



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 13, 2017 – 01:11 am GMT

PDB ID : 3IM2  
Title : Structure of the C-terminal Sec63 unit of yeast Brr2, P41212 Form  
Authors : Pena, V.; Hrmann, R.L.; Wahl, M.C.  
Deposited on : 2009-08-09  
Resolution : 1.99 Å(reported)

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

We welcome your comments at [validation@mail.wwpdb.org](mailto:validation@mail.wwpdb.org)

A user guide is available at

<http://wwpdb.org/validation/2016/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

---

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.7.2 (RC1), CSD as538be (2017)  
Xtriage (Phenix) : 1.9-1692  
EDS : trunk28620  
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)  
Refmac : 5.8.0135  
CCP4 : 6.5.0  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : recalc28949

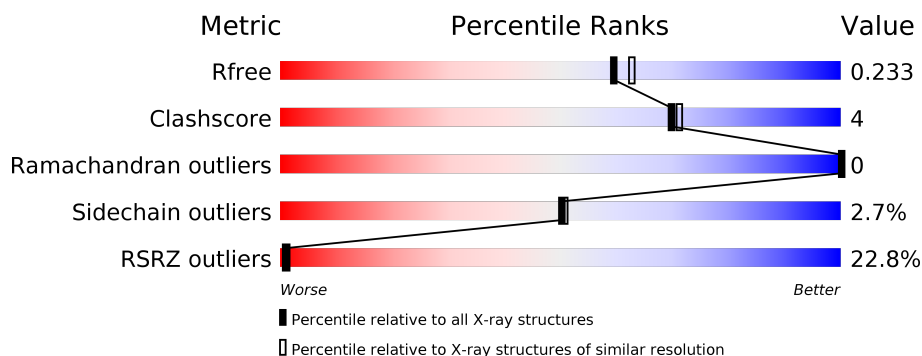
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.99 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))
$R_{free}$	100719	6609 (2.00-2.00)
Clashscore	112137	7775 (2.00-2.00)
Ramachandran outliers	110173	7679 (2.00-2.00)
Sidechain outliers	110143	7678 (2.00-2.00)
RSRZ outliers	101464	6696 (2.00-2.00)

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 on the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria. 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  $\leq 5\%$ . 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	328	<div> <div>22%</div> <div>89%</div> <div>8% ..</div> </div>

## 2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 3070 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.

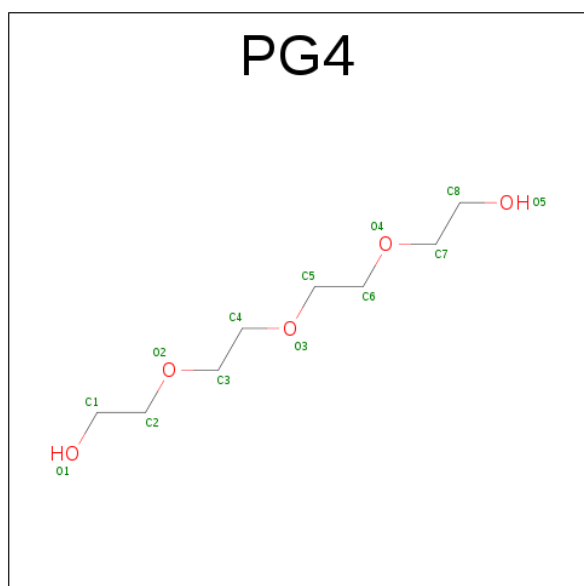
- Molecule 1 is a protein called Pre-mRNA-splicing helicase BRR2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	320	2697	1723	438	529	7	0	16	0

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1836	GLY	-	EXPRESSION TAG	UNP P32639
A	1837	ALA	-	EXPRESSION TAG	UNP P32639
A	1838	MET	-	EXPRESSION TAG	UNP P32639

- Molecule 2 is TETRAETHYLENE GLYCOL (three-letter code: PG4) (formula: C<sub>8</sub>H<sub>18</sub>O<sub>5</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	C	O		
2	A	1	13	8	5	0	0

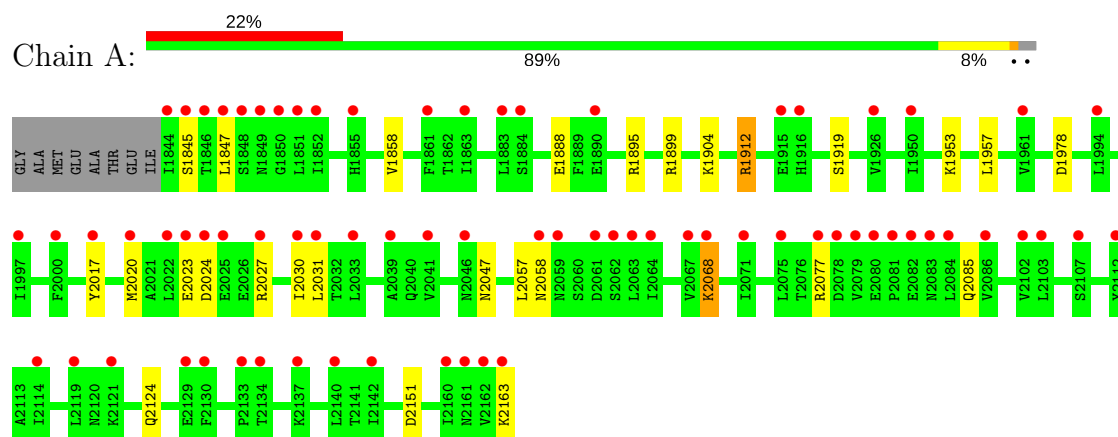
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	352	Total 360	O 360	0	8

### 3 Residue-property plots

These plots are drawn for all protein, RNA and DNA 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 ( $\text{RSRZ} > 2$ ). 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



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 41 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	104.09Å 104.09Å 77.37Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	20.00 – 1.99 19.94 – 1.99	Depositor EDS
% Data completeness (in resolution range)	99.0 (20.00-1.99) 99.0 (19.94-1.99)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.10	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.53 (at 1.99Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, $R_{free}$	0.188 , 0.237 0.187 , 0.233	Depositor DCC
$R_{free}$ test set	1501 reflections (5.37%)	DCC
Wilson B-factor (Å <sup>2</sup> )	32.1	Xtriage
Anisotropy	0.060	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.31 , 42.8	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	3070	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	37.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.55% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  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](#)

Bond lengths and bond angles in the following residue types are not validated in this section: PG4

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	0.70	4/2745 (0.1%)	0.58	0/3725

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	2023[A]	GLU	CD-OE1	17.03	1.44	1.25
1	A	2023[B]	GLU	CD-OE1	17.03	1.44	1.25
1	A	2023[A]	GLU	CB-CG	6.53	1.64	1.52
1	A	2023[B]	GLU	CB-CG	6.53	1.64	1.52

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	2697	0	2696	20	0
2	A	13	0	18	6	0
3	A	360	0	0	10	0
All	All	3070	0	2714	23	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.

All (23) 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)
1:A:1978[B]:ASP:HB2	2:A:1:PG4:H72	1.49	0.91
1:A:1978[B]:ASP:CB	2:A:1:PG4:H72	2.08	0.84
1:A:2151[A]:ASP:OD2	3:A:351[A]:HOH:O	2.06	0.74
1:A:1899[A]:ARG:HE	1:A:1919:SER:HB2	1.51	0.73
2:A:1:PG4:H32	3:A:314:HOH:O	1.93	0.68
1:A:2068:LYS:H	1:A:2068:LYS:HE2	1.64	0.63
1:A:2020:MET:HA	1:A:2027:ARG:HH22	1.66	0.60
1:A:1845:SER:HA	1:A:1847:LEU:HD13	1.86	0.57
2:A:1:PG4:H82	3:A:314:HOH:O	2.07	0.55
1:A:1858:VAL:HG11	1:A:1957:LEU:HD23	1.88	0.54
1:A:1895[B]:ARG:NE	3:A:309[B]:HOH:O	2.30	0.54
1:A:1845:SER:C	1:A:1847:LEU:H	2.12	0.53
1:A:1978[B]:ASP:HB3	2:A:1:PG4:H72	1.91	0.50
1:A:1899[A]:ARG:NH2	3:A:10:HOH:O	2.17	0.50
1:A:2030:ILE:HG13	1:A:2031:LEU:N	2.30	0.47
1:A:1899[A]:ARG:HE	1:A:1919:SER:CB	2.24	0.46
2:A:1:PG4:H51	3:A:122:HOH:O	2.15	0.46
1:A:1912[A]:ARG:NH2	3:A:189:HOH:O	2.50	0.45
1:A:1904:LYS:HE2	3:A:256[A]:HOH:O	2.18	0.44
1:A:2077:ARG:NH2	3:A:240:HOH:O	2.36	0.42
1:A:2085:GLN:NE2	3:A:346:HOH:O	2.53	0.42
1:A:1888:GLU:O	1:A:1953[B]:LYS:HD3	2.19	0.41
1:A:2017[B]:TYR:HA	1:A:2020:MET:HG2	2.03	0.41

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	334/328 (102%)	328 (98%)	6 (2%)	0	100	100



There are no Ramachandran outliers to report.

### 5.3.2 Protein sidechains ⓘ

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	313/302 (104%)	304 (97%)	9 (3%)	48 47

All (9) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	1912[A]	ARG
1	A	1912[B]	ARG
1	A	2024	ASP
1	A	2047	ASN
1	A	2057	LEU
1	A	2058	ASN
1	A	2068	LYS
1	A	2124	GLN
1	A	2163	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

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

## 5.5 Carbohydrates ⓘ

There are no carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

1 ligand is modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the chemical component dictionary. The Link column lists molecule types, if any, to which the group is linked. 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  $|Z| > 2$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z  > 2$	Counts	RMSZ	$\# Z  > 2$
2	PG4	A	1	-	12,12,12	0.49	0	11,11,11	0.62	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the chemical component dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	PG4	A	1	-	-	0/10/10/10	0/0/0/0

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 6 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	1	PG4	6	0

## 5.7 Other polymers [i](#)

There are no such residues in this entry.

## 5.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

## 6 Fit of model and data ⓘ

### 6.1 Protein, DNA and RNA chains ⓘ

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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	320/328 (97%)	1.32	73 (22%) 1 1	22, 35, 49, 52	0

All (73) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	1846	THR	9.4
1	A	1849	ASN	7.4
1	A	2163	LYS	6.1
1	A	2119	LEU	5.7
1	A	2061	ASP	5.7
1	A	2064	ILE	5.5
1	A	2024	ASP	5.4
1	A	1845	SER	5.3
1	A	2031	LEU	4.9
1	A	1844	ILE	4.9
1	A	1851	LEU	4.8
1	A	1847	LEU	4.6
1	A	2137	LYS	4.5
1	A	1850	GLY	4.4
1	A	2083	ASN	4.3
1	A	2161	ASN	4.3
1	A	1848	SER	4.2
1	A	2059	ASN	4.2
1	A	2077	ARG	4.1
1	A	2030	ILE	4.1
1	A	2023[A]	GLU	4.1
1	A	2084	LEU	4.0
1	A	2068	LYS	4.0
1	A	2130	PHE	3.8
1	A	2103	LEU	3.8
1	A	1997	ILE	3.7
1	A	2079	VAL	3.7

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	1915	GLU	3.4
1	A	1994	LEU	3.4
1	A	2134	THR	3.2
1	A	2017[A]	TYR	3.2
1	A	2062	SER	3.2
1	A	2025	GLU	3.1
1	A	2082	GLU	3.1
1	A	2081	PRO	3.0
1	A	2020	MET	3.0
1	A	2067	VAL	3.0
1	A	2162	VAL	2.9
1	A	1861	PHE	2.9
1	A	2121	LYS	2.9
1	A	2063	LEU	2.8
1	A	2075[A]	LEU	2.8
1	A	2022	LEU	2.8
1	A	1916	HIS	2.8
1	A	2080	GLU	2.7
1	A	1950	ILE	2.7
1	A	2140	LEU	2.7
1	A	2160	ILE	2.6
1	A	2058	ASN	2.6
1	A	1852	ILE	2.6
1	A	1855	HIS	2.5
1	A	2041	VAL	2.5
1	A	1961	VAL	2.5
1	A	1890	GLU	2.5
1	A	1926	VAL	2.4
1	A	1863	ILE	2.4
1	A	2071	ILE	2.4
1	A	1883	LEU	2.4
1	A	2129	GLU	2.3
1	A	2078	ASP	2.3
1	A	2112	TYR	2.3
1	A	2000	PHE	2.3
1	A	2027	ARG	2.2
1	A	2046	ASN	2.2
1	A	1884[A]	SER	2.1
1	A	2114	ILE	2.1
1	A	2086	VAL	2.1
1	A	2039	ALA	2.1
1	A	2133	PRO	2.1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	2107	SER	2.1
1	A	2142	ILE	2.1
1	A	2033	LEU	2.1
1	A	2102	VAL	2.0

## 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 carbohydrates in this entry.

## 6.4 Ligands [i](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
2	PG4	A	1	13/13	0.89	0.17	0.91	42,44,48,48	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.