



A noncanonical PWI domain in the N-terminal helicase-associated region of the spliceosomal Brr2 protein

Eva Absmeier,^a Leonie Rosenberger,^a Luise Apelt,^b Christian Becke,^a Karine F. Santos,^a Ulrich Stelzl^b and Markus C. Wahl^{a*}

Received 5 December 2014

Accepted 16 January 2015

Edited by Z. S. Derewenda, University of Virginia, USA

Keywords: Brr2 RNA helicase; pre-mRNA splicing; protein–protein interactions; PWI domain; spliceosome.

PDB reference: PWI-like domain of Brr2, 4rvq

Supporting information: this article has supporting information at journals.iucr.org/d

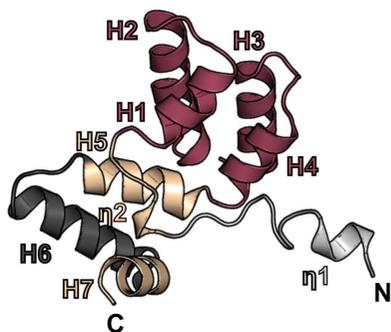
^aLaboratory of Structural Biochemistry, Freie Universität Berlin, Takustrasse 6, 14195 Berlin, Germany, and ^bMolecular Interaction Networks, Max-Planck-Institut für Molekulare Genetik, Ihnestr. 63–73, 14195 Berlin, Germany.

*Correspondence e-mail: mwahl@zedat.fu-berlin.de

The spliceosomal RNA helicase Brr2 is required for the assembly of a catalytically active spliceosome on a messenger RNA precursor. Brr2 exhibits an unusual organization with tandem helicase units, each comprising dual RecA-like domains and a Sec63 homology unit, preceded by a more than 400-residue N-terminal helicase-associated region. Whereas recent crystal structures have provided insights into the molecular architecture and regulation of the Brr2 helicase region, little is known about the structural organization and function of its N-terminal part. Here, a near-atomic resolution crystal structure of a PWI-like domain that resides in the N-terminal region of *Chaetomium thermophilum* Brr2 is presented. CD spectroscopic studies suggested that this domain is conserved in the yeast and human Brr2 orthologues. Although canonical PWI domains act as low-specificity nucleic acid-binding domains, no significant affinity of the unusual PWI domain of Brr2 for a broad spectrum of DNAs and RNAs was detected in band-shift assays. Consistently, the *C. thermophilum* Brr2 PWI-like domain, in the conformation seen in the present crystal structure, lacks an expanded positively charged surface patch as observed in at least one canonical, nucleic acid-binding PWI domain. Instead, in a comprehensive yeast two-hybrid screen against human spliceosomal proteins, fragments of the N-terminal region of human Brr2 were found to interact with several other spliceosomal proteins. At least one of these interactions, with the Prp19 complex protein SPF27, depended on the presence of the PWI-like domain. The results suggest that the N-terminal region of Brr2 serves as a versatile protein–protein interaction platform in the spliceosome and that some interactions require or are reinforced by the PWI-like domain.

1. Introduction

Nuclear precursor-messenger RNA (pre-mRNA) splicing entails the removal of noncoding intervening sequences (introns) and the ligation of neighbouring coding regions (exons). Each splicing event proceeds *via* a two-step transesterification reaction that is carried out by a large and highly dynamic ribonucleoprotein machine, the spliceosome. A spliceosome consists of five small nuclear ribonucleoprotein particles (snRNPs; U1, U2, U4, U5 and U6 in the major spliceosome) and numerous non-snRNP proteins. For each round of splicing, a spliceosome is initially assembled stepwise on a pre-mRNA, is catalytically activated, facilitates the two-step splicing reaction and is subsequently disassembled in an ordered fashion (Wahl *et al.*, 2009). Transitions between stages of each such splicing process are accompanied by compositional and conformational remodelling of the protein–RNA



interaction networks of the spliceosome. The main driving forces for these remodelling events are eight conserved NTPases/RNA helicases (Staley & Guthrie, 1998). The Ski2-like RNA helicase Brr2, a component of the U5 snRNP, acts at the stage of spliceosome activation, during which it is thought to unwind the U4 and U6 snRNAs (the RNA components of the U4 and U6 snRNPs), which are extensively base-paired in the pre-catalytic spliceosome. Brr2 can unwind U4/U6 di-snRNAs *in vitro* (Laggerbauer *et al.*, 1998; Raghunathan & Guthrie, 1998), and mutations in Brr2 interfere with spliceosome activation *in vivo* (Noble & Guthrie, 1996; Kim & Rossi, 1999; Zhao *et al.*, 2009). Before joining the spliceosome, the U5 snRNP assembles with the U4/U6 di-snRNP to form the U4/U6·U5 tri-snRNP. Therefore, Brr2 already encounters its U4/U6 substrate outside of the spliceosome. Furthermore, it remains associated with the spliceosome throughout all steps following catalytic activation. Thus, Brr2 has to be tightly regulated to prevent premature unwinding of U4/U6 and possibly to prevent aberrant RNA rearrangements during the later stages of a splicing cycle. Brr2 is one of the largest spliceosomal proteins, containing an N-terminal region (NTR) of more than 400 amino acids followed by a helicase region with two structurally similar helicase units (cassettes). Each cassette comprises dual RecA-like domains and a Sec63 homology unit (Pena *et al.*, 2009; Zhang *et al.*, 2009; Santos *et al.*, 2012). *In vitro*, only the N-terminal cassette of Brr2 shows ATPase and U4/U6 unwinding activity, whereas the C-terminal cassette acts as an intramolecular regulator of the N-terminal unit (Santos *et al.*, 2012).

While the Brr2 helicase region has been studied both functionally and structurally (Pena *et al.*, 2009; Zhang *et al.*, 2009; Santos *et al.*, 2012), the role of the NTR is presently unknown. Helicases often encompass accessory domains or regions that modulate their function by providing additional nucleic acid-binding, protein-binding or localization signals (Fairman-Williams *et al.*, 2010). Recently, a noncanonical PWI domain was predicted within the NTR of Brr2 (Korneta *et al.*, 2012). Canonical PWI domains are found in nucleic acid-processing proteins and contain a core of four α -helices and a flanking region rich in basic amino acids (Blencowe & Ouzounis, 1999; Szymczyna *et al.*, 2003; Gong *et al.*, 2013). These elements build up an extended, positively charged surface area through which PWI domains bind RNA and DNA with low specificity (Szymczyna *et al.*, 2003; Gong *et al.*, 2013). The name of the domain derives from a conserved proline–tryptophan–isoleucine tripeptide located in the first α -helix of the core fold (Szymczyna *et al.*, 2003; Gong *et al.*, 2013). The PWI-like domains predicted in Brr2 orthologues contain a variant tripeptide sequence, which is also not conserved among Brr2 proteins from different organisms (*e.g.* FFL in yeast, FWL in *Chaetomium thermophilum* and YWL in human).

To explore the potential functions of the predicted non-canonical PWI domain in the NTR of Brr2 proteins, we determined a high-resolution crystal structure of the corresponding region of *C. thermophilum* Brr2. Although the core of the domain adopts a PWI-like fold and is flanked by an

ordered basic region on either side, the PWI-like domains from Brr2 proteins did not interact stably with nucleic acids. Instead, portions of the Brr2 NTR interacted with various other spliceosomal proteins in yeast two-hybrid (Y2H) assays and some of these interactions were apparently reinforced or depended on the presence of the PWI-like domain.

2. Materials and methods

2.1. Cloning, expression and purification

Codon-optimized DNA fragments encoding the predicted PWI-like domains of *Saccharomyces cerevisiae* Brr2 (yBrr2, residues 251–419; yBrr2^{251–419}), *C. thermophilum* Brr2 (cBrr2, residues 287–422; cBrr2^{287–422}) or *Homo sapiens* Brr2 (hBrr2, residues 249–388; hBrr2^{249–388}) and an elongated fragment of yBrr2 (residues 112–441; yBrr2^{112–441}) were cloned *via* NcoI and NotI restriction sites into a pETM-11 vector (EMBL Heidelberg) under the control of a T7 promoter for production of fusion proteins bearing a TEV-cleavable N-terminal His₆ tag. *Escherichia coli* Rosetta2 DE3 cells were transformed with the vectors, cultivated in autoinducing medium (Studier, 2005) at 37°C to an OD₆₀₀ of ~0.6 and subsequently incubated at 20°C. The cells were harvested at an OD₆₀₀ of ~10, resuspended in 40 mM HEPES pH 7.5, 500 mM KCl, 20 mM imidazole, 1 mM DTT supplemented with protease inhibitors (Roche) and lysed by sonication using a Sonopuls HD 3100 ultrasonic homogenizer (Bandelin). The target proteins were captured from the cleared lysate on a 5 ml HisTrap FF column (GE Healthcare) and eluted with a linear gradient from 20 to 500 mM imidazole. The His tags were cleaved with TEV protease during overnight dialysis at 4°C against 40 mM HEPES pH 7.5, 500 mM KCl, 1 mM DTT. The cleaved proteins were loaded onto a 5 ml HisTrap FF column to remove the His-tagged TEV protease, uncleaved proteins and cleaved His tags. The flowthrough fraction was diluted to a concentration of 80 mM KCl and loaded onto a 5 ml Heparin column (GE Healthcare) equilibrated with 40 mM HEPES pH 7.5, 80 mM KCl, 1 mM DTT. The proteins were recovered in the flowthrough, whereas most contaminants remained on the column. The proteins were further purified by gel filtration using a 26/60 Superdex 75 gel-filtration column (GE Healthcare) in 200 mM NaCl, 20 mM Tris–HCl pH 7.5, 1 mM DTT.

2.2. Crystallization

Crystallization of cBrr2^{287–422} was carried out at 8°C using the sitting-drop vapour-diffusion method. Native cBrr2^{287–422} crystals were obtained by mixing 200 nl protein solution (in gel-filtration buffer) at 56 mg ml⁻¹ with 200 nl reservoir solution (2 M ammonium sulfate, 0.1 M Tris–HCl pH 8.5) in a 96-well plate. Co²⁺-derivatized crystals were obtained by mixing 1 μ l protein solution (in gel-filtration buffer) at 56 mg ml⁻¹ with 1 μ l reservoir solution (1.9 M ammonium sulfate, 0.1 M Tris–HCl pH 8.0, 0.01 M CoCl₂) in a 24-well plate and were optimized by micro-seeding. The crystals were cryoprotected by transfer into mother liquor containing 25% ethylene glycol and were flash-cooled in liquid nitrogen.

Table 1

Crystallographic data.

Values in parentheses are for the highest resolution shell.

	Native	Co ²⁺ derivative
Data collection		
Wavelength (Å)	0.91841	1.59797
Space group	C2	C2
Unit-cell parameters		
<i>a</i> (Å)	123.8	117.9
<i>b</i> (Å)	45.8	46.2
<i>c</i> (Å)	27.4	28.4
β (°)	102.5	99.4
Resolution (Å)	50–1.14 (1.20–1.14)	50–2.00 (2.16–2.00)
Unique reflections	54192 (8263)	17991 (2158)
Completeness	97.9 (92.8)	94.7 (71.2)
$\langle I/\sigma(I) \rangle$	13.0 (2.9)	17.3 (2.5)
$R_{\text{meas}}^{\dagger}$ (%)	6.8 (45.8)	6.2 (50.7)
$CC_{1/2}$	99.8 (91.9)	99.8 (84.2)
Multiplicity	3.5 (3.4)	3.6 (2.7)
Phasing		
Resolution		29.7–2.5
No. of sites		1
Phasing power \ddagger (ano acentric)		1.13
R_{Cullis}^{\S}		0.788
FOM \P		
Before DM $\dagger\dagger$		0.35
After DM $\dagger\dagger$		0.97
Refinement		
Resolution (Å)	43–1.14 (1.18–1.14)	
No. of reflections	54190 (4872)	
Reflections in test set (%)	5 (5)	
Final model		
Non-H atoms	1410	
Protein residues	122	
Buffer molecules	8	
Waters	224	
$R_{\text{work}}^{\ddagger\ddagger}$	0.124 (0.180)	
$R_{\text{free}}^{\S\S}$	0.146 (0.188)	
Average <i>B</i> factor (Å ²)	8.4	
R.m.s.d. $\P\P$		
Bond lengths (Å)	0.015	
Bond angles (°)	1.50	
Ramachandran analysis		
Favoured (%)	99	
Outliers (%)	0	
PDB entry	4rvq	

$\dagger R_{\text{meas}} = \sum_{hkl} [N(hkl)/[N(hkl) - 1]]^{1/2} \sum_i |I_i(hkl) - \langle I(hkl) \rangle| / \sum_{hkl} \sum_i I_i(hkl)$, in which $I(hkl)$ is the mean intensity of symmetry-equivalent reflections and $N(hkl)$ is the redundancy. \ddagger Phasing power = $P = \sum_n |F_{\text{H,calc}}| / \sum_n |E| = |F_{\text{PH,obs}}| - |F_{\text{PH,calc}}|$ is the mean lack-of-closure error, where n is the number of observed scattering amplitudes for the derivative, $F_{\text{PH,obs}}$ and $F_{\text{PH,calc}}$ are the observed and calculated structure-factor amplitudes of the derivative, respectively, and $F_{\text{H,calc}}$ is the calculated structure-factor amplitude of the heavy-atom substructure. $\S R_{\text{Cullis}} = \sum_{hkl} ||F_{\text{PH}} \pm F_{\text{P}}| - F_{\text{H,calc}}| / \sum_{hkl} |F_{\text{PH}} \pm F_{\text{P}}|$, where F_{PH} and F_{P} are the observed structure-factor amplitudes of the derivative and native, respectively, $F_{\text{H,calc}}$ are the calculated structure-factor amplitudes of the heavy-atom substructure; + is used if the signs of F_{PH} and F_{P} are equal and – if they are opposite. \P Figure of merit = $m = |F(hkl)_{\text{best}}| / |F(hkl)|$, in which $F(hkl)_{\text{best}} = \sum_{\alpha} [P(\alpha)F_{\text{H,calc}}(\alpha)] / \sum_{\alpha} P(\alpha)$, where P is the phasing power and α is the phase angle. $\dagger\dagger$ Density modification. $\ddagger\ddagger R_{\text{work}} = \sum_{hkl} ||F_{\text{obs}}| - |F_{\text{calc}}|| / \sum_{hkl} |F_{\text{obs}}|$ (working set, no σ cutoff applied). $\S\S R_{\text{free}}$ was calculated using the same expression as for R_{work} but on 5% of the data excluded from refinement. $\P\P$ Root-mean-square deviation from target geometries.

2.3. Diffraction data collection, structure determination and refinement

Diffraction data were collected on beamline 14.2 of the BESSY II storage ring, Berlin, Germany at 100 K and were processed with *XDS* (Kabsch, 2010). Experimental phases of cBrr2^{287–422} were determined by the single anomalous dispersion (SAD) strategy using data collected at the *K* edge

Table 2

Sequences of the nucleic acids used in EMSA.

Nucleic acid	Sequence
ssDNA	5'–CTACTGGCTCCAAATCCAGATCGGCAGG–3'
dsDNA	5'–CTACTGGCTCCAAATCCAGATCGGCAGG–3' 5'–CCTGCCGATCTGGATTGGACCCAGTAG–3'
ssRNA	5'–GGCCAGCUCUAGAAAAACUAUACCCA–3'
dsRNA	5'–GGCCAGCUCUAGAAAAACUAUACCCA–3' 5'–CUAGAGCUGGCC–3'
yU5 snRNA construct	5'–GGGAGCUUACAGAUAUUGCGGAGGGAGG– AAUGGCGGAGGGAGGUCAACAUAAGAACUGU– GGCCUUUUUAUUGCCUUAUAGAACUUUAUACGA– ACAUGGUUCUUGCCUUUUAACGAAACCAUCCG– GGUGUGUCUCCAUAGAAACAGGUAAGCUCC– C–3'
yU4/U6 di-snRNA construct	5'–GGUCAUUUUGAAAAACAAUACAGAGAUGAUCAG– UACAGAGAUGAUCAGCAGUCCCGAAAGGGAA– AUACGCAUAUCGAAAGAUUGUUUUUUGCUGG– UUGAAAUUUUUAUUUAUAAACAGACC–3'

of cobalt (Table 1). Co²⁺ sites were located and initial phases were calculated and improved by density modification using the *SHARP* program suite (Bricogne *et al.*, 2003). Model building was performed using *phenix.autobuild* (Zwart *et al.*, 2008) and *Coot* (Emsley *et al.*, 2010). Refinement was performed with *phenix.refine* (Afonine *et al.*, 2012). The R_{free} factor was calculated using a randomly selected 5% of reflections which were excluded from refinement. The quality of the final model was checked using *MolProbity* (Chen *et al.*, 2010) and *PARVATI* (Merritt, 1999). Secondary-structure content was determined using *DSSP* (Joosten *et al.*, 2011). Data-collection, phasing and refinement statistics are listed in Table 1.

2.4. Electrophoretic mobility shift assays (EMSA)

Short RNA and DNA oligonucleotides were ordered from IBA GmbH and Eurofins, respectively. A tested dsDNA construct contained a central C–C mismatch. As Brr2 initially engages its U4/U6 di-snRNA substrate at a single-stranded region (Mozaffari-Jovin *et al.*, 2012), the tested dsRNA construct contained an additional single-stranded 3'-overhang. RNA encompassing portions of yU5 snRNA and yU4/U6 di-snRNAs were produced by *in vitro* transcription using T7 RNA polymerase. The yU5 snRNA construct consisted of the yU5 snRNA core domain lacking the Sm site. The yU4/U6 di-snRNA construct contained yU4 and yU6 snRNA portions fused by a tetraloop and exhibited a shortened yU4 5' stem-loop (Table 2). Nucleic acids were 5'-end labelled with (γ -³²P)ATP (PerkinElmer) using T4 polynucleotide kinase (NEB). For band-shift assays, nucleic acid concentrations were adjusted to 1 nM and protein was added to a final concentration of 300 μ M in binding buffer (40 mM Tris–HCl pH 7.5, 75 mM NaCl, 2.3 mM MgCl₂, 1 U μ l^{–1} RNasin, 0.1 mg ml^{–1} acetylated BSA). Samples were separated using 6% non-denaturing PAGE. Detection was performed with a Storm phosphorimager (Molecular Dynamics).

2.5. Yeast two-hybrid (Y2H) assays

Y2H screens were carried out as described previously (Hegele *et al.*, 2012). Open reading frames (ORFs) covering

various regions of the hBrr2 NTR (baits) were cloned into a Gateway entry vector (pDONR221) and then shuttled into a Y2H vector under the control of a truncated ADH1 promoter, guaranteeing low production levels for hBrr2 fragments fused

to a LexA DNA-binding domain. Bait vectors were screened against a prey matrix containing ORFs for Gal4 activation domain fusions of 237 human spliceosomal proteins, each represented by several clones (Hegele *et al.*, 2012). Bait strains

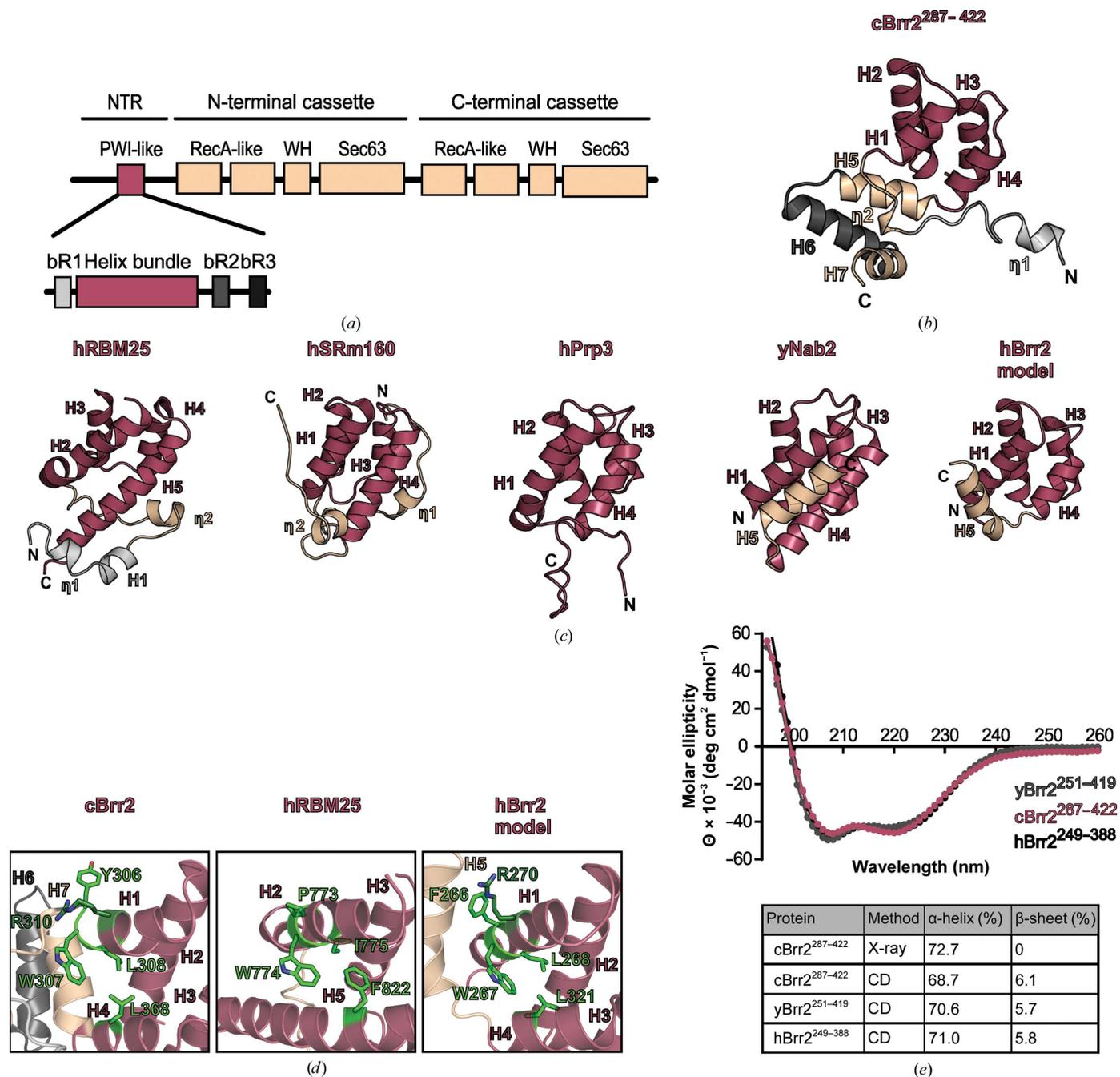


Figure 1 Structural comparison of cBrr2²⁸⁷⁻⁴²² with known PWI domains. (a) Schematic representation of Brr2, with an enlargement of the PWI-like region. PWI-like domain, raspberry; basic regions bR1, bR2 and bR3, light grey, dark grey and black, respectively; N-terminal and C-terminal cassette, wheat. (b) Ribbon plot of cBrr2²⁸⁷⁻⁴²². The four-helix core is shown in raspberry, the flanking basic regions are shown in light grey (bR1, N-terminal) and dark grey (bR2, C-terminal) and additional structural features are shown in wheat. (c) Ribbon plot of known PWI domains in hRBM25 (PDB entry 3v53), hSRm160 (PDB entry 1mp1), hPrp3 (PDB entry 1x4q), the PWI-related domain of yNab2 (PDB entry 2v75) and the model of the PWI-like domain of hBrr2 (Korneta *et al.*, 2012). The four-helix core is shown in raspberry, the flanking basic regions are shown in light grey and additional structural features are shown in wheat. (d) The PWI-like tripeptide sequence of cBrr2²⁸⁷⁻⁴²² in comparison to the PWI tripeptide sequence in hRBM25 and the PWI-like tripeptide sequence of the hBrr2 model. Interacting residues are shown in green; a conserved arginine (Arg310 in cBrr2 and Arg270 hBrr2) is shown in dark green and coloured by atom type (nitrogen, blue; oxygen, red). (e) CD spectrum of yBrr2²⁵¹⁻⁴¹⁹ (grey), cBrr2²⁸⁷⁻⁴²² (raspberry) and hBrr2²⁴⁹⁻³⁸⁸ (black) and comparison of the secondary-structure contents of the respective PWI-like domains deduced from the CD spectra or based on the crystal structure.

Table 3

Tested regions of hBrr2, yBrr2 and cBrr2.

Regions rich in basic residues: 254–259 (bR1), 343–349 (bR2) and 373–386 (bR3) of hBrr2, 288–294 (bR1), 380–394 (bR2) and 412–421 (bR3) of cBrr2 and 255–271 (bR1), 376–380 (bR2) and 390–399 (bR3) of yBrr2.

Organism	Borders	Solubility	Purification	Gene structure
<i>H. sapiens</i>	249–342	+	+	bR1–PWI
<i>H. sapiens</i>	249–351	+	+	bR1–PWI–bR2
<i>H. sapiens</i>	249–388	+	+	bR1–PWI–bR2–bR3
<i>H. sapiens</i>	253–351	+	–	bR1–PWI–bR2
<i>H. sapiens</i>	253–388	+	+	bR1–PWI–bR2–bR3
<i>H. sapiens</i>	255–375	+	+	bR1–PWI–bR2
<i>H. sapiens</i>	256–339	+	+	PWI
<i>H. sapiens</i>	256–341	+	+	PWI
<i>H. sapiens</i>	256–342	+	+	PWI
<i>H. sapiens</i>	256–351	+	–	PWI–bR2
<i>H. sapiens</i>	256–388	+	+	PWI–bR2–bR3
<i>H. sapiens</i>	258–339	–	–	PWI
<i>H. sapiens</i>	258–341	–	–	PWI
<i>H. sapiens</i>	258–342	–	–	PWI
<i>S. cerevisiae</i>	251–366	–	–	bR1–PWI
<i>S. cerevisiae</i>	251–382	+	+	bR1–PWI–bR2
<i>S. cerevisiae</i>	251–419	+	+	bR1–PWI–bR2–bR3
<i>S. cerevisiae</i>	254–382	+	–	bR1–PWI–bR2
<i>S. cerevisiae</i>	254–419	+	+	bR1–PWI–bR2–bR3
<i>S. cerevisiae</i>	271–382	+	–	PWI–bR2
<i>S. cerevisiae</i>	271–419	+	–	PWI–bR2–bR3
<i>S. cerevisiae</i>	273–366	–	–	PWI
<i>S. cerevisiae</i>	273–369	–	–	PWI
<i>S. cerevisiae</i>	282–366	–	–	PWI
<i>S. cerevisiae</i>	282–369	–	–	PWI
<i>C. thermophilum</i>	287–386	+	–	bR1–PWI–bR2
<i>C. thermophilum</i>	287–422	+	+	bR1–PWI–bR2–bR3
<i>C. thermophilum</i>	294–386	+	+	PWI–bR2
<i>C. thermophilum</i>	294–422	+	–	PWI–bR2–bR3

were mated with two independent biological replicas of the prey matrix. Cells were grown at 30°C for 7 d and bait–prey interactions were identified on selective agar plates (Δ Leu-Trp-Ura-His). Only bait–prey pairs that appeared in both replicas were considered in the evaluation. Bait–prey pairs were further evaluated by the percentage of total prey clones found of a protein and the average spot size per clone (rated 1–3, with 1 indicating weak interaction/small yeast spot size and 3 indicating strong interaction/large yeast spot size). When both criteria were ≤ 0.5 an interaction was considered to be weak (+), when both criteria were between 0.5 and 0.8 an interaction was considered to be intermediate (++) and when both criteria were ≥ 0.8 an interaction was considered to be strong (+++).

2.6. Circular-dichroism spectroscopy

Proteins were diluted to a concentration of 0.25 mg ml⁻¹ in CD buffer (10 mM potassium phosphate pH 7.5, 100 mM ammonium sulfate). Far-UV CD spectra (190–260 nm) were recorded at 20°C with a Jasco J-810 spectropolarimeter. CD data were analyzed using the CDNN software.

3. Results

3.1. The NTRs of Brr2 orthologues contain a noncanonical PWI domain

Structure predictions had suggested the presence of a noncanonical PWI domain in the NTR of Brr2 orthologues



Figure 2

Multiple sequence alignment of regions encompassing the PWI-like domains of cBrr2, hBrr2 and yBrr2. The residue numbers above the alignment refer to cBrr2. Background shading represents the conservation of residues, with a darker background corresponding to higher conservation. Green boxes indicate the noncanonical tripeptide and the interacting residue in H4. The green arrow indicates the conserved arginine close to the PWI-like tripeptide. Experimentally determined (cBrr2, residues 287–422) or predicted (hBrr2, residues 249–388; yBrr2, residues 254–419) secondary-structure (ss) elements are depicted as icons below the alignment. Elements of the four-helix core, raspberry; flanking basic regions, light grey (bR1, N-terminal), dark grey (bR2, C-terminal) or black (bR3, C-terminal); additional structural features, wheat.

containing a PWI-like central helical bundle preceded by a single basic region and followed by two basic regions (bR1 and bR2/bR3, respectively; Korneta *et al.*, 2012). To experimentally test for the presence of a PWI-like domain in the Brr2 NTR and to elucidate its atomic structure, we recombinantly produced a large number of hBrr2, yBrr2 and cBrr2 fragments covering the region of interest (Table 3 and Fig. 1*a*). The constructs either contained only the predicted PWI-like core or additionally included the predicted flanking basic regions in various combinations (Table 3 and Fig. 1*a*). The regions encompassing the predicted PWI-like domains share a sequence identity of 18–26% (Larkin *et al.*, 2007) between hBrr2, yBrr2 and cBrr2 (Fig. 2). The first basic region of yBrr2 is distributed over a wider range of residues (Fig. 2), resulting in larger constructs for the yeast variants. Although the majority of the constructs could be efficiently produced and purified, all tested yBrr2 and hBrr2 fragments failed to crystallize. Only a fragment of the Brr2 orthologue from the thermophilic fungus *C. thermophilum* encompassing cBrr2 residues 287–422 (cBrr2^{287–422}) and covering the PWI-like core (residues 304–370) as well as all three predicted flanking basic regions (bR1, residues 288–294; bR2, residues 380–394; bR3, residues 412–422) (Figs. 1*a* and 2) yielded crystals that diffracted to 1.14 Å resolution (Table 1).

As attempts to solve the structure by molecular replacement using known PWI domain structures as search models failed, a SAD experiment was conducted. Diffraction data using an X-ray wavelength near the Co *K* edge were collected from crystals obtained with CoCl₂ as an additive. Anomalous signals in this data set allowed us to locate a single bound Co²⁺ ion and to obtain experimental phases that enabled us to build an initial model that could be transferred by molecular replacement to the high-resolution native data set (Table 1). The native structure was refined to R_{work} and R_{free} values of 12.4 and 14.6%, respectively. Despite noticeable differences in unit-cell parameters (Table 1), we discerned no significant conformational differences between the native and the Co²⁺-derivatized structures. The final model encompasses residues 287–411 of cBrr2. Residues 337–342, constituting a flexible loop, and the C-terminal 11 residues could not be traced.

The portions of cBrr2^{287–422} that could be traced in the electron density form seven α -helices (H1–H7) and two 3_{10} -helices (η 1 and η 2; Fig. 1*b*). The central part of the structure comprises helices H1–H4, which form an antiparallel four-helix bundle with helices H1 and H3 running in the opposite direction to helices H2 and H4 (Fig. 1*b*). The regions N-terminal (η 1, η 2 and connecting loops) and C-terminal (H5–H7 and connecting loops) to this core contain basic regions bR1 and bR2, respectively (Figs. 1*a* and 1*b*). These latter elements run along the bottom of the core fold, with the directions of the helices largely perpendicular to the helices of the central bundle (Fig. 1*b*). The third predicted basic element, bR3, is contained in the C-terminal 11 residues that were disordered in our structure.

A comparison of the complete cBrr2^{287–422} structure with entries in the Protein Data Bank with the DALI server (Holm & Rosenström, 2010) uncovered no significant hits (highest

Z-score of 4.8 to a portion of a cyanobacterial histidyl-tRNA synthetase; PDB entry 3net; Midwest Center for Structural Genomics, unpublished work). However, manual comparisons confirmed that the central helical bundle of cBrr2^{287–422} exhibits the same topology as the corresponding element found in the canonical PWI domains of the proteins hRBM25 (PDB entry 3v53; Gong *et al.*, 2013), hPrp3 (PDB entry 1x4q; RIKEN Structural Genomics/Proteomics Initiative, unpublished work) and hSRm160 (PDB entry 1mp1; Szymczyna *et al.*, 2002) and in the PWI-related domain of yNab2 (PDB entry 2v75; Grant *et al.*, 2008), which features a variant VIV tripeptide sequence [root-mean-square deviations (r.m.s.d.s) ranging from 2.1 to 3.5 Å for 30–52 common C α atoms; Figs. 1*b* and 1*c*]. Helices in these core elements exhibit variable lengths (Figs. 1*b* and 1*c*), possibly explaining our failure to solve the cBrr2^{287–422} structure by molecular replacement. In the canonical PWI domains, the central four-helix bundle is stabilized by hydrophobic interactions between the tryptophan and isoleucine residues of the PWI tripeptide in the first helix with a phenylalanine in the fourth helix (Fig. 1*d*). A similar stabilizing interaction is seen in the structure of cBrr2^{287–422}, with Trp307 and Leu308 from helix H1 clustering with Leu368 from helix H4 (Fig. 1*d*). The proline of the name-giving PWI tripeptide sequence of canonical PWI domains is replaced by a tyrosine (Tyr306) in cBrr2^{287–422} (Fig. 2), which does not participate in this interaction network (Fig. 1*d*). The central helical bundle of our experimental cBrr2^{287–422} structure also closely resembles the bioinformatically predicted structure of a noncanonical PWI domain in hBrr2 (Korneta *et al.*, 2012; r.m.s.d. of 2.2 Å for 52 common C α atoms). The predicted model of hBrr2 contains, additionally to the four-helix core, a fifth helix (H5). This helix is found also in our cBrr2 structure, but it contains no basic region. H5 has a different orientation than in cBrr2, perhaps because H6 and H7 were not included in the modelling of the hBrr2 PWI-like domain (Fig. 1*c*). The stabilizing YWL sequence–Leu368 interaction in cBrr2^{287–422} is replaced by a similarly arranged FWL sequence–Leu321 interaction in the hBrr2 model (Figs. 1*d* and 2).

The canonical PWI domain of RBM25 is the only experimentally determined PWI or PWI-like structure that includes an N-terminal basic region, which is contained within a segment encompassing three short helices that encircle the C-terminal portion of the terminal helix of the core (Fig. 1*c*). The corresponding element of cBrr2^{287–422} comprises two 3_{10} -helices (η 1 and η 2), adopts an extended conformation and is positioned topologically differently with respect to the central helical bundle; it connects to helix H1 along the backside of the central bundle as opposed to running along the front as seen in RBM25 (Figs. 1*b* and 1*c*).

Although the sequence identities between regions containing the PWI-like domains in hBrr2, yBrr2 and cBrr2 only range between 18 and 26%, the secondary-structure predictions for hBrr2 and yBrr2 (PSIPRED; Jones, 1999) are in agreement with the secondary structure derived from the X-ray structure of cBrr2 (Fig. 2). These analyses support the notion that Brr2 orthologues contain a PWI-like domain in their NTRs. To directly test whether the structure of

cBrr2^{287–422} is conserved in other Brr2 orthologues, we conducted comparative CD spectroscopic studies using the corresponding fragments of yeast (yBrr2^{251–419}) and human (hBrr2^{249–388}) Brr2. The equilibrium CD spectra of cBrr2^{287–422}, yBrr2^{251–419} and hBrr2^{249–388} exhibited very similar shapes (Fig. 1e). The secondary-structure contents derived from the spectrum of cBrr2^{287–422} closely corresponded to the secondary-structure contents seen in the crystal structure of this fragment and to the secondary-structure contents derived from the CD spectra of yBrr2^{251–419} and hBrr2^{249–388} (Fig. 1e). Thus, the cBrr2^{287–422} fold is likely to be a conserved feature in the NTRs of all Brr2 orthologues.

3.2. The Brr2 PWI-like domain alone does not stably bind nucleic acids

Previously investigated PWI domains have been observed to bind single-stranded (ss) or double-stranded (ds) RNAs or DNAs largely independently of the nucleic acid sequence and with affinities in the low-micromolar range (Szymczyzna *et al.*, 2003; Gong *et al.*, 2013). The N-terminal basic region of the hRBM25 PWI domain together with amino acids located in the helical core forms an extended positive surface area which has been found to be crucial for nucleic acid binding. Although two basic regions preceding and following the helical core are ordered in the structure of cBrr2^{287–422} (Fig. 1b), the molecular surface of the fragment in the present conformation does not exhibit an extensive electropositive patch (Fig. 3a). To test whether Brr2 PWI-like domains also bind nucleic acids, we performed electrophoretic gel mobility shift assays (EMSAs) using ss and ds RNAs and DNAs and cBrr2^{287–422}, yBrr2^{251–419} and hBrr2^{249–388} (Fig. 3b; only data for yBrr2^{251–419} are shown). We also included RNA constructs resembling yU4/U6 di-snRNAs, which are unwound by yBrr2 during spliceosome catalytic activation (Laggerbauer *et al.*, 1998; Raghunathan & Guthrie, 1998), and yU5 snRNA (Laggerbauer *et al.*, 1998). Variants of yU5 snRNA that lacked an internal loop were found to be synthetically lethal with a R295I exchange of yBrr2 (Nancollis *et al.*, 2013) located in the yBrr2 PWI-like domain. This arginine is found in close proximity to the Brr2 PWI-like tripeptide sequence and is conserved between different organisms (Arg270 in hBrr2, Arg295 in yBrr2 and Arg310 in cBrr2; Figs. 1d and 2). Although we tested various buffer conditions and protein concentrations up to 300 μ M, no binding of any of the PWI-like domains to any of the nucleic acids was detected. Likewise, a yBrr2 fragment that included additional regions of the NTR (residues 112–441) did not show any nucleic acid affinity in these assays (Fig. 3c). Thus, unlike canonical PWI domains, the PWI-like domains of Brr2 orthologues on their own do not seem to act as nucleic acid-binding modules. These results do not rule out the possibility that the PWI-like domain can interact with nucleic acids in the context of full-length Brr2 or within the U5 snRNP, the U4/U6·U5 tri-snRNP or the spliceosome, where it may face high local concentrations of certain RNAs.

3.3. The Brr2 PWI-like domain contributes to Brr2–protein interactions

The noncanonical PWI domain of the yNab2 protein has been shown to act as a protein–protein interaction platform rather than a nucleic acid-binding device (Grant *et al.*, 2008). As we also failed to detect significant nucleic acid binding by Brr2 NTR fragments containing the PWI-like domain, we investigated whether the Brr2 NTR might instead be involved in interactions with other spliceosomal proteins. To this end, we screened various fragments of the NTR of hBrr2 against 237 human spliceosomal proteins in a Y2H approach (Hegele *et al.*, 2012). Whereas the isolated PWI-like domain (hBrr2^{249–388}) did not yield any hits, using larger fragments of the hBrr2 NTR as bait identified previously known Brr2 binding partners as well as potential additional interactors (Fig. 4). A fragment containing the N-terminal 249 amino acids (hBrr2^{1–249}; the region preceding the PWI-like domain) and a fragment containing the first 388 residues of hBrr2 (hBrr2^{1–388}; including the PWI-like domain) showed many Y2H interactions, often represented by large colonies. A fragment starting with the PWI-like domain and extending

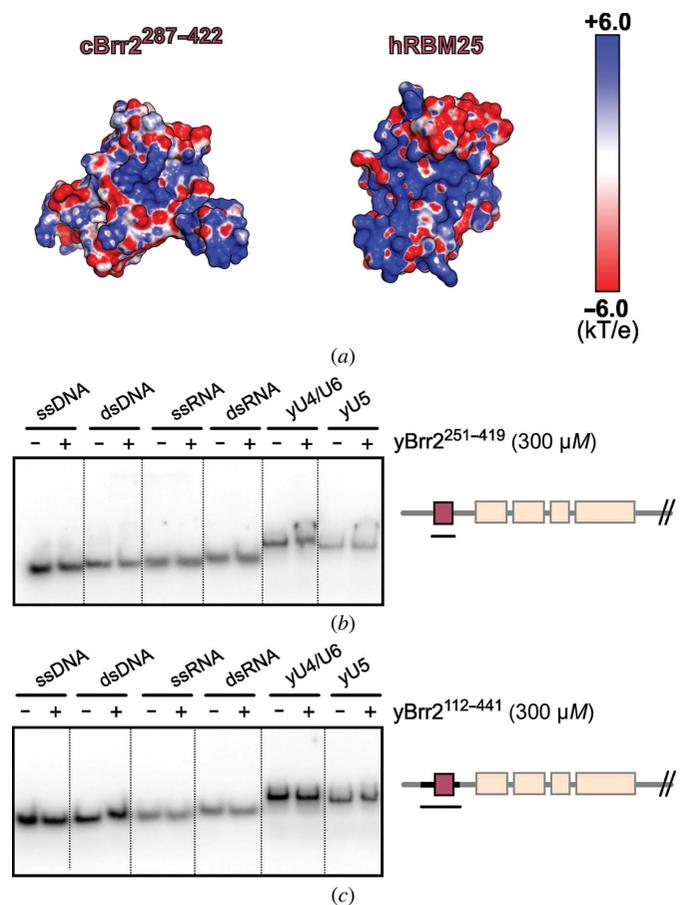


Figure 3 Nucleic acid binding of Brr2 NTR fragments. (a) Electrostatic surface potential of cBrr2^{287–422} (left) and the hRBM25 PWI domain (right). Blue indicates positive charge and red indicates negative charge. (b, c) Electrophoretic mobility shift assays. ss or ds DNA or RNA or yU4/U6 di-snRNA and yU5 snRNA constructs were loaded in the absence (–) or the presence (+) of 300 μ M yBrr2^{251–419} (b) or yBrr2^{112–441} (c).

to the first helicase cassette (hBrr2^{249–424}) yielded fewer Y2H-interaction partners and typically smaller colonies. Apart from self-interactions with Brr2 constructs, Y2H interactions were observed with heterogeneous nuclear (hn) RNP protein A1, the U2 snRNP proteins SF3b145 and SF3b130, the U5 snRNP protein Snu114, the U4/U6 di-snRNP proteins Prp31 and Prp3, the Prp19 complex subunit protein SPF27, the B complex-specific protein RED, the Prp19-related protein SKIP and the C complex protein matrin3 as well as the second-step protein Prp16. Interactions with the Prp19 complex protein SPF27 were only found for N-terminal constructs containing the PWI domain. SPF27 is the smallest subunit of the Prp19 complex, which is crucial for catalytic activation of the spliceosome (Makarova *et al.*, 2004). Whereas interactions between Brr2 and the U2 snRNP protein SF3b145, the U5 snRNP protein Snu114 and the second step splicing factor Prp16 had previously been observed in other analyses (van Nues & Beggs, 2001; Liu, 2006), interactions with hnRNPA1, the U2 snRNP protein SF3b130, the U4/U6

di-snRNP proteins Prp31 and Prp3, the Prp19 protein SPF27, the B complex-specific protein RED, the Prp19-related protein SKIP and the C complex protein matrin3 have not previously been reported. These results suggest that the NTR of Brr2 serves as a versatile protein–protein interaction platform in the spliceosome and that some interactions require or are reinforced by the PWI-like domain.

4. Discussion

Many nucleic acid helicases contain N- or C-terminal extensions in addition to their helicase cores. These regions are thought, and in several cases have been shown, to modulate helicase activity (Fairman-Williams *et al.*, 2010). For example, additional nucleic acid-binding domains might enhance substrate specificity (for example, an OB fold in yPrp43; Walbott *et al.*, 2010); other domains might act as protein–protein interaction platforms (for example, CARD domains in RIG-I; Yoneyama & Fujita, 2008) or direct the recruitment to

complexes (for example, winged-helix domains/regions in RecQ; Shereda *et al.*, 2009). The Brr2 RNA helicase performs its functions as part of a large RNA–protein machine, the spliceosome. Brr2 is a specific subunit of the U5 snRNP and is recruited to the spliceosome as part of the U4/U6-U5 tri-snRNP. Brr2 forms a stable complex with the large regulatory spliceosomal scaffolding protein Prp8 and the EF2-like G protein Snu114, which in turn interact with U5 snRNA (Achsel *et al.*, 1998; van Nues & Beggs, 2001; Liu, 2006) and thus may be the main anchors for Brr2 to the U5 snRNP, the U4/U6-U5 tri-snRNP or the spliceosome. In addition, Brr2 has been shown to interact with a number of other spliceosomal proteins (Achsel *et al.*, 1998; van Nues & Beggs, 2001; Liu, 2006), but the detailed interaction surfaces that support these various contacts are known only in a few cases (Mozaffari-Jovin *et al.*, 2013; Nguyen *et al.*, 2013).

Here, we have used biochemical, genetic and structural analyses to show that the long N-terminal helicase-associated regions of various Brr2 orthologues contain a PWI-like domain, as predicted previously by structural bioinformatics (Korneta *et al.*, 2012). The

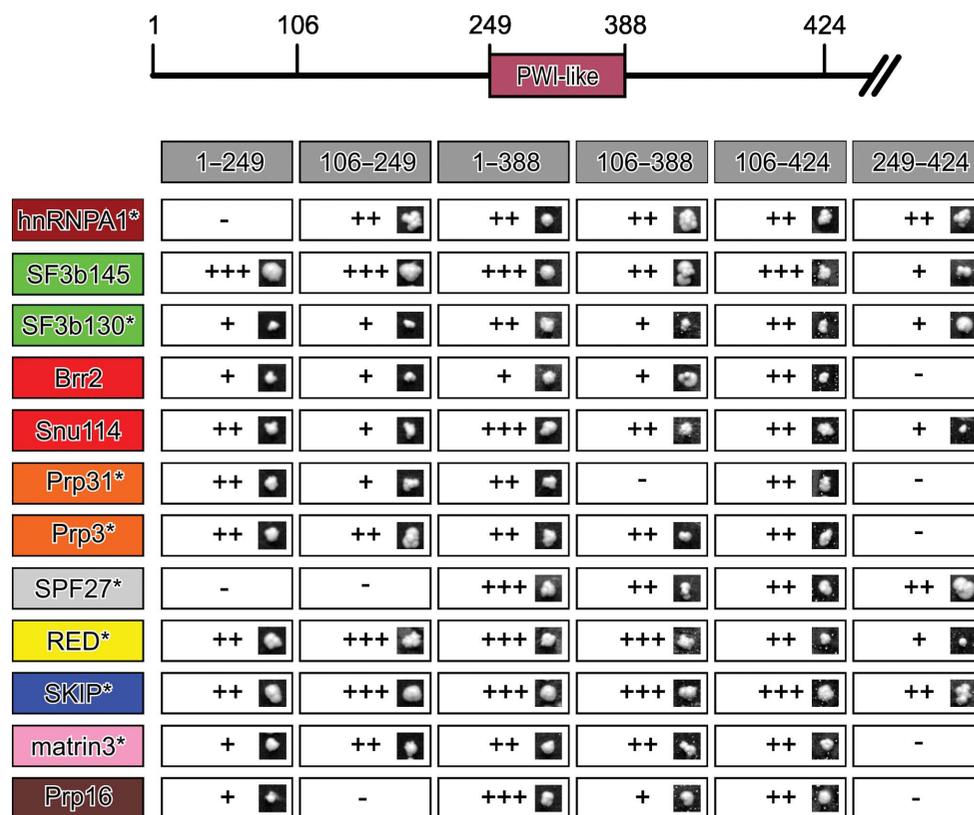


Figure 4

Representative results and quantification of a Y2H screen of N-terminal hBrr2 fragments against human spliceosomal proteins. Top: scheme of the hBrr2 N-terminal helicase-associated region with the position of the PWI-like domain and the borders chosen for hBrr2 constructs indicated. Bottom: residue ranges in hBrr2 constructs (grey boxes; horizontal) tabulated against interacting proteins (coloured boxes; vertical). Interaction partners of hBrr2 are grouped (indicated by similar colours) according to their function or localization in the spliceosome (Hegele *et al.*, 2012). Interacting proteins are listed in the order in which they assemble during splicing. Dark red, hnRNPs; green, U2 snRNP proteins; red, U5 snRNP proteins; orange, U4/U6 di-snRNP proteins; grey, Prp19 complex proteins; yellow, B complex-specific proteins; blue, Prp19-related proteins; light pink, C complex proteins, maroon, second-step factors. Asterisks indicate Brr2 interactions not previously reported. Interactions (–, absent; +, weak; ++, moderate; +++, strong) were judged based on the percentage of total clones of one protein which showed interactions with a hBrr2 construct and the size of the colonies representing a particular protein–protein interaction (for details, see §2). A representative colony is shown for each interaction pair.

domain is characterized by a variation in the name-giving PWI tripeptide (for example, to YWL in cBrr2). The amino acids, nevertheless, stabilize the structure in a similar manner to those in authentic PWI domains. Unlike canonical PWI domains, the Brr2 PWI-like element does not seem to bind nucleic acids, but rather constitutes part of a protein–protein interaction region. Our Y2H analyses involving hBrr2 fragments encompassing this domain revealed interactions with a significant number of other spliceosomal proteins. At least one of these Y2H interactions, with the SPF27 protein of the Prp19 complex, was dependent on the PWI-like domain. Several of these proteins are newly identified Brr2 interactors. Thus, the N-terminal helicase-associated region of Brr2 seems to be a versatile intra-spliceosomal protein–protein interaction region.

Interestingly, the only other known noncanonical PWI domain (with a variant VIV tripeptide sequence) whose structure and functions have been investigated in some detail is that of the yNab2 protein. This domain has also been found to act as a protein–protein interaction platform (Grant *et al.*, 2008). Notably, two other human spliceosomal helicases, Prp2 and Prp22, have also been predicted to contain variant PWI domains in their N-terminal helicase-associated regions (with PWI-like tripeptide sequences RWV and SLV, respectively; Korneta *et al.*, 2012), pointing to the possibility that all of these proteins have acquired a similar domain and neighbouring regions to facilitate their interactions with other proteins in the spliceosome. Unlike Brr2, these other spliceosomal helicases are thought to be recruited to the spliceosome precisely during the stages at which their activities are required. A protein–protein interaction domain that is positioned within a long, presumably intrinsically unstructured region (Korneta & Bujnicki, 2012) may constitute a suitable tool *via* which to achieve fast binding kinetics (Dyson & Wright, 2005) for the incorporation of these proteins at the right times.

Although our results point to the role of the Brr2 PWI-like domain as a protein–protein interaction module, the precise functional consequences of the protein–protein interactions involving the noncanonical PWI domain of Brr2 as well as their molecular basis remain to be determined. An obvious possibility is that the PWI-like domain helps to efficiently recruit Brr2 to the U5 snRNP or facilitates assembly of the tri-snRNP or incorporation of the tri-snRNP into the spliceosome. However, the reverse scenario is also imaginable: Brr2 recruits other proteins or complexes to the spliceosome *via* its PWI-like domain. An obvious candidate here is the Prp19 complex, as the Brr2 PWI-like domain seems to be required for interaction with SPF27, a core Prp19 complex subunit (Grote *et al.*, 2010). The Prp19 complex is recruited to the spliceosome immediately after the tri-snRNP (Hoskins *et al.*, 2011) and is required for spliceosome catalytic activation (Makarova *et al.*, 2004). In humans, Prp19 complex proteins are incorporated into the U5 snRNP upon their recruitment (Makarov *et al.*, 2002). Thus, the interaction of SPF27 and the N-terminus of Brr2 (including the PWI-like domain) might play a role in recruiting the Prp19 complex to the spliceosome

and elicit the formation of a catalytically activated spliceosome.

On the other hand, Brr2 helicase activity is regulated on various levels. Its helicase activity, which is contained solely in its N-terminal helicase cassette, is stimulated by the helicase-inactive C-terminal cassette, and this stimulation may be modulated by factors binding the C-terminal unit (Santos *et al.*, 2012). Brr2 helicase activity seems to be counteracted in the tri-snRNP and directly after incorporation into the spliceosome by a RNaseH-like domain in the Prp8 protein, which sequesters the initial Brr2 binding site on its U4/U6 di-snRNA substrate (Mozaffari-Jovin *et al.*, 2012). Likewise, the C-terminal tail of a Prp8 Jab1/MPN-like domain can inhibit Brr2 helicase activity by blocking the RNA-binding tunnel of the enzyme (Mozaffari-Jovin *et al.*, 2013). However, upon release of this inhibitory tail from the RNA-binding tunnel of Brr2, the Prp8 Jab1/MPN-like domain is converted into a potent activator of Brr2 helicase activity (Mozaffari-Jovin *et al.*, 2013, 2014), presumably to allow the enzyme to unwind U4/U6 di-snRNAs efficiently when catalytic activation of the spliceosome is required. It is conceivable that additional regulatory mechanisms exist to modulate Brr2 helicase activity and that these mechanisms may capitalize on the protein–protein interactions that we have uncovered for the N-terminal helicase-associated region, including the PWI-like domain, of the enzyme. Likewise, other spliceosomal helicases are regulated by *trans*-acting factors. For example, the helicase activities of the yeast Prp2 and Prp43 proteins are stimulated by the G-patch proteins SPP2 (Silverman *et al.*, 2004) and Ntr1 (Tsai *et al.*, 2007), respectively.

Multiple regulatory mechanisms acting on Brr2 may simply exist as mutual backup for Brr2 inhibition or activation at the correct times. However, it has also been shown that regulation of the kinetics with which some of the spliceosomal helicases act on their assembly intermediates can influence alternative splicing decisions (Semlow & Staley, 2012) or be used to implement proofreading mechanisms for splice-site choice (for example, Prp5, Prp16, Prp22, Prp28 and Prp43; Burgess & Guthrie, 1993; Mayas *et al.*, 2006, 2010; Xu & Query, 2007; Yang *et al.*, 2013). Thus, it is conceivable that multiple regulatory mechanisms that influence Brr2 kinetics are also used for splicing regulation.

References

- Achsel, T., Ahrens, K., Brahm, H., Teigelkamp, S. & Lührmann, R. (1998). *Mol. Cell Biol.* **18**, 6756–6766.
- Afonine, P. V., Grosse-Kunstleve, R. W., Echols, N., Headd, J. J., Moriarty, N. W., Mustyakimov, M., Terwilliger, T. C., Urzhumtsev, A., Zwart, P. H. & Adams, P. D. (2012). *Acta Cryst.* **D68**, 352–367.
- Blencowe, B. J. & Ouzounis, C. A. (1999). *Trends Biochem. Sci.* **24**, 179–180.
- Bricogne, G., Vonrhein, C., Flensburg, C., Schiltz, M. & Paciorek, W. (2003). *Acta Cryst.* **D59**, 2023–2030.
- Burgess, S. M. & Guthrie, C. (1993). *Cell*, **73**, 1377–1391.
- Chen, V. B., Arendall, W. B., Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S. & Richardson, D. C. (2010). *Acta Cryst.* **D66**, 12–21.
- Dyson, H. J. & Wright, P. E. (2005). *Nature Rev. Mol. Cell Biol.* **6**, 197–208.

- Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). *Acta Cryst.* **D66**, 486–501.
- Fairman-Williams, M. E., Guenther, U.-P. & Jankowsky, E. (2010). *Curr. Opin. Struct. Biol.* **20**, 313–324.
- Gong, D., Yang, F., Li, F., Qian, D., Wu, M., Shao, Z., Wu, M., Wu, J. & Shi, Y. (2013). *Biochem. J.* **450**, 85–94.
- Grant, R. P., Marshall, N. J., Yang, J.-C., Fasken, M. B., Kelly, S. M., Harreman, M. T., Neuhaus, D., Corbett, A. H. & Stewart, M. (2008). *J. Mol. Biol.* **376**, 1048–1059.
- Grote, M., Wolf, E., Will, C. L., Lemm, I., Agafonov, D. E., Schomburg, A., Fischle, W., Urlaub, H. & Lührmann, R. (2010). *Mol. Cell. Biol.* **30**, 2105–2119.
- Hegele, A., Kamburov, A., Grossmann, A., Sourlis, C., Wowro, S., Weimann, M., Will, C. L., Pena, V., Lührmann, R. & Stelzl, U. (2012). *Mol. Cell*, **45**, 567–580.
- Holm, L. & Rosenström, P. (2010). *Nucleic Acids Res.* **38**, W545–W549.
- Hoskins, A. A., Friedman, L. J., Gallagher, S. S., Crawford, D. J., Anderson, E. G., Wombacher, R., Ramirez, N., Cornish, V. W., Gelles, J. & Moore, M. J. (2011). *Science*, **331**, 1289–1295.
- Jones, D. T. (1999). *J. Mol. Biol.* **292**, 195–202.
- Joosten, R. P., te Beek, T. A. H., Krieger, E., Hekkelman, M. L., Hoof, R. W. W., Schneider, R., Sander, C. & Vriend, G. (2011). *Nucleic Acids Res.* **39**, D411–D419.
- Kabsch, W. (2010). *Acta Cryst.* **D66**, 125–132.
- Kim, D.-H. & Rossi, J. J. (1999). *RNA*, **5**, 959–971.
- Korneta, I. & Bujnicki, J. M. (2012). *PLoS Comput. Biol.* **8**, e1002641.
- Korneta, I., Magnus, M. & Bujnicki, J. M. (2012). *Nucleic Acids Res.* **40**, 7046–7065.
- Laggerbauer, B., Achsel, T. & Lührmann, R. (1998). *Proc. Natl Acad. Sci. USA*, **95**, 4188–4192.
- Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., Valentin, F., Wallace, I. M., Wilm, A., Lopez, R., Thompson, J. D., Gibson, T. J. & Higgins, D. G. (2007). *Bioinformatics*, **23**, 2947–2948.
- Liu, S. (2006). *RNA*, **12**, 1418–1430.
- Makarova, O. V., Makarov, E. M., Urlaub, H., Will, C. L., Gentzel, M., Wilm, M. & Lührmann, R. (2004). *EMBO J.* **23**, 2381–2391.
- Makarova, E. M., Makarova, O. V., Urlaub, H., Gentzel, M., Will, C. L., Wilm, M. & Lührmann, R. (2002). *Science*, **298**, 2205–2208.
- Mayas, R. M., Maita, H., Semlow, D. R. & Staley, J. P. (2010). *Proc. Natl Acad. Sci. USA*, **107**, 10020–10025.
- Mayas, R. M., Maita, H. & Staley, J. P. (2006). *Nature Struct. Mol. Biol.* **13**, 482–490.
- Merritt, E. A. (1999). *Acta Cryst.* **D55**, 1109–1117.
- Mozaffari-Jovin, S., Santos, K. F., Hsiao, H.-H., Will, C. L., Urlaub, H., Wahl, M. C. & Lührmann, R. (2012). *Genes Dev.* **26**, 2422–2434.
- Mozaffari-Jovin, S., Wandersleben, T., Santos, K. F., Will, C. L., Lührmann, R. & Wahl, M. C. (2013). *Science*, **341**, 80–84.
- Mozaffari-Jovin, S., Wandersleben, T., Santos, K. F., Will, C. L., Lührmann, R. & Wahl, M. C. (2014). *RNA Biol.* **11**, 298–312.
- Nancollis, V., Ruckshanthi, J. P. D., Frazer, L. N. & O’Keefe, R. T. (2013). *J. Cell. Biochem.* **114**, 2770–2784.
- Nguyen, T. H. D., Li, J., Galej, W. P., Oshikane, H., Newman, A. J. & Nagai, K. (2013). *Structure*, **21**, 910–919.
- Noble, S. M. & Guthrie, C. (1996). *Genetics*, **143**, 67–80.
- Nues, R. W. van & Beggs, J. D. (2001). *Genetics*, **157**, 1451–1467.
- Pena, V., Jovin, S. M., Fabrizio, P., Orłowski, J., Bujnicki, J. M., Lührmann, R. & Wahl, M. C. (2009). *Mol. Cell*, **35**, 454–466.
- Ragunathan, P. L. & Guthrie, C. (1998). *Curr. Biol.* **8**, 847–855.
- Santos, K. F., Jovin, S. M., Weber, G., Pena, V., Lührmann, R. & Wahl, M. C. (2012). *Proc. Natl Acad. Sci. USA*, **109**, 17418–17423.
- Semlow, D. R. & Staley, J. P. (2012). *Trends Biochem. Sci.* **37**, 263–273.
- Shereda, R. D., Reiter, N. J., Butcher, S. E. & Keck, J. L. (2009). *J. Mol. Biol.* **386**, 612–625.
- Silverman, E. J., Maeda, A., Wei, J., Smith, P., Beggs, J. D. & Lin, R.-J. (2004). *Mol. Cell. Biol.* **24**, 10101–10110.
- Staley, J. P. & Guthrie, C. (1998). *Cell*, **92**, 315–326.
- Studier, F. W. (2005). *Protein Expr. Purif.* **41**, 207–234.
- Szymczyna, B. R., Bowman, J., McCracken, S., Pineda-Lucena, A., Lu, Y., Cox, B., Lambermon, M., Graveley, B. R., Arrowsmith, C. H. & Blencowe, B. J. (2003). *Genes Dev.* **17**, 461–475.
- Szymczyna, B. R., Pineda-Lucena, A., Mills, J. L., Szyperski, T. & Arrowsmith, C. H. (2002). *J. Biomol. NMR*, **22**, 299–300.
- Tsai, R.-T., Tseng, C.-K., Lee, P.-J., Chen, H.-C., Fu, R.-H., Chang, K., Yeh, F.-L. & Cheng, S.-C. (2007). *Mol. Cell. Biol.* **27**, 8027–8037.
- Wahl, M. C., Will, C. L. & Lührmann, R. (2009). *Cell*, **136**, 701–718.
- Walbott, H., Mouffok, S., Capeyrou, R., Lebaron, S., Humbert, O., van Tilbeurgh, H., Henry, Y. & Leulliot, N. (2010). *EMBO J.* **29**, 2194–2204.
- Xu, Y.-Z. & Query, C. C. (2007). *Mol. Cell*, **28**, 838–849.
- Yang, F., Wang, X.-Y., Zhang, Z.-M., Pu, J., Fan, Y.-J., Zhou, J., Query, C. C. & Xu, Y.-Z. (2013). *Nucleic Acids Res.* **41**, 4660–4670.
- Yoneyama, M. & Fujita, T. (2008). *Immunity*, **29**, 178–181.
- Zhang, L., Xu, T., Maeder, C., Bud, L.-O., Shanks, J., Nix, J., Guthrie, C., Pleiss, J. A. & Zhao, R. (2009). *Nature Struct. Mol. Biol.* **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. & Larsson, C. (2009). *Am. J. Hum. Genet.* **85**, 617–627.
- Zwart, P. H., Afonine, P. V., Grosse-Kunstleve, R. W., Hung, L.-W., Ioerger, T. R., McCoy, A. J., McKee, E., Moriarty, N. W., Read, R. J., Sacchettini, J. C., Sauter, N. K., Storoni, L. C., Terwilliger, T. C. & Adams, P. D. (2008). *Methods Mol. Biol.* **426**, 419–435.