Full wwPDB X-ray Structure Validation Report ⓘ

Apr 18, 2026 – 04:48 PM UTC

PDB ID	:	4ILJ / pdb_00004ilj
Title	:	Crystal structure of an Prp8p RNaseH W1911A mutant protein
Authors	:	Weber, G.; Heroven, C.; Santos, K.F.; Wahl, M.C.
Deposited on	:	2012-12-31
Resolution	:	2.00 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*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the ⓘ 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 ⓘ) were used in the production of this report:

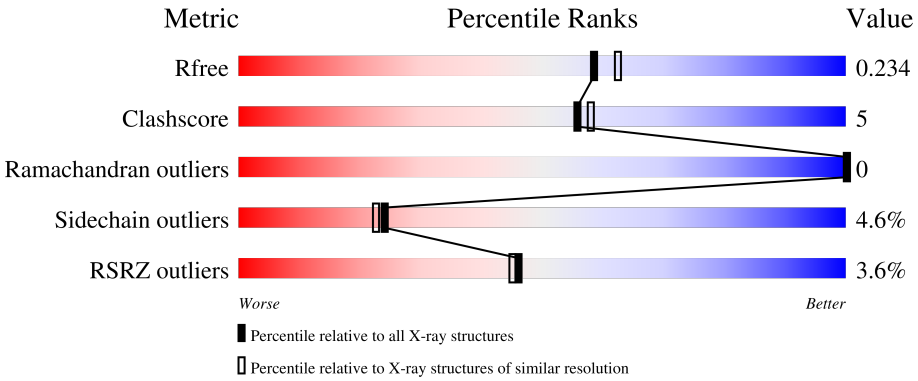
MolProbity	:	4-5-2 with Phenix2.0
Xtriage (Phenix)	:	2.0
EDS	:	3.0
Percentile statistics	:	20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4	:	9.0.010 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.49

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 2.00 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}	180053	10052 (2.00-2.00)
Clashscore	190562	11152 (2.00-2.00)
Ramachandran outliers	187476	11031 (2.00-2.00)
Sidechain outliers	187428	11029 (2.00-2.00)
RSRZ outliers	180081	10067 (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 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	258	5% 88% 10%
1	B	258	2% 85% 10%

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 4650 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 factor 8.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	254	Total	C	N	O	S	1	13	0
			2132	1362	360	401	9			
1	B	254	Total	C	N	O	S	8	15	0
			2152	1375	363	404	10			

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1833	GLY	-	expression tag	UNP P33334
A	1834	ALA	-	expression tag	UNP P33334
A	1835	MET	-	expression tag	UNP P33334
A	1911	ALA	TRP	engineered mutation	UNP P33334
B	1833	GLY	-	expression tag	UNP P33334
B	1834	ALA	-	expression tag	UNP P33334
B	1835	MET	-	expression tag	UNP P33334
B	1911	ALA	TRP	engineered mutation	UNP P33334

- Molecule 2 is water.

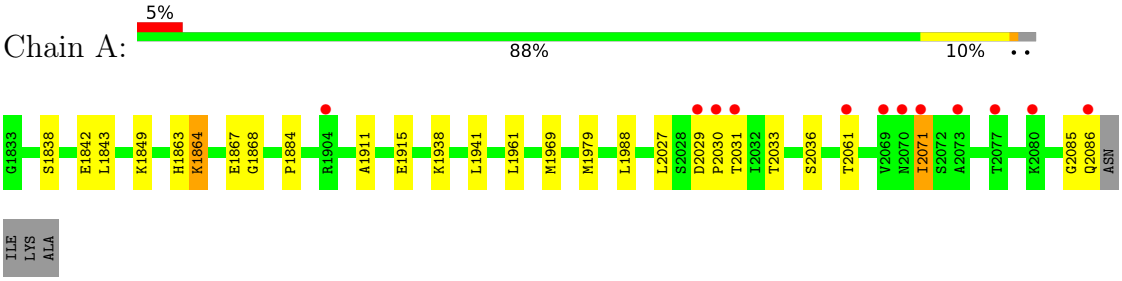
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	169	Total	O	0	0
			169	169		
2	B	197	Total	O	0	0
			197	197		

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.

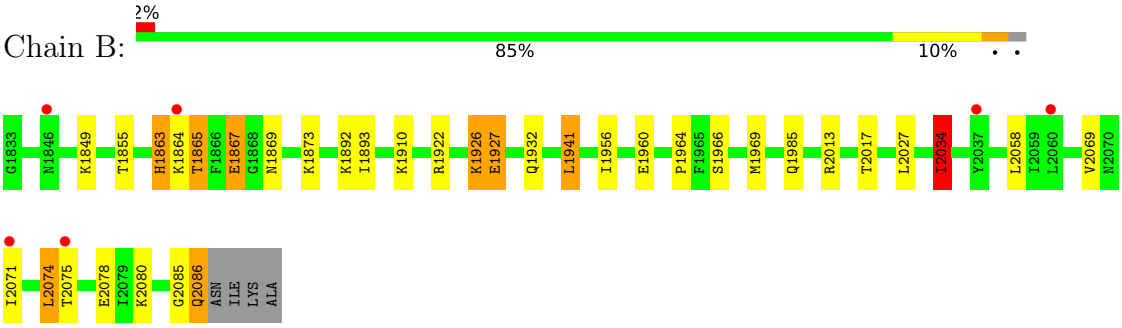
• Molecule 1: Pre-mRNA-splicing factor 8

Chain A:



• Molecule 1: Pre-mRNA-splicing factor 8

Chain B:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	75.60Å 85.02Å 95.95Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	37.80 – 2.00 37.80 – 2.00	Depositor EDS
% Data completeness (in resolution range)	99.7 (37.80-2.00) 99.7 (37.80-2.00)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	0.05	Depositor
$< I/\sigma(I) >$ [1]	2.42 (at 2.00Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.8_1069)	Depositor
R, R_{free}	0.180 , 0.226 0.182 , 0.234	Depositor DCC
R_{free} test set	2118 reflections (4.99%)	wwPDB-VP
Wilson B-factor (Å^2)	31.7	Xtriage
Anisotropy	0.350	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.30 , 50.7	EDS
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	4650	wwPDB-VP
Average B, all atoms (Å^2)	40.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.16% 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

5.1 Standard geometry

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	0.54	0/2180	0.78	1/2954 (0.0%)
1	B	0.56	0/2197	0.81	1/2976 (0.0%)
All	All	0.55	0/4377	0.80	2/5930 (0.0%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	2034	ILE	CB-CA-C	-6.74	100.57	110.62
1	A	2071	ILE	CB-CA-C	-5.49	104.85	111.88

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts

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	2132	0	2177	14	0
1	B	2152	0	2196	27	0
2	A	169	0	0	3	1
2	B	197	0	0	7	1
All	All	4650	0	4373	41	1

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 5.

All (41) 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:B:1865:THR:HG22	1:B:1869:ASN:H	1.40	0.86
1:A:1961:LEU:O	2:A:2184:HOH:O	2.09	0.69
1:B:1910:LYS:NZ	2:B:2218:HOH:O	2.26	0.68
1:A:1863:HIS:NE2	2:A:2246:HOH:O	2.27	0.66
1:B:1922:ARG:NH2	2:B:2111:HOH:O	2.30	0.65
1:B:2086:GLN:NE2	2:B:2183:HOH:O	2.33	0.61
1:B:1892:LYS:HB2	2:B:2274:HOH:O	2.03	0.58
1:B:1855[A]:THR:HG21	1:B:1966:SER:HB2	1.84	0.58
1:B:1960[A]:GLU:OE1	2:B:2199:HOH:O	2.17	0.58
1:A:1864:LYS:HD3	1:A:1868:GLY:HA2	1.86	0.57
1:A:2061:THR:HA	1:A:2071:ILE:HD11	1.88	0.56
1:A:1911:ALA:O	1:A:1915:GLU:HG3	2.09	0.53
1:B:2086:GLN:HE21	1:B:2086:GLN:CA	2.21	0.52
1:A:2085:GLY:O	1:A:2086:GLN:HB2	2.10	0.51
1:B:1863:HIS:CD2	1:B:1873:LYS:HE3	2.46	0.50
1:B:2086:GLN:HE21	1:B:2086:GLN:N	2.09	0.50
1:A:2029:ASP:OD1	1:A:2031:THR:OG1	2.23	0.50
1:B:1849:LYS:HG2	1:B:1932:GLN:HB2	1.92	0.50
1:B:2085:GLY:O	1:B:2086:GLN:HB2	2.12	0.49
1:B:1964:PRO:HG3	1:B:2013[B]:ARG:HG3	1.95	0.48
1:A:1843:LEU:HD22	1:A:1884:PRO:HG3	1.95	0.48
1:A:2031:THR:O	1:A:2033:THR:HG23	2.15	0.47
1:A:2027:LEU:O	1:A:2030:PRO:HB3	2.15	0.47
1:B:1865:THR:HG22	1:B:1869:ASN:N	2.19	0.47
1:A:1969:MET:HA	1:A:1969:MET:HE2	1.97	0.47
1:B:1865:THR:HG23	1:B:1867:GLU:H	1.80	0.47
1:B:2075:THR:HG23	1:B:2078:GLU:H	1.81	0.46
1:B:1893:ILE:HD12	1:B:1985[A]:GLN:HG3	1.98	0.45
1:A:1842[B]:GLU:O	1:A:1849:LYS:NZ	2.36	0.45
1:B:2034:ILE:HD12	2:B:2160:HOH:O	2.17	0.45
1:B:1969:MET:HE2	1:B:1969:MET:HA	1.99	0.44
1:B:1865:THR:CG2	1:B:1867:GLU:HG2	2.48	0.44
1:B:2071:ILE:HA	1:B:2074:LEU:HD23	2.00	0.44
1:A:1979:MET:HG2	2:A:2254:HOH:O	2.17	0.43
1:B:1892:LYS:NZ	2:B:2245:HOH:O	2.52	0.41
1:B:1926[A]:LYS:HD2	1:B:1927[A]:GLU:HG3	2.03	0.41
1:B:1941[B]:LEU:HD23	1:B:1956:ILE:HG22	2.03	0.41
1:B:1865:THR:HG23	1:B:1867:GLU:N	2.36	0.41
1:B:2069:VAL:CG1	1:B:2074:LEU:HD21	2.50	0.41

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the sym-

metry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:A:2188:HOH:O	2:B:2192:HOH:O[4_445]	2.16	0.04

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	265/258 (103%)	258 (97%)	7 (3%)	0	100	100
1	B	267/258 (104%)	263 (98%)	4 (2%)	0	100	100
All	All	532/516 (103%)	521 (98%)	11 (2%)	0	100	100

There are no Ramachandran outliers to report.

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	242/232 (104%)	234 (97%)	8 (3%)	33	34
1	B	244/232 (105%)	227 (93%)	17 (7%)	14	10
All	All	486/464 (105%)	461 (95%)	25 (5%)	24	19

All (25) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	1838	SER

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	1864	LYS
1	A	1867	GLU
1	A	1938	LYS
1	A	1941[A]	LEU
1	A	1941[B]	LEU
1	A	1988	LEU
1	A	2036	SER
1	B	1863	HIS
1	B	1864	LYS
1	B	1865	THR
1	B	1867	GLU
1	B	1926[A]	LYS
1	B	1926[B]	LYS
1	B	1927[A]	GLU
1	B	1927[B]	GLU
1	B	1941[A]	LEU
1	B	1941[B]	LEU
1	B	2017	THR
1	B	2027	LEU
1	B	2034	ILE
1	B	2058	LEU
1	B	2074	LEU
1	B	2080	LYS
1	B	2086	GLN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	1985	GLN
1	B	1863	HIS
1	B	1929	GLN
1	B	2054	GLN
1	B	2086	GLN

5.3.3 RNA

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 oligosaccharides 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)

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^{th} 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	254/258 (98%)	0.01	12 (4%) 36 35	12, 36, 75, 105	14 (5%)
1	B	253/258 (98%)	-0.12	6 (2%) 59 59	11, 36, 65, 80	14 (5%)
All	All	507/516 (98%)	-0.06	18 (3%) 46 45	11, 36, 70, 105	28 (5%)

All (18) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	2060[A]	LEU	6.8
1	A	2077	THR	4.1
1	A	2071	ILE	3.7
1	B	2075	THR	3.2
1	A	1904	ARG	3.1
1	A	2073	ALA	3.0
1	A	2070[A]	ASN	2.8
1	B	2071	ILE	2.5
1	B	1864	LYS	2.5
1	A	2030	PRO	2.5
1	A	2031	THR	2.4
1	A	2086	GLN	2.3
1	A	2029	ASP	2.3
1	A	2069[A]	VAL	2.3
1	A	2061	THR	2.3
1	A	2080	LYS	2.2
1	B	2037	TYR	2.1
1	B	1846	ASN	2.0

6.2 Non-standard residues in protein, DNA, RNA chains

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

6.3 Carbohydrates (i)

There are no oligosaccharides 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.