



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 26, 2014 – 03:46 PM GMT

PDB ID : 4J0X  
Title : Structure of Rrp9  
Authors : Zhang, L.; Lin, J.; Ye, K.  
Deposited on : 2013-01-31  
Resolution : 2.50 Å(reported)

This is a full wwPDB 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/ValidationPDFNotes.html>

---

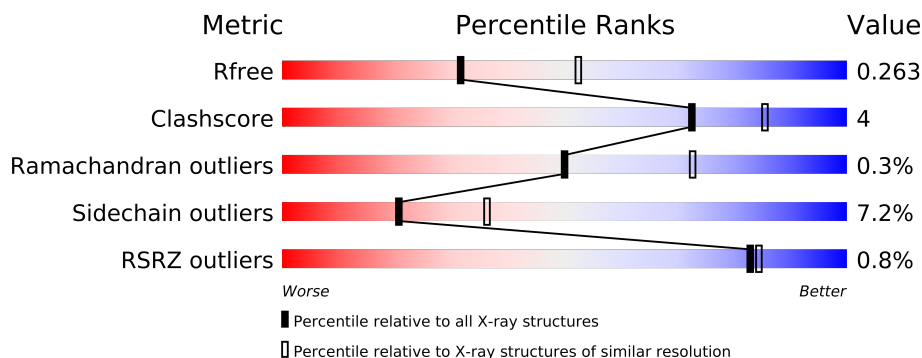
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.15 2013  
Xtriage (Phenix) : dev-1323  
EDS : stable22639  
Percentile statistics : 21963  
Refmac : 5.8.0049  
CCP4 : 6.3.0 (Settle)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)  
Validation Pipeline (wwPDB-VP) : stable22683

# 1 Overall quality at a glance

The reported resolution of this entry is 2.50 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))
$R_{free}$	66092	2784 (2.50-2.50)
Clashscore	79885	3562 (2.50-2.50)
Ramachandran outliers	78287	3480 (2.50-2.50)
Sidechain outliers	78261	3482 (2.50-2.50)
RSRZ outliers	66119	2785 (2.50-2.50)

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. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	451	
1	B	451	

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 5872 atoms, of which 0 are hydrogen and 0 are deuterium.

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 Ribosomal RNA-processing protein 9.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	365	Total	C	N	O	S	0	0	0
			2916	1871	506	529	10			
1	B	366	Total	C	N	O	S	0	0	0
			2924	1875	508	531	10			

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	123	GLY	-	EXPRESSION TAG	UNP Q06506
A	124	PRO	-	EXPRESSION TAG	UNP Q06506
A	125	GLU	-	EXPRESSION TAG	UNP Q06506
A	126	ALA	-	EXPRESSION TAG	UNP Q06506
B	123	GLY	-	EXPRESSION TAG	UNP Q06506
B	124	PRO	-	EXPRESSION TAG	UNP Q06506
B	125	GLU	-	EXPRESSION TAG	UNP Q06506
B	126	ALA	-	EXPRESSION TAG	UNP Q06506

- Molecule 2 is water.

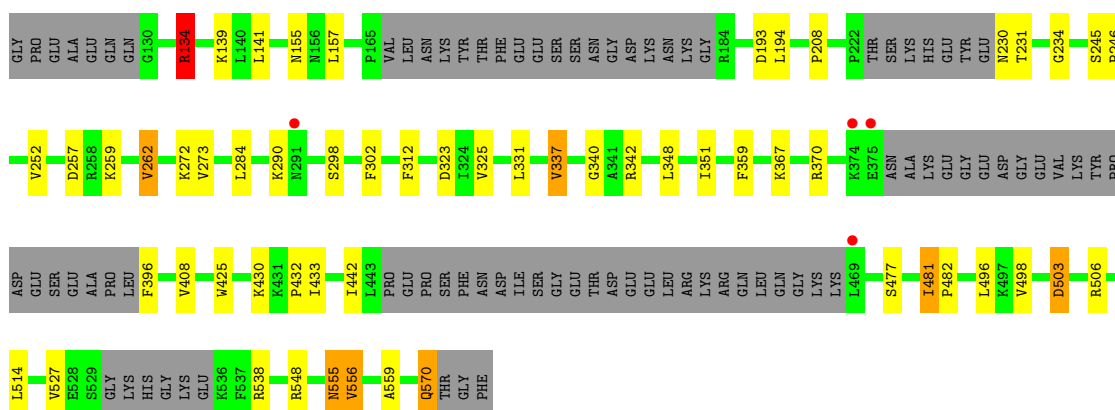
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	11	Total	O	0	0
			11	11		
2	B	21	Total	O	0	0
			21	21		

### 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 errors 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.

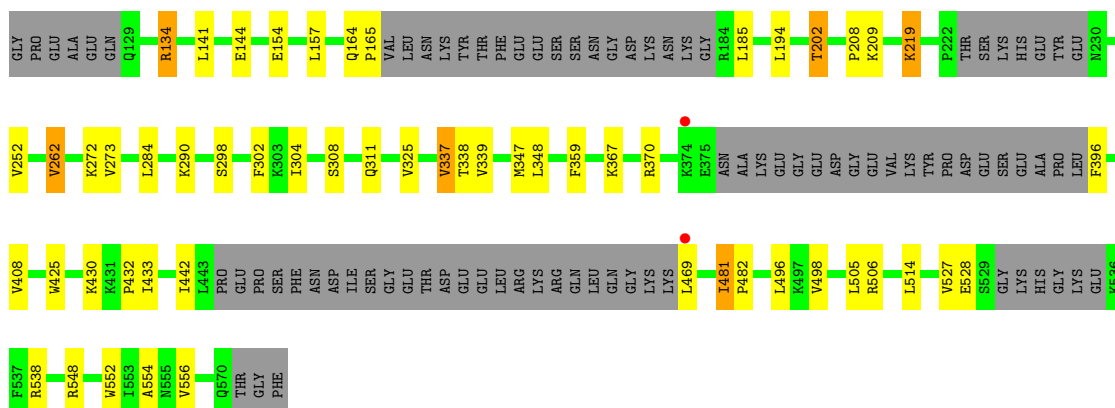
#### • Molecule 1: Ribosomal RNA-processing protein 9

Chain A:



#### • Molecule 1: Ribosomal RNA-processing protein 9

Chain B:



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	48.45Å 78.56Å 80.03Å 61.77° 86.44° 71.99°	Depositor
Resolution (Å)	19.96 – 2.50 19.96 – 2.50	Depositor EDS
% Data completeness (in resolution range)	95.3 (19.96-2.50) 95.2 (19.96-2.50)	Depositor EDS
$R_{merge}$	0.09	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.54 (at 2.50Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.8.1_1168)	Depositor
R, $R_{free}$	0.218 , 0.262 0.219 , 0.263	Depositor DCC
$R_{free}$ test set	1640 reflections (5.08%)	DCC
Wilson B-factor (Å <sup>2</sup> )	31.9	Xtriage
Anisotropy	0.019	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 23.9	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Outliers	0 of 32300 reflections	Xtriage
$F_o, F_c$ correlation	0.91	EDS
Total number of atoms	5872	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	32.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.68% of the height of the origin peak. No significant pseudotranslation is detected.*

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

## 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	0.25	0/2975	0.51	1/4008 (0.0%)
1	B	0.25	0/2983	0.48	0/4018
All	All	0.25	0/5958	0.49	1/8026 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	134	ARG	NE-CZ-NH1	5.76	123.18	120.30

There are no chirality outliers.

There are no planarity outliers.

### 5.2 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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	2916	0	2947	26	0
1	B	2924	0	2951	24	0
2	A	11	0	0	0	0
2	B	21	0	0	0	0
All	All	5872	0	5898	50	0

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 4.

All (50) close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:134:ARG:HH11	1:A:134:ARG:HG2	1.31	0.92
1:A:272:LYS:HG2	1:A:312:PHE:HB3	1.59	0.83
1:B:337:VAL:HG13	1:B:348:LEU:HD23	1.75	0.68
1:B:262:VAL:HG13	1:B:272:LYS:HB3	1.78	0.64
1:A:337:VAL:HG13	1:A:348:LEU:HD23	1.79	0.64
1:A:234:GLY:O	1:A:259:LYS:NZ	2.34	0.58
1:B:185:LEU:HD12	1:B:202:THR:HA	1.84	0.57
1:B:308:SER:HB3	1:B:311:GLN:HB2	1.87	0.56
1:A:262:VAL:HG13	1:A:272:LYS:HB3	1.88	0.54
1:B:165:PRO:HD2	1:B:185:LEU:HD23	1.90	0.54
1:B:185:LEU:H	1:B:202:THR:HG22	1.74	0.53
1:A:284:LEU:O	1:A:548:ARG:NH2	2.41	0.52
1:B:367:LYS:HE3	1:B:370:ARG:HH21	1.75	0.51
1:B:433:ILE:HD11	1:B:506:ARG:HG2	1.93	0.50
1:A:359:PHE:HB3	1:A:425:TRP:CZ3	2.46	0.50
1:B:134:ARG:NH2	1:B:481:ILE:HD11	2.27	0.49
1:A:134:ARG:HG2	1:A:134:ARG:NH1	2.06	0.49
1:B:284:LEU:O	1:B:548:ARG:NH2	2.45	0.49
1:B:164:GLN:NE2	1:B:528:GLU:O	2.47	0.48
1:A:323:ASP:HB3	1:A:340:GLY:O	2.14	0.48
1:A:433:ILE:HD11	1:A:506:ARG:HG2	1.95	0.48
1:B:304:ILE:HG12	1:B:338:THR:HG21	1.95	0.48
1:A:367:LYS:HE3	1:A:370:ARG:HH21	1.77	0.48
1:B:209:LYS:HB2	1:B:209:LYS:HE3	1.67	0.47
1:B:359:PHE:HB3	1:B:425:TRP:CZ3	2.49	0.47
1:B:290:LYS:HE3	1:B:290:LYS:HB3	1.74	0.47
1:B:425:TRP:CH2	1:B:432:PRO:HG3	2.50	0.47
1:A:503:ASP:OD1	1:A:503:ASP:N	2.32	0.47
1:A:290:LYS:HD3	1:A:331:LEU:HA	1.96	0.46
1:A:139:LYS:O	1:A:570:GLN:HB3	2.15	0.46
1:B:298:SER:HB2	1:B:325:VAL:HB	1.96	0.46
1:A:298:SER:HB2	1:A:325:VAL:HB	1.98	0.45
1:A:556:VAL:HG13	1:A:559:ALA:HB2	1.97	0.45
1:B:552:TRP:CH2	1:B:554:ALA:HA	2.52	0.45
1:A:481:ILE:HA	1:A:482:PRO:HD3	1.84	0.44
1:A:323:ASP:OD2	1:A:342:ARG:HB2	2.18	0.44
1:A:230:ASN:ND2	1:A:259:LYS:HG3	2.33	0.43
1:A:302:PHE:N	1:A:302:PHE:CD1	2.86	0.43
1:A:425:TRP:CZ2	1:A:432:PRO:HG3	2.53	0.43
1:A:230:ASN:ND2	1:A:257:ASP:O	2.52	0.43
1:B:481:ILE:HA	1:B:482:PRO:HD3	1.83	0.42
1:A:155:ASN:HB2	1:A:193:ASP:OD1	2.19	0.42

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:B:425:TRP:CZ2	1:B:432:PRO:HG3	2.55	0.41
1:B:302:PHE:CD1	1:B:302:PHE:N	2.87	0.41
1:A:351:ILE:HG12	1:A:351:ILE:H	1.71	0.41
1:A:208:PRO:HD3	1:A:538:ARG:NH2	2.36	0.40
1:B:141:LEU:HB3	1:B:144:GLU:HG2	2.03	0.40
1:B:208:PRO:HD3	1:B:538:ARG:NH2	2.37	0.40
1:B:469:LEU:HD23	1:B:469:LEU:HA	1.94	0.40
1:A:245:SER:HA	1:A:246:PRO:HD3	1.88	0.40

There are no symmetry-related clashes.

## 5.3 Torsion angles

### 5.3.1 Protein backbone ⓘ

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	353/451 (78%)	341 (97%)	11 (3%)	1 (0%)	50	73
1	B	354/451 (78%)	343 (97%)	10 (3%)	1 (0%)	50	73
All	All	707/902 (78%)	684 (97%)	21 (3%)	2 (0%)	50	73

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	555	ASN
1	B	219	LYS

### 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.

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	317/390 (81%)	294 (93%)	23 (7%)	20	36
1	B	318/390 (82%)	295 (93%)	23 (7%)	21	36
All	All	635/780 (81%)	589 (93%)	46 (7%)	21	36

All (46) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	134	ARG
1	A	141	LEU
1	A	157	LEU
1	A	194	LEU
1	A	231	THR
1	A	252	VAL
1	A	262	VAL
1	A	273	VAL
1	A	337	VAL
1	A	396	PHE
1	A	408	VAL
1	A	430	LYS
1	A	442	ILE
1	A	477	SER
1	A	481	ILE
1	A	496	LEU
1	A	498	VAL
1	A	503	ASP
1	A	514	LEU
1	A	527	VAL
1	A	555	ASN
1	A	556	VAL
1	A	570	GLN
1	B	134	ARG
1	B	154	GLU
1	B	157	LEU
1	B	194	LEU
1	B	202	THR
1	B	219	LYS
1	B	252	VAL
1	B	262	VAL
1	B	273	VAL

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	B	337	VAL
1	B	339	VAL
1	B	347	MET
1	B	396	PHE
1	B	408	VAL
1	B	430	LYS
1	B	442	ILE
1	B	481	ILE
1	B	496	LEU
1	B	498	VAL
1	B	505	LEU
1	B	514	LEU
1	B	527	VAL
1	B	556	VAL

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 chains 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 ⓘ

There are no ligands in this entry.

## 5.7 Other polymers ⓘ

There are no such residues in this entry.

## 5.8 Polymer linkage issues

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	365/451 (80%)	-0.18	4 (1%) 77 79	13, 31, 53, 80	0
1	B	366/451 (81%)	-0.19	2 (0%) 88 90	11, 30, 53, 79	0
All	All	731/902 (81%)	-0.19	6 (0%) 83 84	11, 30, 53, 80	0

All (6) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	374	LYS	3.5
1	B	469	LEU	3.2
1	A	291	ASN	2.9
1	A	375	GLU	2.9
1	A	469	LEU	2.6
1	A	374	LYS	2.2

### 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 ⓘ

There are no carbohydrates in this entry.

### 6.4 Ligands ⓘ

There are no ligands in this entry.

## 6.5 Other polymers ⓘ

There are no such residues in this entry.