



wwPDB/EMDataBank EM Map/Model Validation Summary Report ⓘ

Sep 19, 2017 – 02:58 PM EDT

PDB ID : 5XYU
EMDB ID: : EMD-6790
Title : Small subunit of Mycobacterium smegmatis ribosome
Authors : Li, Z.; Zhang, Y.; Zheng, L.; Ge, X.; Sanyal, S.; Gao, N.
Deposited on : unknown
Resolution : 3.45 Å(reported)

This is a wwPDB/EMDataBank EM Map/Model Validation Summary Report
for a publicly released PDB/EMDB entry.

We welcome your comments at validation@mail.wwpdb.org

A user guide is available at

<http://wwpdb.org/validation/2016/EMValidationReportHelp>
with specific help available everywhere you see the ⓘ symbol.

MolProbity : 4.02b-467
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP) : rb-20029824

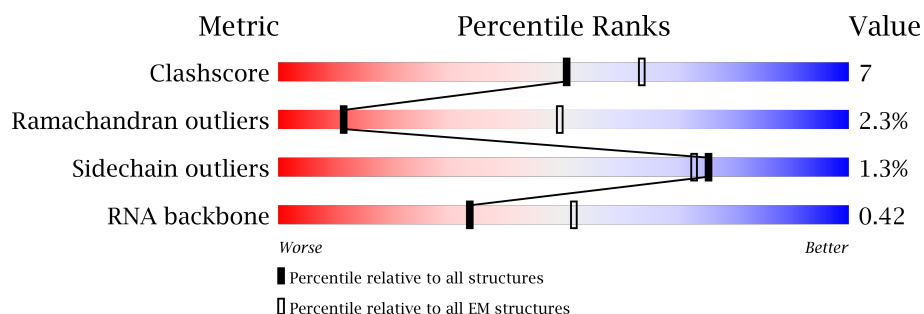
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

ELECTRON MICROSCOPY

The reported resolution of this entry is 3.45 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)	EM structures (#Entries)
Clashscore	125131	1336
Ramachandran outliers	121729	1120
Sidechain outliers	121581	1026
RNA backbone	3398	335

The table below summarises the geometric issues observed across the polymeric chains. The red, orange, yellow and green segments on the bar indicate the fraction of residues that contain outliers for ≥ 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 $\leq 5\%$

Mol	Chain	Length	Quality of chain
1	A	1528	61% 26% • 10%
2	C	275	63% • 33%
3	D	201	56% 11% • 31%
4	E	214	57% 16% • 26%
5	F	90	49% 24% • 26%
6	G	156	83% 9% 8%
7	H	132	78% 20% •
8	I	150	74% 9% 17%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
9	J	101	 82%11% . .
10	K	138	 78% . .17%
11	L	124	 71%27% . .
12	M	124	 65%21% .10%
13	N	61	 67%28% . .
14	O	89	 92% . . .
15	P	156	 54%5% .40%
16	Q	98	 80%11% .7%
17	R	84	 63% .35%
18	S	93	 67%15% .15%
19	T	86	 91%6% .
20	U	33	 70%6%24%

2 Entry composition

There are 21 unique types of molecules in this entry. The entry contains 45157 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, 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 RNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	1379	Total	C	N	O	P	0	0
			29639	13201	5457	9602	1379		

- Molecule 2 is a protein called 30S ribosomal protein S3.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	C	185	Total	C	N	O	S	0	0
			1498	939	287	268	4		

- Molecule 3 is a protein called 30S ribosomal protein S4.

Mol	Chain	Residues	Atoms					AltConf	Trace
3	D	138	Total	C	N	O	S	0	0
			1156	721	226	207	2		

- Molecule 4 is a protein called 30S ribosomal protein S5.

Mol	Chain	Residues	Atoms					AltConf	Trace
4	E	159	Total	C	N	O	S	0	0
			1147	724	214	206	3		

- Molecule 5 is a protein called 30S ribosomal protein S6.

Mol	Chain	Residues	Atoms					AltConf	Trace
5	F	67	Total	C	N	O	S	0	0
			530	335	92	102	1		

- Molecule 6 is a protein called 30S ribosomal protein S7.

Mol	Chain	Residues	Atoms					AltConf	Trace
6	G	144	Total	C	N	O	S	0	0
			1141	708	226	205	2		

- Molecule 7 is a protein called 30S ribosomal protein S8.

Mol	Chain	Residues	Atoms					AltConf	Trace
7	H	129	Total	C	N	O	S	0	0
			990	622	184	183	1		

- Molecule 8 is a protein called 30S ribosomal protein S9.

Mol	Chain	Residues	Atoms					AltConf	Trace
8	I	124	Total	C	N	O		0	0
			974	618	188	168			

- Molecule 9 is a protein called 30S ribosomal protein S10.

Mol	Chain	Residues	Atoms					AltConf	Trace
9	J	97	Total	C	N	O	S	0	0
			778	490	144	141	3		

- Molecule 10 is a protein called 30S ribosomal protein S11.

Mol	Chain	Residues	Atoms					AltConf	Trace
10	K	115	Total	C	N	O	S	0	0
			853	527	169	156	1		

- Molecule 11 is a protein called 30S ribosomal protein S12.

Mol	Chain	Residues	Atoms					AltConf	Trace
11	L	123	Total	C	N	O	S	0	0
			964	597	198	167	2		

- Molecule 12 is a protein called 30S ribosomal protein S13.

Mol	Chain	Residues	Atoms					AltConf	Trace
12	M	112	Total	C	N	O	S	0	0
			868	529	176	160	3		

- Molecule 13 is a protein called 30S ribosomal protein S14 type Z.

Mol	Chain	Residues	Atoms					AltConf	Trace
13	N	60	Total	C	N	O	S	0	0
			477	302	97	73	5		

- Molecule 14 is a protein called 30S ribosomal protein S15.

Mol	Chain	Residues	Atoms				AltConf	Trace
14	O	86	Total	C	N	O	0	0
			698	437	139	122		

- Molecule 15 is a protein called 30S ribosomal protein S16.

Mol	Chain	Residues	Atoms				AltConf	Trace
15	P	94	Total	C	N	O	0	0
			745	477	138	130		

- Molecule 16 is a protein called 30S ribosomal protein S17.

Mol	Chain	Residues	Atoms					AltConf	Trace
16	Q	91	Total	C	N	O	S	0	0
			721	453	137	129	2		

- Molecule 17 is a protein called 30S ribosomal protein S18 2.

Mol	Chain	Residues	Atoms					AltConf	Trace
17	R	55	Total	C	N	O	S	0	0
			432	267	86	78	1		

- Molecule 18 is a protein called 30S ribosomal protein S19.

Mol	Chain	Residues	Atoms					AltConf	Trace
18	S	79	Total	C	N	O	S	0	0
			634	408	116	109	1		

- Molecule 19 is a protein called 30S ribosomal protein S20.

Mol	Chain	Residues	Atoms				AltConf	Trace
19	T	83	Total	C	N	O	0	0
			647	393	137	117		

- Molecule 20 is a protein called Conserved domain protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
20	U	25	Total	C	N	O	S	0	0
			217	133	55	28	1		

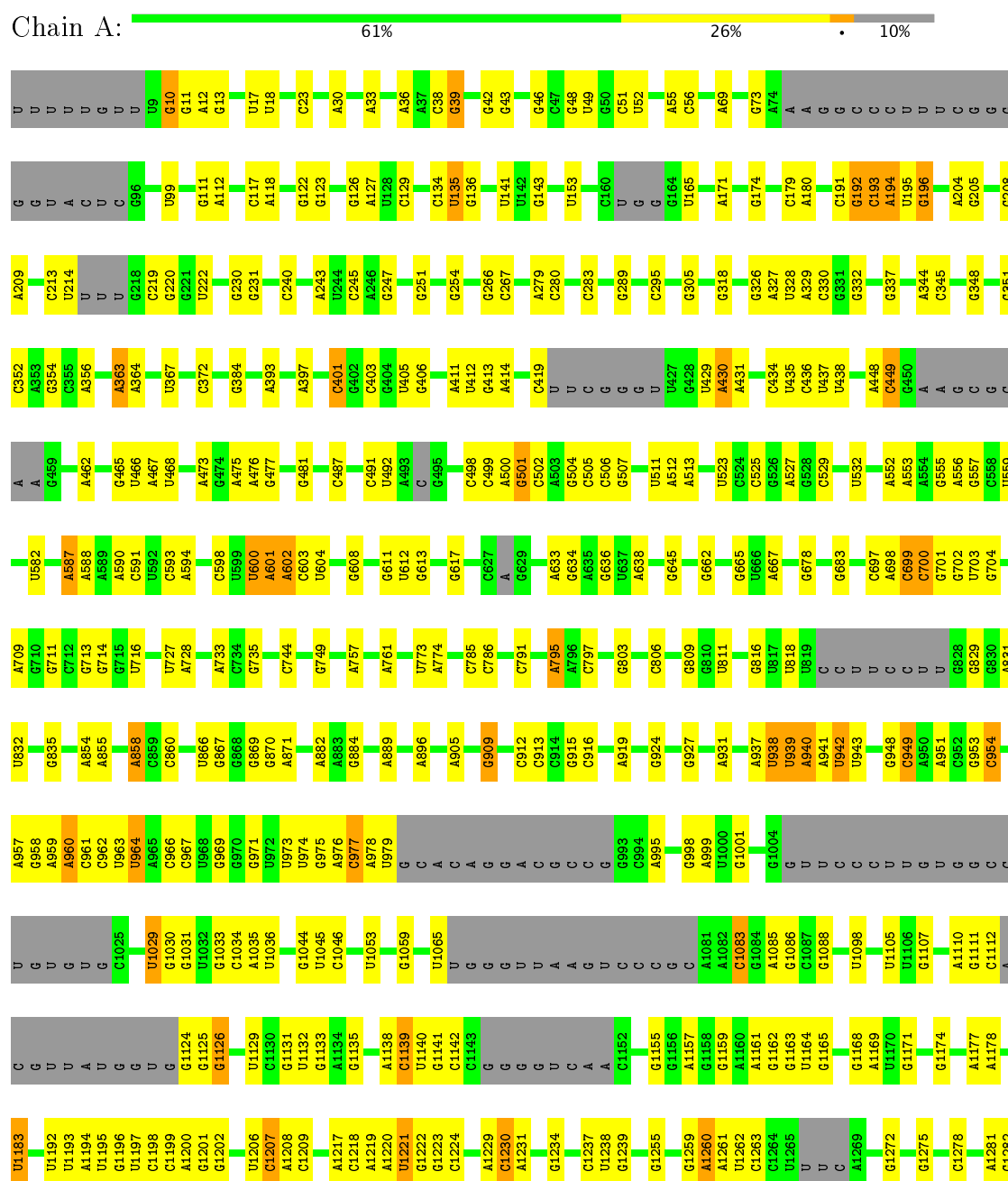
- Molecule 21 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		AltConf
21	A	47	Total 47	Mg 47	0
21	C	1	Total 1	Mg 1	0


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 the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry. 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. 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 RNA




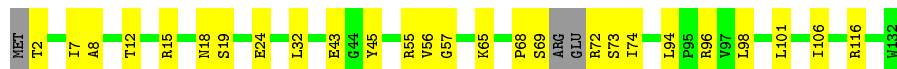
- Molecule 6: 30S ribosomal protein S7

Chain G:  83% 9% 8%



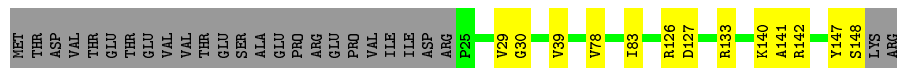
- Molecule 7: 30S ribosomal protein S8

Chain H:  78% 20% .




- Molecule 8: 30S ribosomal protein S9

Chain I:  74% 9% 17%




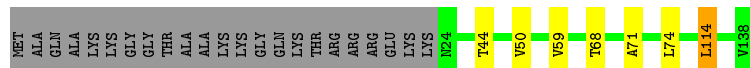
- Molecule 9: 30S ribosomal protein S10

Chain J:  82% 11% . .



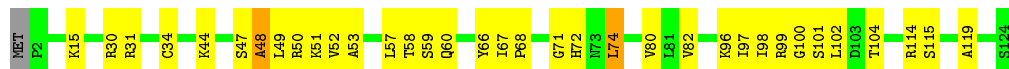
- Molecule 10: 30S ribosomal protein S11

Chain K:  78% . . 17%



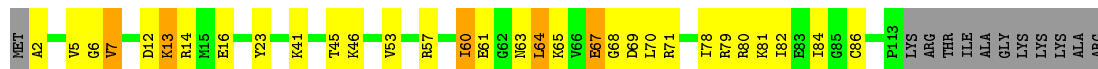
- Molecule 11: 30S ribosomal protein S12

Chain L:  71% 27% . .



- Molecule 12: 30S ribosomal protein S13

Chain M:  65% 21% . 10%



- Molecule 13: 30S ribosomal protein S14 type Z

Chain N:  67% 28% . .



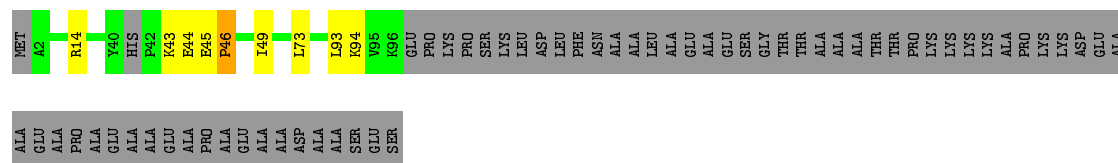
- Molecule 14: 30S ribosomal protein S15

Chain O:  92% . . .




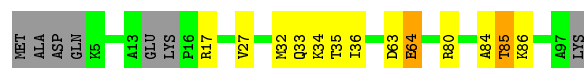
- Molecule 15: 30S ribosomal protein S16

Chain P:  54% 5% . 40%



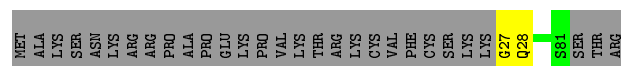
- Molecule 16: 30S ribosomal protein S17

Chain Q:  80% 11% . 7%



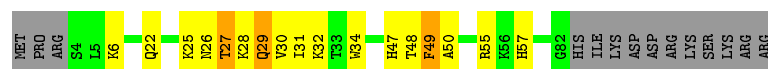
- Molecule 17: 30S ribosomal protein S18 2

Chain R:  63% . 35%



- Molecule 18: 30S ribosomal protein S19

Chain S:  67% 15% . 15%



- Molecule 19: 30S ribosomal protein S20

Chain T:  91% 6% .



- Molecule 20: Conserved domain protein

Chain U:

70%

6%

24%



4 Experimental information

Property	Value	Source
Reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	Depositor
Number of particles used	47338	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	PHASE FLIPPING AND AMPLITUDE CORRECTION	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ($e^-/\text{\AA}^2$)	2	Depositor
Minimum defocus (nm)	Not provided	Depositor
Maximum defocus (nm)	Not provided	Depositor
Magnification	Not provided	Depositor
Image detector	GATAN K2 SUMMIT (4k x 4k)	Depositor

5 Model quality

5.1 Standard geometry

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

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 >2	RMSZ	# Z >2
1	A	0.20	0/33170	0.67	0/51726
10	K	0.35	0/871	0.51	0/1179
11	L	0.38	0/975	0.62	0/1302
12	M	0.38	0/875	0.65	0/1174
13	N	0.42	0/488	0.65	0/650
14	O	0.35	0/707	0.58	0/949
15	P	0.42	0/757	0.60	0/1018
16	Q	0.33	0/731	0.61	0/977
17	R	0.34	0/436	0.59	0/586
18	S	0.36	0/650	0.57	0/875
19	T	0.37	0/650	0.58	0/864
2	C	0.35	0/1518	0.56	0/2034
20	U	0.41	0/217	0.67	0/278
3	D	0.35	0/1169	0.57	0/1561
4	E	0.40	0/1163	0.63	0/1574
5	F	0.39	0/535	0.62	0/724
6	G	0.37	0/1154	0.58	0/1551
7	H	0.38	0/1004	0.59	0/1356
8	I	0.36	0/992	0.55	0/1337
9	J	0.39	0/791	0.59	0/1069
All	All	0.27	0/48853	0.65	0/72784

There are no bond length outliers.

There are no bond angle outliers.

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	29639	0	14926	127	0
2	C	1498	0	1535	15	0
3	D	1156	0	1177	41	0
4	E	1147	0	1206	65	0
5	F	530	0	548	33	0
6	G	1141	0	1201	18	0
7	H	990	0	1026	38	0
8	I	974	0	1024	7	0
9	J	778	0	811	9	0
10	K	853	0	859	2	0
11	L	964	0	1050	43	0
12	M	868	0	873	63	0
13	N	477	0	500	35	0
14	O	698	0	734	3	0
15	P	745	0	793	8	0
16	Q	721	0	768	26	0
17	R	432	0	448	4	0
18	S	634	0	649	41	0
19	T	647	0	696	1	0
20	U	217	0	261	1	0
21	A	47	0	0	0	0
21	C	1	0	0	0	0
All	All	45157	0	31085	514	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 7.

The worst 5 of 514 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)
4:E:182:ARG:NH1	7:H:45:TYR:CZ	1.78	1.42
12:M:81:LYS:HA	12:M:84:ILE:CD1	1.56	1.36
12:M:80:ARG:O	12:M:84:ILE:CG1	1.71	1.35
12:M:80:ARG:C	12:M:84:ILE:HG13	1.47	1.33
5:F:70:ASP:O	5:F:74:ASN:ND2	1.63	1.31

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 PDB entries followed by that with respect to all EM entries.

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	C	177/275 (64%)	161 (91%)	15 (8%)	1 (1%)	28	70
3	D	128/201 (64%)	124 (97%)	3 (2%)	1 (1%)	22	65
4	E	157/214 (73%)	138 (88%)	13 (8%)	6 (4%)	4	30
5	F	63/90 (70%)	58 (92%)	5 (8%)	0	100	100
6	G	138/156 (88%)	132 (96%)	4 (3%)	2 (1%)	13	52
7	H	125/132 (95%)	114 (91%)	10 (8%)	1 (1%)	22	65
8	I	122/150 (81%)	113 (93%)	5 (4%)	4 (3%)	4	34
9	J	93/101 (92%)	82 (88%)	8 (9%)	3 (3%)	5	35
10	K	113/138 (82%)	101 (89%)	10 (9%)	2 (2%)	10	47
11	L	121/124 (98%)	111 (92%)	5 (4%)	5 (4%)	3	29
12	M	110/124 (89%)	101 (92%)	6 (6%)	3 (3%)	6	39
13	N	58/61 (95%)	50 (86%)	6 (10%)	2 (3%)	4	34
14	O	84/89 (94%)	79 (94%)	4 (5%)	1 (1%)	15	56
15	P	90/156 (58%)	78 (87%)	9 (10%)	3 (3%)	4	34
16	Q	87/98 (89%)	78 (90%)	6 (7%)	3 (3%)	4	34
17	R	53/84 (63%)	52 (98%)	1 (2%)	0	100	100
18	S	77/93 (83%)	69 (90%)	5 (6%)	3 (4%)	3	30
19	T	81/86 (94%)	78 (96%)	1 (1%)	2 (2%)	6	40
20	U	23/33 (70%)	22 (96%)	0	1 (4%)	3	27
All	All	1900/2405 (79%)	1741 (92%)	116 (6%)	43 (2%)	11	43

5 of 43 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
4	E	180	ALA
8	I	141	ALA
9	J	36	VAL
10	K	71	ALA
11	L	74	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 PDB entries followed by that with respect to all EM entries.

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	C	155/212 (73%)	154 (99%)	1 (1%)	89	96
3	D	124/176 (70%)	120 (97%)	4 (3%)	44	76
4	E	115/147 (78%)	115 (100%)	0	100	100
5	F	58/79 (73%)	56 (97%)	2 (3%)	42	75
6	G	123/132 (93%)	123 (100%)	0	100	100
7	H	105/108 (97%)	104 (99%)	1 (1%)	80	92
8	I	100/125 (80%)	100 (100%)	0	100	100
9	J	88/90 (98%)	85 (97%)	3 (3%)	42	75
10	K	89/105 (85%)	87 (98%)	2 (2%)	57	83
11	L	104/105 (99%)	102 (98%)	2 (2%)	62	85
12	M	86/104 (83%)	84 (98%)	2 (2%)	56	82
13	N	49/50 (98%)	48 (98%)	1 (2%)	60	84
14	O	74/77 (96%)	74 (100%)	0	100	100
15	P	77/118 (65%)	77 (100%)	0	100	100
16	Q	77/83 (93%)	77 (100%)	0	100	100
17	R	45/72 (62%)	45 (100%)	0	100	100
18	S	70/84 (83%)	67 (96%)	3 (4%)	33	70
19	T	68/70 (97%)	67 (98%)	1 (2%)	70	88
20	U	23/31 (74%)	23 (100%)	0	100	100
All	All	1630/1968 (83%)	1608 (99%)	22 (1%)	75	90

5 of 22 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
9	J	57	LYS
10	K	114	LEU
18	S	49	PHE
9	J	75	ASP
10	K	68	THR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 8 such sidechains are listed below:

Mol	Chain	Res	Type
5	F	74	ASN
18	S	29	GLN
14	O	28	GLN
4	E	111	GLN
11	L	72	HIS

5.3.3 RNA [i](#)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
1	A	1362/1528 (89%)	407 (29%)	41 (3%)

5 of 407 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
1	A	10	G
1	A	11	G
1	A	12	A
1	A	13	G
1	A	17	U

5 of 41 RNA pucker outliers are listed below:

Mol	Chain	Res	Type
1	A	703	U
1	A	942	U
1	A	1374	U
1	A	866	U
1	A	938	U

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 ⓘ

Of 48 ligands modelled in this entry, 48 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 ⓘ

There are no such residues in this entry.

5.8 Polymer linkage issues ⓘ

There are no chain breaks in this entry.