



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 06:25 PM GMT

PDB ID : 4LK2  
Title : Crystal structure of RNA splicing effector Prp5  
Authors : Zhang, Z.-M.; Li, J.; Yang, F.; Xu, Y.; Zhou, J.  
Deposited on : 2013-07-05  
Resolution : 2.12 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](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 (RC4), CSD as536be (2015)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20026688  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
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) : trunk26865

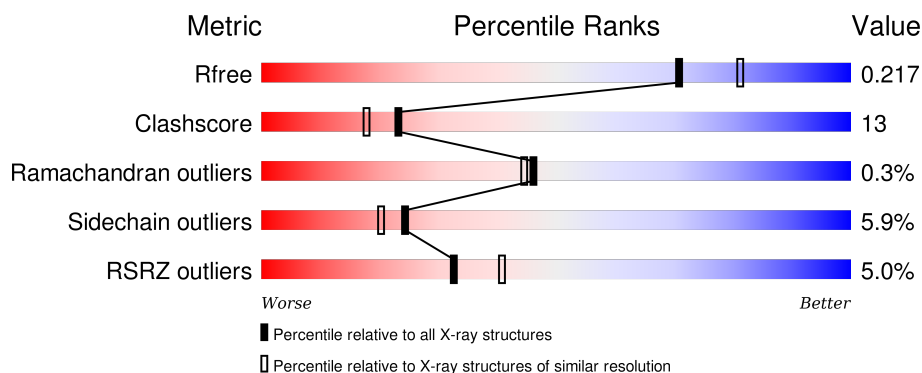
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.12 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))
$R_{free}$	91344	4587 (2.14-2.10)
Clashscore	102246	5132 (2.14-2.10)
Ramachandran outliers	100387	5080 (2.14-2.10)
Sidechain outliers	100360	5081 (2.14-2.10)
RSRZ outliers	91569	4597 (2.14-2.10)

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	493	<div> <div>4%</div> <div>66%</div> <div>23%</div> <div>10%</div> </div>
1	B	493	<div> <div>5%</div> <div>65%</div> <div>21%</div> <div>12%</div> </div>

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 7012 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-processing ATP-dependent RNA helicase PRP5.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	444	Total	C	N	O	S	0	5	0
			3436	2175	591	648	22			
1	B	436	Total	C	N	O	S	0	1	0
			3338	2114	571	632	21			

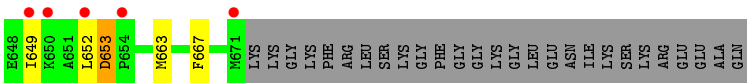
- Molecule 2 is NICKEL (II) ION (three-letter code: NI) (formula: Ni).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	1	Total	Ni	0	0
			1	1		

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	109	Total	O	0	0
			109	109		
3	B	128	Total	O	0	0
			128	128		





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	87.16Å 86.08Å 127.99Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	38.88 – 2.12 38.88 – 2.05	Depositor EDS
% Data completeness (in resolution range)	93.1 (38.88-2.12) 97.3 (38.88-2.05)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.98 (at 2.05Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.6.2_432)	Depositor
R, $R_{free}$	0.212 , 0.255 0.205 , 0.217	Depositor DCC
$R_{free}$ test set	2658 reflections (4.94%)	DCC
Wilson B-factor (Å <sup>2</sup> )	33.3	Xtriage
Anisotropy	0.753	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.35 , 56.5	EDS
Estimated twinning fraction	0.033 for k,h,l	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtriage
Outliers	0 of 59369 reflections	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	7012	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	43.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.47% 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, 0.375 respectively for untwinned datasets, and 0.333, 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: NI

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	0.53	1/3502 (0.0%)	0.63	3/4728 (0.1%)
1	B	0.39	0/3388	0.56	0/4576
All	All	0.46	1/6890 (0.0%)	0.60	3/9304 (0.0%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	271	THR	C-N	20.20	1.80	1.34

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	271	THR	O-C-N	-18.60	92.94	122.70
1	A	271	THR	CA-C-N	6.33	131.12	117.20
1	A	273	LYS	N-CA-C	5.46	125.75	111.00

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	3436	0	3485	93	0
1	B	3338	0	3371	90	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	A	1	0	0	0	0
3	A	109	0	0	4	0
3	B	128	0	0	5	0
All	All	7012	0	6856	183	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 13.

The worst 5 of 183 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:271:THR:C	1:A:273:LYS:N	1.80	1.32
1:A:485:ILE:H	1:A:656:GLN:HE22	1.03	0.94
1:A:484:ARG:HD3	1:A:626:ILE:HD11	1.51	0.91
1:B:555:ILE:HD11	1:B:564:ARG:HA	1.58	0.86
1:A:560:PRO:HG2	1:A:563:GLU:HG2	1.58	0.85

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	445/493 (90%)	431 (97%)	13 (3%)	1 (0%)	52	52
1	B	427/493 (87%)	415 (97%)	10 (2%)	2 (0%)	34	29
All	All	872/986 (88%)	846 (97%)	23 (3%)	3 (0%)	46	44

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	272	GLU

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type
1	A	273	LYS
1	B	274	LEU

### 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	380/437 (87%)	360 (95%)	20 (5%)	28	25
1	B	369/437 (84%)	342 (93%)	27 (7%)	17	13
All	All	749/874 (86%)	702 (94%)	47 (6%)	24	18

5 of 47 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	240	LEU
1	B	404	LEU
1	B	615	ARG
1	B	263	SER
1	B	406	THR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 10 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	451	ASN
1	B	470	ASN
1	B	547	ASN
1	A	656	GLN
1	B	535	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 carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

Of 1 ligands modelled in this entry, 1 is monoatomic - leaving 0 for Mogul analysis.

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.

No monomer is involved in short contacts.

## 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 [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, 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	444/493 (90%)	0.30	18 (4%) 41 49	23, 40, 64, 73	0
1	B	436/493 (88%)	0.30	26 (5%) 25 33	28, 43, 68, 82	0
All	All	880/986 (89%)	0.30	44 (5%) 32 41	23, 42, 65, 82	0

The worst 5 of 44 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	277	GLY	5.7
1	A	620	SER	4.4
1	A	262	LEU	3.8
1	A	327	HIS	3.6
1	A	274	LEU	3.5

### 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	NI	A	701	1/1	0.99	0.13	1.68	32,32,32,32	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.