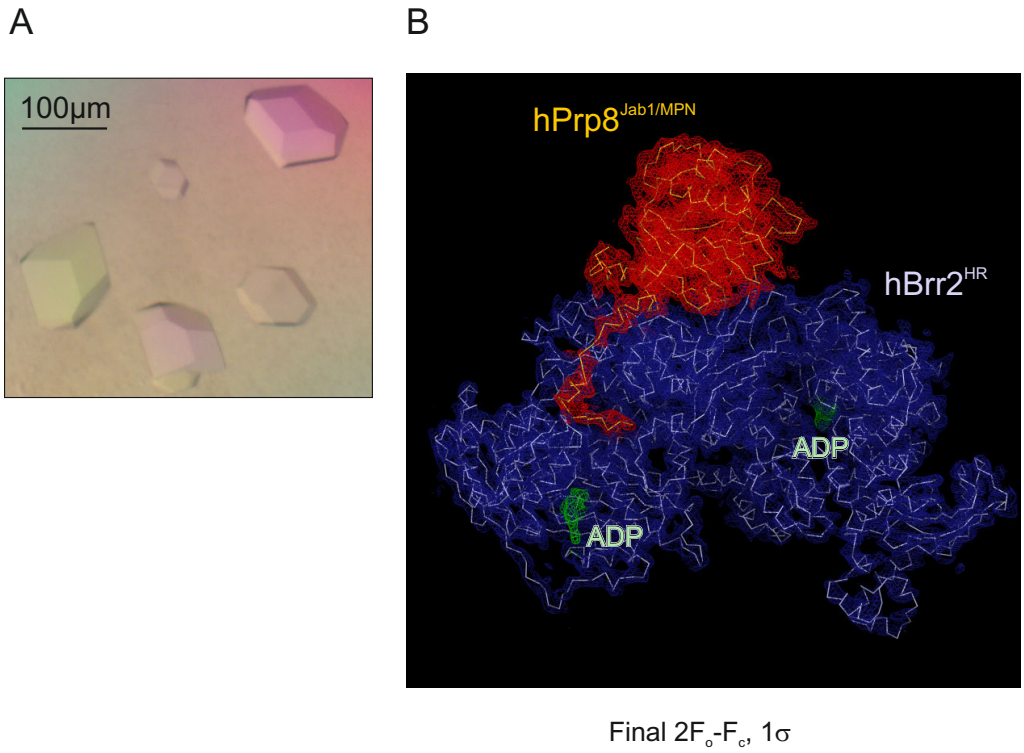

Figure 3.11: Crystals and map of hBrr2^{HR}-hPrp8$^{Jab1/MPN}$ complex. **(A)** Crystals of hBrr2^{HR}-hPrp8$^{Jab1/MPN}$ complex. **(B)** Simulated annealing composite "omit" electron density contoured at the 1σ level. Molecules are shown as Cα traces (hBrr2^{HR} – light blue, hPrp8$^{Jab1/MPN}$ – gold) or sticks (ADP – light green). ADP was added as an additive during crystallization. Density around hBrr2^{HR} – blue, density around hPrp8$^{Jab1/MPN}$ – red, density around nucleotides – green.

replacement was not completely accurate and it required the modification of the model parameters to be able to describe the experimental data as good as possible in a process called refinement. Refinement can be evaluated through the value of the "R-factor" (a measure of the residual differences between the observed and calculated patterns) and for our model it was performed using the refinement tools of the Phenix suite and DEN (Deformable elastic network) of the CNS suite (Crystallography & NMR System). The model was refined to R_{work}/R_{free} values of 19.5%/22.0% with good stereochemistry (Table 3.1).

The obtained map displayed an extra electron density entering the RNA binding tunnel of Brr2 which was used as a guide to build the C-terminal tail of the hPrp8$^{Jab1/MPN}$ that was not visible in all the known structures of the isolated Prp8$^{Jab1/MPN}$ domain. We also could observe density in the expected nucleotide binding pockets of Brr2 where I placed the ADP molecules. In order to reduce the effects of model bias, simulated annealing composite omit maps were calculated. These maps confirmed the placement of the Prp8$^{Jab1/MPN}$ domain over the N-terminal

Table 3.1: Crystallographic statistics for the hBrr2^{HR}-hPrp8$^{Jab1/MPN}$ complex

Data collection	
Wavelenght (Å)	0.976293
Space group	$P4_22_12$
Unit cell (Å)	
a,b	241.0
c	200.0
Resolution (Å)[a]	100-3.4 (3.49-3.40)
Reflections	
Unique	81343 (5960)
Completeness (%)	100 (100)
Redundancy	13.3 (13.0)
$\mathbf{R_{meas}}$	16.3 (313.5)
$\mathbf{I/\sigma}$	16.0 (1.0)

[a] Values for the highest resolution shell in parentheses.

[b] R_{meas}, intensity of the *i-th* measurement of reflection *hkl*; $\langle I(hkl)\rangle$ average value of the intensity of reflection *hkl* for all *i* measurements, *n* is the redundancy.

$$R_{meas} = \left(\frac{n}{n-1}\right)^{1/2} \frac{\sum_{hkl}\sum_i [\ |I_i(hkl) - \langle I(hkl)\rangle|\ ]}{\sum_{hkl}\sum_i I_i(hkl)}$$

Refinement	
Resolution (Å)	57.0-3.4 (3.44-3.4)
Reflections	
Unique	81307 (2649)
Completeness (%)	99.8 (100)
Test set (%)	5
R-factors[c]	
R_{work}	19.5 (33.3)
R_{free}	22.0 (34.1)
Ramachandran plot	
Favoured	94.97
Allowed	4.98
Outlier	0.05
Rmsd[d] **geometry**	
Bonds (Å)	0.006
Angles (°)	1.07
PDB ID:	4KIT

[c] R-factors:

$$R_{work} = \frac{\sum_{hkl}[\ ||F_{obs}| - k|F_{calc}||\ ]}{\sum_{hkl}|F_{obs}|}$$

$$R_{free} = \frac{\sum_{hkl\in T}[\ ||F_{obs}| - k|F_{calc}||\ ]}{\sum_{hkl\in T}|F_{obs}|}$$

$hkl \in T$ - test set; F_{obs}, F_{calc} - observed and calculated (from model) structure factor amplitudes.

[d] Rmsd, Root-mean-square deviation.

cassette of hBrr2^{HR} and also corroborated the extra electron density observed for hPrp8$^{Jab1/MPN}$ C-terminal tail and the ADP molecules (Fig. 3.11B). However, due to the low resolution, we could only generate a rough model with some gaps and without a clear density for many of the amino acid side chains.

3.3.2 hBrr2^{HR}-hPrp8$^{Jab1/MPN}$ complex structural analysis

The generated model showed that, in the complex, hPrp8$^{Jab1/MPN}$ directly interacts with all six domains of the N-terminal hBrr2 cassette, but does not contact the C-terminal cassette (Fig. 3.12), consistent with our binding studies in solution

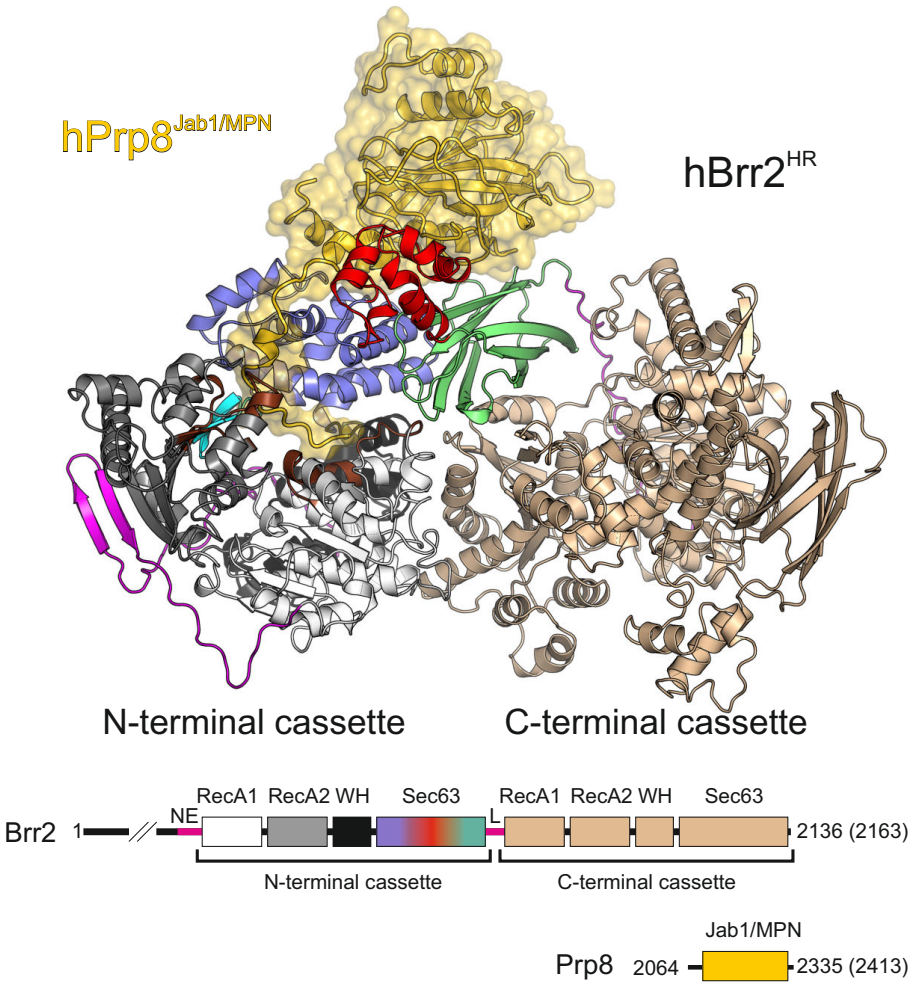

Figure 3.12: Structural overview. The N-terminal cassette of hBrr2^{HR} is coloured by domain. RecA-1 – light grey, RecA-2 – dark grey, WH – black, HB – blue, HLH – red, IG – green, N-terminal extension (NE) and linker – magenta, separator loop – cyan, RNA-binding motifs of the N-terminal RecA-1 and RecA-2 domains – brown, C-terminal cassette – beige. hPrp8$^{Jab1/MPN}$ is shown as a gold ribbon with a semi-transparent surface.

(Fig. 3.9). Similar to the structure of isolated Prp8$^{Jab1/MPN}$ from yeast (PDB ID: 2OG4) and *C. elegans* (PDB ID: 2P87) [Zhang et al., 2007], hPrp8$^{Jab1/MPN}$ contains a β-barrel core with N and C-terminal expansions. One flank of the hPrp8$^{Jab1/MPN}$ domain, including the C-terminal expansion, rests on the exposed β-sheet surface of the N-terminal IG domain of hBrr2^{HR} and an N-terminal helix runs along one edge of the hBrr2^{HR} HB domain (Fig. 3.12). Strikingly, the C-terminal tail of the hPrp8$^{Jab1/MPN}$ domain which was flexible and only partially resolved in the structures of the isolated domains, binds along a cleft between the N-terminal HLH and HB domains, and then turns towards the interior of the N-terminal cassette, where it runs between the RecA-2 and HB domains. It continues along a surface of the RecA-1 domain and ultimately interacts via its very C-terminus with the RecA-1, WH and HB domains (Fig. 3.13A).

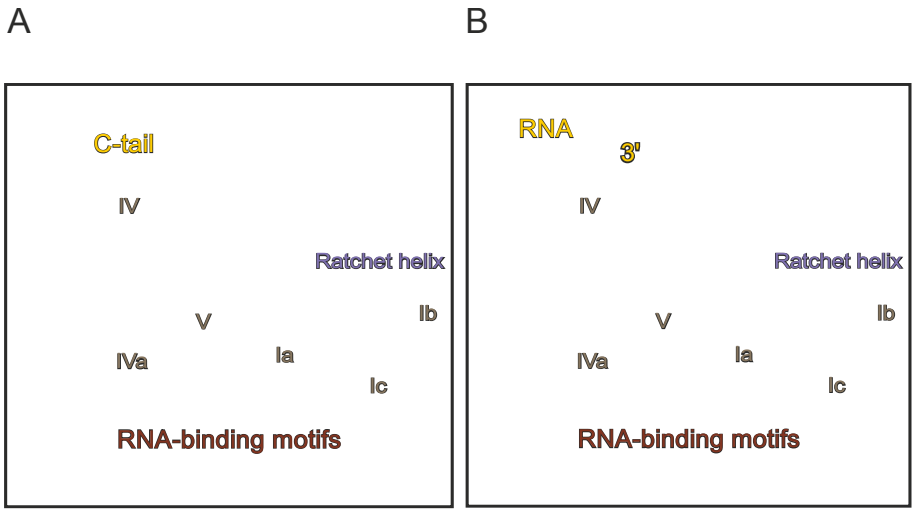

Figure 3.13: (A) Close-up of the C-terminal tail of the hPrp8$^{Jab1/MPN}$ domain running across all canonical RNA-binding motifs of the N-terminal RecA-1 (motifs Ia, Ib and Ic) and RecA-2 (motifs IV, IVa and V) domains. Orientation and color scheme are as in Fig. 3.12. **(B)** Model of an RNA bound in the central tunnel. In isolated hBrr2^{HR}, the central RNA binding tunnel of the N-terminal cassette is non-covalently closed at one side by contacts between the RecA-2 and HB domains, and an extended loop of the RecA-2 domain (the separator loop) extends across its entrance. The RecA-2 and HB domains have to separate intermittently to allow Brr2 to bind an internal single-stranded region of its U4/U6 RNA duplex target in this tunnel next to the separator loop. In the present complex, the last eleven residues of the hPrp8$^{Jab1/MPN}$ C-terminal tail extensively bridge the RecA-2 and HB domains, stabilizing their interaction. Additionally, they cover the canonical RNA-binding surfaces of the two RecA domains, directly contacting five of the six conserved RNA-binding motifs (Ia, Ib, Ic in RecA-1; IV, IVa in RecA-2).

In the present conformation, the C-terminal tail of hPrp8$^{Jab1/MPN}$ is positioned in the RNA binding tunnel of hBrr2^{HR} which contains all the conserved binding motifs for nucleic acid recognition (see Introduction 1.8.1.2) and is expected to directly compete with RNA loading and accommodation (Fig. 3.13B).

The interaction between the proteins is supported by surface charge complementarity, the tail is predominantly negatively-charged which is suitable for binding the RNA tunnel of Brr2, where, under a different situation, the negatively-charged sugar-phosphate backbone of the RNA substrate can be accommodated (Fig. 3.14A and B). 5286 Å^2 of surface area are buried upon complex formation, 2915 Å^2 between hBrr2^{HR} and the globular part of hPrp8$^{Jab1/MPN}$ (residues 2067-2315) and 2371 Å^2 between hBrr2^{HR} and the C-terminal tail of hPrp8$^{Jab1/MPN}$ (residues 2316-2335).

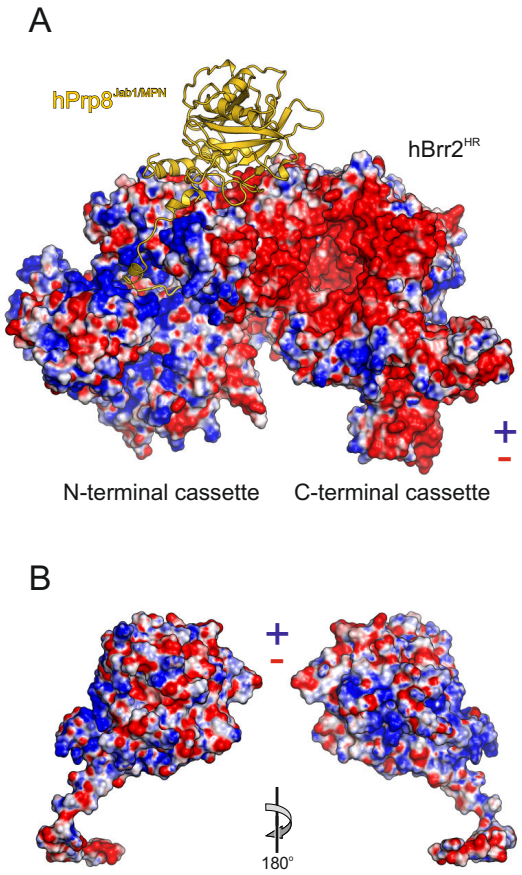

Figure 3.14: (A) hBrr2^{HR} electrostatic surface (positive – blue, negative – red) with the hPrp8$^{Jab1/MPN}$ domain shown as a golden ribbon. Rotated 40° about the horizontal axis (top to front) compared to Fig. 3.12. **(B)** Electrostatic surface of the hPrp8$^{Jab1/MPN}$ domain in two orthogonal views. The left orientation is the same as in (A).

3.3.2.1 Conformational changes upon complex formation

Both proteins undergo conformational changes upon complex formation. However, the extent of the changes observed for hPrp8$^{Jab1/MPN}$ must be evaluated considering that a model for the human unbound protein does not exist and 3.4 Å resolution is not enough to correctly assign the side chains of all the residues. Therefore, the observed deviations are based on the available structures of the domain in different species which share a high similarity with the human protein (74% identity with hPrp8$^{Jab1/MPN}$ domain of *C. elegans* and 45% identity with the yeast ortholog Fig. 3.15). If we use these structures as reference, the only significant deviations were observed in elements that directly contact hBrr2^{HR} (Fig. 3.15 arrows).

In case of hBrr2^{HR} the conformational changes could be best seen using the apo structure of the protein as reference. The observed changes can be divided into two groups, local changes (regions in direct contact with hPrp8$^{Jab1/MPN}$ domain) and global changes.

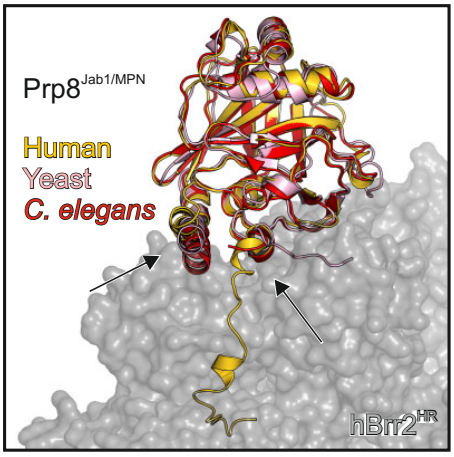

Figure 3.15: Comparison of Prp8$^{Jab1/MPN}$ domains. Prp8$^{Jab1/MPN}$ domains from yeast (pink; PDB ID 2OG4) and C. elegans (red; PDB ID 2P87) superimposed on the hPrp8$^{Jab1/MPN}$ domain (gold) in complex with hBrr2^{HR} (grey semi-transparent surface). Arrows indicate regions that undergo conformational adjustments upon binding to hBrr2^{HR}. Rotated 40° clockwise about the vertical axis and 20° about the horizontal axis (top to front) compared to Fig. 3.12.

Local changes: The individual domains retain similar structures in the complex as in the isolated protein (rmsd's ranging from 0.51 Å for the N-terminal IG domains to 1.13 Å for the N-terminal RecA-2 domains). However, we observe changes in the positioning of certain domains and local conformational adjustments in a number of elements upon hPrp8$^{Jab1/MPN}$ binding. The entire HLH domain, the front end of the HB domain and the upper tip of the RecA-2 domain are pushed away from each other in order to accommodate the C-terminal tail of hPrp8$^{Jab1/MPN}$ (Fig. 3.16 top arrows). In addition, the conformations of the RNA-binding motifs IVa and Ib adapt to the central and terminal portions of the C-terminal tail, respectively (Fig. 3.16 bottom). Thus, hBrr2^{HR} and hPrp8$^{Jab1/MPN}$ bind to each other by a mixed lock-and-key and induced fit interaction mode. Nevertheless, no structural rearrangements are seen in those regions of the N-terminal HB and IG domains of hBrr2^{HR} that are contacted by the globular part of hPrp8$^{Jab1/MPN}$.

Global changes: The most evident change was the entire Brr2 N-terminal cassette motion relative to the C-terminal cassette. Upon hPrp8$^{Jab1/MPN}$ binding, the N-terminal cassette rotates by ca. 5° relative to the C-terminal cassette about a pivot point in the inter-cassette linker (Fig. 3.17A). Due to this rotation, the N-terminal RecA-1 and WH domains are separated from the C-terminal RecA-2 domain and the interface between the cassettes is reduced by ca. 100 Å^2. Several residues in the N-terminal cassette, such as Arg603 and Arg637, which upon mutation to alanine strongly interfere with Brr2-mediated U4/U6 unwinding [Santos et al., 2012], lose contacts to neighboring residues in the C-terminal cassette (Fig. 3.17B and C).

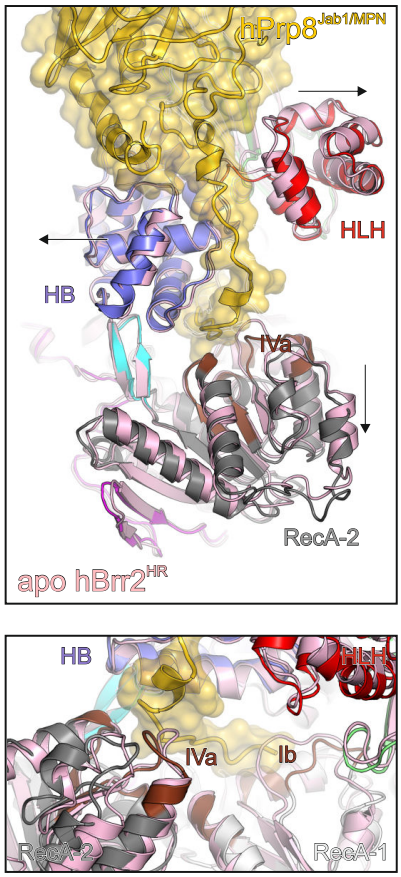

Figure 3.16: Structural changes in hBrr2^{HR} upon binding of hPrp8$^{Jab1/MPN}$. Top panel: domain movements. The hBrr2^{HR}-hPrp8$^{Jab1/MPN}$complex is coloured as before (Fig. 3.12), isolated hBrr2^{HR} is shown in pink. Rotated 70° clockwise about the vertical axis and 50° about the horizontal axis (top to front) compared to Fig. 3.12. Bottom panel: local conformational adjustments. Rotated 30° about the horizontal axis (top to front) compared to Fig. 3.12.

These observations have implications for the mechanism by which the two helicase cassettes cooperate in Brr2 (see Discussion 4.1.1) [Santos et al., 2012].

3.3.3 Functional studies of Brr2^{HR}-Prp8$^{Jab1/MPN}$ interaction†

As we could observe in the complex structure, the C-terminal tail of hPrp8$^{Jab1/MPN}$ is positioned in the RNA binding tunnel of hBrr2^{HR} and is expected to directly compete with RNA loading and accommodation (Fig. 3.13B and C). To test this notion, we monitored Brr2 association with U4/U6 RNA duplex by His-pulldown assays mainly in the yeast system. Efficient pulldown of U4/U6 RNA duplex was observed with yBrr2 alone (Fig. 3.18A). However, in the presence of yPrp8$^{Jab1/MPN}$,

†Conducted by Sina Mozaffari-Jovin

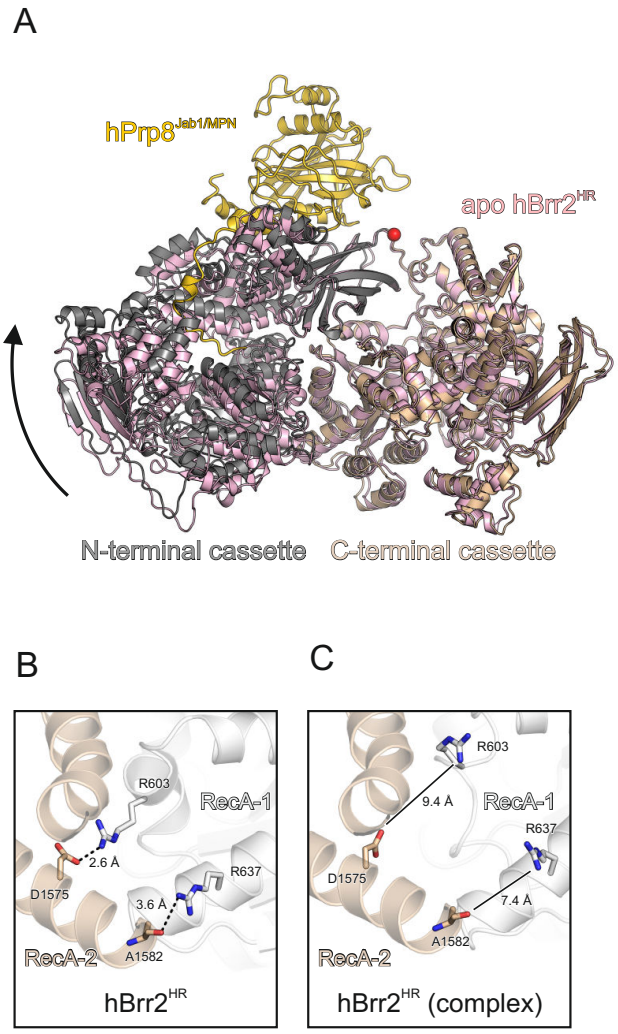

Figure 3.17: **(A)** Superposition of isolated hBrr2^{HR} (pink) and the hBrr2^{HR} - hPrp8$^{Jab1/MPN}$ complex (N-terminal cassette – grey, C-terminal cassette – beige, hPrp8$^{Jab1/MPN}$ – gold) according to the C-terminal cassettes of the hBrr2^{HR} molecules. The arrow indicates the movement of the N-terminal cassette about a pivot point in the linker (red sphere) relative to the rigid C-terminal cassette. Same orientation as in Fig. 3.12. **(B)** Close-up of contacts between the N-terminal RecA-1 domain (light grey) and the C-terminal RecA-2 domain (beige) which upon mutation led to severely reduced helicase activity in hBrr2^{HR}. Dashed lines indicate hydrogen bonds or salt bridges. Contacting residues are labeled and contact distances are indicated. Rotated 180° about the vertical axis compared to (A). **(C)** The same region of hBrr2^{HR} in the hPrp8$^{Jab1/MPN}$ complex. Corresponding distances as in left box are indicated. Rotated 180° about the vertical axis compared to (A).

only a very low amount of U4/U6 was co-precipitated with yBrr2, without (lane 3) or with added ATPγS (lane 5), demonstrating that yPrp8$^{Jab1/MPN}$ inhibits Brr2 interaction with U4/U6. A variant of yPrp8$^{Jab1/MPN}$ lacking the last 16 amino acids (yPrp8$^{Jab1/MPN\text{-}\Delta C16}$) still bound to Brr2, but U4/U6 RNA duplex was efficiently co-precipitated (Fig. 3.18A).

We next tested whether Prp8$^{Jab1/MPN}$ also inhibits Brr2's helicase activity,

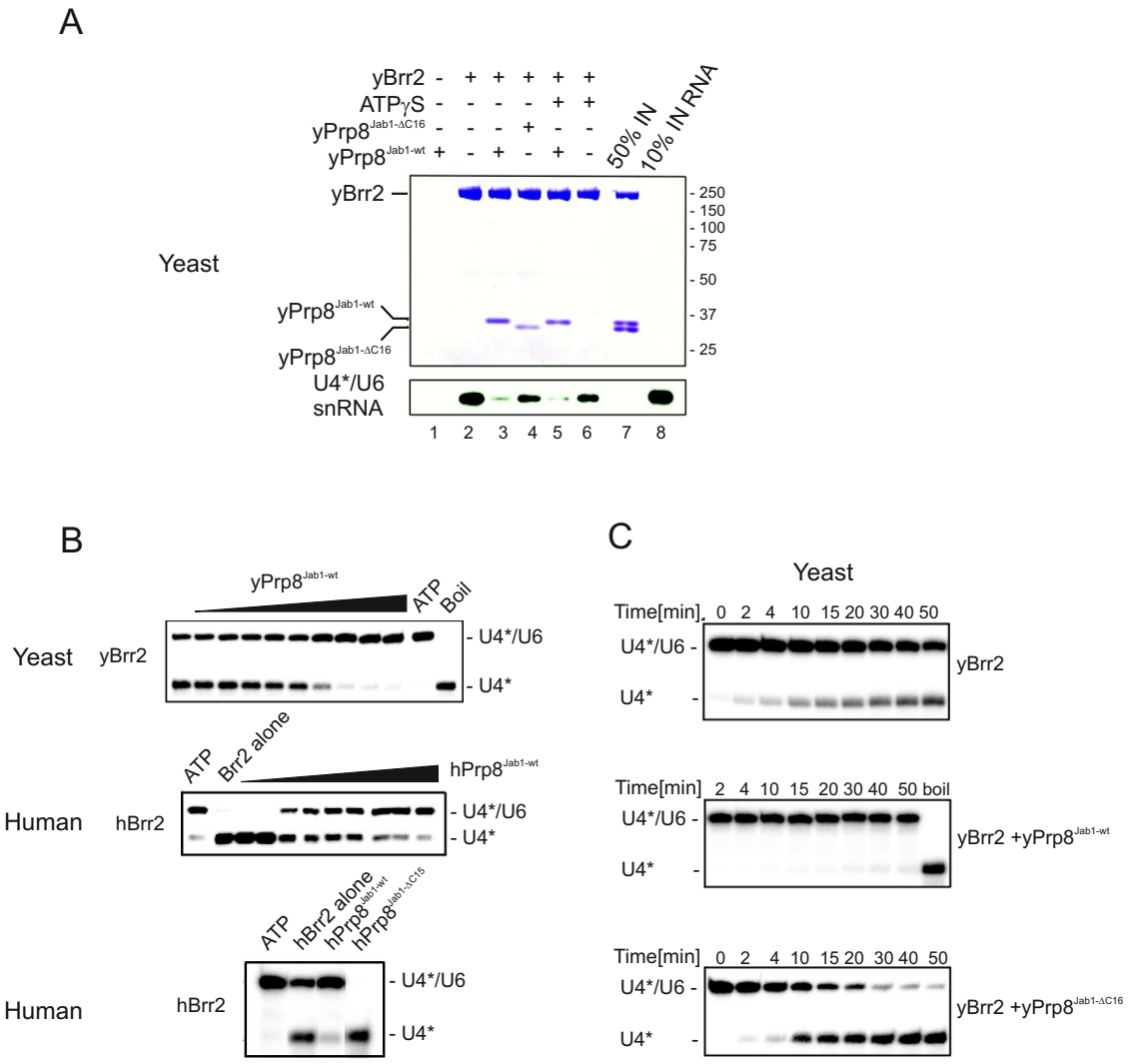

Figure 3.18: RNA binding and unwinding studies. **(A)** His-Brr2 pulldown assay using yeast proteins in the absence (lanes 1-4) or presence (lanes 5 and 6) of 2 mM ATPγS. Lane 7 - 50% input of all proteins. Lane 8 - 10% of input RNA. **(B)** yBrr2 (200 nM; upper panel), $hBrr2^{HR}$ (100 nM; middle panel) were preincubated with 0-2 μM $Prp8^{Jab1/MPN}$. U4/U6 duplex (0.5 nM, U4 labelled) was subsequently added, followed by 1 mM ATP/MgCl2 to start the reaction, which was then incubated at 20°C for 40 min. U4/U6 unwinding was analyzed by 6% native PAGE and visualised by autoradiography. Lower panel, single point unwinding assays using the human proteins, 100 mM Brr2, 1 μM $Prp8^{Jab1/MPN}$, 0.5 nM RNA duplex for 20 min at 20°C.**(C)** Time course of U4/U6 duplex unwinding under multiple turnover conditions by Brr2 in the absence (upper panel) or presence of $yPrp8^{Jab1/MPN}$ (middle panel) or $Prp8^{Jab1/MPN\text{-}\Delta C16}$ (lower panel). 200 nM Brr2, 0.5 nM RNA duplex, U4 radiolabelled. The reaction was initiated by the addition of ATP and incubated for the indicated times at 20°C.

assaying U4/U6 unwinding by Brr2 in the presence of increasing concentrations of $Prp8^{Jab1/MPN}$ for proteins of both species (Fig. 3.18B). Unwinding was also tested as a function of time in the presence of $Prp8^{Jab1/MPN}$ or $Prp8^{Jab1/MPN\text{-}\Delta C16}$ (Fig. 3.18C) which confirmed the tail's inhibition and the ability of the globular part to stimulate Brr2's helicase activity. When the activity was quantified we could observe that, if

yBrr2 (100 nM) was in large excess over the U4/U6 RNA duplex (0.5 nM), U4/U6 was less efficiently unwound in the presence of Prp8$^{Jab1/MPN}$ compared to Brr2 alone (Fig. 3.19A). Nearly identical results were obtained with hBrr2^{HR} and hPrp8$^{Jab1/MPN}$ (Fig. 3.18B). In contrast, yPrp8$^{Jab1/MPN\text{-}\Delta C16}$ stimulated yBrr2's U4/U6 unwinding activity (Fig. 3.18C and 3.19A). These results indicate that the C-terminal amino acids of Prp8$^{Jab1/MPN}$ can negatively regulate Brr2 helicase activity by preventing accommodation of its RNA substrate. Removal of the Prp8$^{Jab1/MPN}$ tail not only enhances the affinity of Brr2 for RNA but it might also restore functional contacts between Brr2's N- and C-terminal cassettes, relieving the potentially unfavourable intercassette conformation that arises upon interaction of the Prp8$^{Jab1/MPN}$ C-terminal tail.

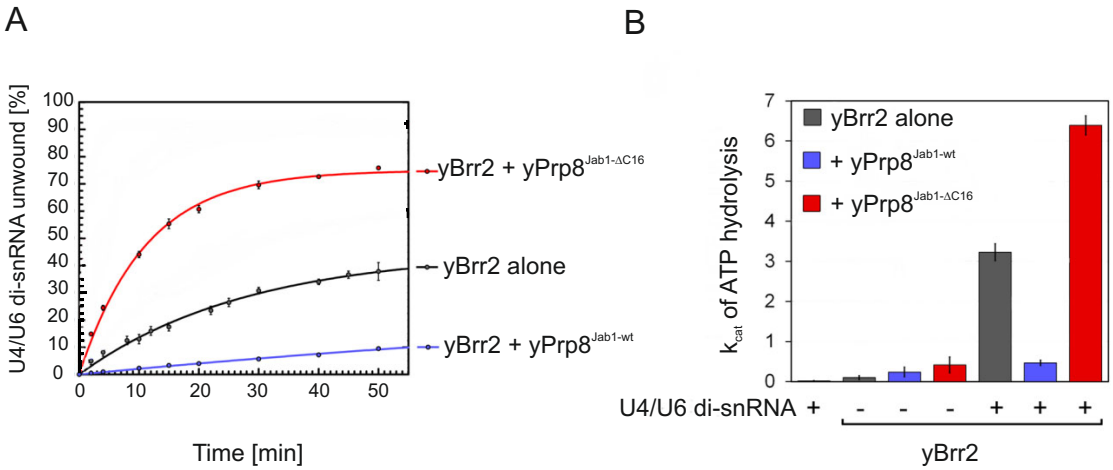

Figure 3.19: Helicase and ATPase activity of yBrr2. **(A)** Quantified data from (3.18) fitted to a single exponential equation: % duplex unwound = A1-exp(-k_u t); A, amplitude of the reaction; k_u, apparent first-order rate constant for unwinding; t, time. Kinetic parameters: k_u(Brr2) = 0.070 ± 0.009 min^{-1}, A(Brr2) = 32.1 ± 1.3; k_u(Brr2/Prp8$^{Jab1/MPN}$) = 0.064 ± 0.002 min^{-1}, A(Brr2/Prp8$^{Jab1/MPN}$) = 63.3 ± 0.9; k_u(Brr2/Prp8$^{Jab1/MPN\text{-}\Delta C16}$) = 0.59 ± 0.04 min^{-1}, A(Brr2/$^{Jab1/MPN\text{-}\Delta C16}$) = 83.7 ± 1.3. **(B)** Intrinsic (- U4/U6 RNA duplex) and RNA-stimulated (+ U4/U6 di-snRNA) steady state ATPase activity of Brr2 alone (black bars) or in the presence of Prp8$^{Jab1/MPN}$ (blue bars) or Prp8$^{Jab1/MPN\text{-}\Delta C16}$ (red bars). Error bars represent SEMs of at least two independent experiments. Turnover (k_{cat}) for U4/U6-stimulated ATPase: k_{cat}(Brr2) ≈ 3; k_{cat}(Brr2-$^{Jab1/MPN}$) ≈ 0.4; k_{cat}(Brr2-$^{Jab1/MPN\text{-}\Delta C16}$) ≈ 6.

To further investigate how Prp8$^{Jab1/MPN}$ modulates Brr2, we performed Brr2 ATPase assays under steady-state conditions. In the absence of U4/U6 di-snRNA, Brr2 exhibited very low ATPase activity on which Prp8$^{Jab1/MPN}$ had little effect (Fig. 3.19B). The addition of U4/U6 di-snRNA strongly stimulated Brr2's ATPase activity and Prp8$^{Jab1/MPN}$ reduced the rate of RNA-stimulated ATP hydrolysis ca. 8-fold (Fig. 3.19B). Given that Brr2's ATPase activity is largely driven by the interaction of RNA with the conserved helicase motifs inside the Brr2 central tunnel, these

results are consistent with the reduced Brr2 RNA binding observed in the presence of Prp8$^{Jab1/MPN}$. Consistently, Prp8$^{Jab1/MPN\text{-}\Delta C16}$ enhanced RNA-stimulated ATP hydrolysis by Brr2 ca. 2-fold (Fig. 3.19A).

3.3.4 Prp8$^{Jab1/MPN}$ RP related mutations

Based on the crystal structure we could determine the molecular basis for the RP phenotype and divide the residues into three groups: group I or globular part (residues 2064-2309), group II or proximal tail (residues 2310-2320) and group III or distal tail (residues 2321-end) (Fig. 3.5 and Fig. 3.20).

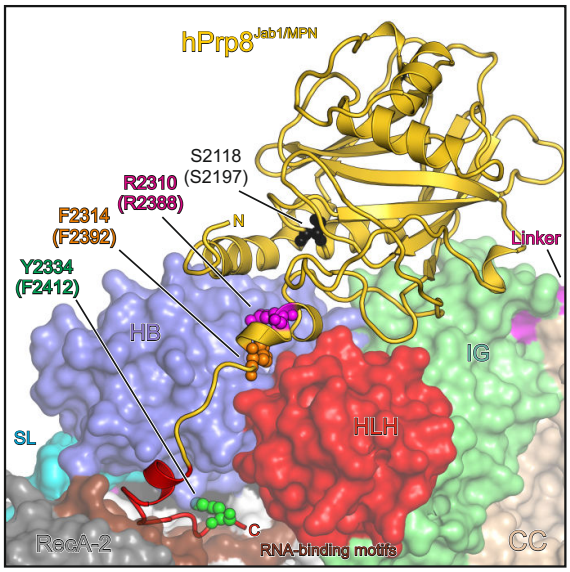

Figure 3.20: Overview of the location of RP13-linked residues in hPrp8$^{Jab1/MPN}$ in the context of hBrr2^{HR}. Domain colouring as before. RP13-linked residues investigated here are shown as spheres and coloured as the corresponding elements in previous Figures: S2118 (2197) – black, R2310 (2388) – magenta, F2314 (2392) – orange, Y2334 (2412) – green. Human (yeast) numbering. The distal C-terminal tail, removed by the Q2321stop (A2399stop) RP13 mutation, is shown in red. Rotated 30° about the horizontal axis (top to front) compared to Fig. 3.12.

The residues of group I lie in the globular Prp8$^{Jab1/MPN}$ region (S2118, P2301, F2304 or H2309; Fig. 3.20, 3.21A) and maintain structurally important interactions within the hPrp8$^{Jab1/MPN}$ domain but do not directly contact hBrr2^{HR}. Upon mutation of these residues in yeast (H2309P, H2387P) severe growth defects have been observed [Boon et al., 2007; Maeder et al., 2008] and also *in vivo* splicing defects at 37 °C [Mozaffari-Jovin et al., 2013]. Thus, the observed growth defects likely arise from defects in pre-mRNA splicing triggered by reduced solubility of the hPrp8$^{Jab1/MPN}$ mutants which consistently lead to reduced formation of U4/U6.U5 tri-snRNPs [Mozaffari-Jovin et al., 2013]. The reduced solubility of the mutants was

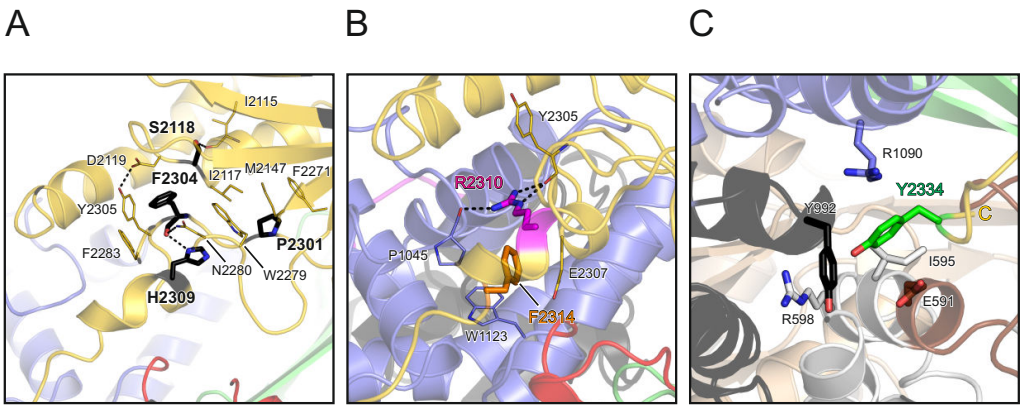

Figure 3.21: **(A)** Environment of the RP13-linked group I residues S2118, P2301, F2304 and H2309 (S2197, P2379, F2382 and H2387 in yeast), which are part of the globular core of the hPrp8$^{Jab1/MPN}$ domain. Dashed lines, hydrogen bonds or salt bridges. RP13-linked residues, thick sticks; neighboring amino acids, thin sticks. Coloured by atom type (carbon, residue-specific; nitrogen, blue; oxygen, red). Rotated 40° about the horizontal axis (top to front) compared to Fig. 3.12. **(B)** Environment of RP13-linked group II residues R2310 and F2314 (R2388 and F2392 in yeast), which engage in intramolecular interactions and in contacts to hBrr2^{HR}. Rotated 60° about the horizontal axis (top to front) compared to Fig. 3.12. **(C)** Environment of RP13-linked group III residue Y2334 (F2412 in yeast) contacting hBrr2^{HR} deep inside the central tunnel of the N-terminal cassette. Rotated 90° counter-clockwise about the vertical axis compared to Fig. 3.12.

also observed in this work during the initial attempts of expression and purification of the RP mutants (see section 3.1.1.5).

The residues of group II are contained in the proximal part of the C-terminal tail (R2310 and F2314; Fig. 3.20, 3.21B), which interact within hPrp8$^{Jab1/MPN}$ and with hBrr2 along the cleft between the HLH and HB domains. Regarding the role of the R2310 residue, it seems to mediate structurally important intra-Jab1 contacts and engage in direct interactions with the N-terminal HB domain of hBrr2^{HR}. This role appears to be shared by residue F2314 (Fig. 3.21). Mutation of the corresponding residues in yeast, R2388K and F2392L, led to reduced yeast growth at elevated temperature (37 °C) [Boon et al., 2007; Maeder et al., 2008], reduced formation of U4/U6.U5 tri-snRNPs [Mozaffari-Jovin et al., 2013] and, in case of the R2388K mutation, reduced *in vivo* splicing [Mozaffari-Jovin et al., 2013]. Mutations in group II residues also interfered with Jab1-mediated Brr2 modulation [Mozaffari-Jovin et al., 2013]. Thus, misregulation of Brr2 likely still contributes to the RP13 phenotypes in cases where lower amounts of tri-snRNPs are still formed.

Group III includes the residues of the distal part of the tail (Q2321stop and Y2334; Fig. 3.20, 3.21C), which interact with Brr2's RNA binding tunnel. The only point mutation of this group in yeast (F2412N) exhibited yeast growth and *in vivo* splicing defects, but did not interfere with U4/U6.U5 tri-snRNP formation [Mozaffari-Jovin

et al., 2013]. It also had no apparent effect on Prp8$^{Jab1/MPN}$ solubility. Mutations in the distal tail seem to reduce the negative effect on Brr2's U4/U6 unwinding as observed in the functional studies with yPrp8$^{Jab1/MPN\text{-}\Delta C16}$, hPrp8$^{Jab1/MPN\text{-}\Delta C15}$ (see section 3.3.3) and the yeast mutant F2412N [Mozaffari-Jovin et al., 2013].

3.3.4.1 Effects of RP hPrp8$^{Jab1/MPN}$ point mutations on the interaction with hBrr2^{HR}

To investigate the mutated hPrp8$^{Jab1/MPN}$'s behaviour upon hBrr2^{HR} binding, I conducted interaction studies through analytical gel filtration with each of the soluble mutants. I could not observe any differences in the binding capacity of the mutated proteins in the used buffer (20mM Tris, 200 mM NaCl, 1mM DTT pH 7.5) (Fig. 3.22). Therefore, we tested increasing salt concentrations. Only at 1 M NaCl, it was possible to see a clear effect in one of the mutants. R2310G lost binding to hBrr2^{HR}, while the other mutants were still able to bind (Fig. 3.23). Based on this observation, it is possible to conclude that mutant R2310G has a lower affinity than wild-type hPrp8$^{Jab1/MPN}$. Strikingly, the yeast protein carrying the mutation R2388K (equivalent to R2310K) showed reduced interaction with hBrr2^{HR} in gel filtration even at milder salt condition (200 mM NaCl) [Mozaffari-Jovin et al., 2013].

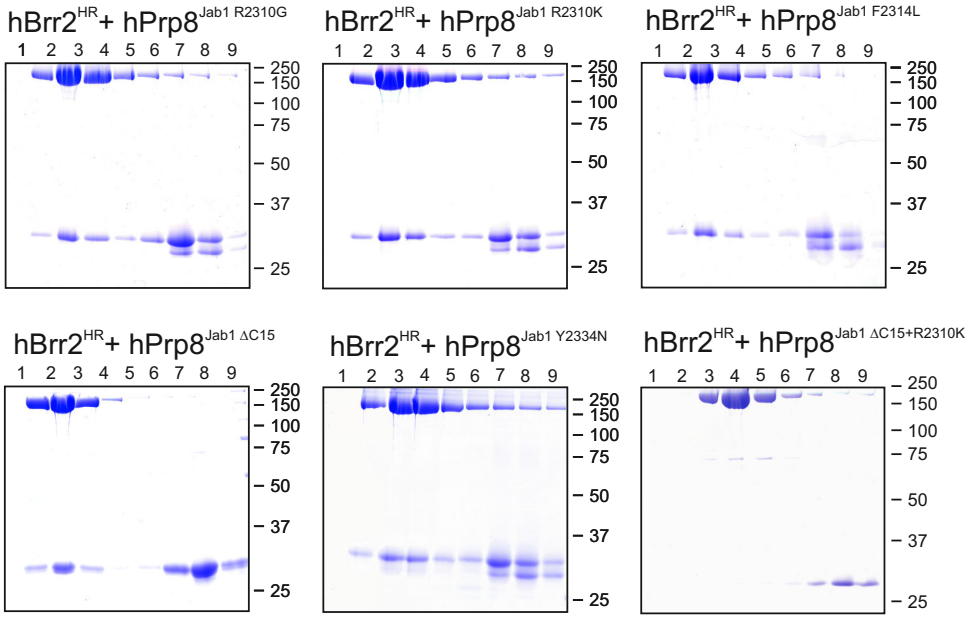

Figure 3.22: Interaction of various hPrp8$^{Jab1/MPN}$ mutants with hBrr2^{HR} monitored by gel filtration. Nine fractions representing the same elution volume were analyzed for each chromatogram. Numbers on the right correspond to the molecular weight standards in kDa.

The Prp8 protein is extremely conserved across species and the residues critical for the interaction with Brr2 are present in both ortholog Prp8$^{Jab1/MPN}$ domains.

Nevertheless, a mild mutation can produce a dramatic reduction in interaction between the yeast proteins, while the human proteins still binds. This result suggests a slightly different interaction interface between $Prp8^{Jab1/MPN}$ and Brr2 of the two species, which is not apparent when the yeast and human complex structures are compared. As a proof of the general conservation of the binding mode, $hPrp8^{Jab1/MPN}$ is able to interact with yBrr2 (Fig. 3.24 upper left panel). However, $yPrp8^{Jab1/MPN}$ does not cross-interact with $hBrr2^{HR}$ (Fig. 3.24 upper right panel). The Brr2 proteins of each organism are less conserved (sharing 38% identity between human and yeast) and they may present important differences in residues responsible for stabilizing the complex interaction.

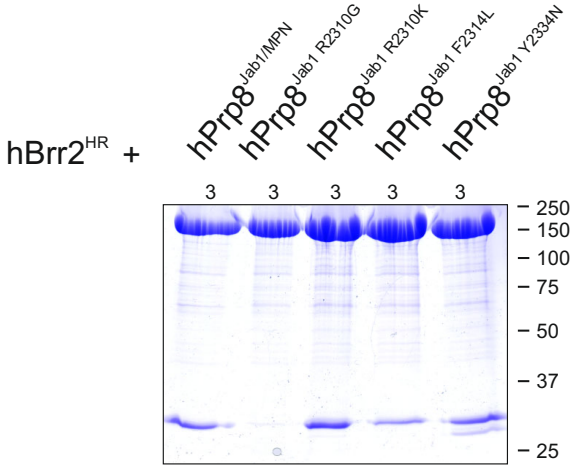

Figure 3.23: Comparative SDS-PAGE showing the main peak fraction of gel filtration runs for each of the $hPrp8^{Jab1/MPN}$ mutants together with $hBrr2^{HR}$ conducted in presence of 1 M NaCl. Numbers on the right correspond to the molecular weight standards in kDa.

3.3.4.2 Analysis of the effects of the $hPrp8^{Jab1/MPN}$ C-terminal tail

Using the more sensitive yBrr2-$hPrp8^{Jab1/MPN}$ cross-interaction system, we could observe that the removal of the tail considerably reduced the binding to yBrr2 and the mutation on residue 2310 abrogated it (Fig. 3.24 lower panels), as observed previously for the yeast system. As a rough way of delimiting the region which favoured increased interaction in the human system, we produced a $hPrp8^{Jab1/MPN}$ mutant containing the R2310K mutation and lacking the last 15 residues of the C-terminal tail. The double mutant was not able to bind hBrr2 as shown in the analytical gel filtration run (Fig. 3.22 lower right panel). This result suggests that the distal tail is responsible for the additional contacts. Moreover, the last 15 residues harbor the major differences between the $Prp8^{Jab1/MPN}$ domains of the two species. Nevertheless, further analyses are required to prove this hypothesis.

Another surprising effect was detected with the mutation at the penultimate

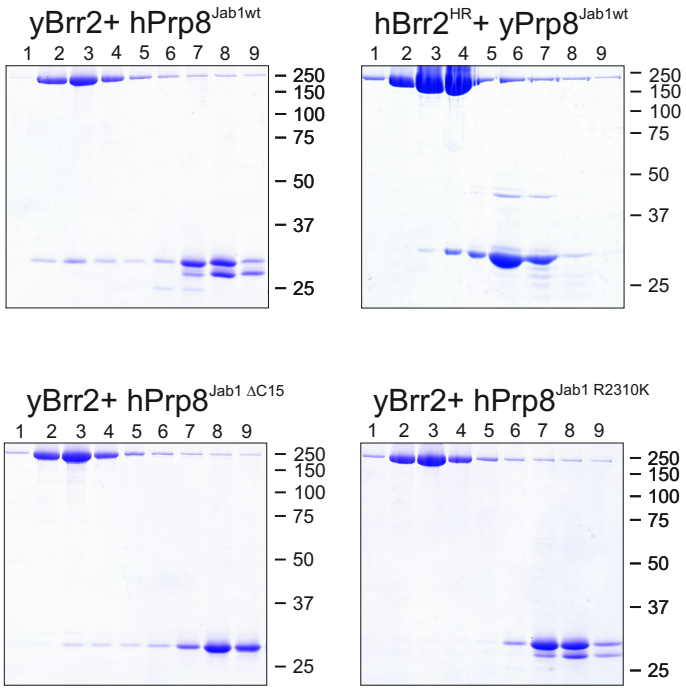

Figure 3.24: Cross-interaction analysis of the human and yeast proteins monitored by gel filtration. Nine fractions representing the same elution volume were analyzed for each chromatogram. Numbers on the right correspond to the molecular weight standards in kDa.

residue, Y2334N. Upon complex formation with hBrr2^{HR}, this mutant showed no enhanced binding of the longer form of the Prp8$^{Jab1/MPN}$ domain and it was possible to observe both fragments, large and short, bound to hBrr2^{HR} no matter the excess of hPrp8$^{Jab1/MPN}$ used (Fig. 3.23). This result suggested that the increased affinity given by the tail was largely reduced by the Y2334N mutation and hence the short fragment was able to compete for the hBrr2^{HR} binding site. Thus, we can infer that this residue is determinant for hPrp8$^{Jab1/MPN}$ tail binding and most probably it involves the interaction of the aromatic group of the tyrosine with the hydrophobic environment deep inside the central tunnel provided by Y992 in the hBrr2^{HR} molecule (Fig. 3.21C). The hydroxyl group of the tyrosine may also contribute to the interactions (e.g. forming a salt bridge with R598 Fig. 3.21C).

3.3.5 hBrr2^{HR}-hPrp8^{CTF} complex crystallization and first low resolution model

Once we unveiled the structure of the hBrr2^{HR}-hPrp8$^{Jab1/MPN}$ complex, many new questions related to Brr2-Prp8 interaction arised. How is the Prp8$^{Jab1/MPN}$ tail release regulated? What triggers the exchange of the tail for the RNA substrate? Which role does the Prp8$^{RNase\ H\text{-}like}$ play domain in this regulation? Does the

Prp8$^{\text{RNase H-like}}$ domain directly interact with Brr2 or the Prp8$^{\text{Jab1/MPN}}$ domain? Most of the questions posed above cannot be answered with the data we have at the moment, but we tried to find an answer for the last of them by crystallizing the hBrr2$^{\text{HR}}$-hPrp8$^{\text{CTF}}$ complex.

This part of the project was conducted in collaboration with the bachelor student Jia Hui Li. We used the co-expressed and purified human complex described in section 3.1.1.4 for initial crystallization screens which yielded crystals in two different conditions, one at 18°C and another one at 4°C (Fig. 3.25A and B). We were able to reproduce the crystals and they were assessed at the synchrotron beamline 14.3 of BESSY II. The best crystal diffracted to 7.0 Å. After these initial results, we used several techniques aimed to improve resolution, but only seeding proved to be successful. The new crystals were tested at beamline P14, Petra III DESY, using a more intense micro-focused beam and a pilatus detector. A complete data set was collected to a resolution of 4.5 Å (Fig. 3.25C). We processed the data using the XDS software, but we were not able to properly assign the space group. Therefore, we ran molecular replacement using all the different space groups related to the initial offered point group (P422) and the coordinates of hBrr2$^{\text{HR}}$-hPrp8$^{\text{Jab1/MPN}}$ complex structure as search model. The outcome was a clear solution with reasonable R-factors in the space group $P4_22_12$. The final space group and cell constants were almost the same as for the hBrr2$^{\text{HR}}$-hPrp8$^{\text{Jab1/MPN}}$ complex with one complex per asymmetric unit.

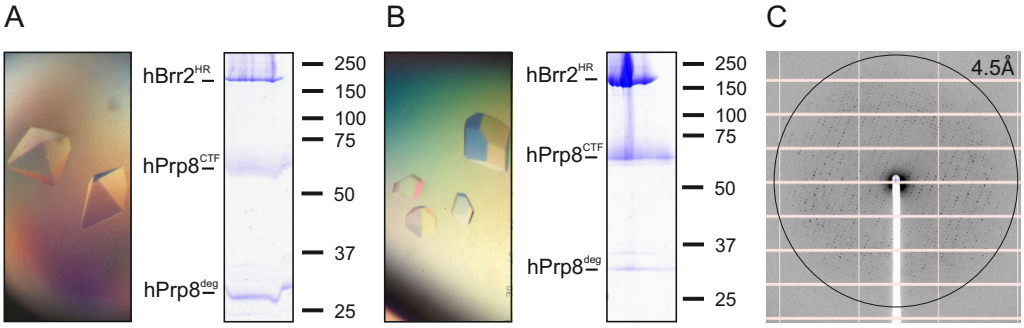

Figure 3.25: (A) Crystals of hBrr2$^{\text{HR}}$-hPrp8$^{\text{CTF}}$ complex condition at 18°C with corresponding SDS gel of washed crystals. hPrp8$^{\text{deg}}$ stands for hPrp8 degradation. **(B)** Crystals of hBrr2$^{\text{HR}}$-hPrp8$^{\text{CTF}}$ complex condition at 4°C and the SDS gel of the crystal. The molecular weight marker is shown on the right (sizes in kDa). **(C)** Diffraction pattern of the crystals on a pilatus detector. The circle indicates a resolution limit of 4.5 Å

We were able to localize hBrr2$^{\text{HR}}$ and hPrp8$^{\text{Jab1/MPN}}$ components in the model. However, no extra continuous electron density was observed that could help us to place the RNase H-like domain of Prp8 (Fig. 3.26 and Table 3.2).

Moreover, any attempt of molecular replacement using hPrp8$^{\text{RNase H}}$ (PDB ID:

Table 3.2: Crystallographic statistics for the hBrr2^{HR}-hPrp8^{CTF} complex

Data collection	
Wavelenght (Å)	1.24
Space group	$P4_22_12$
Unit cell (Å)	
a,b	239.3
c	204.5
Resolution (Å)[a]	120-4.5 (4.6-4.5)
Reflections	
Unique	35862 (2594)
Completeness (%)	99.9 (99.6)
Redundancy	9.0 (8.5)
$\mathbf{R_{meas}}$	24.1 (241.3)
$\mathbf{I/\sigma}$	6.66 (1.0)

Refinement	
Resolution (Å)	107-4.6 (4.7-4.6)
Reflections	
Unique	33636 (2604)
Completeness (%)	99.9 (100.0)
Test set (%)	5
R-factors[c]	
R_{work}	23.5 (33.1)
R_{free}	27.5 (39.1)

a Values for the highest resolution shell in parentheses.

b R_{meas}, intensity of the *i-th* measurement of reflection *hkl*; $\langle I(hkl)\rangle$ average value of the intensity of reflection *hkl* for all i measurements, n is the redundancy.

$$R_{meas} = \left(\frac{n}{n-1}\right)^{1/2} \frac{\sum_{hkl}\sum_i [\ |I_i(hkl) - \langle I(hkl)\rangle|\ ]}{\sum_{hkl}\sum_i I_i(hkl)}$$

c R-factors:

$$R_{work} = \frac{\sum_{hkl} [\ ||F_{obs}| - k|F_{calc}||\ ]}{\sum_{hkl} |F_{obs}|}$$

$$R_{free} = \frac{\sum_{hkl\in T} [\ ||F_{obs}| - k|F_{calc}||\ ]}{\sum_{hkl\in T} |F_{obs}|}$$

$hkl \in T$ - test set; F_{obs}, F_{calc} - observed and calculated (from model) structure factor amplitudes.

3E9L) as a search model failed. In order to confirm the presence of the full Prp8^{CTF} molecule in the crystals, we loaded several of them on an SDS-PAGE (Fig. 3.25A and B). The gel of the crystals showed that the Prp8^{CTF} molecule was present in the two conditions tested, but we could also observe a band of lower molecular weight that most probably corresponds to the separate RNase H-like and Jab1/MPN domains. This degradation product might be one of the reasons for the partial occupancy of the RNase H-like domain and, therefore, for the lack of continuous density. However, we cannot discard that this smaller band is an artefact of the SDS-PAGE sample treatment.

Nevertheless, by placing the symmetry molecules around our model, we could observe that the crystal contained big solvent channels and enough space to harbor an additional molecule (Fig. 3.27A and B). Looking for insights into the possible conformation of the Prp8$^{RNase\ H}$ in our structure, we superposed the structure of

yPrp8-yAar2 complexes available on the PDB (PDB ID: 4I43 [Galej et al., 2013] and PDB ID: 4ILG [Weber et al., 2013]), but the Prp8$^{\text{RNase H}}$ in the given conformation clashes with Brr2 hence the binding of this domain to Brr2 and Aar2 are mutually exclusive. Bringing it all together, we can conjecture that the hPrp8$^{\text{RNase H}}$ domain might be present in the structure. The lack of electron density for the hPrp8$^{\text{RNase H}}$ fragment could also be related to the flexible 60-residue linker that connects both domains in the hPrp8$^{\text{CTF}}$ molecule. This linker most probably contributes to the motion of the domain which has insufficient anchoring points in the crystal lattice.

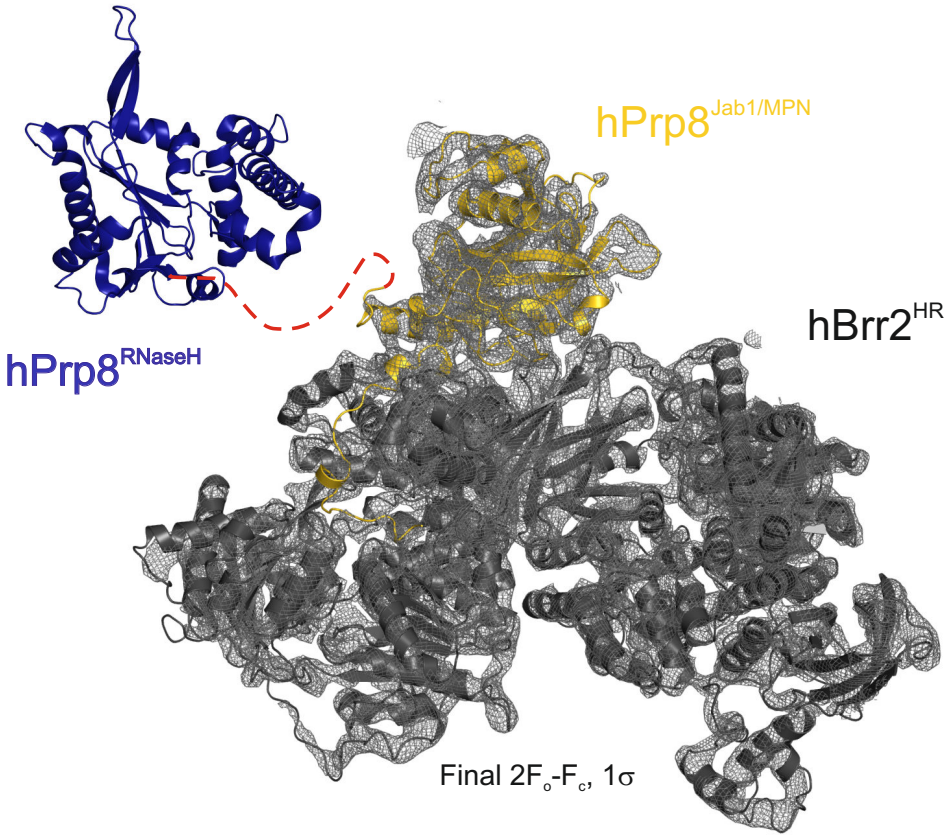

Figure 3.26: Electron density map of the model after molecular replacement using hBrr2$^{\text{HR}}$-hPrp8$^{\text{Jab1/MPN}}$ complex as search model. hPrp8$^{\text{RNase H}}$ structure was represented in the figure next to the hBrr2$^{\text{HR}}$-hPrp8$^{\text{Jab1/MPN}}$ complex connected by dashed red line as the flexible linker between hPrp8$^{\text{RNase H}}$ and hPrp8$^{\text{Jab1/MPN}}$ domains. hBrr2$^{\text{HR}}$ – dark grey, hPrp8$^{\text{RNase H}}$ – blue, hPrp8$^{\text{Jab1/MPN}}$ – gold. Electron density map contoured at the 1σ level.

3.4 Yeast complexes

Most of the functional analyses available for the Brr2-Prp8 interaction have been performed in the yeast system, therefore it would be of great interest to have

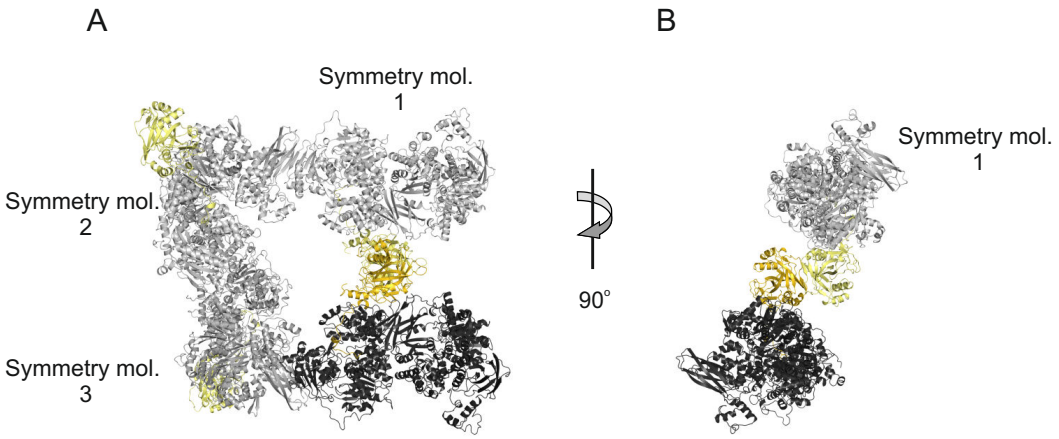

Figure 3.27: **(A)** Molecular packing in the crystal lattice of the hBrr2^{HR}-hPrp8^{CTF} complex after molecular replacement using hBrr2^{HR}-hPrp8$^{Jab1/MPN}$ complex as search model. Central molecule in dark gray and gold, symmetry molecules in gray and yellow. The central molecule shows the same orientation as in Fig. 3.26. **(B)** Image rotated 90° to the left about the horizontal axis compared to Fig. 3.26 and displaying only the closest symmetry molecule.

structural insights of this interaction. Recently, a structure of a yBrr2 construct in complex with yPrp8$^{Jab1/MPN}$ was published and we could observe that the yeast system presents an analogous interaction interface to the human complex structure [Nguyen et al., 2013]. Nevertheless, we worked with a longer construct of the yBrr2 that includes a very long N-terminal extension. The extension entails a region that resembles a PWI domain in homology studies [Korneta et al., 2012] and is not included in the published structures. Additionally, we tried to crystallize the yBrr2^{enHR}-yPrp8^{CTF} complex.

3.4.1 yBrr2^{enHR}-yPrp8^{CTF} complex crystallization

This part of the project was part of the bachelor thesis of Jia Hui Li (FU Berlin, Germany). We managed to obtain crystals of yBrr2^{enHR}-yPrp8^{CTF} complexes with both fragments of the Prp8^{CTF} used in this thesis. The longer fragment includes the C-terminal tail covering the very end of the molecule (residues 1836-2314) and the shorter fragment lacks the last 16 residues (residues 1836-2398). Both complexes yielded very small needle-like crystals (Fig. 3.28A). These needles only grew in few similar PEG based conditions and exclusively in MRC 96-well plates. We failed to reproduce these crystals in a grid screen and, consequently, we could not improve their quality. Anyway, a couple of crystals were tested in diffraction experiments using 30% ethylenglycol as cryoprotectant. During fishing, the crystals became

flexible and bent over the micro-loop. The obtained resolution for the crystals was never better than 20 Å.

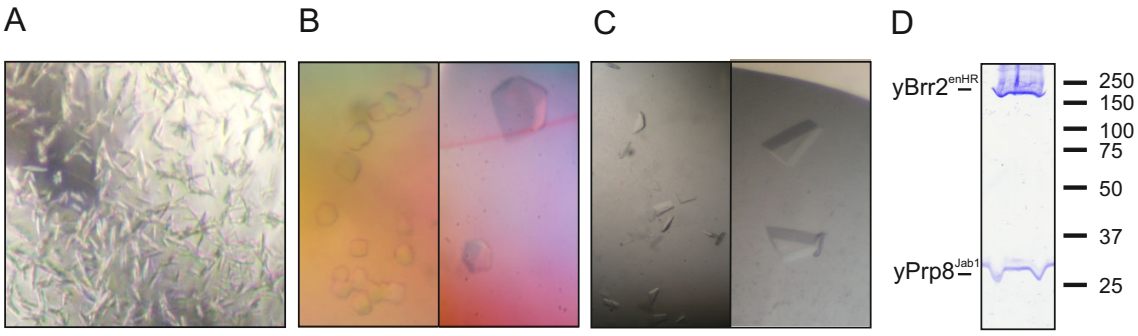

Figure 3.28: **(A)** Crystals of yBrr2^{enHR}-yPrp8^{CTF} complex. **(B)** Crystals of yBrr2^{enHR}-yPrp8$^{Jab1/MPN}$ complex condition without Mg. **(C)** Crystals of yBrr2^{enHR}-yPrp8$^{Jab1/MPN}$ complex condition with Mg. **(D)** SDS-PAGE of the crystals in **(C)**. The molecular weight marker is shown on the right (sizes in kDa).

3.4.2 yBrr2^{enHR}-yPrp8$^{Jab1/MPN}$ complex crystallization and data collection

We were also able to crystallize the yBrr2^{enHR} in complex with yPrp8$^{Jab1/MPN}$. This yPrp8$^{Jab1/MPN}$ construct covered the very C-terminus of Prp8 (residues 2147-2413) including the 16-residue tail that was not included in the yPrp8$^{Jab1/MPN}$ X-ray structure [Pena et al., 2007; Zhang et al., 2007]. After several grid screens, we selected two conditions as the most promising. The first one only contained HEPES buffer pH 7.5 and 8% PEG 8000, but the resolution of the crystals was limited to 8 Å (Fig. 3.28B). The second condition was more acidic, with a pH of 6.2, and contained Mg. The crystals from the second condition had a more pronounced contour with sharp edges (Fig. 3.28C). Unfortunately, the crystals still diffracted poorly to 6.2 Å. The presence of both proteins in the crystals was confirmed by washing and loading several of them on SDS-PAGE (Fig. 3.28D).

We had problems assigning the correct space group due to low quality of the dataset or putative twinning. The twinning was not apparent in the crystal or the diffraction pattern (Fig. 3.29A). Twinning was suggested after data processing in our first attempt of molecular replacement using phaser. Since twinning may affect the determination of the right space group, generating an apparent higher point group symmetry (e.g. P422 might be in reality P4) we decided to process the data in both point groups (P4 and P422) and check them using the Xtriage tool of the Phenix suite. Xtriage suggested twinning in both point groups. The point group P4 presented one possible merohedral twin operator with a twin fraction (α) of 0.468

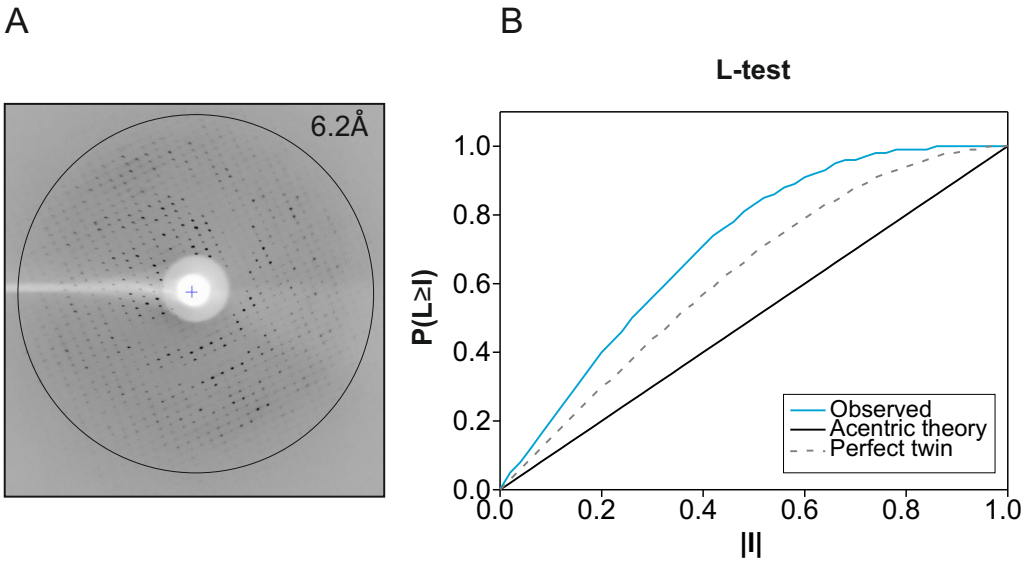

Figure 3.29: Diffraction pattern and twinning analysis. **(A)** Diffraction pattern at 6.2 Å of the yBrr2$^{\text{enHR}}$-yPrp8$^{\text{Jab1/MPN}}$ complex. **(B)** L-test of Padilla & Yeates for acentric data. L = $(I^1_{obs} - I^2_{obs})/(I^1_{obs} + I^2_{obs})$. I= intensity, $|I|$= mean intensity. The plots for both point groups tested presented the same tendency.

(which is too high to attempt de-twinning of the data) and a twin law = h, -k, -l. However, Xtriage also includes a step called "Exploring higher metric symmetry" which identified the point group P422 as the most likely. Unfortunately, there are no twin laws possible for P422 related space groups, because all axial pairs are already equivalent as a result of the crystal symmetry (Fig. 3.29B). Nevertheless, we cannot discard the low quality of the dataset as a source of the observed problems.

Since molecular replacement is still able to work with twinned data, we used Phaser to perform molecular replacement. The existence of screw axes was tested during the replacement allowing the software to choose the most suitable space group related to the chosen point group (P422). For the molecular replacement we divided the hBrr2$^{\text{HR}}$ pdb file (PDBID:4F91) into the N-terminal and C-terminal cassettes and performed separate searches for each of them. Once the Brr2 component was found, we used the obtained result as template for the next molecular replacement with the coordinates of the unbound yPrp8$^{\text{Jab1/MPN}}$ (PDB ID: 2OG4 [Pena et al., 2007]) structure as search model. The statistics of the data collection and the first refinement attempt are summarized in Table 3.3.

The molecular replacement software was able to place Brr2 and yPrp8$^{\text{Jab1/MPN}}$ domain onto a patch of continuous density, but the extra density blobs were not enough for model building of the PWI domain and the extensions of the yBrr2 helical region. After refinement, the difference between the R_{work}/R_{free} was too large to

Table 3.3: Crystallographic statistics for the yBrr2^{enHR}-yPrp8$^{Jab1/MPN}$ complex

Data collection	
Wavelenght (Å)	0.9184
Space group	$P4_3\ 2_12$
Unit cell (Å)	
a,b	180.4
c	374.9
Resolution (Å)[a]	50-6.2 (6.4-6.2)
Reflections	
Unique	14412 (1003)
Completeness (%)	99.4 (95.1)
Redundancy	6.9 (6.6)
$\mathbf{R_{meas}}$	14.1 (159.4)
$\mathbf{I/\sigma}$	10.20 (1.3)

[a] Values for the highest resolution shell in parentheses.

[b] R_{meas}, intensity of the *i-th* measurement of reflection *hkl*; $\langle I(hkl)\rangle$ average value of the intensity of reflection *hkl* for all i measurements, n is the redundancy.

$$R_{meas} = \left(\frac{n}{n-1}\right)^{1/2} \frac{\sum_{hkl}\sum_i [\ |I_i(hkl) - \langle I(hkl)\rangle|\ ]}{\sum_{hkl}\sum_i I_i(hkl)}$$

Refinement	
Resolution (Å)	48.6-6.4 (6.9-6.4)
Reflections	
Unique	13351 (1289)
Completeness (%)	99.8 (99.6)
Test set (%)	5
R-factors[c]	
R_{work}	33.9 (30.0)
R_{free}	42.7 (44.8)

[c] R-factors:

$$R_{work} = \frac{\sum_{hkl} [\ ||F_{obs}| - k|F_{calc}||\ ]}{\sum_{hkl} |F_{obs}|}$$

$$R_{free} = \frac{\sum_{hkl\in T} [\ ||F_{obs}| - k|F_{calc}||\ ]}{\sum_{hkl\in T} |F_{obs}|}$$

$hkl \in T$ - test set; F_{obs}, F_{calc} - observed and calculated (from model) structure factor amplitudes.

be acceptable (33.9/42.7) and hence it was not possible to continue refining the structure. Besides all the problems, the obtained solution had an architecture very similar to the human complex and the yPrp8$^{Jab1/MPN}$ solely contacted the N-terminal cassette of Brr2 (Fig. 3.30) as was also observed in the recently published structure of yBrr2^{enHR}-yPrp8$^{Jab1/MPN}$ [Nguyen et al., 2013].

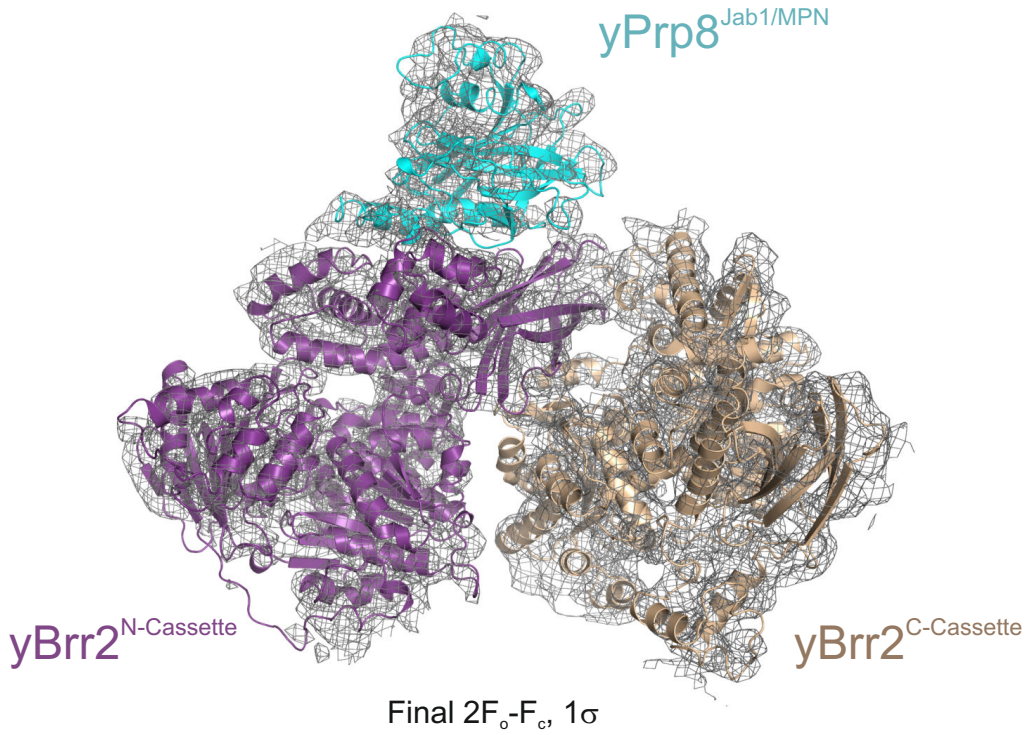

Figure 3.30: Electron density map of the model yBrr2^{enHR}-yPrp8$^{Jab1/MPN}$ complex after molecular replacement using the cassettes of hBrr2^{HR} and yPrp8$^{Jab1/MPN}$ as search models. N-terminal cassette – purple, C-terminal cassette – wheat, yPrp8$^{Jab1/MPN}$ – cyan. Electron density map contoured at the 1σ level.

Chapter 4

Discussion

In this thesis we established for first time the recombinant expression and purification of the hPrp8$^{Jab1/MPN}$ domain and also the co-production of the hBrr2^{HR}-hPrp8^{CTF} complex. In both cases we successfully used the baculovirus/insect cell expression platform (MultiBac).

We were able to crystallize several complexes of different fragments of Brr2 and Prp8 splicing factors from two organisms, human and yeast (hBrr2^{HR}-hPrp8$^{Jab1/MPN}$, hBrr2^{HR}-hPrp8^{CTF}, yBrr2^{enHR}-yPrp8$^{Jab1/MPN}$ and yBrr2^{enHR}-yPrp8^{CTF}). However, it was the hBrr2^{HR}-hPrp8$^{Jab1/MPN}$ complex which allowed us to solve the structure with the highest resolution (3.4 Å). This model offered us for the first time the possibility to study the molecular basis of regulation of a member of the Ski2-like family of helicases by a protein cofactor.

We also produced the hPrp8$^{Jab1/MPN}$ associated RP mutants which, by means of interaction and functional studies, gave us insight into the principles of Brr2 misregulation. The effects of those point mutations in the human protein were investigated through interaction studies and compared to previous biochemical studies conducted with the yeast proteins.

4.1 Brr2-Prp8$^{Jab1/MPN}$ interaction

Despite the link between Brr2 and Prp8 was identified in human over a decade ago [Achsel et al., 1998], little was known about the specific interaction interface. By yeast two-hybrid analyses (Y2H), using full length and fragments of yBrr2 and yPrp8 proteins, the interaction region was narrowed down to two segments: the N-terminal region of Prp8 (residues 1-263) selected a C-terminal fragment of Brr2

(residues 1282-1749) and full length Brr2 was able to recognize fragments at the very C-terminus of Prp8 (residues 2010-end) as interaction partners [van Nues and Beggs, 2001]. The proposed N-terminal region of Prp8 harbors a proline-rich tract which has only been observed in fungal proteins and it is absent in proteins from other organisms. The C-terminal region of Prp8, encompassing the RNaseH-like and Jab1/MPN domains, was further characterized and confirmed as a main binding area for Brr2 [Liu et al., 2006; Pena et al., 2007; Maeder et al., 2008]. Additionally, this C-terminal region proved to be directly involved in Brr2's activity regulation [Maeder et al., 2008]. The boundaries for the protein-protein interaction region in Prp8 were refined again by gel filtration analyses using isolated protein domains, proving that Prp8$^{Jab1/MPN}$ domain was solely required for Brr2 binding [Weber et al., 2011].

When I started this PhD thesis, the knowledge of the Brr2 binding region for Prp8$^{Jab1/MPN}$ was very limited. Using the isolated cassettes of hBrr2 and gel filtration it was possible to identify the N-terminal cassette as the unique interaction partner of the hPrp8$^{Jab1/MPN}$ domain (Fig. 3.9). Even though both cassettes presented a very similar fold, they share only 27% sequence indentity which explains the specificity of the Prp8$^{Jab1/MPN}$ domain for just one cassette. However, this result was conflicting with previous Y2H analyses which suggested that the Prp8$^{Jab1/MPN}$ domain interacted with the C-terminal cassette of Brr2 [Liu et al., 2006]. The apparent contradictory outcome may lie in the borders of the yBrr2 fragments used as bait in the Y2H experiments. Since no Brr2 structure was available at that time, the fragments for the Y2H assay were designed based on the domain information deduced from the sequence. The chosen borders of the fragment that was identified as putative interaction region (residues 1301-1816) encompassed the intercassette linker and the two RecA domains of the C-terminal cassette of Brr2. In our structure, the hPrp8$^{Jab1/MPN}$ contacts all six domains of the N-terminal cassette and none of the C-terminal cassette. The fragment used in Y2H lacks the Sec63 homology unit which, in case of the N-terminal cassette, harbors the major contact area required for hPrp8$^{Jab1/MPN}$ interaction and might be determinant for binding specificity. Thus, this truncated C-terminal cassette could be able to unspecifically bind Prp8$^{Jab1/MPN}$ using solely the RecA domains as contact points which are analogous to the RecAs in the N-terminal cassette. Another possible explanation for the false positive result is that fragment might be incorrectly fold which may have made it prone to unspecific interactions.

Upon binding of hPrp8$^{Jab1/MPN}$ to the N-terminal cassette, few conformational changes were observed (section 3.3.2.1). One of the clearest effects was the rotation of the N-terminal cassette relative to the C-terminal cassette (Fig. 3.17A). This

motion induced the separation of residues in the intercassette region whose contacts were required for Brr2's unwinding activity (Fig. 3.17B and C) [Santos et al., 2012]. These conformational changes have implications for the mechanism by which the two helicase cassettes cooperate in Brr2.

4.1.1 Implications for inter-cassette communication in Brr2

The structural changes observed upon complex formation provide a straightforward explanation for the severe effects of mutations in the linker, the intercassette RecA-contacts and the ATP pocket of the C-terminal cassette on hBrr2^{HR} ATPase and helicase activity [Santos et al., 2012]. During cycles of RNA duplex unwinding, the intercassette RecA-contacts may have to be intermittently broken and re-formed, e.g. to allow repositioning of the RNA bound at the N-terminal cassette. Close association of the linker with either cassette would allow it to tightly fasten the cassettes at one side, which in turn would enable a portion of the linker between the cassette's contact regions to act like a hinge during opening and closing motions on the RecA-side. A bound nucleotide would stabilize the C-terminal cassette, allowing it to act as a monolithic unit relative to which the N-terminal cassette can move. To confirm the offered model of inter-cassette interaction some tests are required which could be done using single-molecule Förster resonance energy transfer (smFRET) [Weiss, 1999; Deniz et al., 2000]. The most common approach for this method is to engineer two cysteines in the regions of interest, which are then reacted with maleimide derivatives of single molecule suitable dyes. This technique is very sensitive, being able to detect distances between the dyes in a range of 30-100Å and hence conformational changes of the target molecule [Weiss, 1999; Deniz et al., 2000].

4.1.2 hPrp8$^{Jab1/MPN}$ C-terminal tail occludes the RNA binding tunnel

Another novel feature observed in our model was the binding site for the C-terminal tail of hPrp8$^{Jab1/MPN}$. The unexpected positioning of the tail along the RNA binding cleft (Fig. 3.13) (which possesses all the highly conserved residues responsible for the RNA recognition in all SF1 and SF2 helicases) has no precedent in the known regulative cofactors for this kind of enzymes. The tail completely occludes the RNA binding motifs and competes with the loading of the nucleic acid, as it has been shown in our pull-down assays conducted with the yeast proteins (Fig. 3.18A).

Since the proteins are highly conserved in nature, the human and yeast systems are

considered comparable. Most of the functional studies for this thesis were conducted in the yeast system, because the proteins were available in higher quantity and they were also easier to produce. Besides, the ability of hPrp8$^{Jab1/MPN}$ to cross-interact with yBrr2 (Fig. 3.24) and the recent structure of yBrr2^{enHR}-yPrp8$^{Jab1/MPN}$ complex [Nguyen et al., 2013], which presented the same architecture as the human one, are evidences of the high similitude of both systems. Additionally, the hPrp8$^{Jab1/MPN}$ construct covering the very C-terminus of the molecule tends to produce a degradation product which most probably lacks the distal C-terminal tail of hPrp8 (residues 2321-2335) and our results showed that this tail is determinant for the switch between activation and inhibition of hBrr2. Thus, the presence of the shorter contaminant could interfere with the analysis of hPrp8$^{Jab1/MPN}$ effect on Brr2. Nevertheless, some assays were corroborated using the human proteins with equivalent outcomes.

Based on the results of our biochemical analyses, we can conclude that the distal C-terminal tail of Prp8$^{Jab1/MPN}$ not only blocks the interaction between the RNA and Brr2, but also inhibits Brr2 ATPase and helicase activity (Fig. 3.19). The removal of 16 residues at the very C-terminus of yPrp8$^{Jab1/MPN}$, on the other hand, restored RNA binding. Strikingly, the truncated variant no longer inhibited, but rather enhanced Brr2 activity. The RNA-stimulated ATP hydrolysis increased ca. 2-fold and U4/U6 unwinding was also stimulated. Brr2 is considered a very poor helicase that unwinds its substrate with low efficiency [Raghunathan et al., 1998; Maeder et al., 2008]. Therefore, the increase in its activity suggests that hPrp8$^{Jab1/MPN}$ has dual regulatory roles, being able to inhibit and activate the helicase.

Although a structure of the activated complex is presently lacking, in our hBrr2^{HR}-hPrp8$^{Jab1/MPN}$ structure the proximal part (residues 2310-2320 in human) of the C-terminal tail of Prp8$^{Jab1/MPN}$, which is still contained in hPrp8$^{Jab1/MPN\text{-}\Delta C}$, runs between the Brr2 HB and HLH domains (Fig. 3.12). While the HB domain is thought to provide a ratchet function during RNA translocation [Büttner et al., 2007], the HLH domain provides a surface via which the unwound RNA exits Brr2 [Büttner et al., 2007; Santos et al., 2012]. Direct interactions between these domains are missing in the Brr2 apo structure [Santos et al., 2012], but have been shown to be important in related helicases [Woodman et al., 2007]. HB-HLH interaction mediated by Prp8$^{Jab1/MPN}$ could thus help to couple RNA-driven ATP hydrolysis to the ratcheting movement of Brr2 on the RNA substrate. Furthermore, the core of the Prp8$^{Jab1/MPN}$ domain is close to the presumed RNA exit path via the Brr2 HLH domain and may foster additional direct contacts to the unwound strand.

4.2 Brr2 Regulation

From our results we can conclude that Prp8$^{Jab1/MPN}$ uses its C-terminal tail to inhibit the activity of Brr2 and hence avoids premature unwinding of the substrate when the U4/U6.U5 tri-snRNP is formed. Recent studies suggested that the C-terminal tail of Prp8 might not be alone in this initial activity blockage. Another domain of Prp8, the RNaseH-like domain, may contribute to the inhibition, competing with Brr2 by loading onto the single-stranded region of U4 snRNA preceding U4/U6 stem I [Hahn et al., 2012; Mozaffari-Jovin et al., 2012]. Once the tri-snRNP enters the spliceosome, Brr2 has to be activated to be able to unwind U4/U6 RNA duplex. At this point two presumed events should occur, the sequestered U4/U6 di-snRNA must be transferred from Prp8^{RNAseH} to the RNA binding tunnel of Brr2 and the Prp8 C-terminal tail must be displaced. The RNA transfer might be induced by pre-mRNA incorporation. It is known from early studies that the RNaseH-like domain can be cross-linked with the 5'SS [Reyes et al., 1999; Turner, 2006] and it has been postulated that it might be involved in the handover of the 5' SS from the U1 snRNA to the ACAGAGA box of the U6 snRNA [Kuhn et al., 1999]. Thus, when the RNaseH-like domain recognizes the pre-mRNA 5'SS, it possibly triggers the release of Brr2's binding site on U4 snRNA which can be then transferred to the helicase.

The presence of pre-mRNA may have an extra effect by increasing the local RNA concentration in the protein's surroundings and promoting the displacement of the Prp8$^{Jab1/MPN}$ C-terminal tail. This effect has been suggested based on previous and recent studies which showed that an excess of RNA induced an increase in Brr2 helicase activity in presence of Prp8$^{Jab1/MPN}$ wild type [Maeder et al., 2008; Mozaffari-Jovin et al., 2013]. However, the increment in Brr2's activity is not as high as for the truncated Prp8$^{Jab1/MPN\text{-}\Delta C}$ [Mozaffari-Jovin et al., 2013], implying that the release of the tail might be partial when the construct including the C-terminus is present and, for complete activation, the action of a trigger is required (section 4.2.1).

After spliceosomal activation, since Brr2 presents low substrate specificity (being able to unwind RNA duplexes *in vitro* that are unrelated to U4 and U6 snRNAs [Laggerbauer et al., 1998]) it most probably requires to be kept in an inactive state until it is used again for the last steps of splicing. We have suggested that the Prp8 C-terminal tail could return to its initial position and be employed as inhibitor once more. Nevertheless, there is no direct evidence to prove this statement.

Bringing all the gathered information together, we have postulated a putative

model for Brr2 regulation that includes the coordinated action of the two C-terminal domains of Prp8 (Fig. 4.1).

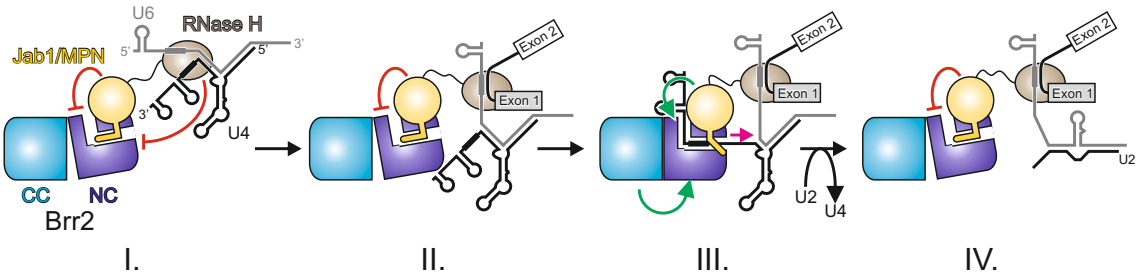

Figure 4.1: Working model for the regulation of Brr2 by the Prp8$^{\text{RNase H}}$ and Prp8$^{\text{Jab1/MPN}}$ domains during pre-mRNA splicing. Red lines, inhibitory effects; green arrows, stimulatory effects. (I.) Inhibited state in the U4/U6.U5 tri-snRNP and during early spliceosome assembly achieved by competitive binding of a single-stranded U4 snRNA region upstream of stem I (thick black line) via the Prp8$^{\text{RNase H}}$ domain and by blocking the RNA-binding tunnel and disruption of the N-terminal cassette (NC) C-terminal cassette (CC) interactions via the Prp8$^{\text{Jab1/MPN}}$ domain. (II.) State after release of U4 snRNA from RNase H with Brr2 still blocked via the Prp8$^{\text{Jab1/MPN}}$ domain. The conserved U6 ACAGAG box (thick grey line) has taken over base pairing of the 5' splice site from U1 snRNA. It is likely but not certain that the block by the Prp8$^{\text{RNase H}}$ domain is released before the block by the Prp8$^{\text{Jab1/MPN}}$ domain. (III.) Brr2 stimulation during spliceosome catalytic activation via the globular part and proximal tail of Prp8$^{\text{Jab1/MPN}}$ domain and by direct interaction of N- and C-terminal cassettes. Magenta arrow, movement of Brr2 on U4 snRNA. (IV.) The Prp8$^{\text{Jab1/MPN}}$ domain may again block Brr2 after spliceosome catalytic activation.

4.2.1 Possible triggers for Prp8$^{\text{Jab1/MPN}}$ inhibitory state release

The release of the C-terminal tail from Brr2's RNA-binding channel could be achieved by altering the posttranslational modification status of Prp8. Previous studies have suggested that ubiquitination of Prp8 negatively regulates U4/U6 unwinding by Brr2 [Bellare et al., 2008]. The Prp8$^{\text{Jab1/MPN}}$ domain evolved from a Zn^{+2}-dependent isopeptidase, as employed by the proteasome to remove ubiquitin chains from proteins targeted for degradation, and has retained its ability to bind ubiquitin [Bellare et al., 2005]. In yPrp8$^{\text{Jab1/MPN}}$, residues V2184 and L2185 have been implicated in ubiquitin binding, and the corresponding residues of hPrp8 (I2105 and L2106) are located on the backside of the Prp8$^{\text{Jab1/MPN}}$ domain that is unobstructed in the hBrr2$^{\text{HR}}$-hPrp8$^{\text{Jab1/MPN}}$ complex (Fig. 4.2). Ubiquitin bound at this site could concomitantly interact with the neighboring N-terminal IG domain or the intercassette linker, thereby eliciting conformational changes that could enhance the negative regulation of Brr2 by Prp8$^{\text{Jab1/MPN}}$ either by stabilizing the Prp8$^{\text{Jab1/MPN}}$ C-terminal tail in the RNA-binding tunnel of Brr2 and/or disrupting intercassette contacts

important for Brr2 helicase activity. Thus, deubiquitination of Prp8 constitutes a possible signal *in vivo* that may trigger release of the Prp8$^{Jab1/MPN}$ C-terminal peptide from Brr2 and thereby allow Brr2 to load onto its RNA substrate. As our *in vitro* Brr2 unwinding assays were performed with Prp8$^{Jab1/MPN}$ lacking posttranslational modifications, inhibition of Brr2 unwinding activity can be achieved in their absence, at least in the purified, ternary system used in these studies. However, in the complex environment of the spliceosome, posttranslational modifications, as well as other spliceosomal proteins, likely play important, coordinated roles in switching Brr2 activity on and off. Indeed, it is also conceivable that the action of the GTPase Snu114, which also regulates Brr2 activity [Small et al., 2006], might trigger the release of the C-terminal tail of Prp8 from Brr2 and/or the release of Prp8's RNase H domain from U4 snRNA, and thereby allow Brr2 interaction with its substrate.

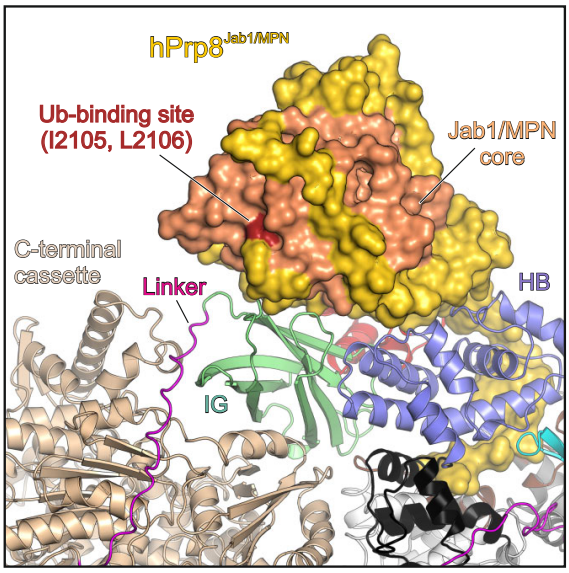

Figure 4.2: Ubiquitin binding site on hPrp8$^{Jab1/MPN}$. Core of the hPrp8$^{Jab1/MPN}$ domain, peach; residues I2105 and L2106 of hPrp8$^{Jab1/MPN}$, brick red. Other coloring as before. Rotated 180° about the vertical axis compared to Fig. 3.12.

4.2.2 Comparison with other helicase's protein cofactors

Helicases need to function at the appropriate time and on the appropriate substrate. Hence their activities need to be regulated. In addition, most of monomeric SF1 and SF2 enzymes are relatively poor helicases *in vitro* and many of them require the interaction with accessory proteins to stimulate their activity [Lohman et al., 2008]. Several proteins have been suggested as putative interaction partners for different helicases, but just few confirmed as effective regulators of their activity [Silverman et al., 2003] and even less have their interaction interface characterized.

Most of the identified cofactors interact with members of the DEAD-box helicase family.

The *E. coli* RhlB DEAD-box helicase is a component of the RNA degradosome and its activity is enhanced by the action of the non-catalytic region (residues 698-762) of RNase E (enzyme which belongs to the same machinery). The C-terminal domain of RhlB was identified as its interaction region and, based on a homology model, it was placed far from the catalytic site of the helicase, therefore, the exact activation mechanism is not clear [Chandran et al., 2007].

The eukaryotic translation initiation factor eIF4A is an ATP-dependent RNA helicase whose activity is known to be stimulated by two factors in a mutually exclusive manner, eIF4B and eIF4H. But for overall function a third protein is required, eIF4G. The four proteins are members of the same multiprotein complex. eIF4H binds to the C-terminal domain of the helicase, far from the catalytic center, and their stable association requires ATP. The mode of action is based in conformational changes between "closed" and "open" state [Rozovsky et al., 2008; Marintchev et al., 2009]. In the closed active conformation, the two RecA-like domains come in contact with each other, forming the ATP binding pocket, and a contiguous RNA-binding surface. eIF4H stabilizes the closed conformation and increases the affinity of the helicase for the nucleotide. Once the ATP is hydrolysed the conformation changes again to a "less closed" conformation that could in turn affect the formation of a contiguous RNA binding surface and also reduce the affinity of eIF4A for eIF4H [Marintchev et al., 2009].

The eukaryotic initiation factor 4III (eIF4AIII) is an ATP-dependent RNA helicase member of the exon junction complex (EJC). This enzyme is known to be regulated by three protein factors: MLN51, MAGOH and Y14. MLN51 increases the RNA binding efficiency of the helicase by direct contact with the nucleic acid. MAGOH and Y14 form a heterodimer that inhibits eIF4AIII ATPase activity by stabilizing the closed conformation of the enzyme bound to RNA and interacting MLN51 [Andersen et al., 2006; Bono et al., 2006].

The helicase Dbp5 is involved in nuclear export of mRNA and is known to interact with the protein Gle1 which stimulates its ATPase activity when bound to the small molecule inositol hexakisphosphate. The interaction is mediated by the C-terminal domain of the enzyme and has been suggested a similar interaction interface as the eIF4A-eIF4H complex [Dossani et al., 2009].

The *modus operandi* of all protein cofactors discussed here is based in the allosteric regulation of their helicases inducing conformational changes which promote the

transition between "closed" and "open" states. Prp8$^{Jab1/MPN}$, on the other hand, binds very close to the catalytic center and its C-terminal tail directly interferes with the nucleic acid interaction. Furthermore, the globular part of the domain is also able to stimulate Brr2's activity. This dual role is unique in the known unwindases so far and it is the first regulation model postulated for a member of the Ski2-like family.

4.3 Molecular basis of the RP13 mutations phenotype

The majority of studies to date on RP13 mutations had been conducted in yeast and only few analyses could be carried out with human cells. One study reported the effects observed on spliceosomes and splicing for patients harboring two hPrp8 mutations, R2310G and Y2334N [Tanackovic et al., 2011]. The isolated spliceosomes of these patients showed alterations in the stoichiometry of spliceosome components with reduction in the amounts of snRNAs compared to the healthy control group ($\sim$ 30-40% of reduction of the U4, U5 and U6 snRNAs). The snRNP associated proteins were also affected displaying $\sim$ 10% hBrr2 content and $\sim$ 50% hSnu114 content relative to the control. These differences were reflected in difficulties for spliceosome formation and splicing. The nuclear extracts from patient cells can form spliceosomes *in vitro*, although the process has reduced kinetics with delays in complex A accumulation and thus inefficient conversion rate for the subsequent complexes. Additionally, pre-mRNA splicing analyses from RP patient's cell extracts showed also a reduction in splicing activity. Even though, the splicing process itself was not abolished (patients are heterozygous and can produce a wild type copy of the protein), its efficiency was reduced [Tanackovic et al., 2011]. It has been also suggested that the mutations may affect more dramatically alternative splicing of a subset of pre-mRNAs important for the retina and, therefore, the effects in other cells are sub-pathological [Ivings et al., 2008; Tanackovic et al., 2011].

The yeast system, on the other hand, allowed the analysis of a broader number of mutations and displayed clearer and more dramatic effects. It has been observed that all reported RP13-linked Prp8 mutations led to inhibition of yeast growth and *in vivo* splicing [Boon et al., 2007; Maeder et al., 2008]. Nevertheless, there is a correlation between the severity of growth phenotypes in haploid yeast and the severity of retinal degeneration in patients with the corresponding Prp8 mutations, at least in the case of the three most common RP13 mutations, confirming the yeast

system as a comparable model.

In our work we observed that the mutations localized on the globular part (group I) of Prp8$^{Jab1/MPN}$ affected the overall fold stability of the molecule and hence the mutants tended to precipitate. In yeast, apart from the splicing defects, it was shown that these mutants also inhibited U4/U6.U5 tri-snRNP formation [Boon et al., 2007; Maeder et al., 2008].

The group II mutants in the proximal tail seems to be related to reduced affinity for Brr2, with the effects being more clear in the yeast than in the human system. Human complexes were more difficult to disrupt and probably the functional effects are not as drastic as for the yeast complexes, where one point mutation on residue R2388K (R2310K in human) completely abolishes interaction with yBrr2 in gel filtration. Nevertheless, the mutated proteins displayed impairment in U4/U6.U5 tri-snRNP formation for both systems [Boon et al., 2007; Maeder et al., 2008; Tanackovic et al., 2011].

Our results also indicated that the extra affinity contact points for the human complex are contained in the distal tail region, because a double mutant containing a stop codon at residue 2320 and a point mutant at R2310K disrupted hPrp8$^{Jab1/MPN}$-hBrr2 interaction (Fig. 3.22). Additionally, cross-interaction studies showed a clear reduction of the hPrp8$^{Jab1/MPN\Delta C15}$ affinity for yBrr2 compared to the hPrp8$^{Jab1/MPN}$ wild type (Fig. 3.24). This difference between the tail's affinities of both organisms is also evident in the recent complex structure of the helical region of yBrr2 with yPrp8$^{Jab1/MPN}$ where (even when the resolution is slightly higher than our structure (3.1 Å)) it is not possible to observe any density for the distal C-terminal tail of Prp8 entering Brr2 [Nguyen et al., 2013]. The C-terminal tail (residues 2310-2334 in human) contains a large fraction of charged residues (8 of the 24 C-terminal residues in human) and lacks regular secondary structure elements except for a single 3_{10} helical turn, supporting the idea that it constitutes an intrinsically unstructured element in isolation. Thus, immobilization of this portion of Prp8 on Brr2 must be accompanied by a large loss in conformational entropy. This entropic loss will in part compensate the expected large positive interaction enthalpy, likely rendering binding of the Prp8 C-terminal tail on Brr2 only marginally thermodynamically stable but at the same time highly specific. Displacement of the tail would then require only a small perturbation of the binding thermodynamics, which could be provided by posttranslational modifications or additional protein cofactors as previously discussed.

The point mutation in the penultimate residue that is part of the group III mutants (F2412N in yeast) did not interfere with U4/U6.U5 tri-snRNP formation

and also had no apparent effect on yPrp8$^{Jab1/MPN}$ solubility or interaction with yBrr2 [Mozaffari-Jovin et al., 2013]. However, yPrp8$^{Jab1/MPN\text{-}F2412N}$ enhanced yBrr2's RNA affinity, repressed yBrr2's RNA-stimulated ATPase activity less efficiently compared to wildtype yPrp8$^{Jab1/MPN}$ and it de-repressed yBrr2's helicase activity similar to yPrp8$^{Jab1/MPN\Delta C16}$ [Mozaffari-Jovin et al., 2013]. Thus, this RP13 mutation substantially alleviates the negative effect of the Prp8$^{Jab1/MPN}$ C-terminal tail on Brr2 unwinding activity, presumably by destabilizing the tail's interaction with the RNA-binding tunnel of Brr2. This observation was corroborated in the human system, because upon mutation on the ortholog residue (Y2334), the higher affinity displayed by the long hPrp8$^{Jab1/MPN}$ was lost and the shorter degradation fragment was able to compete for Brr2 binding site.

Based on the results of this work, we were able to establish the importance of the Prp8$^{Jab1/MPN}$ tail and its effect on Brr2 regulation. The mutations studied here indicated that any reduction on the Brr2-Prp8$^{Jab1/MPN}$ interaction affinity leads to misregulation of Brr2's activity and hence it interferes with spliceosome formation.

4.4 Outlook

Brr2's activity is essential for splicing, being involved in the RNA arrangements required for spliceosome activation. Despite its substrate was identified long ago [Laggerbauer et al., 1998; Raghunathan et al., 1998], little was known about the mechanism of action and regulation. Initially, Brr2 was identified as a member of the DExD/H- box family of RNA-dependent ATPases, based on the consensus sequence of the conserved helicase motifs [Laggerbauer et al., 1998; Raghunathan et al., 1998]. Afterwards, it was grouped into the Ski2-like helicases [Bleichert and Baserga, 2007]. These enzymes bind their substrate in a single stranded region at the 3' terminus and unwind the nucleic acid with a 3'$\longrightarrow$ 5' directionality. Additionally, the structure of a member of this family (Hel308), bound to a double stranded DNA with 3' overhang, suggested that the central helix of the HB domain can work as a ratchet for directional transport while a prominent β-hairpin loop acts as a wedge separating the two strands [Büttner et al., 2007; Woodman et al., 2007]. The recent structure of hBrr2^{HR} [Santos et al., 2012] showed a similar domain distribution for the individual cassettes compared to Hel308, but the interaction between the HB and HLH domains required for the ATPase and helicase activity coupling is missing [Woodman et al., 2007; Santos et al., 2012]. Consequently, Brr2 displays low efficiency [Maeder et al., 2008; Santos et al., 2012]. Thus, the existence of an activator was expected. Besides, Brr2 encounters its substrate before joining the spliceosome and,

to avoid the premature unwinding, its activity must be also inhibited.

Previous studies postulated Prp8 as a regulator [Maeder et al., 2008], but the nature of this regulation was unknown. In this thesis we demonstrated the dual role of $Prp8^{Jab1/MPN}$ as activator and inhibitor of Brr2's activity and we were able to establish the structural basis for this regulation. Nevertheless, there are many unsolved questions: How is the RNA loaded? How do the cassettes interact upon substrate binding? What happens with the tail when the helicase is activated? How is the release of the tail triggered? To answer part of these questions a structure of Brr2 bound to RNA in presence and absence of $Prp8^{Jab1/MPN}$ might be required. Additionally, the action of Snu114 (another regulator of Brr2's activity [Small et al., 2006]) should be analyzed taking into account that this protein forms a very stable complex with Prp8 [Achsel et al., 1998] and may induce important conformational changes which might influence Prp8's function.

Bibliography

Achsel, T., Ahrens, K., Brahms, H., Teigelkamp, S. and Lührmann, R. (1998). The human U5-220kD protein (hPrp8) forms a stable RNA-free complex with several U5-specific proteins, including an RNA unwindase, a homologue of ribosomal elongation factor EF-2, and a novel WD-40 protein. Molecular and cellular biology *18*, 6756–6766.

Adams, P. D., Grosse-Kunstleve, R. W., Hung, L. W., Ioerger, T. R., McCoy, A. T., Moriarty, N. W., Read, R. J., Sacchettini, J. C., Sauter, N. K. and Terwilliger, T. C. (2002). Acta Cryst *D58*, 1948–1954.

Andersen, C. B., Ballut, L., Johansen, J. S., Chamieh, H., Nielsen, K. H., Oliveira, C. L., Pedersen, J. S., Sraphin, B., Le Hir, H. and Andersen, G. R. (2006). Structure of the exon junction core complex with a trapped DEAD-box ATPase bound to RNA. Science *313*, 19681972.

Anderson, G., Bach, M., Lührmann, R. and Beggs, J. (1989). Conservation between yeast and man of a protein associated with U5 small nuclear ribonucleoprotein. Nature *342*, 819–821.

Baetu, T. M. (2012). Genes after the human genome project. Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences *43*, 191–201.

Bartels, C., Klatt, C., Lührmann, R. and Fabrizio, P. (2002). The ribosomal translocase homologue Snu114p is involved in unwinding U4/U6 RNA during activation of the spliceosome. EMBO Rep *3*, 875–880.

Battye, T. G. G., Kontogiannis, L., Johnson, O., Powell, H. R. and Leslie, A. G. W. (2011). iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM. Acta Crystallographica Section D Biological Crystallography *67*, 271–281.

Bellare, P., Small, E. C., Huang, X., Wohlschlegel, J. A., Staley, J. P. and Sontheimer,

E. J. (2005). Ubiquitin binding by a variant Jab1/MPN domain in the essential pre-mRNA splicing factor Prp8p. RNA *12*, 292–302.

Bellare, P., Small, E. C., Huang, X., Wohlschlegel, J. A., Staley, J. P. and Sontheimer, E. J. (2008). A role for ubiquitin in the spliceosome assembly pathway. Nature Structural & Molecular Biology *15*, 444–451.

Benaglio, P., McGee, T. L., Capelli, L. P., Harper, S., Berson, E. L. and Rivolta, C. (2011). Next generation sequencing of pooled samples reveals new SNRNP200 mutations associated with retinitis pigmentosa. Human Mutation *32*, E2246–E2258.

Bleichert, F. and Baserga, S. J. (2007). The Long Unwinding Road of RNA Helicases. Molecular Cell *27*, 339–352.

Bono, F., Ebert, J., Lorentzen, E. and Conti, E. (2006). The Crystal Structure of the Exon Junction Complex Reveals How It Maintains a Stable Grip on mRNA. Cell *126*, 713–725.

Boon, K.-L. (2005). Prp8p dissection reveals domain structure and protein interaction sites. RNA *12*, 198–205.

Boon, K.-L., Grainger, R. J., Ehsani, P., Barrass, J. D., Auchynnikava, T., Inglehearn, C. F. and Beggs, J. D. (2007). Prp8 mutations that cause human retinitis pigmentosa lead to a U5 snRNP maturation defect in yeast. Nature Structural & Molecular Biology *14*, 1077–1083.

Brantl, S. (2002). Antisense-RNA regulation and RNA interference. Biochimica et Biophysica Acta (BBA)-Gene Structure and Expression *1575*, 15–25.

Brenner, T. J. and Guthrie, C. (2006). Assembly of Snu114 into U5 snRNP requires Prp8 and a functional GTPase domain. RNA *12*, 862–871.

Brünger, A., Adams, P., Clore, G., DeLano, W., Gros, P., Grosse-Kunstleve, R.W.and Jiang, J., Kuszewski, J., Nilges, M. and Pannu, N. (1998). Crystallography & NMR system: a new software suite for macromolecular structure determination.

Buratti, E. and Baralle, D. (2010). Novel roles of U1 snRNP in alternative splicing regulation. RNA Biology *7*, 412–419.

Burge, C. B., Padgett, R. A. and Sharp, P. A. (1998). Evolutionary fates and origins of U12-type introns. Molecular cell *2*, 773–785.

Burgess, S. M. and Guthrie, C. (1993). A mechanism to enhance mRNA splicing fidelity: the RNA-dependent ATPase Prp16 governs usage of a discard pathway for aberrant lariat intermediates. Cell *73*, 1377–1391.

Büttner, K., Nehring, S. and Hopfner, K.-P. (2007). Structural basis for DNA duplex separation by a superfamily-2 helicase. Nature Structural & Molecular Biology *14*, 647–652.

Calvin, K. and Li, H. (2008). RNA-splicing endonuclease structure and function. Cellular and Molecular Life Sciences *65*, 1176–1185.

Caprara, M. G. and Nilsen, T. W. (2000). RNA: Versatility in form and function. Nature Structural Biology *7*, 831–833.

Chakarova, C. F., Hims, M. M., Bolz, H., Abu-Safieh, L., Patel, R. J., Papaioannou, M. G., Inglehearn, C. F., Keen, T. J., Willis, C., Moore, A. T., Rosenberg, T., Webster, A. R., Bird, A. C., Gal, A., Hunt, D., Vithana, E. N. and Bhattacharya, S. S. (2002). Mutations in HPRP3, a third member of pre-mRNA splicing factor genes, implicated in autosomal dominant retinitis pigmentosa. Human Molecular Genetics *11*, 87–92.

Chan, S.-P. and Cheng, S.-C. (2005). The Prp19-associated Complex Is Required for Specifying Interactions of U5 and U6 with Pre-mRNA during Spliceosome Activation. Journal of Biological Chemistry *280*, 31190–31199.

Chan, S.-P., Kao, D.-I., Tsai, W.-Y. and Cheng, S.-C. (2003). The Prp19p-Associated Complex in Spliceosome Activation. Science *302*, 279–282.

Chandran, V., Poljak, L., Vanzo, N. F., Leroy, A., Miguel, R. N., Fernandez-Recio, J., Parkinson, J., Burns, C., Carpousis, A. J. and Luisi, B. F. (2007). Recognition and Cooperation Between the ATP-dependent RNA Helicase RhlB and Ribonuclease RNase E. Journal of Molecular Biology *367*, 113–132.

Choudhary, C., Kumar, C., Gnad, F., Nielsen, M. L., Rehman, M., Walther, T. C., Olsen, J. V. and Mann, M. (2009). Lysine Acetylation Targets Protein Complexes and Co-Regulates Major Cellular Functions. Science *325*, 834–840.

Combs, D. J., Nagel, R. J., Ares Jr, M. and Stevens, S. W. (2006). Prp43p is a DEAH-box spliceosome disassembly factor essential for ribosome biogenesis. Molecular and cellular biology *26*, 523–534.

Contreras, R., Rogiers, R., de Voorde, A. V. and Fiers, W. (1977). Overlapping of the VP2-VP3 gene and the VP1 gene in the SV40 genome. Cell *12*, 529–538.

Cordin, O., Hahn, D. and Beggs, J. D. (2012). Structure, function and regulation of spliceosomal RNA helicases. Current Opinion in Cell Biology *24*, 431–438.

Crick, F. (1958). On protein synthesis. Symp Soc Exp Biol *12*.

Crick, F. (1970). Central Dogma of Molecular Biology. Nature *227*, 561–563.

Dango, S., Mosammaparast, N., Sowa, M. E., Xiong, L.-J., Wu, F., Park, K., Rubin, M., Gygi, S., Harper, J. W. and Shi, Y. (2011). DNA Unwinding by ASCC3 Helicase Is Coupled to ALKBH3-Dependent DNA Alkylation Repair and Cancer Cell Proliferation. Molecular Cell *44*, 373–384.

De Erkenez, A., Berson, E. and Dryja, T. (2002). Novel Mutations in the PRPC8 Gene, Encoding a Pre-mRNA Splicing Factor in Patients with Autosomal Dominant Retinitis Pigmentosa. Invest. Ophthalmol. Vis. Sci. *43*, E–abstract 791.

Deniz, A. A., Laurence, T. A., Beligere, G. S., Dahan, M., Martin, A. B., Chemla, D. S., Dawson, P. E., Schultz, P. G. and Weiss, S. (2000). Single-molecule protein folding: diffusion fluorescence resonance energy transfer studies of the denaturation of chymotrypsin inhibitor 2. Proceedings of the National Academy of Sciences *97*, 5179–5184.

Dephoure, N., Zhou, C., Villn, J., Beausoleil, S. A., Bakalarski, C. E., Elledge, S. J. and Gygi, S. P. (2008). A quantitative atlas of mitotic phosphorylation. Proceedings of the National Academy of Sciences *105*, 1076210767.

Dlakic, M. and Mushegian, A. (2011). Prp8, the pivotal protein of the spliceosomal catalytic center, evolved from a retroelement-encoded reverse transcriptase. RNA *17*, 799–808.

Dossani, Z. Y., Weirich, C. S., Erzberger, J. P., Berger, J. M. and Weis, K. (2009). Structure of the C-terminus of the mRNA export factor Dbp5 reveals the interaction surface for the ATPase activator Gle1. Proceedings of the National Academy of Sciences *106*, 1625116256.

Dybkov, O., Will, C. L., Deckert, J., Behzadnia, N., Hartmuth, K. and Lührmann, R. (2006). U2 snRNA-protein contacts in purified human 17S U2 snRNPs and in spliceosomal A and B complexes. Molecular and cellular biology *26*, 2803–2816.

Dziembowski, A., Ventura, A. P., Rutz, B., Caspary, F., Faux, C., Halgand, F., Laprvote, O. and Sraphin, B. (2004). Proteomic analysis identifies a new complex required for nuclear pre-mRNA retention and splicing. The EMBO journal *23*, 4847–4856.

Emsley, P. and Cowtan, K. (2004). Coot: model-building tools for molecular graphics. Acta Cryst. *D60*, 2126–32.

Fabrizio, P., Dannenberg, J., Dube, P., Kastner, B., Stark, H., Urlaub, H. and Lührmann, R. (2009). The Evolutionarily Conserved Core Design of the Catalytic Activation Step of the Yeast Spliceosome. Molecular Cell *36*, 593–608.

Fairbanks, G., Steck, T. and Wallach, D. (1971). Electrophoretic analysis of the major polypeptides of the human erythrocyte membrane. Biochemistry *10*, 2606–17.

Fairman-Williams, M. E., Guenther, U.-P. and Jankowsky, E. (2010). SF1 and SF2 helicases: family matters. Current Opinion in Structural Biology *20*, 313–324.

Fedorova, O. and Zingler, N. (2007). Group II introns: structure, folding and splicing mechanism. Biological Chemistry *388*, 665–678.

Fitzgerald, D. J., Berger, P., Schaffitzel, C., Yamada, K., Richmond, T. J. and Berger, I. (2006). Protein complex expression by using multigene baculoviral vectors. Nat Methods *3*, 1021–32.

Fourmann, J.-B., Schmitzova, J., Christian, H., Urlaub, H., Ficner, R., Boon, K.-L., Fabrizio, P. and Luhrmann, R. (2013). Dissection of the factor requirements for spliceosome disassembly and the elucidation of its dissociation products using a purified splicing system. Genes & Development *27*, 413–428.

Fox-Walsh, K. L., Dou, Y., Lam, B. J., Hung, S., Baldi, P. F. and Hertel, K. J. (2005). The architecture of pre-mRNAs affects mechanisms of splice-site pairing. Proceedings of the National Academy of Sciences of the United States of America *102*, 16176–16181.

Galej, W. P., Oubridge, C., Newman, A. J. and Nagai, K. (2013). Crystal structure of Prp8 reveals active site cavity of the spliceosome. Nature *493*, 638–643.

Gerstein, M. B., Bruce, C., Rozowsky, J. S., Zheng, D., Du, J., Korbel, J. O., Emanuelsson, O., Zhang, Z. D., Weissman, S. and Snyder, M. (2007). What is a gene, post-ENCODE? History and updated definition. Genome Research *17*, 669–681.

Golas, M. M., Sander, B., Will, C. L., Lührmann, R. and Stark, H. (2003). Molecular Architecture of the Multiprotein Splicing Factor SF3b. Science *300*, 980–984.

Golas, M. M., Sander, B., Will, C. L., Lührmann, R. and Stark, H. (2005). Major Conformational Change in the Complex SF3b upon Integration into the Spliceosomal U11/U12 di-snRNP as Revealed by Electron Cryomicroscopy. Molecular Cell *17*, 869–883.

Grainger, R. J., Barrass, J. D., Jacquier, A., Rain, J.-C. and Beggs, J. D. (2009). Physical and genetic interactions of yeast Cwc21p, an ortholog of human SRm300/SRRM2, suggest a role at the catalytic center of the spliceosome. RNA *15*, 2161–2173.

Grainger, R. J. and Beggs, J. (2005). Prp8 protein: At the heart of the spliceosome. RNA *11*, 533–557.

Guthrie, C. and Patterson, B. (1988). Spliceosomal snRNAs. Annual review of genetics *22*, 387419.

Häcker, I., Sander, B., Golas, M. M., Wolf, E., Karagöz, E., Kastner, B., Stark, H., Fabrizio, P. and Lührmann, R. (2008). Localization of Prp8, Brr2, Snu114 and U4/U6 proteins in the yeast tri-snRNP by electron microscopy. Nature Structural & Molecular Biology *15*, 1206–1212.

Hahn, D., Kudla, G., Tollervey, D. and Beggs, J. D. (2012). Brr2p-mediated conformational rearrangements in the spliceosome during activation and substrate repositioning. Genes & Development *26*, 2408–2421.

Hartong, D. T., Berson, E. L. and Dryja, T. P. (2006). Retinitis pigmentosa. The Lancet *368*, 1795–1809.

Haugen, P., Simon, D. M. and Bhattacharya, D. (2005). The natural history of group I introns. Trends in Genetics *21*, 111–119.

He, Y., Andersen, G. R. and Nielsen, K. H. (2010). Structural basis for the function of DEAH helicases. EMBO Rep *11*, 180–186.

Heras, B. n. and Martin, J. L. (2005). Post-crystallization treatments for improving diffraction quality of protein crystals. Acta Crystallographica Section D Biological Crystallography *61*, 1173–80.

Herold, N., Will, C. L., Wolf, E., Kastner, B., Urlaub, H. and Lührmann, R. (2008). Conservation of the Protein Composition and Electron Microscopy Structure of Drosophila melanogaster and Human Spliceosomal Complexes. Molecular and Cellular Biology *29*, 281–301.

Hodges, P., Jackson, S., Brown, J. and Beggs, J. (1995). Extraordinary sequence conservation of the PRP8 splicing factor. Yeast *11*, 337–42.

Ivings, L., Towns, K. V., Matin, M. A., Taylor, C., Ponchel, F., Grainger, R. J., Ramesar, R. S., Mackey, D. A. and Inglehearn, C. F. (2008). Evaluation of splicing

efficiency in lymphoblastoid cell lines from patients with splicing-factor retinitis pigmentosa. Molecular vision *14*, 2357.

Jackson, S. P., Lossky, M. and Beggs, J. D. (1988). Cloning of the RNA8 gene of Saccharomyces cerevisiae, detection of the RNA8 protein, and demonstration that it is essential for nuclear pre-mRNA splicing. Molecular and cellular biology *8*, 1067–1075.

Jankowsky, E. and Bowers, H. (2006). Remodeling of ribonucleoprotein complexes with DExH/D RNA helicases. Nucleic Acids Research *34*, 4181–4188.

Jurica, M. S. and Moore, M. J. (2003). Pre-mRNA splicing: awash in a sea of proteins. Molecular cell *12*, 5–14.

Kabsch, W. (2010). XDS. Acta Crystallographica Section D *66*, 125–132.

Keen, T., Hims, M., McKie, A., Moore, A., Doran, R., Mackey, D., Mansfield, D., Mueller, R., Bhattacharya, S., Bird, A., Markham, A. and Inglehearn, C. (2002). Mutations in a protein target of the Pim-1 kinase associated with the RP9 form of autosomal dominant retinitis pigmentosa. Eur J Hum Genet *10*, 245–249.

Kim, D. H. O. and Rossi, J. J. (1999). The first ATPase domain of the yeast 246-kDa protein is required for in vivo unwinding of the U4/U6 duplex. RNA *5*, 959–971.

Kim, R. Y., Fitzke, F. W., Moore, A. T., Jay, M., Inglehearn, C., Arden, G. B., Bhattacharya, S. S. and Bird, A. C. (1995). Autosomal dominant retinitis pigmentosa mapping to chromosome 7p exhibits variable expression. The British Journal of Ophthalmology *79*, 23–27.

Kim, S. H. and Lin, R. J. (1996). Spliceosome activation by PRP2 ATPase prior to the first transesterification reaction of pre-mRNA splicing. Molecular and cellular biology *16*, 6810–6819.

Kim, S. H., Smith, J., Claude, A. and Lin, R. J. (1992). The purified yeast pre-mRNA splicing factor PRP2 is an RNA-dependent NTPase. The EMBO journal *11*, 2319–2326.

Kistler, A. L. and Guthrie, C. (2001). Deletion of MUD2, the yeast homolog of U2AF65, can bypass the requirement for sub2, an essential spliceosomal ATPase. Genes & development *15*, 42–49.

Kohtz, J. D., Jamison, S. F., Will, C. L., Zuo, P., Lührmann, R., Garcia-Blanco, M. A. and Manley, J. L. (1994). Proteinprotein interactions and 5'-splice-site recognition in mammalian mRNA precursors. Nature *368*, 119–124.

Konarska, M. M., Padgett, R. A. and Sharp, P. A. (1985). Trans splicing of mrna precursors in vitro. Cell *42*, 165–171.

Kondo, H., Tahira, T., Mizota, A., Adachi-Usami, E., Oshima, K. and Hayashi, K. (2003). Diagnosis of Autosomal Dominant Retinitis Pigmentosa by Linkage-Based Exclusion Screening with Multiple Locus-Specific Microsatellite Markers. Investigative Ophthalmology & Visual Science *44*, 1275–1281.

Kopec, J. and Schneider, G. (2011). Comparison of fluorescence and light scattering based methods to assess formation and stability of proteinprotein complexes. Journal of Structural Biology *175*, 216–223.

Korneta, I., Magnus, M. and Bujnicki, J. M. (2012). Structural bioinformatics of the human spliceosomal proteome. Nucleic Acids Research *40*, 7046–7065.

Kosowski, T. R., Keys, H. R., Quan, T. K. and Ruby, S. W. (2009). DExD/H-box Prp5 protein is in the spliceosome during most of the splicing cycle. RNA *15*, 1345–1362.

Krol, A., Westhof, E., Bach, M., Lührmann, R., Ebel, J. P. and Carbon, P. (1990). Solution structure of human U1 snRNA. Derivation of a possible three-dimensional model. Nucleic acids research *18*, 3803–3811.

Kuhn, A. N., Li, Z. and Brow, D. A. (1999). Splicing factor Prp8 governs U4/U6 RNA unwinding during activation of the spliceosome. Molecular cell *3*, 65–75.

Kuhn, A. N., Reichl, E. M. and Brow, D. A. (2002). Distinct domains of splicing factor Prp8 mediate different aspects of spliceosome activation. Proceedings of the National Academy of Sciences *99*, 9145–9149.

Laemmli, U. K. (1970). Cleavage of Structural Proteins during the Assembly of the Head of Bacteriophage T4. Nature *227*, 680–685.

Laggerbauer, B., Achsel, T. and Lührmann, R. (1998). The human U5-200kD DEXH-box protein unwinds U4/U6 RNA duplices in vitro. Proceedings of the National Academy of Sciences *95*, 4188–4192.

Laggerbauer, B., Lauber, J. and Lührmann, R. (1996). Identification of an RNA-dependent ATPase activity in mammalian U5 snRNPs. Nucleic acids research *24*, 868–875.

Lambowitz, A. M. and Zimmerly, S. (2010). Group II Introns: Mobile Ribozymes that Invade DNA. Cold Spring Harbor Perspectives in Biology *3*, a003616–a003616.

Lasda, E. L. and Blumenthal, T. (2011). Trans-splicing. Wiley Interdisciplinary Reviews: RNA *2*, 417–434.

Lebaron, S., Froment, C., Fromont-Racine, M., Rain, J. C., Monsarrat, B., Caizergues-Ferrer, M. and Henry, Y. (2005). The splicing ATPase prp43p is a component of multiple preribosomal particles. Molecular and cellular biology *25*, 9269–9282.

Lebaron, S., Papin, C., Capeyrou, R., Chen, Y. L., Froment, C., Monsarrat, B., Caizergues-Ferrer, M., Grigoriev, M. and Henry, Y. (2009). The ATPase and helicase activities of Prp43p are stimulated by the G-patch protein Pfa1p during yeast ribosome biogenesis. The EMBO journal *28*, 3808–3819.

Leeds, N. B., Small, E. C., Hiley, S. L., Hughes, T. R. and Staley, J. P. (2006). The splicing factor Prp43p, a DEAH box ATPase, functions in ribosome biogenesis. Molecular and cellular biology *26*, 513–522.

Leung, A., Nagai, K. and Li, J. (2011). Structure of the spliceosomal U4 snRNP core domain and its implication for snRNP biogenesis. Nature *473*, 536–539.

Lewin, B. (2003). Genes VIII. 8th edition, Benjamin Cummings.

Liu, S., Ghalei, H., Lührmann, R. and Wahl, M. C. (2011). Structural basis for the dual U4 and U4atac snRNA-binding specificity of spliceosomal protein hPrp31. RNA *17*, 1655–1663.

Liu, S., Li, P., Dybkov, O., Nottrott, S., Hartmuth, K., Lührmann, R., Carlomagno, T. and Wahl, M. C. (2007). Binding of the Human Prp31 Nop Domain to a Composite RNA-Protein Platform in U4 snRNP. Science *316*, 115–120.

Liu, S., Rauhut, R., Vornlocher, H.-P. and Lührmann, R. (2006). The network of protein-protein interactions within the human U4/U6.U5 tri-snRNP. RNA *12*, 1418–1430.

Lohman, T. M. and Bjornson, K. P. (1996). Mechanisms of Helicase-Catalyzed DNA Unwinding. Annual Review of Biochemistry *65*, 169–214.

Lohman, T. M., Tomko, E. J. and Wu, C. G. (2008). Non-hexameric DNA helicases and translocases: mechanisms and regulation. Nature Reviews Molecular Cell Biology *9*, 391–401.

Lossky, M., Anderson, G. J., Jackson, S. P. and Beggs, J. (1987). Identification of a yeast snRNP protein and detection of snRNP-snRNP interactions. Cell *51*, 1019–1026.

Luckow, V. A. and Summers, M. D. (1988). Trends in the Development of Baculovirus Expression Vectors. Nature Biotechnology *6*, 47–55.

Luco, R. F., Allo, M., Schor, I. E., Kornblihtt, A. R. and Misteli, T. (2011). Epigenetics in Alternative Pre-mRNA Splicing. Cell *144*, 16–26.

Luo, M.-J., Zhou, Z., Magni, K., Christoforides, C., Rappsilber, J., Mann, M. and Reed, R. (2001). Pre-mRNA splicing and mRNA export linked by direct interactions between UAP56 and Aly. Nature *413*, 644–647.

Lustig, A. J., Lin, R. J. and Abelson, J. (1986). The yeast RNA gene products are essential for mRNA splicing in vitro. Cell *47*, 953–963.

Lutzelberger, M., Bottner, C. A., Schwelnus, W., Zock-Emmenthal, S., Razanau, A. and Kaufer, N. F. (2009). The N-terminus of Prp1 (Prp6/U5-102 K) is essential for spliceosome activation in vivo. Nucleic Acids Research *38*, 1610–1622.

Maeder, C., Kutach, A. K. and Guthrie, C. (2008). ATP-dependent unwinding of U4/U6 snRNAs by the Brr2 helicase requires the C terminus of Prp8. Nature Structural & Molecular Biology *16*, 42–48.

Maita, H., Kitaura, H., Ariga, H. and Iguchi-Ariga, S. M. (2005). Association of PAP-1 and Prp3p, the products of causative genes of dominant retinitis pigmentosa, in the tri-snRNP complex. Experimental Cell Research *302*, 61–68.

Maita, H., Kitaura, H., Keen, T. J., Inglehearn, C. F., Ariga, H. and Iguchi-Ariga, S. M. (2004). PAP-1, the mutated gene underlying the RP9 form of dominant retinitis pigmentosa, is a splicing factor. Experimental Cell Research *300*, 283–296.

Makarov, E. M., Makarova, O. V., Achsel, T. and Lührmann, R. (2000). The human homologue of the yeast splicing factor Prp6p contains multiple TPR elements and is stably associated with the U5 snRNP via protein-protein interactions. Journal of Molecular Biology *298*, 567–575.

Makarova, O. V., Makarov, E. M., Urlaub, H., Will, C. L., Gentzel, M., Wilm, M. and Lührmann, R. (2004). A subset of human 35S U5 proteins, including Prp19, function prior to catalytic step 1 of splicing. The EMBO journal *23*, 2381–2391.

Marintchev, A., Edmonds, K. A., Marintcheva, B., Hendrickson, E., Oberer, M., Suzuki, C., Herdy, B., Sonenberg, N. and Wagner, G. (2009). Topology and Regulation of the Human eIF4A/4G/4H Helicase Complex in Translation Initiation. Cell *136*, 447–460.

Martegani, E., Vanoni, M., Mauri, I., Rudoni, S., Saliola, M. and Alberghina, L. (1997). Identification of Gene encoding a Putative RNA-Helicase, Homologous to SKI2, in Chromosome VII of Saccharomyces cerevisiae. Yeast *13*, 391–397.

Martinez-Gimeno, M., Gamundi, M., Hernan, I., Maseras, M., Milla, E., Ayuso, C., García-Sandoval, B., Beneyto, M., Vilela, C., Baiget, M., Antinolo, G. and Carballo, M. (2003). Mutations in the Pre-mRNA Splicing-Factor Genes PRPF3, PRPF8, and PRPF31 in Spanish Families with Autosomal Dominant Retinitis Pigmentosa. Investigative Ophthalmology & Visual Science *44*, 2171–2177.

Mathew, R., Hartmuth, K., Möhlmann, S., Urlaub, H., Ficner, R. and Lührmann, R. (2008). Phosphorylation of human PRP28 by SRPK2 is required for integration of the U4/U6-U5 tri-snRNP into the spliceosome. Nature structural & molecular biology *15*, 435–443.

Mayas, R. M., Maita, H. and Staley, J. P. (2006). Exon ligation is proofread by the DExD/H-box ATPase Prp22p. Nature Structural & Molecular Biology *13*, 482–490.

Maytal-Kivity, V., Reis, N., Hofmann, K. and Glickman, M. H. (2002). MPN+, a putative catalytic motif found in a subset of MPN domain proteins from eukaryotes and prokaryotes, is critical for Rpn11 function. BMC biochemistry *3*, 28–39.

McCoy, A., Grosse-Kunstleve, R., Adams, P., Winn, M., Storoni, L. C. and Read, R. (2007).

McKie, A. B., McHale, J. C., Keen, T. J., Tarttelin, E. E., Goliath, R., Lith-Verhoeven, J. J. C. v., Greenberg, J., Ramesar, R. S., Hoyng, C. B., Cremers, F. P. M., Mackey, D. A., Bhattacharya, S. S., Bird, A. C., Markham, A. F. and Inglehearn, C. F. (2001). Mutations in the pre-mRNA splicing factor gene PRPC8 in autosomal dominant retinitis pigmentosa (RP13). Human Molecular Genetics *10*, 1555–1562.

Melnikov, S., Ben-Shem, A., Garreau de Loubresse, N., Jenner, L., Yusupova, G. and Yusupov, M. (2012). One core, two shells: bacterial and eukaryotic ribosomes. Nature Structural & Molecular Biology *19*, 560–567.

Montes, M., Becerra, S., Snchez-lvarez, M. and Su, C. (2012). Functional coupling of transcription and splicing. Gene *501*, 104–117.

Moore, P. and Steitz, T. (2011). The roles of RNA in the synthesis of protein. Cold Spring Harb Perspect Biol *3*, a003780.

Mordes, D., Luo, X., Kar, A., Kuo, D., Xu, L., Fushimi, K., Yu, G., Sternberg Jr, P. and Wu, J. Y. (2006). Pre-mRNA splicing and retinitis pigmentosa. Molecular vision *12*, 1259–1271.

Mozaffari-Jovin, S., Santos, K. F., Hsiao, H.-H., Will, C. L., Urlaub, H., Wahl, M. C. and Lührmann, R. (2012). The Prp8 RNase H-like domain inhibits Brr2-mediated U4/U6 snRNA unwinding by blocking Brr2 loading onto the U4 snRNA. Genes & Development *26*, 2422–2434.

Mozaffari-Jovin, S., Wandersleben, T., Santos, K. F., Will, C. L., Luhrmann, R. and Wahl, M. C. (2013). Inhibition of RNA Helicase Brr2 by the C-Terminal Tail of the Spliceosomal Protein Prp8. Science .

Nakagawa, T. and Kolodner, R. D. (2002). The MER3 DNA Helicase Catalyzes the Unwinding of Holliday Junctions. Journal of Biological Chemistry *277*, 28019–28024.

Nakai, K. and Horton, P. (1999). PSORT: a program for detecting sorting signals in proteins and predicting their subcellular localization. Trends in Biochemical Sciences *24*, 34–35.

Newman, A. J. (1997). The role of U5 snRNP in pre-mRNA splicing. The EMBO journal *16*, 5797–5800.

Nguyen, T. H. D., Li, J., Galej, W. P., Oshikane, H., Newman, A. J. and Nagai, K. (2013). Structural Basis of Brr2-Prp8 Interactions and Implications for U5 snRNP Biogenesis and the Spliceosome Active Site. Structure *21*, 910–919.

Nilsen, T. W. (1994). RNA-RNA interactions in the spliceosome: Unraveling the ties that bind. Cell *78*, 1–4.

Noble, S. M. and Guthrie, C. (1996). Identification of novel genes required for yeast pre-mRNA splicing by means of cold-sensitive mutations. Genetics *143*, 67–80.

Normark, S., Bergstrom, S., Edlund, T., Grundstrom, T., Jaurin, B., Lindberg, F. P. and Olsson, O. (1983). Overlapping Genes. Annual Review of Genetics *17*, 499–525.

Nottrott, S., Urlaub, H. and Lührmann, R. (2002). Hierarchical, clustered protein interactions with U4/U6 snRNA: a biochemical role for U4/U6 proteins. The EMBO journal *21*, 5527–5538.

O'Day, C. L., Dalbadie-McFarland, G. and Abelson, J. (1996). The Saccharomyces cerevisiae Prp5 protein has RNA-dependent ATPase activity with specificity for U2 small nuclear RNA. Journal of Biological Chemistry *271*, 33261–33267.

Olsen, J. V., Vermeulen, M., Santamaria, A., Kumar, C., Miller, M. L., Jensen, L. J., Gnad, F., Cox, J., Jensen, T. S. and Nigg, E. A. (2010). Quantitative phosphoproteomics reveals widespread full phosphorylation site occupancy during mitosis. Science signaling *3*, ra3 1–15.

Oubridge, C., Ito, N., Evans, P., Teo, C. 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.

Pasman, Z., Wang, J., Will, C., Lührmann, R., Manley, J. L. and Garcia-Blanco, M. A. (1995). U1 snRNP-ASF/SF2 interaction and 5 splice site recognition: characterization of required elements. Nucleic acids research *23*, 3260–3267.

Patel, S. B. and Bellini, M. (2008). The assembly of a spliceosomal small nuclear ribonucleoprotein particle. Nucleic Acids Research *36*, 6482–6493.

Pena, V., Jovin, S. M., Fabrizio, P., Orlowski, J., Bujnicki, J. M., Lührmann, R. and Wahl, M. C. (2009). Common Design Principles in the Spliceosomal RNA Helicase Brr2 and in the Hel308 DNA Helicase. Molecular Cell *35*, 454–466.

Pena, V., Liu, S., Bujnicki, J. M., Lührmann, R. and Wahl, M. C. (2007). Structure of a Multipartite Protein-Protein Interaction Domain in Splicing Factor Prp8 and Its Link to Retinitis Pigmentosa. Molecular Cell *25*, 615–624.

Pena, V., Rozov, A., Fabrizio, P., Lührmann, R. and Wahl, M. C. (2008). Structure and function of an RNase H domain at the heart of the spliceosome. The EMBO Journal *27*, 2929–2940.

Perriman, R., Barta, I., Voeltz, G. K., Abelson, J. and Ares, M. (2003). ATP requirement for Prp5p function is determined by Cus2p and the structure of U2 small nuclear RNA. Proceedings of the National Academy of Sciences *100*, 13857–13862.

Perriman, R. J. and Ares, M. (2007). Rearrangement of competing U2 RNA helices within the spliceosome promotes multiple steps in splicing. Genes & development *21*, 811–820.

Pessa, H. K. (2006). The abundance of the spliceosomal snRNPs is not limiting the splicing of U12-type introns. RNA *12*, 1883–1892.

Pinto, A. L. and Steitz, J. A. (1989). The mammalian analogue of the yeast PRP8 splicing protein is present in the U4/5/6 small nuclear ribonucleoprotein particle and the spliceosome. Proceedings of the National Academy of Sciences *86*, 8742–8746.

Pomeranz Krummel, D. A., Oubridge, C., Leung, A. K. W., Li, J. and Nagai, K. (2009). Crystal structure of human spliceosomal U1 snRNP at 5.5 resolution. Nature *458*, 475–480.

Ponting, C. P. (2000). Proteins of the endoplasmic-reticulum-associated degradation pathway: domain detection and function prediction. Biochemical Journal *351*, 527–535.

Potterton, L., McNicholas, S., Krissinel, E., Gruber, J., Cowtan, K., Emsley, P., Murshudov, G. N., Cohen, S., Perrakis, A. and Noble, M. (2004). Developments in the CCP4 molecular-graphics project. Acta Cryst. *D60*, 2288–94.

Price, S., Evans, P. and Nagai, K. (1998). Crystal structure of the spliceosomal U2B"-U2A' protein complex bound to a fragment of U2 small nuclear RNA. Nature *394*, 645–650.

Query, C. C. and Konarska, M. M. (2004). Suppression of Multiple Substrate Mutations by Spliceosomal prp8 Alleles Suggests Functional Correlations with Ribosomal Ambiguity Mutants. Molecular cell *14*, 343–354.

Raghunathan, P. L., Guthrie, C. et al. (1998). RNA unwinding in U4/U6 snRNPs requires ATP hydrolysis and the DEIH-box splicing factor Brr2. Current biology *8*, 847–855.

Reed, R. (2000). Mechanisms of fidelity in pre-mRNA splicing. Current opinion in cell biology *12*, 340–345.

Reuter, K., Nottrott, S., Fabrizio, P., Lührmann, R. and Ficner, R. (1999). Identification, characterization and crystal structure analysis of the human spliceosomal U5 snRNP-specific 15 kd protein. Journal of Molecular Biology *294*, 515–525.

Reyes, J. L., Gustafson, E. H., Luo, H. R., Moore, M. J. and Konarska, M. M. (1999). The C-terminal region of hPrp8 interacts with the conserved GU dinucleotide at the 5'splice site. RNA *5*, 167–179.

Ritchie, D. B., Schellenberg, M. J., Gesner, E. M., Raithatha, S. A., Stuart, D. T. and MacMillan, A. M. (2008). Structural elucidation of a PRP8 core domain from the heart of the spliceosome. Nature Structural & Molecular Biology *15*, 1199–1205.

Rozovsky, N., Butterworth, A. C. and Moore, M. J. (2008). Interactions between eIF4AI and its accessory factors eIF4B and eIF4H. RNA *14*, 2136–2148.

Sambrook, J. and Russell, D. (2001). Molecular Cloning: A Laboratory Manual. Number v. 1 in Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory Press.

Sander, B., Golas, M. M., Makarov, E. M., Brahms, H., Kastner, B., Lührmann, R. and Stark, H. (2006). Organization of Core Spliceosomal Components U5 snRNA Loop I and U4/U6 Di-snRNP within U4/U6.U5 Tri-snRNP as Revealed by Electron Cryomicroscopy. Molecular Cell *24*, 267–278.

Santos, K. F. (2012). Structural and functional studies of the spliceosomal RNP remodeling enzyme Brr2. PhD thesis, Division of Mathematics and Natural Sciences of the Georg August Universität Göttingen.

Santos, K. F., Jovin, S. M., Weber, G., Pena, V., Lührmann, R. and Wahl, M. C. (2012). Structural basis for functional cooperation between tandem helicase cassettes in Brr2-mediated remodeling of the spliceosome. Proceedings of the National Academy of Sciences *109*, 17418–17423.

Schellenberg, M. J., Edwards, R. A., Ritchie, D. B., Kent, O. A., Golas, M. M., Stark, H., Lührmann, R., Glover, J. N. M. and MacMillan, A. M. (2006). Crystal structure of a core spliceosomal protein interface. Proceedings of the National Academy of Sciences of the United States of America *103*, 1266–1271.

Schneider, C., Will, C. L., Makarova, O. V., Makarov, E. M. and Lührmann, R. (2002). Human U4/U6. U5 and U4atac/U6atac. U5 tri-snRNPs exhibit similar protein compositions. Molecular and cellular biology *22*, 3219–3229.

Schwer, B. (2008). A conformational rearrangement in the spliceosome sets the stage for Prp22-dependent mRNA release. Molecular cell *30*, 743–754.

Schwer, B. and Gross, C. H. (1998). Prp22, a DExH-box RNA helicase, plays two distinct roles in yeast pre-mRNA splicing. The EMBO journal *17*, 2086–2094.

Schwer, B. and Guthrie, C. (1991). PRP16 is an RNA-dependent ATPase that interacts transiently with the spliceosome. , Published online: 07 February 1991; | doi:10.1038/349494a0 *349*, 494–499.

Searfoss, A. M. and Wickner, R. B. (2000). 3 poly (A) is dispensable for translation. Proceedings of the National Academy of Sciences *97*, 9133–9137.

Serganov, A. and Patel, D. J. (2007). Ribozymes, riboswitches and beyond: regulation of gene expression without proteins. Nature Reviews Genetics *8*, 776–790.

Sharp, P. A. and Burge, C. B. (1997). Classification of Introns: U2-Type or U12-Type. Cell *91*, 875–879.

Shen, J., Zhang, L. and Zhao, R. (2007). Biochemical characterization of the ATPase and helicase activity of UAP56, an essential pre-mRNA splicing and mRNA export factor. Journal of Biological Chemistry *282*, 22544–22550.

Silverman, E., Edwalds-Gilbert, G. and Lin, R.-J. (2003). DExD/H-box proteins and their partners: helping RNA helicases unwind. Gene *312*, 1–16.

Silverman, E. J., Maeda, A., Wei, J., Smith, P., Beggs, J. D. and Lin, R. J. (2004). Interaction between a G-patch protein and a spliceosomal DEXD/H-box ATPase that is critical for splicing. Molecular and cellular biology *24*, 10101–10110.

Small, E. C., Leggett, S. R., Winans, A. A. and Staley, J. P. (2006). The EF-G-like GTPase Snu114p Regulates Spliceosome Dynamics Mediated by Brr2p, a DExD/H Box ATPase. Molecular Cell *23*, 389–399.

Smith, C. W. J. and Valcárcel, J. (2000). Alternative pre-mRNA splicing: the logic of combinatorial control. Trends in biochemical sciences *25*, 381–388.

Smith, G. E., Summers, M. D. and Fraser, M. J. (1983). Production of human beta interferon in insect cells infected with a baculovirus expression vector. Molecular and Cellular Biology *3*, 2156–2165.

Solnick, D. (1985). Trans splicing of mRNA precursors. Cell *42*, 157–164.

Spingola, M., Grate, L., Haussler, D. and Ares Jr, M. (1999). Genome-wide bioinformatic and molecular analysis of introns in Saccharomyces cerevisiae. RNA *5*, 221–234.

Spirin, A. S. (2002). Omnipotent rna. FEBS letters *530*, 4–8.

Staley, J. P. and Guthrie, C. (1998a). Mechanical Devices of the Spliceosome: Motors, Clocks, Springs, and Things. Cell *92*, 315–326.

Staley, J. P. and Guthrie, C. (1998b). Mechanical devices of the spliceosome: review motors, clocks, springs, and things. Cell *92*, 315–326.

Staley, J. P., Guthrie, C. et al. (1999). An RNA switch at the 5'splice site requires ATP and the DEAD box protein Prp28p. Molecular cell *3*, 55–64.

Stark, H., Dube, P., Lührmann, R. and Kastner, B. (2001). Arrangement of RNA and proteins in the spliceosomal U1 small nuclear ribonucleoprotein particle. Nature *409*, 539–542.

Steitz, T. A. (2008). A structural understanding of the dynamic ribosome machine. Nat Rev Mol Cell Biol *9*, 242–253.

Stevens, S. W. and Abelson, J. (1999). Purification of the yeast U4/U6 U5 small nuclear ribonucleoprotein particle and identification of its proteins. Proceedings of the National Academy of Sciences *96*, 7226–7231.

Strässer, K. and Hurt, E. (2001). Splicing factor Sub2p is required for nuclear mRNA export through its interaction with Yra1p. Nature *413*, 648–652.

Strauss, E. J. and Guthrie, C. (1994). PRP28, a DEAD-boxprotein, is required for the first step of mRNA splicing in vitro. Nucleic acids research *22*, 3187–3193.

Sullivan, L. S., Bowne, S. J., Birch, D. G., Hughbanks-Wheaton, D., Heckenlively, J. R., Lewis, R. A., Garcia, C. A., Ruiz, R. S., Blanton, S. H., Northrup, H., Gire, A. I., Seaman, R., Duzkale, H., Spellicy, C. J., Zhu, J., Shankar, S. P. and Daiger, S. P. (2006). Prevalence of Disease-Causing Mutations in Families with Autosomal Dominant Retinitis Pigmentosa: A Screen of Known Genes in 200 Families. Investigative Ophthalmology & Visual Science *47*, 3052–3064.

Swaroop, A. and Zack, D. (2002). Transcriptome analysis of the retina. Genome Biology *3*, reviews1022.1–reviews1022.4.

Takemura, R., Takeiwa, T., Taniguchi, I., McCloskey, A. and Ohno, M. (2011). Multiple factors in the early splicing complex are involved in the nuclear retention of pre-mRNAs in mammalian cells. Genes to Cells *16*, 1035–1049.

Tanackovic, G., Ransijn, A., Thibault, P., Abou Elela, S., Klinck, R., Berson, E. L., Chabot, B. and Rivolta, C. (2011). PRPF mutations are associated with generalized defects in spliceosome formation and pre-mRNA splicing in patients with retinitis pigmentosa. Human Molecular Genetics *20*, 2116–2130.

Tanaka, N., Aronova, A. and Schwer, B. (2007). Ntr1 activates the Prp43 helicase to trigger release of lariat-intron from the spliceosome. Genes & Development *21*, 2312–2325.

Tanaka, N. and Schwer, B. (2005). Characterization of the NTPase, RNA-Binding, and RNA Helicase Activities of the DEAH-Box Splicing Factor Prp22. Biochemistry *44*, 9795–9803.

Tanaka, N. and Schwer, B. (2006). Mutations in PRP43 that uncouple RNA-dependent NTPase activity and pre-mRNA splicing function. Biochemistry *45*, 6510–6521.

Tange, T., Shibuya, T., Jurica, M. S. and Moore, M. J. (2005). Biochemical analysis of the EJC reveals two new factors and a stable tetrameric protein core. RNA *11*, 1869–1883.

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.

Temin, H. M. and Mizutani, S. (1970). Viral RNA-dependent DNA Polymerase: RNA-dependent DNA Polymerase in Virions of Rous Sarcoma Virus. Nature *226*, 1211–1213.

Toor, N., Keating, K. S., Taylor, S. D. and Pyle, A. M. (2008). Crystal Structure of a Self-Spliced Group II Intron. Science *320*, 77–82.

Towns, K. V., Kipioti, A., Long, V., McKibbin, M., Maubaret, C., Vaclavik, V., Ehsani, P., Springell, K., Kamal, M., Ramesar, R. S., Mackey, D. A., Moore, A. T., Mukhopadhyay, R., Webster, A. R., Black, G. C., O'Sullivan, J., Bhattacharya, S. S., Pierce, E. A., Beggs, J. D. and Inglehearn, C. F. (2010). Prognosis for splicing factor PRPF8 retinitis pigmentosa, novel mutations and correlation between human and yeast phenotypes. Human Mutation *31*, E1361–E1376.

Trowitzsch, S., Bieniossek, C., Nie, Y., Garzoni, F. and Berger, I. (2010). New baculovirus expression tools for recombinant protein complex production. Journal of Structural Biology *172*, 45 – 54.

Tsai, R., Tseng, C., Lee, P., Chen, H., Fu, R., Chang, K., Yeh, F. and Cheng, S. (2007). Dynamic Interactions of Ntr1-Ntr2 with Prp43 and with U5 Govern the Recruitment of Prp43 To Mediate Spliceosome Disassembly. Molecular and Cellular Biology *27*, 8027–8037.

Tsai, R. T., Fu, R. H., Yeh, F. L., Tseng, C. K., Lin, Y. C., Huang, Y. and Cheng, S. C. (2005). Spliceosome disassembly catalyzed by Prp43 and its associated components Ntr1 and Ntr2. Genes & development *19*, 2991–3003.

Tseng, C. K., Liu, H. L. and Cheng, S. C. (2011). DEAH-box ATPase Prp16 has dual roles in remodeling of the spliceosome in catalytic steps. RNA *17*, 145–154.

Turner, I., Norman, C., Churcher, M. and Newman, A. (2004). Roles of the U5 snRNP in spliceosome dynamics and catalysis. Biochemical Society Transactions *32*, 928–931.

Turner, I. A. (2006). Dissection of Prp8 protein defines multiple interactions with crucial RNA sequences in the catalytic core of the spliceosome. RNA *12*, 375–386.

Vagin, A. and Teplyakov, A. (1997). MOLREP: an automated program for molecular replacement. J. Appl. Cryst. *30*, 1022–1025.

Valadkhan, S. and Jaladat, Y. (2010). The spliceosomal proteome: At the heart of the largest cellular ribonucleoprotein machine. PROTEOMICS *10*, 4128–4141.

van Nues, R. W. and Beggs, J. D. (2001). Functional contacts with a range of splicing proteins suggest a central role for Brr2p in the dynamic control of the order of events in spliceosomes of Saccharomyces cerevisiae. Genetics *157*, 1451–1467.

Villa, T. and Guthrie, C. (2005). The Isy1p component of the NineTeen complex interacts with the ATPase Prp16p to regulate the fidelity of pre-mRNA splicing. Genes & development *19*, 1894–1904.

Vithana, E. N., Abu-Safieh, L., Allen, M. J., Carey, A., Papaioannou, M., Chakarova, C., Al-Maghtheh, M., Ebenezer, N. D., Willis, C. and Moore, A. T. (2001). A Human Homolog of Yeast Pre-mRNA Splicing Gene, PRP31, Underlies Autosomal Dominant Retinitis Pigmentosa on Chromosome 19q13.4 RP11. Molecular cell *8*, 375–381.

Wahl, M. C., Will, C. L. and Lührmann, R. (2009). The Spliceosome: Design Principles of a Dynamic RNP Machine. Cell *136*, 701–718.

Wain, H. M., Bruford, E. A., Lovering, R. C., Lush, M. J., Wright, M. W. and Povey, S. (2002). Guidelines for Human Gene Nomenclature. Genomics *79*, 464–470.

Wan, Y., Kertesz, M., Spitale, R. C., Segal, E. and Chang, H. Y. (2011). Understanding the transcriptome through RNA structure. Nature Reviews Genetics *12*, 641–655.

Wang, E. T., Sandberg, R., Luo, S., Khrebtukova, I., Zhang, L., Mayr, C., Kingsmore, S. F., Schroth, G. P. and Burge, C. B. (2008). Alternative isoform regulation in human tissue transcriptomes. Nature *456*, 470–476.

Wang, Y., Wagner, J. D. O. and Guthrie, C. (1998). The DEAH-box splicing factor Prp16 unwinds RNA duplexes in vitro. Current biology *8*, 441–451.

Weber, G., Cristõ, V. F., Santos, K. F., Jovin, S. M., Heroven, A. C., Holton, N., Lührmann, R., Beggs, J. D. and Wahl, M. C. (2013). Structural basis for dual roles of Aar2p in U5 snRNP assembly. Genes Dev *5*, 525–540.

Weber, G., Cristo, V. F., de L. Alves, F., Santos, K. F., Holton, N., Rappsilber, J., Beggs, J. D. and Wahl, M. C. (2011). Mechanism for Aar2p function as a U5 snRNP assembly factor. Genes & Development *25*, 1601–1612.

Weber, G., Trowitzsch, S., Kastner, B., Lührmann, R. and Wahl, M. C. (2010). Functional organization of the Sm core in the crystal structure of human U1 snRNP. The EMBO Journal *29*, 4172–4184.

Weidenhammer, E. M., Ruiz-Noriega, M. and Woolford, J. L. (1997). Prp31p promotes the association of the U4/U6 x U5 tri-snRNP with prespliceosomes to form spliceosomes in Saccharomyces cerevisiae. Molecular and Cellular Biology *17*, 3580–3588.

Weidenhammer, E. M., Singh, M., Ruiz-Noriega, M. and Woolford, J. L. (1996). The PRP31 gene encodes a novel protein required for pre-mRNA splicing in Saccharomyces cerevisiae. Nucleic Acids Research *24*, 1164–1170.

Weiss, S. (1999). Fluorescence Spectroscopy of Single Biomolecules. Science *283*, 1676–1683.

Widner, W. R. and Wickner, R. B. (1993). Evidence that the SKI antiviral system of Saccharomyces cerevisiae acts by blocking expression of viral mRNA. Molecular and cellular biology *13*, 4331–4341.

Will, C. L. (2004). The human 18S U11/U12 snRNP contains a set of novel proteins not found in the U2-dependent spliceosome. RNA *10*, 929–941.

Will, C. L. and Lührmann, R. (2001). Spliceosomal UsnRNP biogenesis, structure and function. Current Opinion in Cell Biology *13*, 290–301.

Will, C. L. and Lührmann, R. (2005). Splicing of a rare class of introns by the U12-dependent spliceosome. Biological Chemistry *386*, 713–724.

Will, C. L. and Lührmann, R. (2010). Spliceosome Structure and Function. Cold Spring Harbor Perspectives in Biology *3*, a003707–a003707.

Will, C. L., Urlaub, H., Achsel, T., Gentzel, M., Wilm, M. and Lührmann, R. (2002). Characterization of novel SF3b and 17S U2 snRNP proteins, including a human Prp5p homologue and an SF3b DEAD-box protein. The EMBO journal *21*, 4978–4988.

Woodman, I. L., Briggs, G. S. and Bolt, E. L. (2007). Archaeal Hel308 Domain V Couples DNA Binding to ATP Hydrolysis and Positions DNA for Unwinding Over the Helicase Ratchet. Journal of Molecular Biology *374*, 1139–1144.

Xu, Y. Z. and Query, C. C. (2007). Competition between the ATPase Prp5 and branch region-U2 snRNA pairing modulates the fidelity of spliceosome assembly. Molecular cell *28*, 838–849.

Yang, K., Zhang, L., Xu, T., Heroux, A. and Zhao, R. (2008). Crystal structure of the β-finger domain of Prp8 reveals analogy to ribosomal proteins. Proceedings of the National Academy of Sciences *105*, 13817–13822.

Zhang, L., Shen, J., Guarnieri, M. T., Heroux, A., Yang, K. and Zhao, R. (2007). Crystal structure of the C-terminal domain of splicing factor Prp8 carrying retinitis pigmentosa mutants. Protein Science *16*, 1024–1031.

Zhang, L., Xu, T., Maeder, C., Bud, L.-O., Shanks, J., Nix, J., Guthrie, C., Pleiss, J. A. and Zhao, R. (2009). Structural evidence for consecutive Hel308-like modules in the spliceosomal ATPase Brr2. Nature Structural & Molecular Biology *16*, 731–739.

Zhao, C., Bellur, D. L., Lu, S., Zhao, F., Grassi, M. A., Bowne, S. J., Sullivan, L. S., Daiger, S. P., Chen, L. J., Pang, C. P., Zhao, K., Staley, J. P. and Larsson, C. (2009). Autosomal-Dominant Retinitis Pigmentosa Caused by a Mutation in SNRNP200, a Gene Required for Unwinding of U4/U6 snRNAs. The American Journal of Human Genetics *85*, 617–627.

Ziviello, C., Simonelli, F., Testa, F., Anastasi, M., Marzoli, S. B., Falsini, B., Ghiglione, D., Macaluso, C., Manitto, M. P., Garre, C., Ciccodicola, A., Rinaldi, E. and Banfi, S. (2005). Molecular genetics of autosomal dominant retinitis pigmentosa (ADRP): a comprehensive study of 43 Italian families. Journal of Medical Genetics *42*, e43–e47.

Related Publications

Mozzafari-Jovin, S.*, Wandersleben, T.*, Santos, K. F.*, Will C. L., Lührmann, R., Wahl, M. (2013) Inhibition of RNA Helicase Brr2 by the C-Terminal Tail of the Spliceosomal Protein Prp8. Science.

* These authors contributed equally to this work.