wwPDB EM Validation Summary Report

Jun 22, 2021 – 06:13 AM BST

PDB ID	:	7O5H
EMDB ID	:	EMD-12736
Title	:	Ribosomal methyltransferase KsgA bound to small ribosomal subunit
Authors	:	Stephan, N.C.; Ries, A.B.; Boehringer, D.; Ban, N.
Deposited on	:	2021-04-08
Resolution	:	3.10 Å(reported)

This is a wwPDB EM Validation Summary Report for a publicly released PDB entry.

We welcome your comments at *validation@mail.wwpdb.org*
A user guide is available at
https://www.wwpdb.org/validation/2017/EMValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The following versions of software and data (see references (1)) were used in the production of this report:

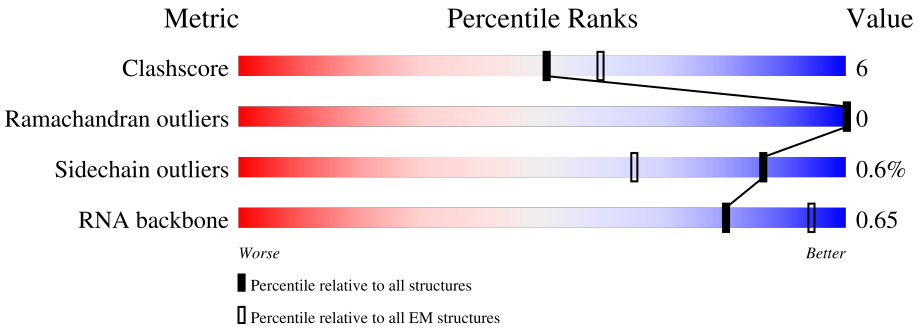
EMDB validation analysis	:	0.0.0.dev75
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.20

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
ELECTRON MICROSCOPY

The reported resolution of this entry is 3.10 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	158937	4297
Ramachandran outliers	154571	4023
Sidechain outliers	154315	3826
RNA backbone	4643	859

The table below summarises the geometric issues observed across the polymeric chains and their fit to the map. The red, orange, yellow and green segments of the bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria respectively. 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 <=5%

Mol	Chain	Length	Quality of chain
1	V	252	84% 15%
2	A	964	69% 27% .
3	B	225	79% 20% .
4	D	205	82% 18%
5	E	158	79% 21%
6	F	106	75% 22% .
7	H	129	80% 20%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
8	K	117	79% 21%
9	L	123	79% 21%
10	O	88	83% 16%
11	P	82	82% 18%
12	Q	80	88% 11%
13	R	55	84% 16%
14	T	86	93% 7%
15	U	56	48% 5% 46%

2 Entry composition

There are 16 unique types of molecules in this entry. The entry contains 34436 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 protein called Ribosomal RNA small subunit methyltransferase A.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	V	251	Total	C	N	O	S	0	0
			1955	1247	336	359	13		

- Molecule 2 is a RNA chain called 16S rRNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	A	964	Total	C	N	O	P	0	0
			20726	9239	3830	6693	964		

- Molecule 3 is a protein called 30S ribosomal protein S2.

Mol	Chain	Residues	Atoms					AltConf	Trace
3	B	225	Total	C	N	O	S	0	0
			1760	1113	316	323	8		

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

Mol	Chain	Residues	Atoms					AltConf	Trace
4	D	205	Total	C	N	O	S	0	0
			1643	1026	315	298	4		

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

Mol	Chain	Residues	Atoms					AltConf	Trace
5	E	158	Total	C	N	O	S	0	0
			1165	725	220	214	6		

- Molecule 6 is a protein called 30S ribosomal protein S6, fully modified isoform.

Mol	Chain	Residues	Atoms					AltConf	Trace
6	F	103	Total	C	N	O	S	0	0
			839	530	151	151	7		

- 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
			979	616	173	184	6		

- Molecule 8 is a protein called 30S ribosomal protein S11.

Mol	Chain	Residues	Atoms					AltConf	Trace
8	K	117	Total	C	N	O	S	0	0
			877	540	174	160	3		

- Molecule 9 is a protein called 30S ribosomal protein S12.

Mol	Chain	Residues	Atoms					AltConf	Trace
9	L	123	Total	C	N	O	S	0	0
			955	590	196	165	4		

- Molecule 10 is a protein called 30S ribosomal protein S15.

Mol	Chain	Residues	Atoms					AltConf	Trace
10	O	87	Total	C	N	O	S	0	0
			702	433	140	128	1		

- Molecule 11 is a protein called 30S ribosomal protein S16.

Mol	Chain	Residues	Atoms					AltConf	Trace
11	P	82	Total	C	N	O	S	0	0
			649	406	128	114	1		

- Molecule 12 is a protein called 30S ribosomal protein S17.

Mol	Chain	Residues	Atoms					AltConf	Trace
12	Q	79	Total	C	N	O	S	0	0
			641	406	120	112	3		

- Molecule 13 is a protein called 30S ribosomal protein S18.

Mol	Chain	Residues	Atoms				AltConf	Trace
13	R	55	Total	C	N	O	0	0
			456	288	86	82		

- Molecule 14 is a protein called 30S ribosomal protein S20.

Mol	Chain	Residues	Atoms					AltConf	Trace
14	T	86	Total 670	C 414	N 138	O 115	S 3	0	0

- Molecule 15 is a protein called 30S ribosomal protein S21.

Mol	Chain	Residues	Atoms				AltConf	Trace
15	U	30	Total 252	C 156	N 54	O 42	0	0

- Molecule 16 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

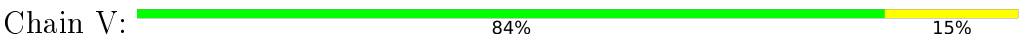
Mol	Chain	Residues	Atoms		AltConf
16	A	161	Total 161	Mg 161	0
16	D	1	Total 1	Mg 1	0
16	E	1	Total 1	Mg 1	0
16	F	1	Total 1	Mg 1	0
16	H	1	Total 1	Mg 1	0
16	K	1	Total 1	Mg 1	0
16	U	1	Total 1	Mg 1	0

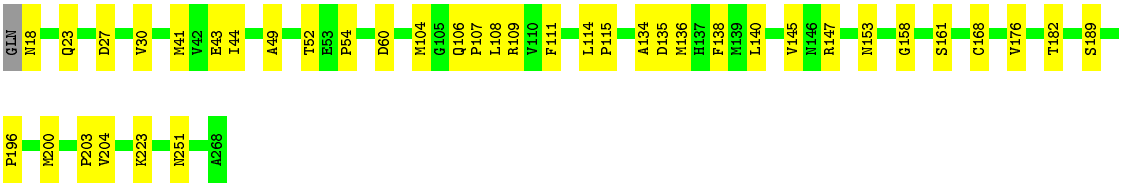
3 Residue-property plots (i)

These plots are drawn for all protein, RNA, DNA and oligosaccharide 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: Ribosomal RNA small subunit methyltransferase A

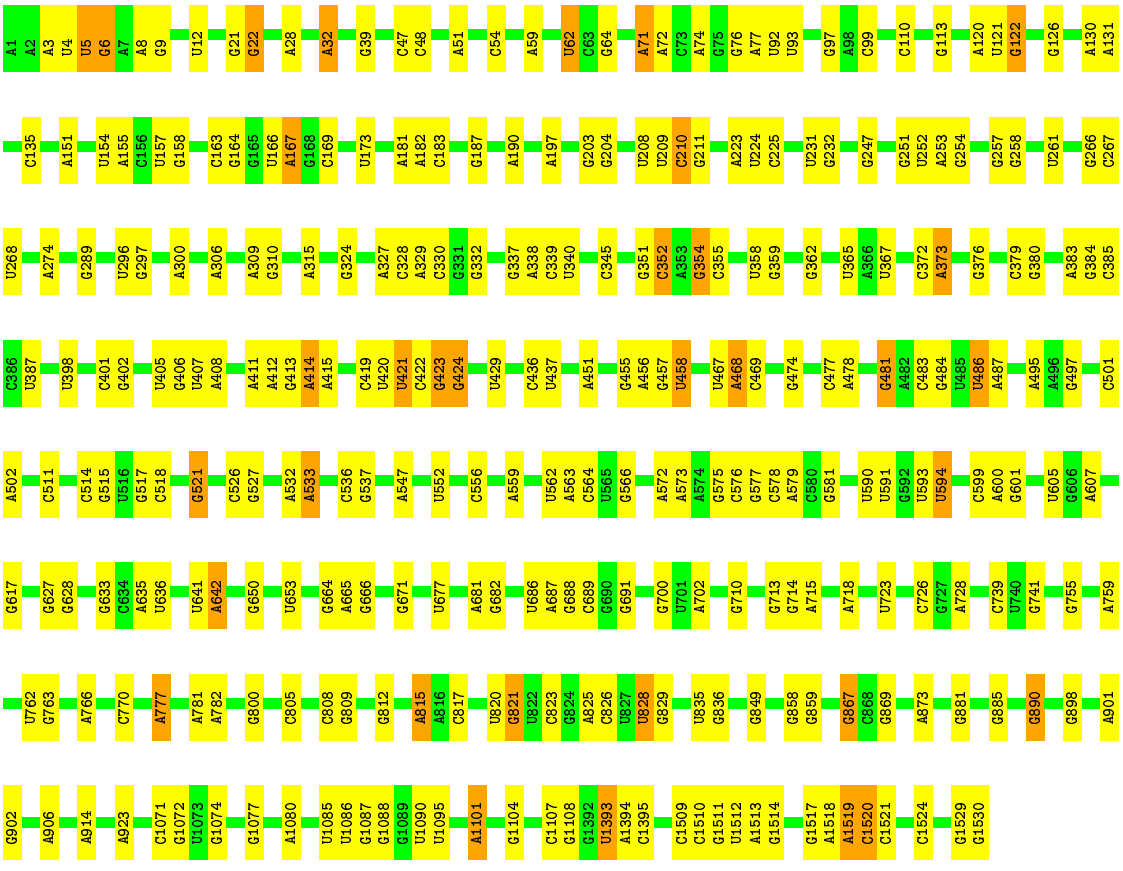
Chain V:





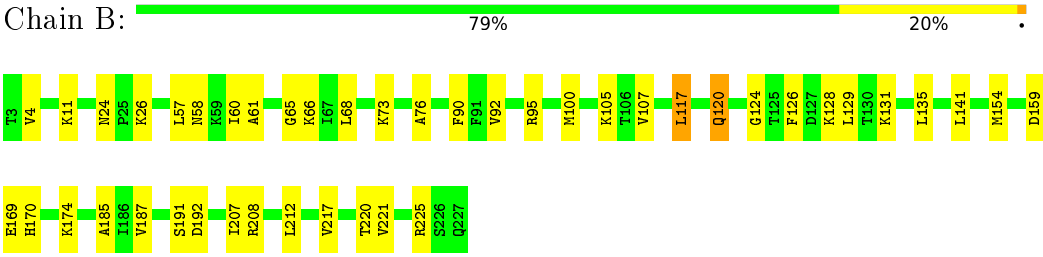
- Molecule 2: 16S rRNA

Chain A:



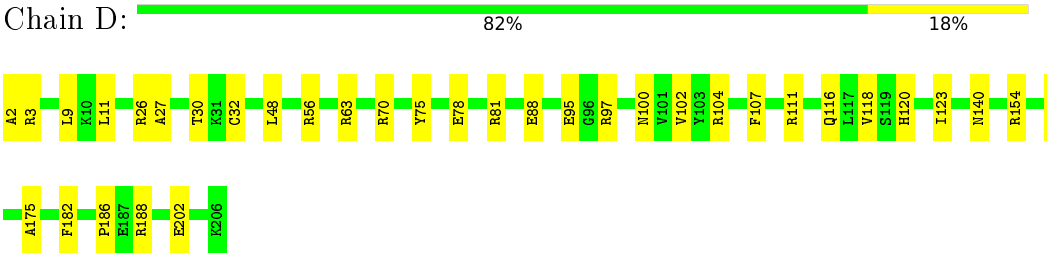
• Molecule 3: 30S ribosomal protein S2

Chain B:



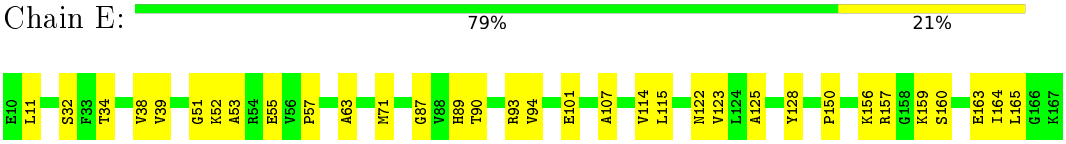
• Molecule 4: 30S ribosomal protein S4

Chain D:



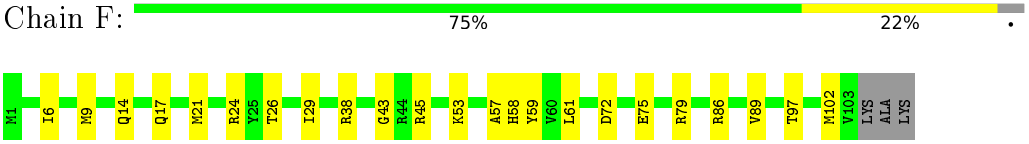
• Molecule 5: 30S ribosomal protein S5

Chain E:



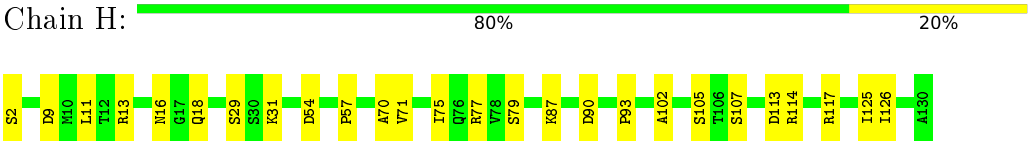
• Molecule 6: 30S ribosomal protein S6, fully modified isoform

Chain F:



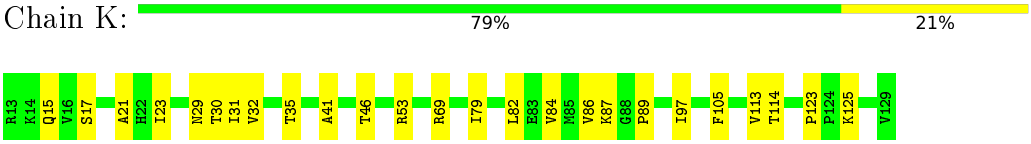
• Molecule 7: 30S ribosomal protein S8

Chain H:



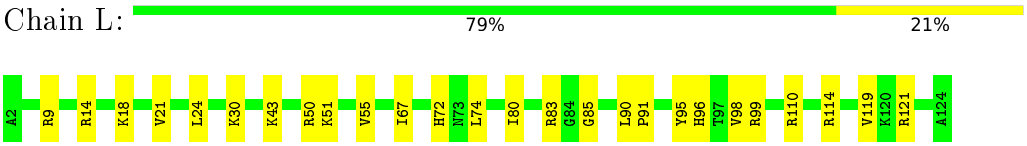
• Molecule 8: 30S ribosomal protein S11

Chain K:



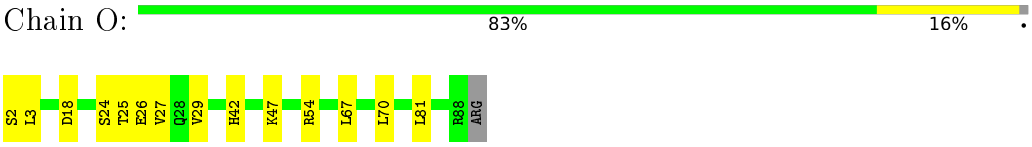
• Molecule 9: 30S ribosomal protein S12

Chain L:



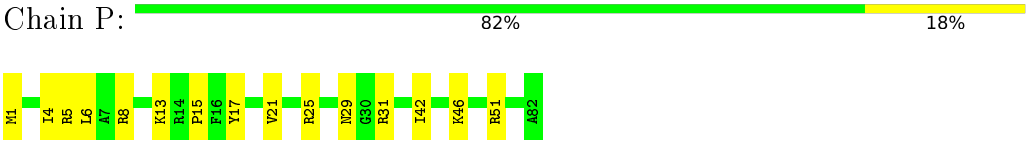
• Molecule 10: 30S ribosomal protein S15

Chain O:



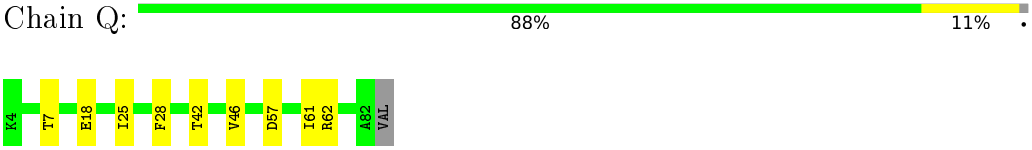
• Molecule 11: 30S ribosomal protein S16

Chain P:



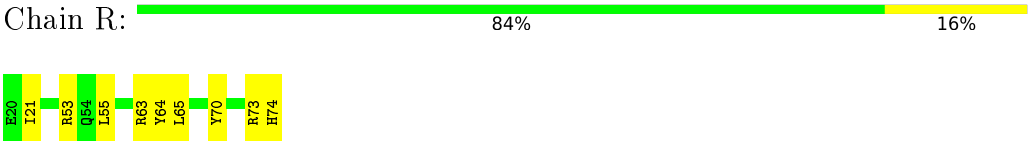
• Molecule 12: 30S ribosomal protein S17

Chain Q:



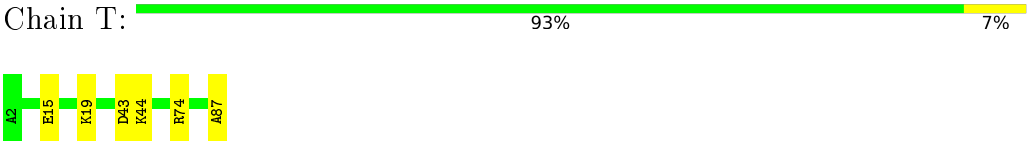
• Molecule 13: 30S ribosomal protein S18

Chain R:



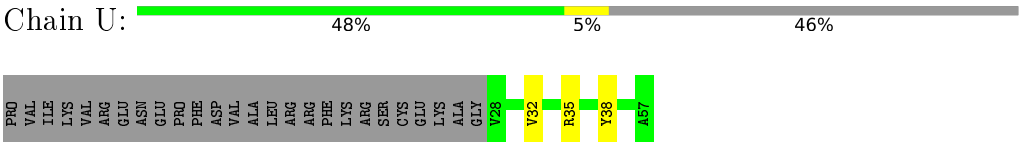
• Molecule 14: 30S ribosomal protein S20

Chain T:



• Molecule 15: 30S ribosomal protein S21

Chain U:



4 Experimental information (i)

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	165073	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 ($e^-/Å^2$)	35	Depositor
Minimum defocus (nm)	Not provided	
Maximum defocus (nm)	Not provided	
Magnification	Not provided	
Image detector	FEI FALCON III (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	#$\|Z\|>5$	RMSZ	#$\|Z\|>5$
1	V	0.25	0/1998	0.45	0/2718
2	A	0.21	0/23209	0.67	0/36199
3	B	0.24	0/1791	0.45	0/2413
4	D	0.24	0/1665	0.49	0/2227
5	E	0.25	0/1178	0.49	0/1584
6	F	0.24	0/858	0.47	0/1160
7	H	0.25	0/989	0.47	0/1326
8	K	0.24	0/893	0.52	0/1205
9	L	0.25	0/969	0.56	0/1300
10	O	0.22	0/710	0.48	0/950
11	P	0.24	0/659	0.54	0/884
12	Q	0.24	0/650	0.51	0/871
13	R	0.24	0/463	0.51	0/621
14	T	0.23	0/676	0.43	0/895
15	U	0.23	0/255	0.55	0/338
All	All	0.22	0/36963	0.61	0/54691

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	V	1955	0	1969	23	0
2	A	20726	0	10422	148	0
3	B	1760	0	1787	28	0
4	D	1643	0	1707	25	0
5	E	1165	0	1212	18	0
6	F	839	0	833	16	0
7	H	979	0	1031	18	0
8	K	877	0	886	19	0
9	L	955	0	1016	19	0
10	O	702	0	721	9	0
11	P	649	0	666	12	0
12	Q	641	0	682	5	0
13	R	456	0	478	7	0
14	T	670	0	719	4	0
15	U	252	0	267	3	0
16	A	161	0	0	0	0
16	D	1	0	0	0	0
16	E	1	0	0	0	0
16	F	1	0	0	0	0
16	H	1	0	0	0	0
16	K	1	0	0	0	0
16	U	1	0	0	0	0
All	All	34436	0	24396	307	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 6.

The worst 5 of 307 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)
2:A:437:U:HO2'	4:D:120:HIS:HD1	1.17	0.92
2:A:823:C:HO2'	7:H:2:SER:N	1.73	0.87
2:A:71:A:N6	2:A:99:C:O2	2.20	0.75
2:A:483:C:O2	11:P:13:LYS:NZ	2.22	0.72
8:K:23:ILE:HG22	8:K:32:VAL:HG12	1.70	0.72

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 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	
1	V	249/252 (99%)	246 (99%)	3 (1%)	0	100	100
3	B	223/225 (99%)	212 (95%)	11 (5%)	0	100	100
4	D	203/205 (99%)	200 (98%)	3 (2%)	0	100	100
5	E	156/158 (99%)	155 (99%)	1 (1%)	0	100	100
6	F	101/106 (95%)	101 (100%)	0	0	100	100
7	H	127/129 (98%)	125 (98%)	2 (2%)	0	100	100
8	K	115/117 (98%)	110 (96%)	5 (4%)	0	100	100
9	L	121/123 (98%)	115 (95%)	6 (5%)	0	100	100
10	O	85/88 (97%)	84 (99%)	1 (1%)	0	100	100
11	P	80/82 (98%)	77 (96%)	3 (4%)	0	100	100
12	Q	77/80 (96%)	76 (99%)	1 (1%)	0	100	100
13	R	53/55 (96%)	53 (100%)	0	0	100	100
14	T	84/86 (98%)	84 (100%)	0	0	100	100
15	U	28/56 (50%)	28 (100%)	0	0	100	100
All	All	1702/1762 (97%)	1666 (98%)	36 (2%)	0	100	100

There are no Ramachandran outliers to report.

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
1	V	215/216 (100%)	212 (99%)	3 (1%)	67 86
3	B	187/187 (100%)	185 (99%)	2 (1%)	73 89
4	D	172/172 (100%)	171 (99%)	1 (1%)	86 94
5	E	120/120 (100%)	119 (99%)	1 (1%)	81 92
6	F	90/92 (98%)	90 (100%)	0	100 100
7	H	104/104 (100%)	104 (100%)	0	100 100
8	K	90/90 (100%)	90 (100%)	0	100 100
9	L	103/103 (100%)	102 (99%)	1 (1%)	76 90
10	O	75/76 (99%)	75 (100%)	0	100 100
11	P	65/65 (100%)	65 (100%)	0	100 100
12	Q	73/74 (99%)	72 (99%)	1 (1%)	67 86
13	R	48/48 (100%)	48 (100%)	0	100 100
14	T	65/65 (100%)	65 (100%)	0	100 100
15	U	25/48 (52%)	25 (100%)	0	100 100
All	All	1432/1460 (98%)	1423 (99%)	9 (1%)	86 94

5 of 9 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
9	L	18	LYS
12	Q	28	PHE
3	B	117	LEU
3	B	120	GLN
4	D	160	GLU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 6 such sidechains are listed below:

Mol	Chain	Res	Type
9	L	5	ASN
13	R	74	HIS
14	T	48	GLN
6	F	14	GLN
1	V	106	GLN

5.3.3 RNA (i)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
2	A	956/964 (99%)	126 (13%)	2 (0%)

5 of 126 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
2	A	5	U
2	A	6	G
2	A	9	G
2	A	22	G
2	A	32	A

All (2) RNA pucker outliers are listed below:

Mol	Chain	Res	Type
2	A	411	A
2	A	1107	C

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 monosaccharides in this entry.

5.6 Ligand geometry (i)

Of 167 ligands modelled in this entry, 167 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)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
2	A	7

The worst 5 of 7 chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	A	1108:G	O3'	1392:G	P	42.05
1	A	1396:A	O3'	1507:A	P	29.66
1	A	924:C	O3'	1068:G	P	23.32
1	A	840:C	O3'	846:G	P	16.77
1	A	77:A	O3'	92:U	P	16.40

6 Map visualisation

This section contains visualisations of the EMDB entry EMD-12736. These allow visual inspection of the internal detail of the map and identification of artifacts.

Images derived from a raw map, generated by summing the deposited half-maps, are presented below the corresponding image components of the primary map to allow further visual inspection and comparison with those of the primary map.

6.1 Orthogonal projections

This section was not generated.

6.2 Central slices

This section was not generated.

6.3 Largest variance slices

This section was not generated.

6.4 Orthogonal surface views

This section was not generated.

6.5 Mask visualisation

This section was not generated. No masks/segmentation were deposited.

7 Map analysis

This section contains the results of statistical analysis of the map.

7.1 Map-value distribution

This section was not generated.

7.2 Volume estimate versus contour level

This section was not generated.

7.3 Rotationally averaged power spectrum

This section was not generated. The rotationally averaged power spectrum had issues being displayed.

8 Fourier-Shell correlation

This section was not generated. No FSC curve or half-maps provided.

9 Map-model fit

This section was not generated.