



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 1, 2016 – 05:01 AM GMT

PDB ID : 2OZB  
Title : Structure of a human Prp31-15.5K-U4 snRNA complex  
Authors : Liu, S.; Luehrmann, R.; Wahl, M.C.  
Deposited on : 2007-02-26  
Resolution : 2.60 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 (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

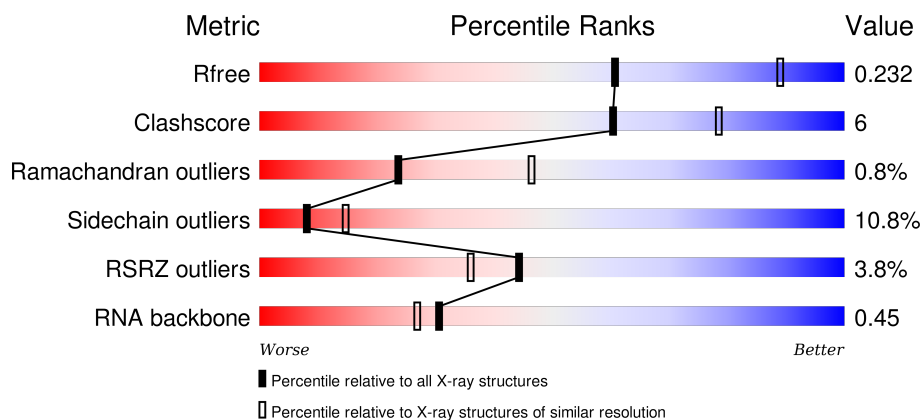
# 1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.60 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	2328 (2.60-2.60)
Clashscore	102246	2679 (2.60-2.60)
Ramachandran outliers	100387	2635 (2.60-2.60)
Sidechain outliers	100360	2635 (2.60-2.60)
RSRZ outliers	91569	2334 (2.60-2.60)
RNA backbone	2183	1022 (3.00-2.20)

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	C	33	<div> <div>6%</div> <div> <div></div> <div>45%</div> <div>36%</div> <div>9%</div> <div>9%</div> </div> </div>
1	F	33	<div> <div>15%</div> <div> <div></div> <div>42%</div> <div>36%</div> <div>18%</div> <div>•</div> </div> </div>
2	A	130	<div> <div>4%</div> <div> <div></div> <div>80%</div> <div>15%</div> <div>•••</div> </div> </div>
2	D	130	<div> <div>5%</div> <div> <div></div> <div>72%</div> <div>21%</div> <div>••</div> </div> </div>

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	B	260	<div> <div>3%</div> <div>71%</div> <div>18%</div> <div>8%</div> </div>
3	E	260	<div> <div>0%</div> <div>78%</div> <div>12%</div> <div>8%</div> </div>

## 2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 7375 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 RNA chain called RNA comprising the 5' Stem-Loop RNA of U4snRNA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	C	33	Total	C	N	O	P	0	0	0
			703	315	127	229	32			
1	F	33	Total	C	N	O	P	0	0	0
			703	315	127	229	32			

- Molecule 2 is a protein called U4/U6.U5 tri-snRNP 15.5 kDa protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	A	126	Total	C	N	O	S	0	0	0
			977	616	173	183	5			
2	D	125	Total	C	N	O	S	0	0	0
			968	611	172	180	5			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-1	GLY	-	CLONING ARTIFACT	UNP P55769
A	0	SER	-	CLONING ARTIFACT	UNP P55769
D	-1	GLY	-	CLONING ARTIFACT	UNP P55769
D	0	SER	-	CLONING ARTIFACT	UNP P55769

- Molecule 3 is a protein called U4/U6 small nuclear ribonucleoprotein Prp31.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	B	239	Total	C	N	O	S	0	2	0
			1896	1187	328	371	10			
3	E	240	Total	C	N	O	S	0	0	0
			1886	1182	327	367	10			

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	74	GLY	-	CLONING ARTIFACT	UNP Q8WWY3
B	75	PRO	-	CLONING ARTIFACT	UNP Q8WWY3
B	76	LEU	-	CLONING ARTIFACT	UNP Q8WWY3
B	77	GLY	-	CLONING ARTIFACT	UNP Q8WWY3
E	74	GLY	-	CLONING ARTIFACT	UNP Q8WWY3
E	75	PRO	-	CLONING ARTIFACT	UNP Q8WWY3
E	76	LEU	-	CLONING ARTIFACT	UNP Q8WWY3
E	77	GLY	-	CLONING ARTIFACT	UNP Q8WWY3

- Molecule 4 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	B	1	Total Ca 1 1	0	0
4	A	1	Total Ca 1 1	0	0
4	E	1	Total Ca 1 1	0	0

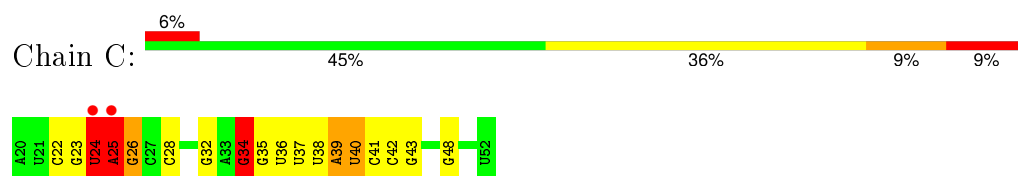
- Molecule 5 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	C	20	Total O 21 21	0	1
5	F	14	Total O 14 14	0	0
5	A	28	Total O 28 28	0	0
5	B	81	Total O 87 87	0	6
5	D	12	Total O 13 13	0	1
5	E	73	Total O 76 76	0	3

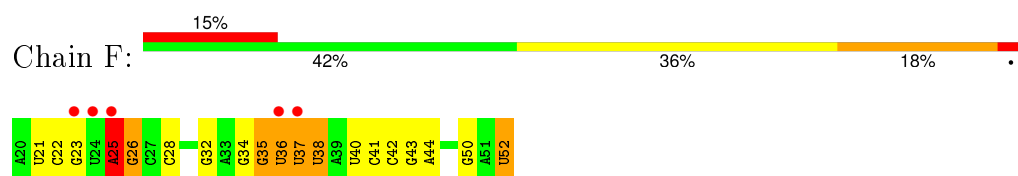
### 3 Residue-property plots [i](#)

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.

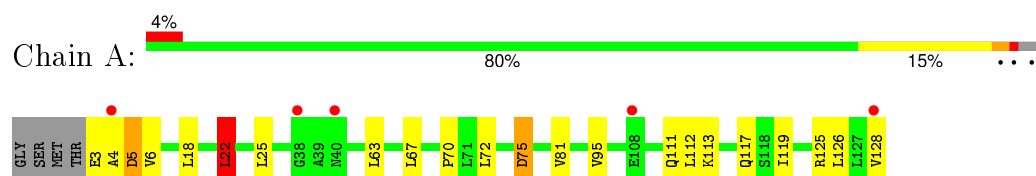
- Molecule 1: RNA comprising the 5' Stem-Loop RNA of U4snRNA



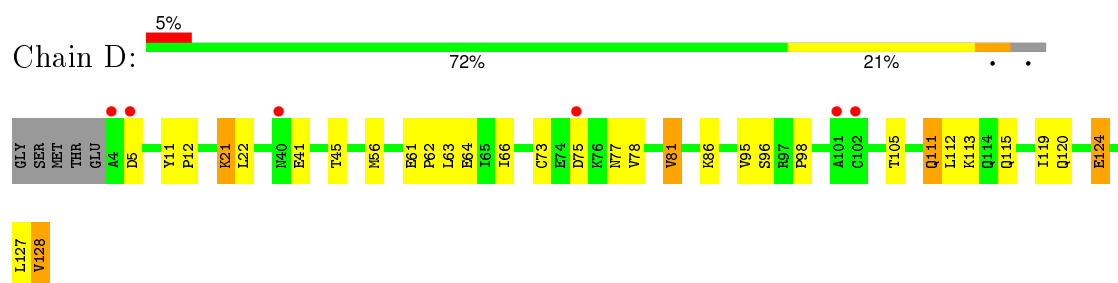
- Molecule 1: RNA comprising the 5' Stem-Loop RNA of U4snRNA



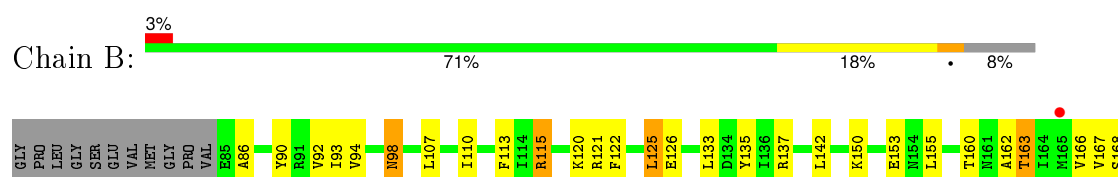
- Molecule 2: U4/U6.U5 tri-snRNP 15.5 kDa protein

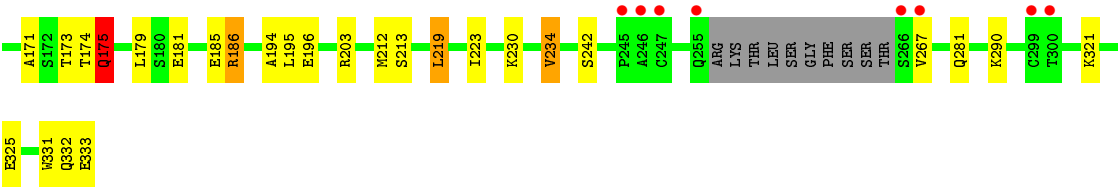


- Molecule 2: U4/U6.U5 tri-snRNP 15.5 kDa protein

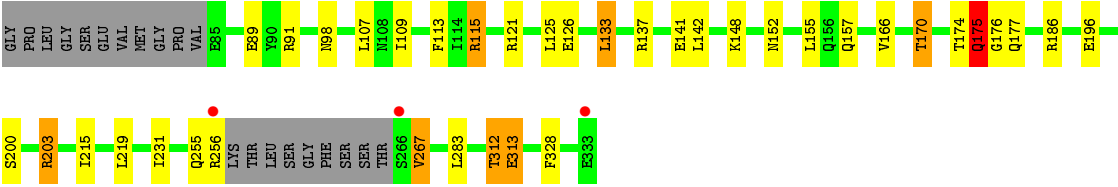
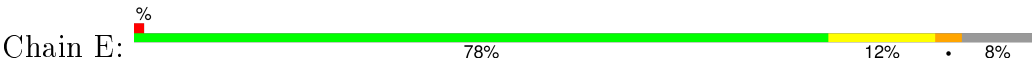


- Molecule 3: U4/U6 small nuclear ribonucleoprotein Prp31





● Molecule 3: U4/U6 small nuclear ribonucleoprotein Prp31



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	83.56 Å 99.90 Å 169.65 Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	30.00 – 2.60 29.99 – 2.60	Depositor EDS
% Data completeness (in resolution range)	99.7 (30.00-2.60) 99.7 (29.99-2.60)	Depositor EDS
$R_{merge}$	0.11	Depositor
$R_{sym}$	0.08	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.76 (at 2.61 Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, $R_{free}$	0.206 , 0.248 0.195 , 0.232	Depositor DCC
$R_{free}$ test set	2227 reflections (5.30%)	DCC
Wilson B-factor (Å <sup>2</sup> )	58.4	Xtriage
Anisotropy	0.175	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.31 , 60.6	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Outliers	1 of 44257 reflections (0.002%)	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	7375	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	53.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.29% 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

### 5.1 Standard geometry

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

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	C	0.93	0/786	1.67	15/1224 (1.2%)
1	F	1.06	1/786 (0.1%)	1.74	21/1224 (1.7%)
2	A	0.55	0/989	0.69	1/1335 (0.1%)
2	D	0.55	0/980	0.68	1/1323 (0.1%)
3	B	0.61	0/1922	0.69	0/2594
3	E	0.60	0/1912	0.70	0/2582
All	All	0.69	1/7375 (0.0%)	1.03	38/10282 (0.4%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	F	50	G	N9-C8	5.09	1.41	1.37

All (38) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	34	G	O4'-C1'-N9	14.26	119.61	108.20
1	C	34	G	N9-C1'-C2'	-12.66	97.54	114.00
1	F	34	G	O4'-C1'-N9	11.70	117.56	108.20
1	F	40	U	O4'-C1'-N1	10.33	116.46	108.20
1	F	38	U	O4'-C1'-N1	8.79	115.23	108.20
1	F	35	G	C4'-C3'-C2'	-8.11	94.49	102.60
1	F	36	U	O4'-C1'-N1	7.88	114.50	108.20
1	F	40	U	N1-C1'-C2'	-7.61	103.63	112.00
1	F	40	U	P-O3'-C3'	7.20	128.34	119.70
1	C	34	G	O4'-C1'-C2'	-7.17	98.63	105.80
1	C	32	G	O4'-C1'-N9	7.10	113.88	108.20
1	C	25	A	P-O3'-C3'	6.99	128.08	119.70
1	F	25	A	C3'-C2'-C1'	-6.89	95.98	101.50
1	F	28	C	O4'-C1'-N1	6.72	113.58	108.20

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	22	C	O4'-C1'-N1	6.65	113.52	108.20
1	F	35	G	C3'-C2'-C1'	-6.59	96.23	101.50
1	C	34	G	C5'-C4'-C3'	-6.54	105.53	116.00
1	C	24	U	P-O3'-C3'	6.47	127.46	119.70
1	F	43	G	O4'-C1'-N9	6.20	113.16	108.20
1	F	21	U	O4'-C1'-N1	5.87	112.89	108.20
1	C	43	G	O4'-C1'-N9	5.76	112.81	108.20
1	C	41	C	C1'-O4'-C4'	-5.72	105.33	109.90
1	F	23	G	O4'-C1'-N9	5.70	112.76	108.20
1	C	42	C	C4'-C3'-C2'	-5.57	97.03	102.60
1	F	52	U	C3'-C2'-C1'	5.56	105.95	101.50
1	C	39	A	N9-C1'-C2'	-5.55	105.89	112.00
1	F	44	A	P-O3'-C3'	5.54	126.35	119.70
1	C	40	U	N3-C4-C5	5.53	117.92	114.60
1	F	22	C	O4'-C1'-N1	5.44	112.55	108.20
1	C	28	C	C4'-C3'-C2'	-5.28	97.32	102.60
1	F	50	G	C8-N9-C4	-5.26	104.29	106.40
1	F	32	G	O4'-C1'-N9	5.25	112.40	108.20
1	F	36	U	C1'-O4'-C4'	-5.21	105.73	109.90
1	C	48	G	O4'-C1'-N9	5.15	112.32	108.20
1	F	35	G	C8-N9-C4	-5.13	104.35	106.40
2	A	22	LEU	CA-CB-CG	5.09	127.01	115.30
1	F	42	C	O4'-C4'-C3'	-5.08	98.92	104.00
2	D	81	VAL	CB-CA-C	-5.03	101.84	111.40

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	C	703	0	357	6	0
1	F	703	0	357	2	0
2	A	977	0	1023	10	0
2	D	968	0	1017	11	0
3	B	1896	0	1906	30	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	E	1886	0	1903	19	0
4	A	1	0	0	0	0
4	B	1	0	0	0	0
4	E	1	0	0	0	0
5	A	28	0	0	0	0
5	B	87	0	0	3	1
5	C	21	0	0	3	0
5	D	13	0	0	1	0
5	E	76	0	0	2	1
5	F	14	0	0	0	0
All	All	7375	0	6563	76	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 6.

All (76) 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:C:34:G:H1'	5:C:67:HOH:O	1.42	1.17
3:E:176:GLY:HA2	3:E:177:GLN:HB2	1.29	1.12
2:A:5:ASP:HB2	2:A:6:VAL:HG23	1.22	1.11
3:E:174:THR:O	3:E:175:GLN:HB2	1.54	1.05
3:E:255:GLN:HG3	3:E:267:VAL:HG21	1.43	0.96
2:A:5:ASP:CB	2:A:6:VAL:HG23	2.12	0.73
3:B:160:THR:HB	3:B:163:THR:HG23	1.69	0.73
3:B:160:THR:HB	3:B:163:THR:CG2	2.20	0.71
3:E:152:ASN:HB3	3:E:155:LEU:HB2	1.73	0.70
3:E:113:PHE:HZ	3:E:186:ARG:HD3	1.57	0.70
3:E:176:GLY:CA	3:E:177:GLN:HB2	2.16	0.66
1:C:34:G:C1'	5:C:67:HOH:O	2.16	0.66
3:E:196:GLU:HG3	5:E:525:HOH:O	1.96	0.65
3:B:93:ILE:HG21	3:B:212:MET:HE2	1.78	0.65
5:B:507[B]:HOH:O	3:E:200:SER:HB2	1.97	0.64
3:E:312:THR:HG22	3:E:313:GLU:OE2	1.98	0.64
1:C:37:U:H5	5:C:53:HOH:O	1.81	0.64
2:D:111:GLN:HB3	5:D:135:HOH:O	1.97	0.63
3:B:115:ARG:HD3	3:B:126:GLU:HG3	1.80	0.62
2:A:3:GLU:N	2:A:4:ALA:HA	2.15	0.62
3:B:171:ALA:O	3:B:174:THR:HB	1.99	0.62
2:A:18:LEU:HG	2:A:22:LEU:HD22	1.83	0.59
2:D:127:LEU:O	2:D:128:VAL:HB	2.03	0.58

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:B:92:VAL:HG22	3:E:109:ILE:HG22	1.86	0.57
3:B:93:ILE:HG21	3:B:212:MET:CE	2.34	0.57
2:D:21:LYS:HB3	2:D:119:ILE:HD11	1.86	0.57
3:B:174:THR:O	3:B:175:GLN:HB2	2.05	0.56
3:B:196:GLU:HG2	3:E:203:ARG:HD3	1.86	0.56
2:D:120:GLN:O	2:D:124:GLU:HG2	2.07	0.54
2:A:72:LEU:O	2:A:75:ASP:HB2	2.08	0.53
3:B:332:GLN:O	3:B:333:GLU:CD	2.46	0.53
3:B:175:GLN:HA	3:B:175:GLN:HE21	1.71	0.53
2:A:25:LEU:HD23	2:A:119:ILE:HG13	1.91	0.52
1:C:25:A:O2'	1:C:26:G:OP1	2.25	0.52
3:B:160:THR:HG23	5:B:543[A]:HOH:O	2.10	0.52
3:B:181:GLU:O	3:B:185:GLU:HG3	2.10	0.52
2:D:86:LYS:HG3	2:D:98:PRO:HB3	1.91	0.51
2:A:126:LEU:C	2:A:128:VAL:H	2.14	0.51
3:B:321:LYS:O	3:B:325:GLU:HG2	2.12	0.50
3:B:113:PHE:HZ	3:B:186:ARG:HD2	1.77	0.50
1:F:37:U:O2	1:F:37:U:H2'	2.12	0.50
2:A:4:ALA:O	2:A:5:ASP:O	2.29	0.49
3:B:163:THR:O	3:B:167:VAL:HG13	2.13	0.48
3:E:113:PHE:CZ	3:E:186:ARG:HD3	2.42	0.48
2:D:56:MET:HB3	2:D:66:ILE:HD12	1.96	0.48
3:B:133:LEU:HD13	3:B:137:ARG:NH1	2.29	0.47
3:E:176:GLY:HA2	3:E:177:GLN:CB	2.17	0.46
3:B:110:ILE:HG21	3:B:194:ALA:HB2	1.97	0.46
3:E:255:GLN:HG3	3:E:267:VAL:CG2	2.31	0.46
3:B:120:LYS:HD3	3:B:179:LEU:HD23	1.97	0.46
3:B:290:LYS:HE2	3:B:331:TRP:CZ2	2.49	0.46
3:E:166:VAL:O	3:E:170:THR:CG2	2.64	0.46
3:B:93:ILE:CG2	3:B:212:MET:CE	2.93	0.46
2:D:73:CYS:HB3	2:D:78:VAL:O	2.15	0.45
3:B:230:LYS:O	3:B:234:VAL:HG12	2.17	0.45
3:E:89:GLU:HG2	3:E:215:ILE:HD11	1.98	0.45
2:A:67:LEU:O	2:A:70:PRO:HD2	2.17	0.44
1:C:25:A:OP1	1:C:25:A:H8	2.01	0.44
3:B:219:LEU:HD22	3:B:223:ILE:HD11	1.98	0.44
1:C:23:G:C6	1:C:24:U:C2	3.06	0.44
1:F:25:A:H2'	1:F:26:G:OP1	2.17	0.43
3:B:160:THR:HG22	3:B:162:ALA:N	2.34	0.43
2:D:41:GLU:O	2:D:45:THR:HG23	2.19	0.42
3:E:133:LEU:O	3:E:137:ARG:HG3	2.19	0.42

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:B:174:THR:O	3:B:175:GLN:CB	2.68	0.41
3:B:90:TYR:O	3:B:94:VAL:HG23	2.20	0.41
2:D:11:TYR:HA	2:D:12:PRO:C	2.41	0.41
2:D:64:GLU:H	2:D:64:GLU:CD	2.24	0.41
3:B:93:ILE:CG2	3:B:212:MET:HE3	2.50	0.41
3:B:125:LEU:HD13	3:B:135:TYR:CE1	2.55	0.41
3:E:283:LEU:HD11	3:E:328:PHE:CE2	2.56	0.41
2:A:113:LYS:HA	2:A:113:LYS:HD3	1.84	0.41
3:B:98:ASN:ND2	5:B:575[B]:HOH:O	2.53	0.40
2:D:61:GLU:HA	2:D:62:PRO:HA	1.84	0.40
3:E:115:ARG:NH1	5:E:543:HOH:O	2.38	0.40
3:B:160:THR:HG22	3:B:162:ALA:H	1.85	0.40

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
5:B:503:HOH:O	5:E:502:HOH:O[3_555]	2.14	0.06

## 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	
2	A	124/130 (95%)	119 (96%)	4 (3%)	1 (1%)	24	46
2	D	123/130 (95%)	121 (98%)	2 (2%)	0	100	100
3	B	236/260 (91%)	229 (97%)	4 (2%)	3 (1%)	15	30
3	E	236/260 (91%)	227 (96%)	7 (3%)	2 (1%)	24	46
All	All	719/780 (92%)	696 (97%)	17 (2%)	6 (1%)	24	46

All (6) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	A	5	ASP
3	B	267	VAL
3	B	86	ALA
3	B	175	GLN
3	E	148	LYS
3	E	175	GLN

### 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	
2	A	109/112 (97%)	100 (92%)	9 (8%)	14	27
2	D	108/112 (96%)	92 (85%)	16 (15%)	4	6
3	B	209/224 (93%)	185 (88%)	24 (12%)	7	12
3	E	208/224 (93%)	188 (90%)	20 (10%)	10	20
All	All	634/672 (94%)	565 (89%)	69 (11%)	8	14

All (69) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	A	22	LEU
2	A	63	LEU
2	A	75	ASP
2	A	81	VAL
2	A	95	VAL
2	A	111	GLN
2	A	112	LEU
2	A	117	GLN
2	A	125	ARG
3	B	98	ASN
3	B	107	LEU
3	B	115	ARG
3	B	121	ARG
3	B	122	PHE
3	B	125	LEU
3	B	142	LEU

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
3	B	150	LYS
3	B	153	GLU
3	B	155	LEU
3	B	163	THR
3	B	166	VAL
3	B	168	SER
3	B	173	THR
3	B	175	GLN
3	B	186	ARG
3	B	195	LEU
3	B	203[A]	ARG
3	B	203[B]	ARG
3	B	213	SER
3	B	219	LEU
3	B	234	VAL
3	B	242	SER
3	B	281	GLN
2	D	5	ASP
2	D	21	LYS
2	D	22	LEU
2	D	63	LEU
2	D	75	ASP
2	D	77	ASN
2	D	81	VAL
2	D	95	VAL
2	D	96	SER
2	D	105	THR
2	D	111	GLN
2	D	112	LEU
2	D	113	LYS
2	D	115	GLN
2	D	124	GLU
2	D	128	VAL
3	E	91	ARG
3	E	98	ASN
3	E	107	LEU
3	E	115	ARG
3	E	121	ARG
3	E	125	LEU
3	E	126	GLU
3	E	133	LEU
3	E	141	GLU

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
3	E	142	LEU
3	E	157	GLN
3	E	170	THR
3	E	175	GLN
3	E	203	ARG
3	E	219	LEU
3	E	231	ILE
3	E	256	ARG
3	E	267	VAL
3	E	312	THR
3	E	313	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (15) such sidechains are listed below:

Mol	Chain	Res	Type
2	A	31	ASN
2	A	111	GLN
3	B	175	GLN
3	B	281	GLN
2	D	17	HIS
2	D	28	GLN
2	D	31	ASN
2	D	77	ASN
2	D	111	GLN
2	D	115	GLN
3	E	105	ASN
3	E	157	GLN
3	E	178	GLN
3	E	281	GLN
3	E	332	GLN

### 5.3.3 RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
1	C	32/33 (96%)	7 (21%)	4 (12%)
1	F	32/33 (96%)	8 (25%)	1 (3%)
All	All	64/66 (96%)	15 (23%)	5 (7%)

All (15) RNA backbone outliers are listed below:



Mol	Chain	Res	Type
1	C	24	U
1	C	25	A
1	C	26	G
1	C	35	G
1	C	36	U
1	C	39	A
1	C	40	U
1	F	25	A
1	F	26	G
1	F	35	G
1	F	36	U
1	F	37	U
1	F	38	U
1	F	41	C
1	F	52	U

All (5) RNA pucker outliers are listed below:

Mol	Chain	Res	Type
1	C	24	U
1	C	25	A
1	C	34	G
1	C	38	U
1	F	35	G

## 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 3 ligands modelled in this entry, 3 are 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 ⓘ

### 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	C	33/33 (100%)	0.52	2 (6%) 25 18	46, 59, 100, 107	0
1	F	33/33 (100%)	0.76	5 (15%) 3 1	45, 61, 98, 105	0
2	A	126/130 (96%)	0.10	5 (3%) 42 34	43, 49, 63, 74	0
2	D	125/130 (96%)	0.17	6 (4%) 34 27	48, 53, 60, 70	0
3	B	239/260 (91%)	0.05	9 (3%) 44 36	36, 48, 65, 71	0
3	E	240/260 (92%)	-0.05	3 (1%) 79 75	34, 47, 70, 76	0
All	All	796/846 (94%)	0.10	30 (3%) 44 36	34, 50, 72, 107	0

All (30) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	A	128	VAL	7.9
2	D	4	ALA	4.4
2	A	4	ALA	4.2
3	E	266	SER	4.1
3	B	266	SER	3.8
1	C	25	A	3.6
3	E	256	ARG	3.5
1	F	36	U	3.4
1	F	25	A	3.3
3	B	300	THR	3.1
3	B	246	ALA	3.1
2	D	5	ASP	2.9
3	B	165	MET	2.8
1	F	37	U	2.7
3	B	245	PRO	2.6
1	F	23	G	2.6
1	C	24	U	2.5
2	D	102	CYS	2.4
2	D	75	ASP	2.4

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
3	E	333	GLU	2.3
2	D	40	ASN	2.3
3	B	247	CYS	2.2
3	B	267	VAL	2.2
2	A	38	GLY	2.2
3	B	299	CYS	2.2
2	A	40	ASN	2.2
2	A	108	GLU	2.1
2	D	101	ALA	2.1
1	F	24	U	2.0
3	B	255	GLN	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(Å <sup>2</sup> )	Q<0.9
4	CA	B	502	1/1	0.80	0.09	-	108,108,108,108	0
4	CA	A	503	1/1	0.80	0.07	-	106,106,106,106	0
4	CA	E	501	1/1	0.98	0.08	-	57,57,57,57	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.