



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 05:58 PM GMT

PDB ID : 4JV5  
Title : Crystal structures of pseudouridinilated stop codons with ASLs  
Authors : Fernandez, I.S.; Ng, C.L.; Kelley, A.C.; Guowei, W.; Yu, Y.T.; Ramakrishnan, V.  
Deposited on : 2013-03-25  
Resolution : 3.16 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

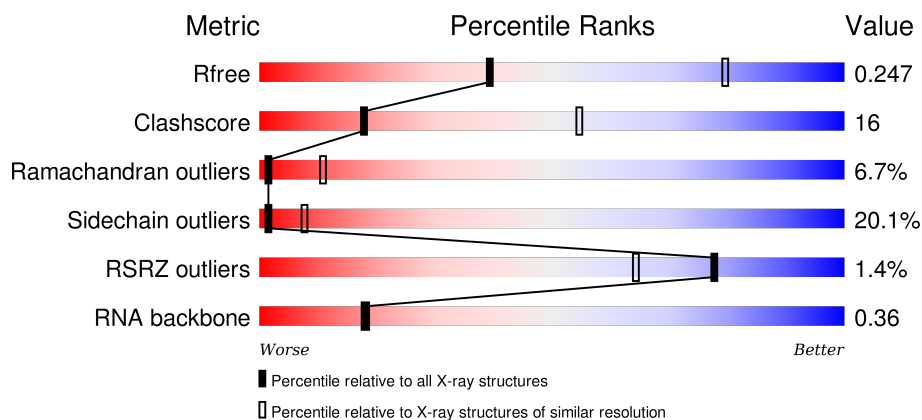
# 1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 3.16 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	1112 (3.20-3.12)
Clashscore	102246	1249 (3.20-3.12)
Ramachandran outliers	100387	1222 (3.20-3.12)
Sidechain outliers	100360	1221 (3.20-3.12)
RSRZ outliers	91569	1117 (3.20-3.12)
RNA backbone	2183	1046 (3.62-2.70)

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	1517	<div> <div> <div>3%</div> <div>43%</div> <div>40%</div> <div>15%</div> <div>•</div> </div> </div>
2	B	234	<div> <div> <div>3%</div> <div>45%</div> <div>39%</div> <div>14%</div> <div>•</div> </div> </div>
3	C	206	<div> <div> <div>•</div> <div>45%</div> <div>43%</div> <div>11%</div> <div>•</div> </div> </div>
4	D	208	<div> <div> <div>2%</div> <div>50%</div> <div>34%</div> <div>13%</div> <div>•</div> </div> </div>

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
5	E	150	
6	F	101	
7	G	155	
8	H	138	
9	I	127	
10	J	98	
11	K	119	
12	L	125	
13	M	120	
14	N	60	
15	O	88	
16	P	83	
17	Q	99	
18	R	70	
19	S	78	
20	T	99	
21	U	24	
22	X	5	
23	Y	10	

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
24	MG	A	1607	-	-	-	X
24	MG	A	1608	-	-	-	X
24	MG	A	1610	-	-	-	X
24	MG	A	1614	-	-	-	X

## 2 Entry composition

There are 25 unique types of molecules in this entry. The entry contains 51923 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 16S ribosomal RNA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	1511	Total	C	N	O	P	0	0	0
			32468	14453	6008	10497	1510			

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	80	A	G	CONFLICT	GB 55771382

- Molecule 2 is a protein called 30S ribosomal protein S2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	234	Total	C	N	O	S	0	0	0
			1900	1213	341	341	5			

- Molecule 3 is a protein called 30S ribosomal protein S3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	206	Total	C	N	O	S	0	0	0
			1612	1016	314	281	1			

- Molecule 4 is a protein called 30S ribosomal protein S4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	D	208	Total	C	N	O	S	0	0	0
			1703	1066	339	291	7			

- Molecule 5 is a protein called 30S ribosomal protein S5.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	E	150	Total	C	N	O	S	0	0	0
			1146	724	217	201	4			

- Molecule 6 is a protein called 30S ribosomal protein S6.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
6	F	101	Total	C	N	O	S	0	0	0
			843	531	155	154	3			

- Molecule 7 is a protein called 30S ribosomal protein S7.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
7	G	155	Total	C	N	O	S	0	0	0
			1257	781	252	218	6			

- Molecule 8 is a protein called 30S ribosomal protein S8.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
8	H	138	Total	C	N	O	S	0	0	0
			1116	705	215	193	3			

- Molecule 9 is a protein called 30S ribosomal protein S9.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
9	I	127	Total	C	N	O	0	0	0
			1011	639	198	174			

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
I	58	ARG	HIS	CONFLICT	UNP P80374

- Molecule 10 is a protein called 30S ribosomal protein S10.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
10	J	98	Total	C	N	O	S	0	0	0
			794	499	156	138	1			

- Molecule 11 is a protein called 30S ribosomal protein S11.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
11	K	119	Total	C	N	O	S	0	0	0
			885	549	168	165	3			

- Molecule 12 is a protein called 30S ribosomal protein S12.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
12	L	125	Total	C	N	O	S	0	0	0
			975	614	196	164	1			

- Molecule 13 is a protein called 30S ribosomal protein S13.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
13	M	120	Total	C	N	O	S	0	0	0
			955	591	197	165	2			

- Molecule 14 is a protein called 30S ribosomal protein S14.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
14	N	60	Total	C	N	O	S	0	0	0
			492	312	104	72	4			

- Molecule 15 is a protein called 30S ribosomal protein S15.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
15	O	88	Total	C	N	O	S	0	0	0
			734	459	147	126	2			

- Molecule 16 is a protein called 30S ribosomal protein S16.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
16	P	83	Total	C	N	O	S	0	0	0
			700	443	139	117	1			

- Molecule 17 is a protein called 30S ribosomal protein S17.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
17	Q	99	Total	C	N	O	S	0	0	0
			823	528	151	142	2			

- Molecule 18 is a protein called 30S ribosomal protein S18.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
18	R	70	Total	C	N	O	0	0	0
			574	367	112	95			

- Molecule 19 is a protein called 30S ribosomal protein S19.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
19	S	78	Total	C	N	O	S	0	0	0
			629	403	114	110	2			

- Molecule 20 is a protein called 30S ribosomal protein 20.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
20	T	99	Total	C	N	O	S	0	0	0
			763	470	162	129	2			

- Molecule 21 is a protein called 30S ribosomal protein THX.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
21	U	24	Total	C	N	O	0	0	0
			208	128	50	30			

- Molecule 22 is a RNA chain called mRNA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
22	X	5	Total	C	N	O	P	0	0	0
			104	48	19	33	4			

- Molecule 23 is a RNA chain called ASL-tRNA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
23	Y	10	Total	C	N	O	P	0	0	0
			213	96	38	69	10			

- Molecule 24 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
24	X	1	Total	Mg	0	0
			1	1		
24	A	15	Total	Mg	0	0
			15	15		
24	Y	1	Total	Mg	0	0
			1	1		

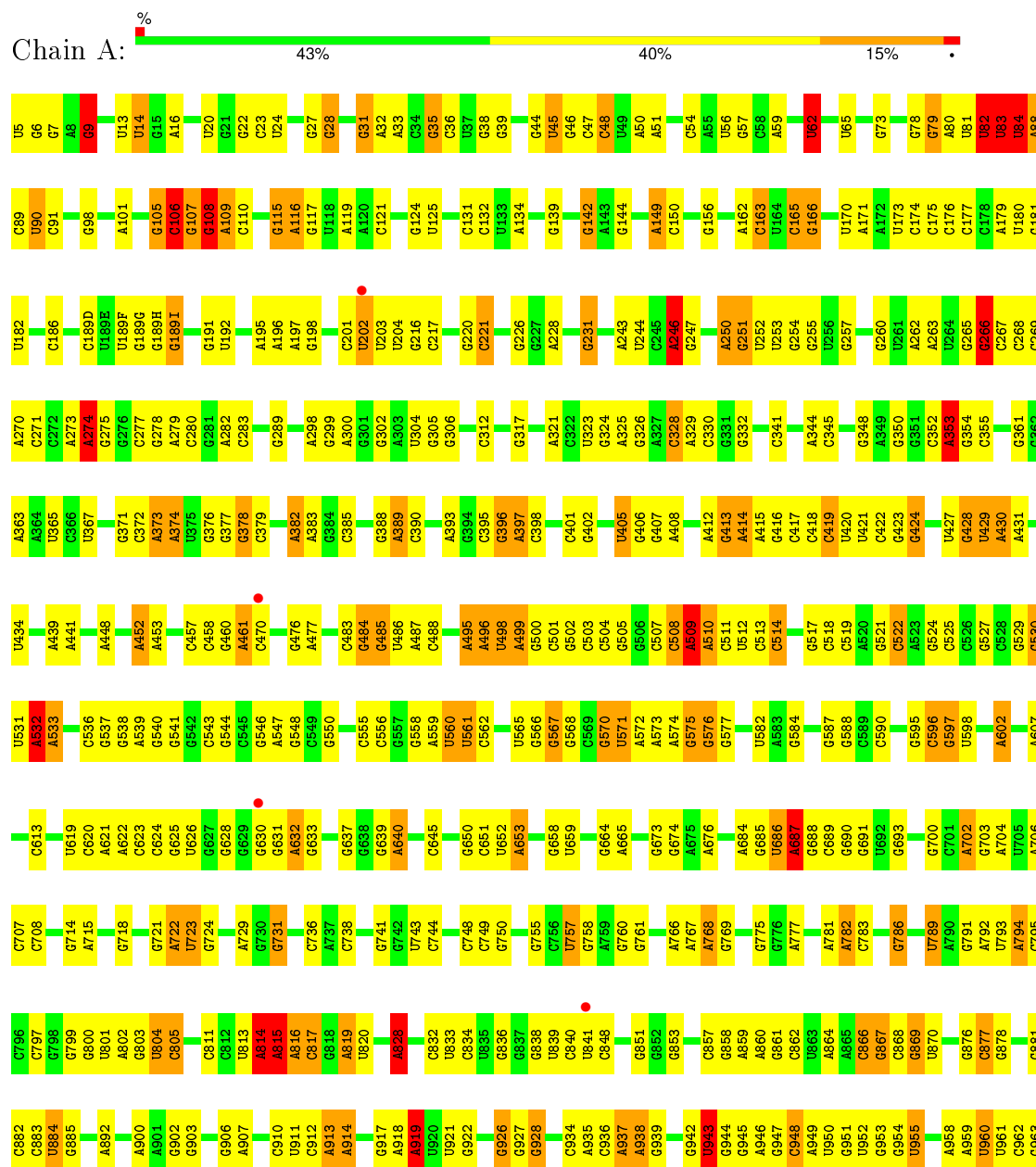
- Molecule 25 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
25	D	1	Total	Zn	0	0
			1	1		

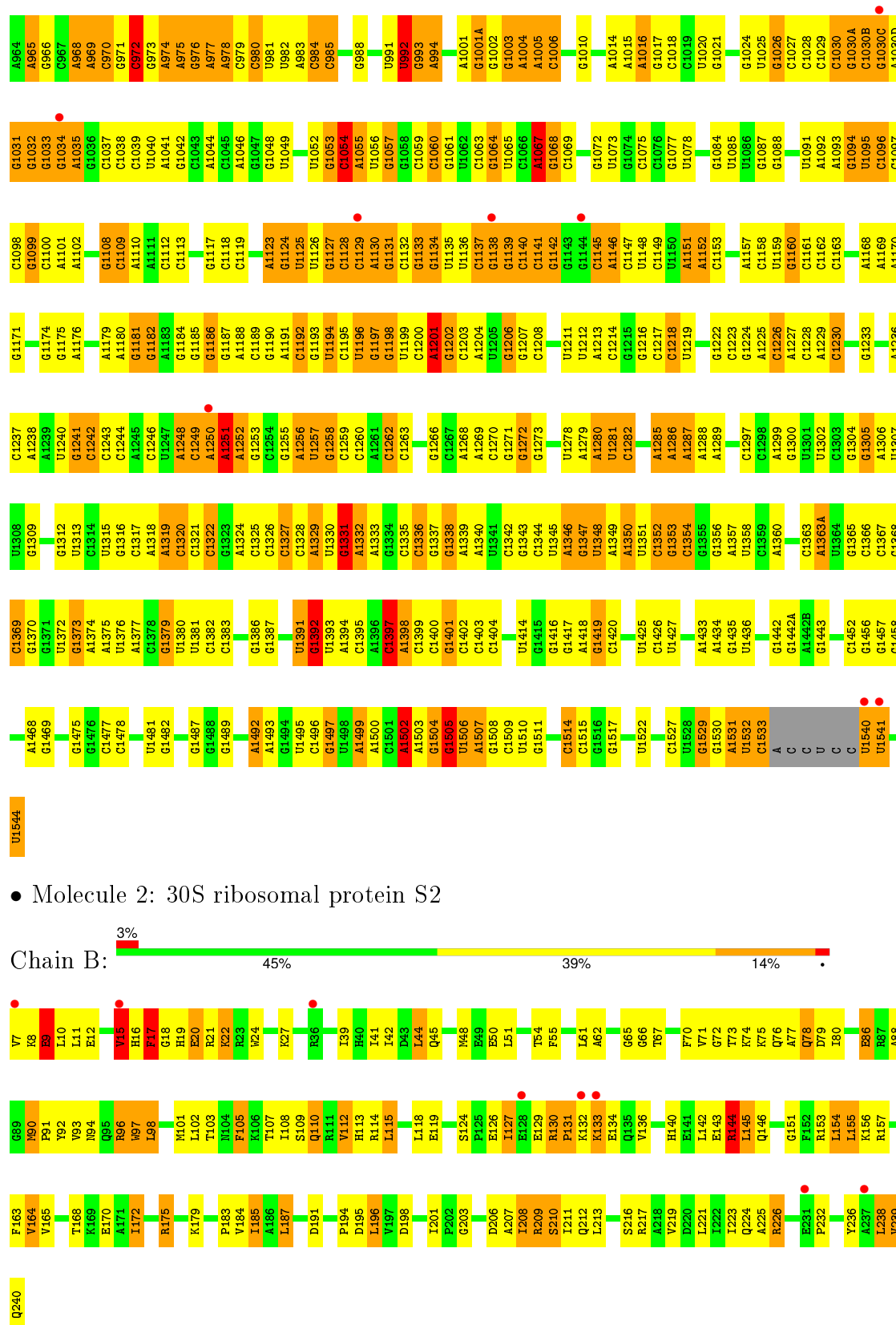
### 3 Residue-property plots

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 ( $\text{RSRZ} > 2$ ). 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.

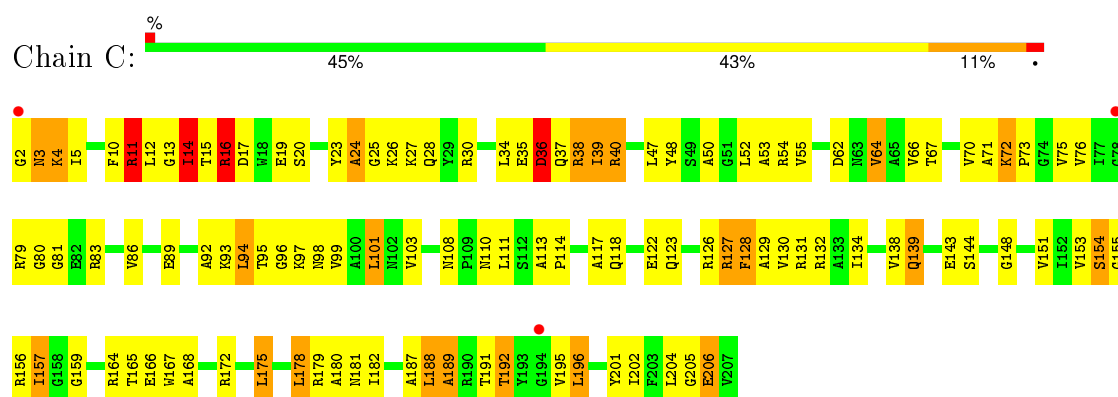
- Molecule 1: 16S ribosomal RNA



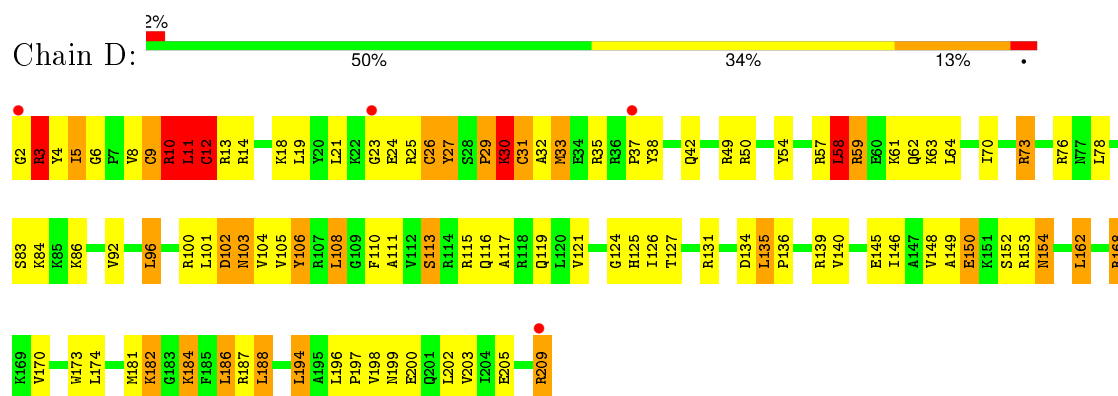




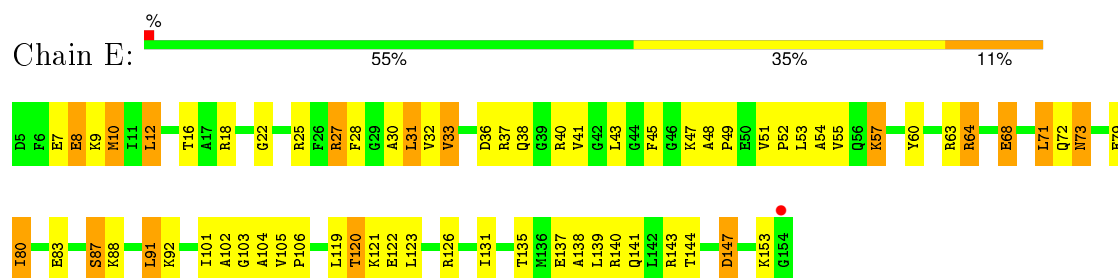
• Molecule 3: 30S ribosomal protein S3



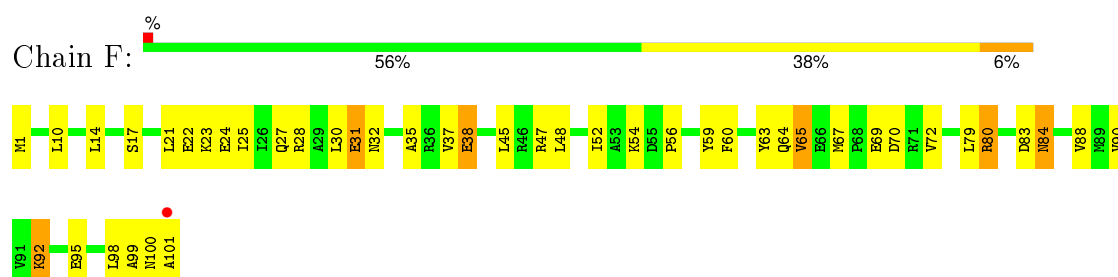
• Molecule 4: 30S ribosomal protein S4



• Molecule 5: 30S ribosomal protein S5

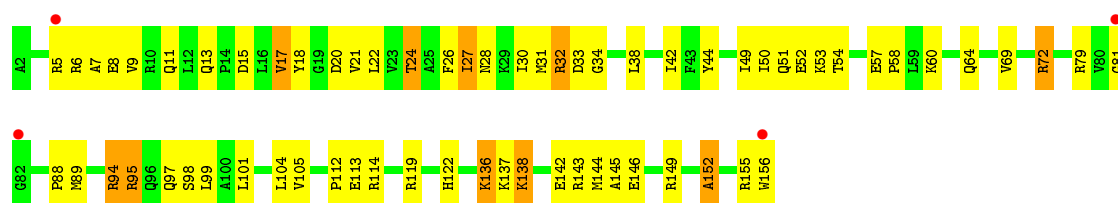


• Molecule 6: 30S ribosomal protein S6

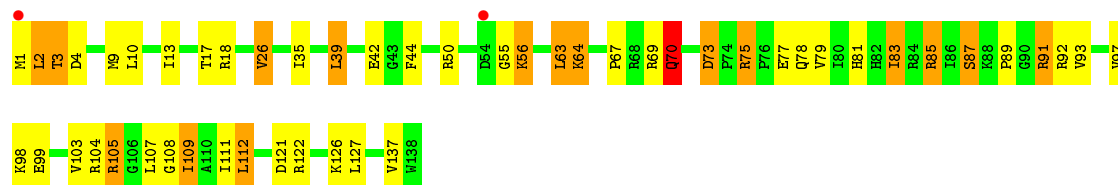


• Molecule 7: 30S ribosomal protein S7

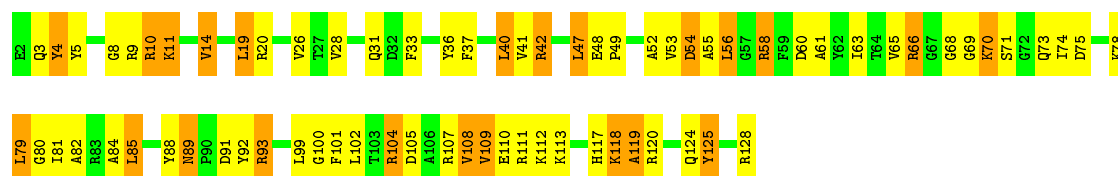




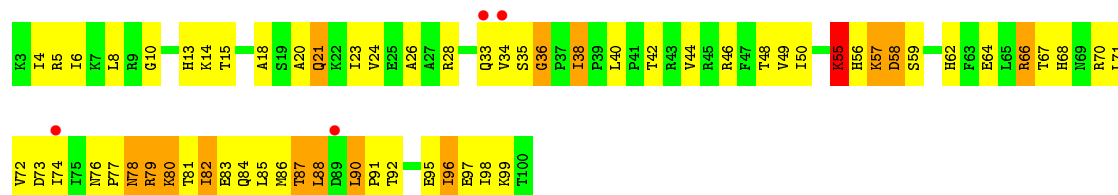
• Molecule 8: 30S ribosomal protein S8



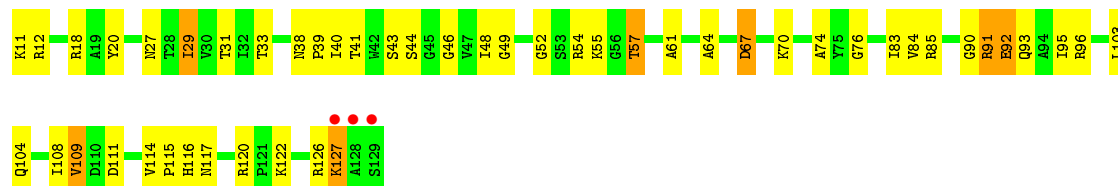
• Molecule 9: 30S ribosomal protein S9



• Molecule 10: 30S ribosomal protein S10

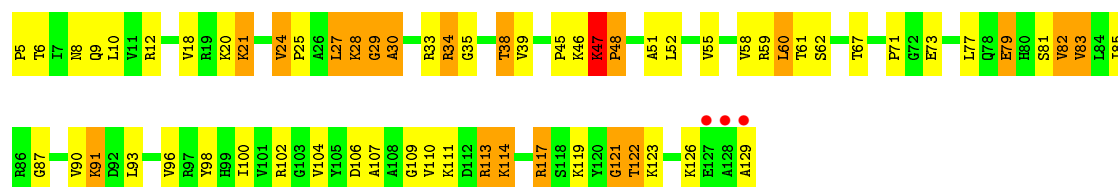


• Molecule 11: 30S ribosomal protein S11

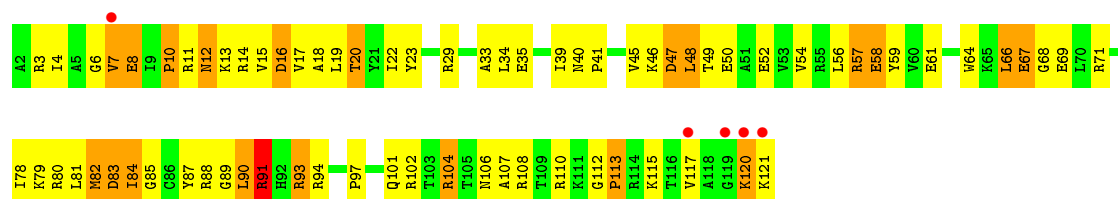
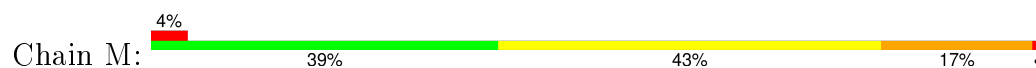


• Molecule 12: 30S ribosomal protein S12

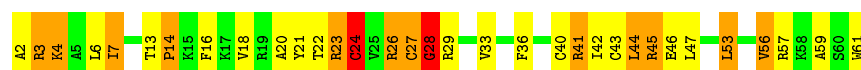




• Molecule 13: 30S ribosomal protein S13



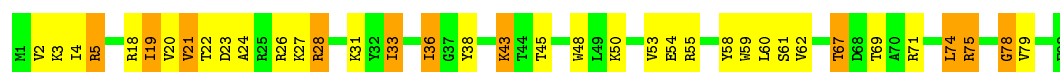
• Molecule 14: 30S ribosomal protein S14



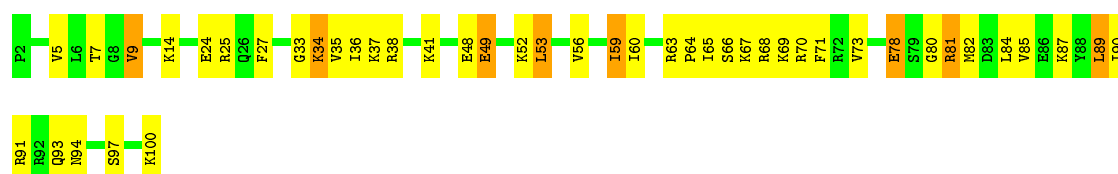
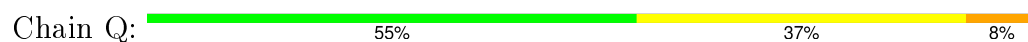
• Molecule 15: 30S ribosomal protein S15



• Molecule 16: 30S ribosomal protein S16

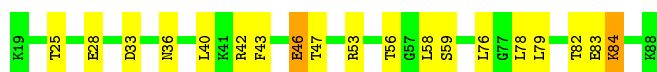


• Molecule 17: 30S ribosomal protein S17

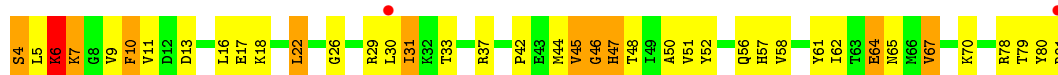


• Molecule 18: 30S ribosomal protein S18





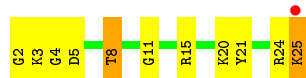
- Molecule 19: 30S ribosomal protein S19



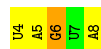
- Molecule 20: 30S ribosomal protein 20



- Molecule 21: 30S ribosomal protein THX



- Molecule 22: mRNA



- Molecule 23: ASL-tRNA



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 41 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	401.00Å 401.00Å 176.00Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	29.88 – 3.16 29.88 – 3.16	Depositor EDS
% Data completeness (in resolution range)	95.4 (29.88-3.16) 95.6 (29.88-3.16)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.67 (at 3.18Å)	Xtriage
Refinement program	REFMAC 5.8.0016	Depositor
R, $R_{free}$	0.197 , 0.249 0.199 , 0.247	Depositor DCC
$R_{free}$ test set	11558 reflections (5.26%)	DCC
Wilson B-factor (Å <sup>2</sup> )	81.3	Xtriage
Anisotropy	0.028	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.28 , 68.4	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.46$ , $\langle L^2 \rangle = 0.29$	Xtriage
Outliers	0 of 231145 reflections	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	51923	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	87.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 1.85% 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: ZN, MG, PSU

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	0.50	7/36342 (0.0%)	0.92	105/56718 (0.2%)
2	B	0.55	0/1935	0.86	2/2609 (0.1%)
3	C	0.56	0/1636	0.86	4/2205 (0.2%)
4	D	0.60	1/1733 (0.1%)	0.96	9/2318 (0.4%)
5	E	0.67	0/1162	0.91	2/1564 (0.1%)
6	F	0.45	0/856	0.76	0/1154
7	G	0.47	0/1276	0.79	0/1709
8	H	0.65	0/1136	0.92	0/1527
9	I	0.52	0/1029	0.81	0/1378
10	J	0.55	0/807	0.79	0/1085
11	K	0.55	0/900	0.84	0/1213
12	L	0.61	0/991	0.92	0/1327
13	M	0.57	0/965	0.89	1/1292 (0.1%)
14	N	0.70	0/501	1.11	5/664 (0.8%)
15	O	0.54	0/745	0.83	0/992
16	P	0.60	0/716	0.95	2/963 (0.2%)
17	Q	0.66	0/836	0.95	1/1117 (0.1%)
18	R	0.52	0/579	0.80	0/768
19	S	0.58	0/642	0.84	0/865
20	T	0.57	0/765	0.87	0/1007
21	U	0.62	0/212	0.91	0/277
22	X	0.44	0/98	0.97	0/153
23	Y	1.01	2/237 (0.8%)	0.94	1/364 (0.3%)
All	All	0.53	10/56099 (0.0%)	0.90	132/83269 (0.2%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
2	B	0	1
3	C	0	2
4	D	0	2
9	I	0	2
11	K	0	1
14	N	0	1
20	T	0	5
All	All	0	14

The worst 5 of 10 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
23	Y	34	I	C4'-O4'	13.09	1.67	1.41
1	A	82	U	O3'-P	9.84	1.73	1.61
1	A	83	U	O3'-P	8.17	1.71	1.61
1	A	814	A	O3'-P	6.26	1.68	1.61
1	A	766	A	P-OP2	5.96	1.59	1.49

The worst 5 of 132 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	574	A	O5'-P-OP1	-15.93	91.36	105.70
1	A	1198	G	O5'-P-OP2	11.08	124.00	110.70
1	A	266	G	C2'-C3'-O3'	10.13	131.79	109.50
1	A	574	A	O5'-P-OP2	9.70	122.34	110.70
1	A	1078	U	O5'-P-OP1	-9.63	97.03	105.70

There are no chirality outliers.

5 of 14 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	B	191	ASP	Peptide
3	C	14	ILE	Peptide
3	C	25	GLY	Peptide
4	D	11	LEU	Peptide
4	D	30	LYS	Peptide

## 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	32468	0	16391	666	1
2	B	1900	0	1951	85	0
3	C	1612	0	1677	98	0
4	D	1703	0	1763	75	0
5	E	1146	0	1207	52	0
6	F	843	0	857	26	0
7	G	1257	0	1296	27	0
8	H	1116	0	1177	43	0
9	I	1011	0	1043	59	0
10	J	794	0	840	43	0
11	K	885	0	904	26	0
12	L	975	0	1062	42	0
13	M	955	0	1021	44	0
14	N	492	0	529	44	0
15	O	734	0	771	21	0
16	P	700	0	720	21	0
17	Q	823	0	891	35	0
18	R	574	0	644	11	0
19	S	629	0	652	25	0
20	T	763	0	861	30	0
21	U	208	0	221	7	0
22	X	104	0	54	8	0
23	Y	213	0	107	19	0
24	A	15	0	0	0	0
24	X	1	0	0	0	0
24	Y	1	0	0	0	0
25	D	1	0	0	0	0
All	All	51923	0	36639	1390	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 16.

The worst 5 of 1390 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)
23:Y:34:I:O4'	23:Y:34:I:C4'	1.67	1.36
23:Y:34:I:C5'	23:Y:34:I:H8	1.53	1.20
4:D:31:CYS:O	4:D:33:MET:N	1.77	1.17
23:Y:34:I:C8	23:Y:34:I:H5'	1.82	1.15
3:C:154:SER:OG	3:C:155:GLY:N	1.74	1.12

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)
1:A:79:G:O2'	1:A:1340:A:O2'[3_445]	2.15	0.05

## 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	B	232/234 (99%)	167 (72%)	51 (22%)	14 (6%)	2	14
3	C	204/206 (99%)	139 (68%)	45 (22%)	20 (10%)	1	4
4	D	206/208 (99%)	164 (80%)	26 (13%)	16 (8%)	1	7
5	E	148/150 (99%)	135 (91%)	9 (6%)	4 (3%)	6	37
6	F	99/101 (98%)	83 (84%)	13 (13%)	3 (3%)	5	34
7	G	153/155 (99%)	113 (74%)	29 (19%)	11 (7%)	1	9
8	H	136/138 (99%)	117 (86%)	14 (10%)	5 (4%)	4	27
9	I	125/127 (98%)	86 (69%)	29 (23%)	10 (8%)	1	7
10	J	96/98 (98%)	78 (81%)	11 (12%)	7 (7%)	1	9
11	K	117/119 (98%)	95 (81%)	17 (14%)	5 (4%)	3	23
12	L	123/125 (98%)	96 (78%)	14 (11%)	13 (11%)	0	3
13	M	118/120 (98%)	85 (72%)	22 (19%)	11 (9%)	1	4
14	N	58/60 (97%)	31 (53%)	18 (31%)	9 (16%)	0	1
15	O	86/88 (98%)	75 (87%)	9 (10%)	2 (2%)	8	42
16	P	81/83 (98%)	68 (84%)	12 (15%)	1 (1%)	16	58
17	Q	97/99 (98%)	86 (89%)	8 (8%)	3 (3%)	5	32
18	R	68/70 (97%)	63 (93%)	4 (6%)	1 (2%)	13	53
19	S	76/78 (97%)	53 (70%)	16 (21%)	7 (9%)	1	4
20	T	97/99 (98%)	66 (68%)	16 (16%)	15 (16%)	0	1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
21	U	22/24 (92%)	19 (86%)	2 (9%)	1 (4%)	3	22
All	All	2342/2382 (98%)	1819 (78%)	365 (16%)	158 (7%)	1	11

5 of 158 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	17	PHE
2	B	20	GLU
2	B	78	GLN
2	B	130	ARG
2	B	226	ARG

### 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	B	202/202 (100%)	153 (76%)	49 (24%)	1	3
3	C	160/160 (100%)	126 (79%)	34 (21%)	1	6
4	D	180/180 (100%)	141 (78%)	39 (22%)	1	6
5	E	115/115 (100%)	94 (82%)	21 (18%)	2	10
6	F	90/90 (100%)	76 (84%)	14 (16%)	3	14
7	G	126/126 (100%)	101 (80%)	25 (20%)	1	8
8	H	119/119 (100%)	97 (82%)	22 (18%)	2	10
9	I	98/98 (100%)	78 (80%)	20 (20%)	1	7
10	J	88/88 (100%)	66 (75%)	22 (25%)	1	3
11	K	90/90 (100%)	73 (81%)	17 (19%)	2	10
12	L	104/104 (100%)	81 (78%)	23 (22%)	1	6
13	M	96/96 (100%)	64 (67%)	32 (33%)	0	0
14	N	49/49 (100%)	41 (84%)	8 (16%)	3	13
15	O	79/79 (100%)	72 (91%)	7 (9%)	12	42
16	P	72/72 (100%)	57 (79%)	15 (21%)	1	7

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
17	Q	94/94 (100%)	81 (86%)	13 (14%)	4	20
18	R	61/61 (100%)	53 (87%)	8 (13%)	5	23
19	S	69/69 (100%)	54 (78%)	15 (22%)	1	6
20	T	76/76 (100%)	64 (84%)	12 (16%)	3	14
21	U	19/19 (100%)	16 (84%)	3 (16%)	3	14
All	All	1987/1987 (100%)	1588 (80%)	399 (20%)	1	8

5 of 399 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
7	G	149	ARG
9	I	128	ARG
18	R	84	LYS
8	H	39	LEU
8	H	126	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 17 such sidechains are listed below:

Mol	Chain	Res	Type
4	D	42	GLN
6	F	7	ASN
13	M	12	ASN
3	C	123	GLN
14	N	49	HIS

### 5.3.3 RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
1	A	1510/1517 (99%)	437 (28%)	98 (6%)
22	X	4/5 (80%)	2 (50%)	0
23	Y	9/10 (90%)	3 (33%)	1 (11%)
All	All	1523/1532 (99%)	442 (29%)	99 (6%)

5 of 442 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
1	A	9	G
1	A	14	U

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	20	U
1	A	22	G
1	A	28	G

5 of 99 RNA pucker outliers are listed below:

Mol	Chain	Res	Type
1	A	723	U
1	A	962	C
1	A	1380	U
1	A	760	G
1	A	819	A

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

1 non-standard protein/DNA/RNA residue 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
22	PSU	X	4	22,23	14,18,22	2.22	2 (14%)	18,26,33	4.23	12 (66%)

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
22	PSU	X	4	22,23	-	0/6/22/26	0/2/2/2

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
22	X	4	PSU	C5-C1'	-6.59	1.46	1.52
22	X	4	PSU	C2'-C1'	-3.98	1.50	1.53

The worst 5 of 12 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed( $^{\circ}$ )	Ideal( $^{\circ}$ )
22	X	4	PSU	N1-C2-N3	-10.36	121.72	128.33
22	X	4	PSU	C5-C1'-C2'	-5.50	105.75	115.52
22	X	4	PSU	C5-C6-N1	-3.68	119.20	124.39
22	X	4	PSU	O4'-C4'-C3'	-2.88	99.34	105.15
22	X	4	PSU	C3'-C2'-C1'	-2.11	99.35	101.79

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
22	X	4	PSU	3	0

## 5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

Of 18 ligands modelled in this entry, 18 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 [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	1511/1517 (99%)	-0.30	12 (0%) 87 79	37, 75, 144, 237	0
2	B	234/234 (100%)	-0.09	8 (3%) 49 32	55, 108, 168, 190	0
3	C	206/206 (100%)	-0.25	3 (1%) 76 62	67, 102, 146, 182	0
4	D	208/208 (100%)	-0.27	4 (1%) 70 54	61, 92, 129, 145	0
5	E	150/150 (100%)	-0.40	1 (0%) 89 82	50, 72, 97, 154	0
6	F	101/101 (100%)	-0.38	1 (0%) 84 74	76, 108, 133, 147	0
7	G	155/155 (100%)	-0.30	4 (2%) 59 44	67, 99, 144, 204	0
8	H	138/138 (100%)	-0.47	2 (1%) 78 64	48, 67, 94, 113	0
9	I	127/127 (100%)	-0.26	0 100 100	61, 113, 140, 170	0
10	J	98/98 (100%)	0.10	4 (4%) 41 25	65, 135, 176, 194	0
11	K	119/119 (100%)	-0.35	3 (2%) 61 45	50, 80, 118, 140	0
12	L	125/125 (100%)	-0.32	3 (2%) 62 45	42, 81, 119, 162	0
13	M	120/120 (100%)	-0.04	5 (4%) 40 24	56, 96, 141, 171	0
14	N	60/60 (100%)	-0.20	0 100 100	72, 94, 123, 130	0
15	O	88/88 (100%)	-0.52	0 100 100	54, 82, 120, 160	0
16	P	83/83 (100%)	-0.39	0 100 100	53, 72, 94, 127	0
17	Q	99/99 (100%)	-0.51	0 100 100	45, 71, 101, 113	0
18	R	70/70 (100%)	-0.39	0 100 100	63, 89, 122, 137	0
19	S	78/78 (100%)	-0.20	2 (2%) 59 44	77, 117, 151, 176	0
20	T	99/99 (100%)	-0.46	0 100 100	51, 77, 124, 145	0
21	U	24/24 (100%)	0.04	1 (4%) 40 24	67, 80, 113, 139	0
22	X	4/5 (80%)	-0.19	0 100 100	78, 91, 109, 125	0
23	Y	9/10 (90%)	0.42	0 100 100	73, 84, 116, 137	0
All	All	3906/3914 (99%)	-0.29	53 (1%) 78 64	37, 85, 145, 237	0

The worst 5 of 53 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
13	M	119	GLY	7.5
13	M	120	LYS	6.1
1	A	1034	G	5.2
1	A	1540	U	5.1
7	G	81	GLY	5.1

## 6.2 Non-standard residues in protein, DNA, RNA chains [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
22	PSU	X	4	17/21	0.90	0.17	-	37,57,73,75	0

## 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
24	MG	A	1608	1/1	0.98	0.68	43.40	48,48,48,48	0
24	MG	A	1607	1/1	0.91	0.30	16.20	41,41,41,41	0
24	MG	A	1614	1/1	0.93	0.29	12.74	36,36,36,36	0
24	MG	A	1610	1/1	0.97	0.27	10.66	39,39,39,39	0
25	ZN	D	301	1/1	0.98	0.21	-1.14	119,119,119,119	0
24	MG	X	101	1/1	0.99	0.09	-2.80	43,43,43,43	0
24	MG	A	1611	1/1	0.94	0.26	-	48,48,48,48	0
24	MG	A	1603	1/1	0.83	0.39	-	60,60,60,60	0

*Continued on next page...*



*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
24	MG	A	1606	1/1	0.87	0.22	-	52,52,52,52	0
24	MG	A	1605	1/1	0.93	0.43	-	35,35,35,35	0
24	MG	A	1615	1/1	0.98	0.44	-	43,43,43,43	0
24	MG	Y	101	1/1	0.95	0.30	-	70,70,70,70	0
24	MG	A	1609	1/1	0.93	0.33	-	47,47,47,47	0
24	MG	A	1604	1/1	0.96	0.65	-	35,35,35,35	0
24	MG	A	1602	1/1	0.96	0.57	-	36,36,36,36	0
24	MG	A	1613	1/1	0.94	0.48	-	48,48,48,48	0
24	MG	A	1601	1/1	0.95	0.52	-	59,59,59,59	0
24	MG	A	1612	1/1	0.96	0.41	-	41,41,41,41	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.