wwPDB X-ray Structure Validation Summary Report (i)

Jun 16, 2024 – 12:01 PM EDT

PDB ID	:	5DCA
Title	:	Crystal structure of yeast full length Brr2 in complex with Prp8 Jab1 domain
Authors	:	Absmeier, E.; Wollenhaupt, J.; Santos, K.F.; Wahl, M.C.
Deposited on	:	2015-08-23
Resolution	:	2.80 Å(reported)

This is a wwPDB X-ray Structure Validation Summary Report for a publicly released PDB entry.

We welcome your comments at *validation@mail.wwpdb.org*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

The following versions of software and data (see references (1)) were used in the production of this report:

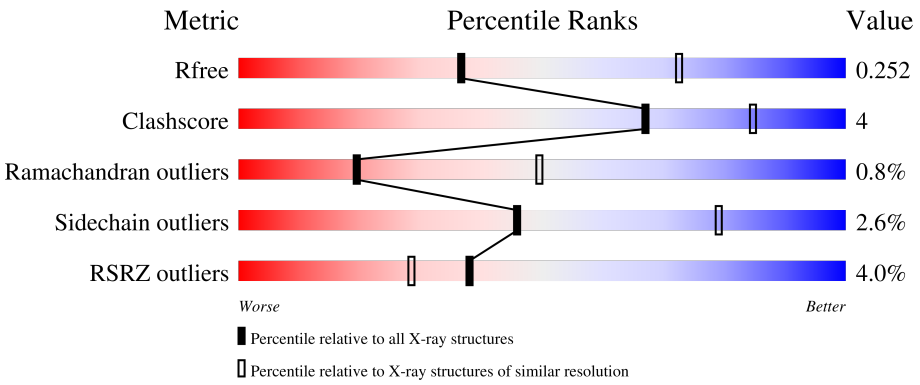
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.37.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.37.1

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.80 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	3140 (2.80-2.80)
Clashscore	141614	3569 (2.80-2.80)
Ramachandran outliers	138981	3498 (2.80-2.80)
Sidechain outliers	138945	3500 (2.80-2.80)
RSRZ outliers	127900	3078 (2.80-2.80)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments of the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions <=5% The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	1948	4% 88% 11% •
2	J	251	% 90% 10%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 17842 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called Pre-mRNA-splicing helicase BRR2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	1948	15639	10006	2602	2973	58	0	0	0

There are 103 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	?	-	THR	deletion	UNP P32639
A	?	-	ASP	deletion	UNP P32639
A	?	-	TYR	deletion	UNP P32639
A	?	-	GLU	deletion	UNP P32639
A	?	-	THR	deletion	UNP P32639
A	?	-	HIS	deletion	UNP P32639
A	?	-	PRO	deletion	UNP P32639
A	?	-	ASP	deletion	UNP P32639
A	?	-	ASN	deletion	UNP P32639
A	?	-	SER	deletion	UNP P32639
A	?	-	ASN	deletion	UNP P32639
A	?	-	LYS	deletion	UNP P32639
A	?	-	GLN	deletion	UNP P32639
A	?	-	ALA	deletion	UNP P32639
A	?	-	VAL	deletion	UNP P32639
A	?	-	ALA	deletion	UNP P32639
A	?	-	ILE	deletion	UNP P32639
A	?	-	LEU	deletion	UNP P32639
A	?	-	ALA	deletion	UNP P32639
A	?	-	ASP	deletion	UNP P32639
A	?	-	ASP	deletion	UNP P32639
A	?	-	GLU	deletion	UNP P32639
A	?	-	LYS	deletion	UNP P32639
A	?	-	SER	deletion	UNP P32639
A	?	-	ASP	deletion	UNP P32639
A	?	-	GLU	deletion	UNP P32639
A	?	-	GLU	deletion	UNP P32639

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	?	-	GLU	deletion	UNP P32639
A	?	-	VAL	deletion	UNP P32639
A	?	-	THR	deletion	UNP P32639
A	?	-	GLU	deletion	UNP P32639
A	?	-	MET	deletion	UNP P32639
A	?	-	SER	deletion	UNP P32639
A	?	-	ASN	deletion	UNP P32639
A	?	-	ASN	deletion	UNP P32639
A	?	-	ALA	deletion	UNP P32639
A	?	-	ASN	deletion	UNP P32639
A	?	-	VAL	deletion	UNP P32639
A	?	-	LEU	deletion	UNP P32639
A	?	-	GLY	deletion	UNP P32639
A	?	-	GLY	deletion	UNP P32639
A	?	-	GLU	deletion	UNP P32639
A	?	-	ILE	deletion	UNP P32639
A	?	-	ASN	deletion	UNP P32639
A	?	-	ASP	deletion	UNP P32639
A	?	-	ASN	deletion	UNP P32639
A	?	-	GLU	deletion	UNP P32639
A	?	-	ASP	deletion	UNP P32639
A	?	-	ASP	deletion	UNP P32639
A	?	-	ASP	deletion	UNP P32639
A	?	-	GLU	deletion	UNP P32639
A	?	-	GLU	deletion	UNP P32639
A	?	-	TYR	deletion	UNP P32639
A	?	-	ASP	deletion	UNP P32639
A	?	-	TYR	deletion	UNP P32639
A	?	-	ASN	deletion	UNP P32639
A	?	-	ASP	deletion	UNP P32639
A	?	-	VAL	deletion	UNP P32639
A	?	-	GLU	deletion	UNP P32639
A	?	-	VAL	deletion	UNP P32639
A	?	-	ASN	deletion	UNP P32639
A	?	-	SER	deletion	UNP P32639
A	?	-	LYS	deletion	UNP P32639
A	?	-	LYS	deletion	UNP P32639
A	?	-	LYS	deletion	UNP P32639
A	?	-	THR	deletion	UNP P32639
A	?	-	THR	deletion	UNP P32639
A	?	-	HIS	deletion	UNP P32639
A	?	-	SER	deletion	UNP P32639

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	?	-	LYS	deletion	UNP P32639
A	?	-	ARG	deletion	UNP P32639
A	?	-	GLU	deletion	UNP P32639
A	?	-	LEU	deletion	UNP P32639
A	?	-	ASP	deletion	UNP P32639
A	?	-	SER	deletion	UNP P32639
A	?	-	GLY	deletion	UNP P32639
A	?	-	ASP	deletion	UNP P32639
A	?	-	ASP	deletion	UNP P32639
A	?	-	GLN	deletion	UNP P32639
A	?	-	PRO	deletion	UNP P32639
A	?	-	GLN	deletion	UNP P32639
A	?	-	SER	deletion	UNP P32639
A	?	-	SER	deletion	UNP P32639
A	?	-	GLU	deletion	UNP P32639
A	?	-	ALA	deletion	UNP P32639
A	?	-	LYS	deletion	UNP P32639
A	?	-	ARG	deletion	UNP P32639
A	?	-	THR	deletion	UNP P32639
A	?	-	LYS	deletion	UNP P32639
A	?	-	PHE	deletion	UNP P32639
A	?	-	SER	deletion	UNP P32639
A	?	-	VAL	deletion	UNP P32639
A	?	-	THR	deletion	UNP P32639
A	?	-	ALA	deletion	UNP P32639
A	?	-	GLU	deletion	UNP P32639
A	?	-	VAL	deletion	UNP P32639
A	?	-	ASN	deletion	UNP P32639
A	?	-	GLY	deletion	UNP P32639
A	?	-	GLY	deletion	UNP P32639
A	?	-	ASP	deletion	UNP P32639
A	?	-	ASP	deletion	UNP P32639
A	?	-	GLU	deletion	UNP P32639
A	?	-	ALA	deletion	UNP P32639

- Molecule 2 is a protein called Pre-mRNA-splicing factor 8.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	J	251	2012	1298	325	383	6	0	0	0

- Molecule 3 is water.

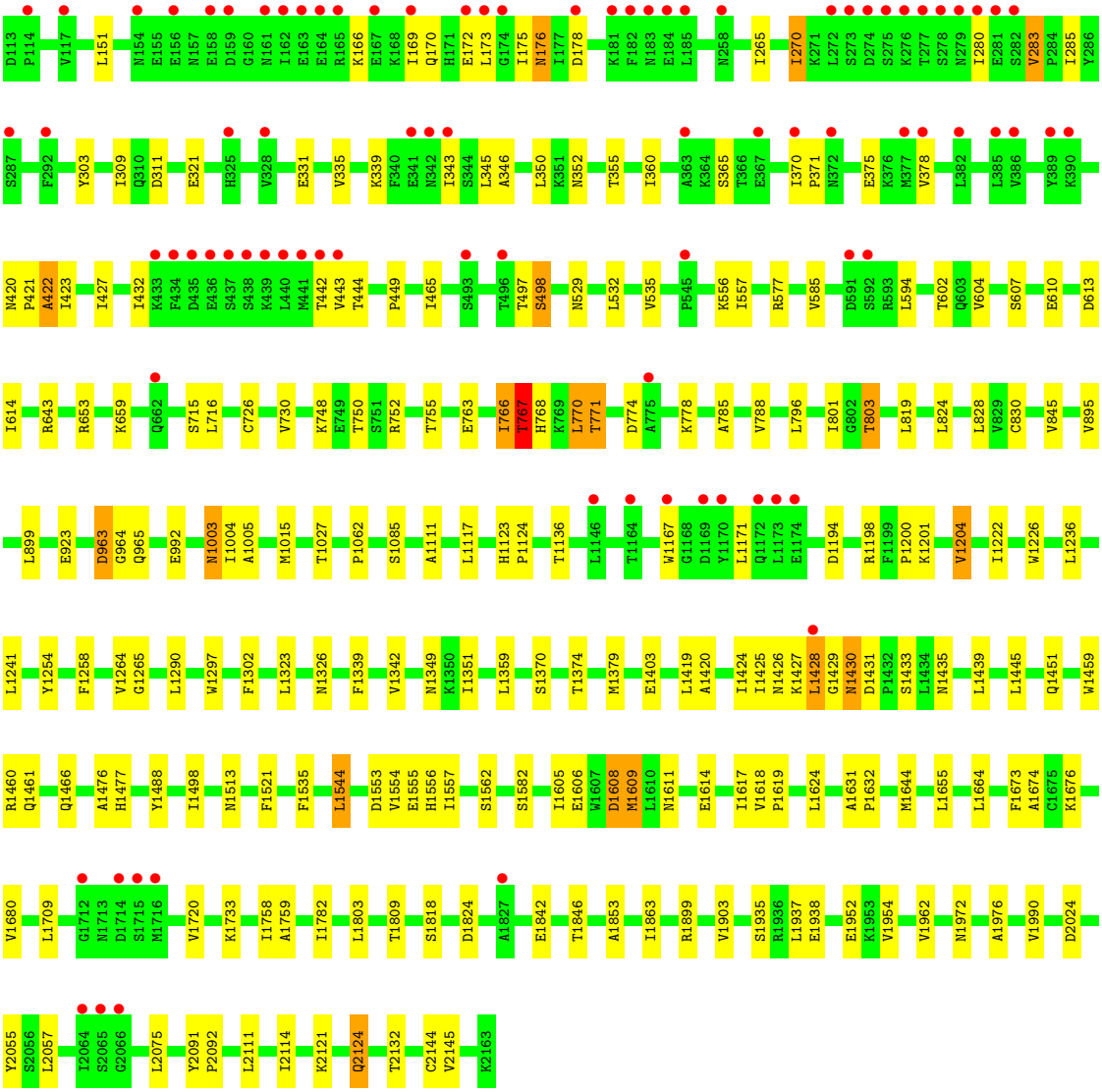
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
3	A	169	169	169	0	0
			Total	O		
3	J	22	22	22	0	0

3 Residue-property plots (i)

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density (RSRZ > 2). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

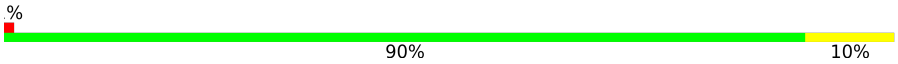
• Molecule 1: Pre-mRNA-splicing helicase BRR2

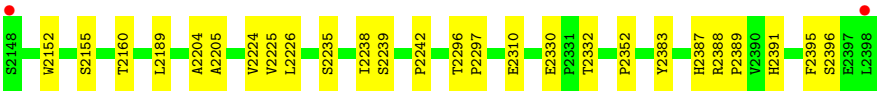
Chain A: 4% 88% 11%



• Molecule 2: Pre-mRNA-splicing factor 8

Chain J:


%
90%
10%


S2148
W2152
S2155
T2160
L2189
A2204
A2205
V2224
V2225
L2226
S2235
I2238
S2239
P2242
T2296
P2297
E2310
E2330
P2331
T2332
P2352
Y2383
H2387
R2388
P2389
V2390
H2391
F2395
S2396
E2397
L2398

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	107.85Å 178.85Å 181.11Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	50.00 – 2.80 178.85 – 2.73	Depositor EDS
% Data completeness (in resolution range)	98.5 (50.00-2.80) 95.5 (178.85-2.73)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.10	Depositor
$< I/\sigma(I) >$ [1]	1.13 (at 2.73Å)	Xtriage
Refinement program	REFMAC 5.8.0073	Depositor
R, R_{free}	0.202 , 0.255 0.205 , 0.252	Depositor DCC
R_{free} test set	4481 reflections (4.99%)	wwPDB-VP
Wilson B-factor (Å^2)	85.3	Xtriage
Anisotropy	0.245	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.29 , 53.5	EDS
L-test for twinning[2]	$< \|L\| > = 0.47$, $< L^2 > = 0.30$	Xtriage
Estimated twinning fraction	0.020 for -h,l,k	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	17842	wwPDB-VP
Average B, all atoms (Å^2)	102.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.21% of the height of the origin peak. No significant pseudotranslation is detected.*

[1]Intensities estimated from amplitudes.

[2]Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality (i)

5.1 Standard geometry (i)

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 5$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	$\#\|Z\| >5$	RMSZ	$\#\|Z\| >5$
1	A	0.32	0/15960	0.52	0/21620
2	J	0.32	0/2060	0.51	0/2792
All	All	0.32	0/18020	0.52	0/24412

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts (i)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry-related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	15639	0	15665	123	0
2	J	2012	0	1972	17	0
3	A	169	0	0	3	0
3	J	22	0	0	0	0
All	All	17842	0	17637	136	0

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 4.

The worst 5 of 136 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:J:2387:HIS:C	2:J:2388:ARG:N	2.05	1.08

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:1439:LEU:HA	1:A:1445:LEU:HD11	1.64	0.80
1:A:1555:GLU:C	1:A:1556:HIS:N	2.41	0.74
1:A:767:THR:O	1:A:771:THR:N	2.25	0.70
1:A:2121:LYS:O	1:A:2124:GLN:NE2	2.24	0.69

There are no symmetry-related clashes.

5.3 Torsion angles (i)

5.3.1 Protein backbone (i)

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
1	A	1937/1948 (99%)	1823 (94%)	97 (5%)	17 (1%)	17 46
2	J	247/251 (98%)	231 (94%)	16 (6%)	0	100 100
All	All	2184/2199 (99%)	2054 (94%)	113 (5%)	17 (1%)	19 49

5 of 17 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	270	ILE
1	A	498	SER
1	A	767	THR
1	A	172	GLU
1	A	321	GLU

5.3.2 Protein sidechains (i)

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	1762/1762 (100%)	1714 (97%)	48 (3%)	44	78
2	J	223/223 (100%)	220 (99%)	3 (1%)	69	91
All	All	1985/1985 (100%)	1934 (97%)	51 (3%)	46	79

5 of 51 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	1460	ARG
1	A	1608	ASP
2	J	2235	SER
1	A	1461	GLN
1	A	1544	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 16 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	2161	ASN
1	A	1864	GLN
1	A	984	GLN
1	A	1426	ASN
1	A	739	GLN

5.3.3 RNA (i)

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains (i)

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

There are no ligands in this entry.

5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
1	A	5
2	J	1

The worst 5 of 6 chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	A	192:ILE	C	258:ASN	N	48.54
1	A	393:GLU	C	420:ASN	N	23.97
1	A	1828:GLU	C	1841:THR	N	9.71
1	A	1556:HIS	C	1557:ILE	N	2.56
1	A	1555:GLU	C	1556:HIS	N	2.41

6 Fit of model and data (i)

6.1 Protein, DNA and RNA chains (i)

In the following table, the column labelled '#RSRZ> 2' contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95th percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled 'Q< 0.9' lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2			OWAB(Å^2)	Q<0.9
1	A	1948/1948 (100%)	0.10	87 (4%)	33	23	57, 90, 184, 237	0
2	J	251/251 (100%)	-0.23	2 (0%)	86	81	59, 88, 151, 194	0
All	All	2199/2199 (100%)	0.06	89 (4%)	38	28	57, 89, 182, 237	0

The worst 5 of 89 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	279	ASN	13.4
1	A	440	LEU	12.8
1	A	342	ASN	12.3
1	A	441	MET	12.1
1	A	173	LEU	9.2

6.2 Non-standard residues in protein, DNA, RNA chains (i)

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates (i)

There are no monosaccharides in this entry.

6.4 Ligands (i)

There are no ligands in this entry.

6.5 Other polymers (i)

There are no such residues in this entry.