



# wwPDB/EMDataBank EM Map/Model Validation Summary Report ⓘ

Jan 22, 2020 – 10:42 AM EST

PDB ID : 6O7K  
EMDB ID: : EMD-0643  
Title : 30S initiation complex  
Authors : Frank, J.; Gonzalez Jr., R.L.; kaledhonkar, S.; Fu, Z.; Caban, K.; Li, W.;  
Chen, B.; Sun, M.  
Deposited on : 2019-03-08  
Resolution : 4.20 Å(reported)  
Based on PDB ID : 2AVY

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](mailto: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 ⓘ symbol.

---

MolProbity : 4.02b-467  
Percentile statistics : 20171227.v01 (using entries in the PDB archive December 27th 2017)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.4

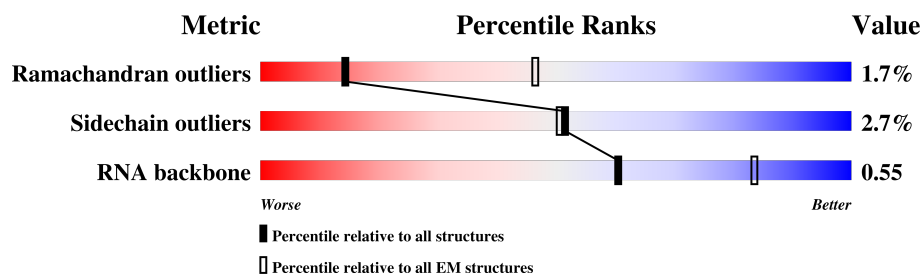
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*ELECTRON MICROSCOPY*

The reported resolution of this entry is 4.20 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.



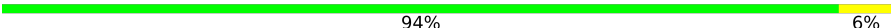











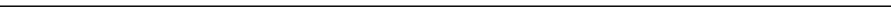


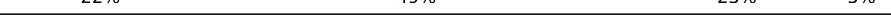
Metric	Whole archive (#Entries)	EM structures (#Entries)
Ramachandran outliers	132723	1663
Sidechain outliers	132532	1531
RNA backbone	3747	458

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  $\geq 3$ , 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  $\leq 5\%$

Mol	Chain	Length	Quality of chain
1	5	71	
2	f	509	
3	g	1539	
4	P	80	
5	r	98	
6	q	117	
7	t	123	
8	s	114	
9	w	100	

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
10	u	88	 94% 6%
11	y	82	 91% 9%
12	1	55	 87% 9% .
13	z	79	 92% 8%
14	j	218	 95% . .
15	3	85	 94% 6%
16	2	51	 82% 16% .
17	h	206	 93% 6% .
18	l	205	 90% 9%
19	k	150	 93% 5% .
20	n	100	 92% 8%
21	m	151	 94% 6%
22	p	129	 92% 7% .
23	o	127	 83% 14% . .
24	v	77	 22% 49% 23% 5%
25	N	6	 17% 50% 33%

## 2 Entry composition [i](#)

There are 25 unique types of molecules in this entry. The entry contains 57748 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 Translation initiation factor IF-1.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	5	71	Total	C	N	O	S	0	0
			570	362	103	103	2		

- Molecule 2 is a protein called Translation initiation factor IF-2.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	f	509	Total	C	N	O	S	0	0
			3847	2409	675	748	15		

- Molecule 3 is a RNA chain called 16S ribosomal RNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
3	g	1539	Total	C	N	O	P	0	0
			33012	14725	6052	10697	1538		

- Molecule 4 is a protein called 30S ribosomal protein S17.

Mol	Chain	Residues	Atoms					AltConf	Trace
4	P	80	Total	C	N	O	S	0	0
			649	411	121	114	3		

- Molecule 5 is a protein called 30S ribosomal protein S10.

Mol	Chain	Residues	Atoms					AltConf	Trace
5	r	98	Total	C	N	O	S	0	0
			787	493	150	143	1		

- Molecule 6 is a protein called 30S ribosomal protein S11.

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

- Molecule 7 is a protein called 30S ribosomal protein S12.

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

- Molecule 8 is a protein called 30S ribosomal protein S13.

Mol	Chain	Residues	Atoms					AltConf	Trace
8	s	114	Total	C	N	O	S	0	0
			884	546	178	157	3		

- Molecule 9 is a protein called 30S ribosomal protein S14.

Mol	Chain	Residues	Atoms					AltConf	Trace
9	w	96	Total	C	N	O	S	0	0
			774	483	160	128	3		

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

Mol	Chain	Residues	Atoms					AltConf	Trace
10	u	88	Total	C	N	O	S	0	0
			714	439	144	130	1		

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

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

- Molecule 12 is a protein called 30S ribosomal protein S18.

Mol	Chain	Residues	Atoms				AltConf	Trace
12	1	55	Total	C	N	O	0	0
			456	288	86	82		

- Molecule 13 is a protein called 30S ribosomal protein S19.

Mol	Chain	Residues	Atoms					AltConf	Trace
13	z	79	Total	C	N	O	S	0	0
			638	408	120	108	2		

- Molecule 14 is a protein called 30S ribosomal protein S2.

Mol	Chain	Residues	Atoms					AltConf	Trace
14	j	218	Total	C	N	O	S	0	0
			1705	1081	305	312	7		

- Molecule 15 is a protein called 30S ribosomal protein S20.

Mol	Chain	Residues	Atoms					AltConf	Trace
15	3	85	Total	C	N	O	S	0	0
			665	411	137	114	3		

- Molecule 16 is a protein called 30S ribosomal protein S21.

Mol	Chain	Residues	Atoms					AltConf	Trace
16	2	51	Total	C	N	O	S	0	0
			426	265	86	74	1		

- Molecule 17 is a protein called 30S ribosomal protein S3.

Mol	Chain	Residues	Atoms					AltConf	Trace
17	h	206	Total	C	N	O	S	0	0
			1625	1028	305	289	3		

- Molecule 18 is a protein called 30S ribosomal protein S4.

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

- Molecule 19 is a protein called 30S ribosomal protein S5.

Mol	Chain	Residues	Atoms					AltConf	Trace
19	k	150	Total	C	N	O	S	0	0
			1106	687	211	202	6		

- Molecule 20 is a protein called 30S ribosomal protein S6.

Mol	Chain	Residues	Atoms					AltConf	Trace
20	n	100	Total	C	N	O	S	0	0
			818	515	148	149	6		

- Molecule 21 is a protein called 30S ribosomal protein S7.

Mol	Chain	Residues	Atoms					AltConf	Trace
21	m	151	Total	C	N	O	S	0	0
			1182	735	227	216	4		

- Molecule 22 is a protein called 30S ribosomal protein S8.

Mol	Chain	Residues	Atoms					AltConf	Trace
22	p	129	Total	C	N	O	S	0	0
			979	616	173	184	6		

- Molecule 23 is a protein called 30S ribosomal protein S9.

Mol	Chain	Residues	Atoms					AltConf	Trace
23	o	127	Total	C	N	O	S	0	0
			1022	634	206	179	3		

- Molecule 24 is a RNA chain called tRNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
24	v	77	Total	C	N	O	P	0	0
			1639	732	297	534	76		

- Molecule 25 is a RNA chain called mRNA.

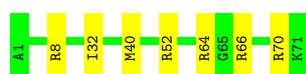
Mol	Chain	Residues	Atoms					AltConf	Trace
25	N	6	Total	C	N	O	P	0	0
			126	58	24	39	5		

### 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.

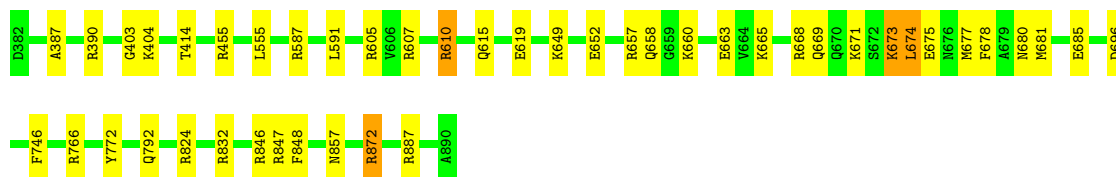
- Molecule 1: Translation initiation factor IF-1

Chain 5:  90% 10%



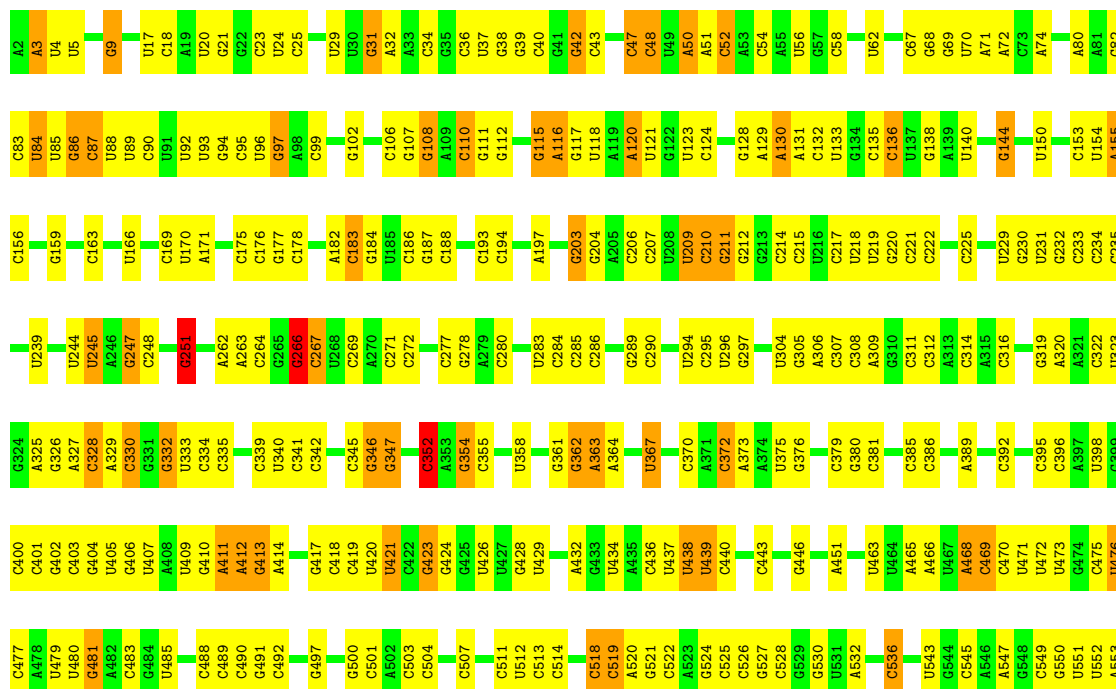
- Molecule 2: Translation initiation factor IF-2

Chain f:  91% 8%

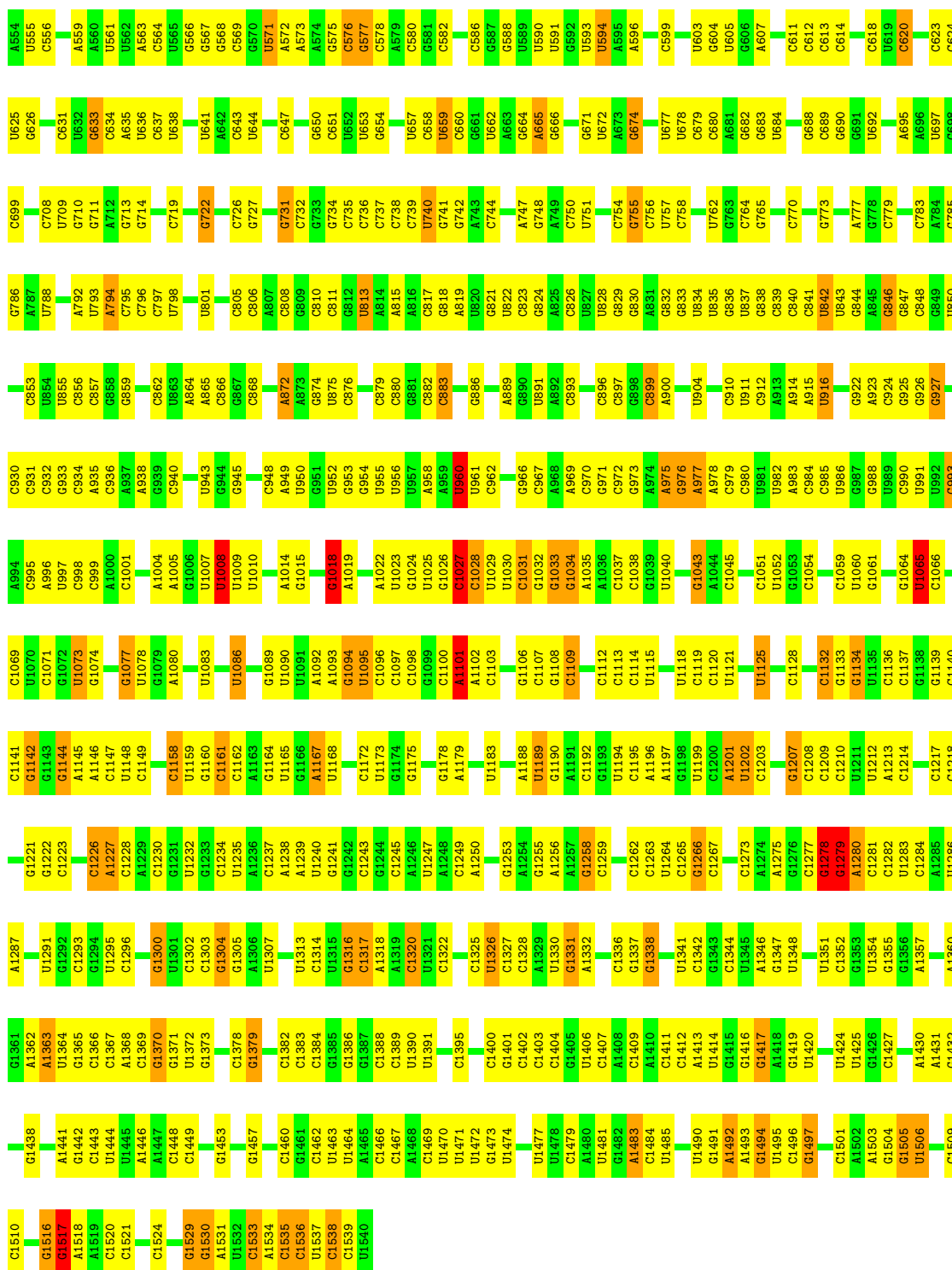



- Molecule 3: 16S ribosomal RNA

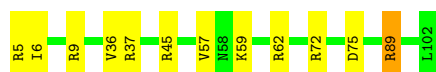
Chain g:  44% 47% 9%







Chain r:  88% 11% .




- Molecule 6: 30S ribosomal protein S11

Chain q:  91% 9%




- Molecule 7: 30S ribosomal protein S12

Chain t:  90% 7% .




- Molecule 8: 30S ribosomal protein S13

Chain s:  89% 11% .



- Molecule 9: 30S ribosomal protein S14

Chain w:  85% 11% .



- Molecule 10: 30S ribosomal protein S15

Chain u:  94% 6%



- Molecule 11: 30S ribosomal protein S16

Chain y:  91% 9%



- Molecule 12: 30S ribosomal protein S18

Chain 1:  87% 9% .



- Molecule 13: 30S ribosomal protein S19

Chain z: 92% 8%



- Molecule 14: 30S ribosomal protein S2

Chain j: 95% . .



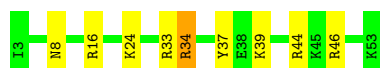
- Molecule 15: 30S ribosomal protein S20

Chain 3: 94% 6%



- Molecule 16: 30S ribosomal protein S21

Chain 2: 82% 16% .



- Molecule 17: 30S ribosomal protein S3

Chain h: 93% 6% .



- Molecule 18: 30S ribosomal protein S4

Chain l: 90% 9%



- Molecule 19: 30S ribosomal protein S5

Chain k: 93% 5% .



- Molecule 20: 30S ribosomal protein S6

Chain n:  92% 8%



- Molecule 21: 30S ribosomal protein S7

Chain m:  94% 6%




- Molecule 22: 30S ribosomal protein S8

Chain p:  92% 7%



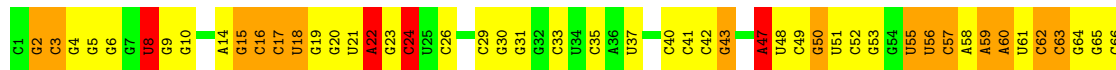
- Molecule 23: 30S ribosomal protein S9

Chain o:  83% 14%



- Molecule 24: tRNA

Chain v:  22% 49% 23% 5%



- Molecule 25: mRNA

Chain N:  17% 50% 33%



## 4 Experimental information

Property	Value	Source
Reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, C1	Depositor
Number of particles used	86367	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	PHASE FLIPPING AND AMPLITUDE CORRECTION	Depositor
Microscope	FEI TECNAI F30	Depositor
Voltage (kV)	300	Depositor
Electron dose ( $e^-/\text{\AA}^2$ )	35	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 ⓘ

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	5	0.70	0/580	1.14	5/782 (0.6%)
10	u	0.69	0/722	1.13	5/964 (0.5%)
11	y	0.70	0/659	1.13	3/884 (0.3%)
12	1	0.73	0/463	1.34	8/621 (1.3%)
13	z	0.68	0/653	1.14	6/877 (0.7%)
14	j	0.67	0/1736	1.01	5/2338 (0.2%)
15	3	0.67	0/671	1.09	5/888 (0.6%)
16	2	0.78	0/431	1.33	4/570 (0.7%)
17	h	0.68	0/1652	1.07	11/2225 (0.5%)
18	l	0.71	0/1665	1.15	13/2227 (0.6%)
19	k	0.67	0/1119	1.06	5/1504 (0.3%)
2	f	0.66	0/3895	1.01	17/5264 (0.3%)
20	n	0.69	0/836	1.12	7/1128 (0.6%)
21	m	0.70	0/1196	1.16	12/1602 (0.7%)
22	p	0.67	0/989	1.08	7/1326 (0.5%)
23	o	0.73	0/1034	1.37	23/1375 (1.7%)
24	v	1.21	2/1831 (0.1%)	1.74	70/2853 (2.5%)
25	N	1.37	0/141	2.23	8/218 (3.7%)
3	g	1.07	1/36963 (0.0%)	1.45	759/57662 (1.3%)
4	P	0.69	0/658	1.10	4/881 (0.5%)
5	r	0.67	0/797	1.15	7/1077 (0.6%)
6	q	0.68	0/893	1.09	5/1205 (0.4%)
7	t	0.69	0/969	1.21	12/1300 (0.9%)
8	s	0.68	0/893	1.21	10/1193 (0.8%)
9	w	0.69	0/785	1.29	12/1043 (1.2%)
All	All	0.95	3/62231 (0.0%)	1.36	1023/92007 (1.1%)

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.

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	#Chirality outliers	#Planarity outliers
11	y	0	1
12	1	0	2
14	j	0	1
16	2	0	3
17	h	0	3
18	l	0	5
19	k	0	4
2	f	0	3
21	m	0	1
22	p	0	2
23	o	0	4
24	v	0	15
25	N	0	2
3	g	0	108
4	P	0	1
5	r	0	1
6	q	0	2
7	t	0	2
8	s	0	3
All	All	0	163

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
24	v	22	A	N9-C4	-7.46	1.33	1.37
24	v	22	A	C3'-C2'	5.83	1.59	1.52
3	g	1228	C	P-O5'	-5.20	1.54	1.59

The worst 5 of 1023 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
24	v	56	U	P-O3'-C3'	16.86	139.93	119.70
25	N	14	A	P-O3'-C3'	13.63	136.06	119.70
3	g	325	A	P-O3'-C3'	10.67	132.50	119.70
3	g	1201	A	P-O3'-C3'	10.60	132.42	119.70
24	v	24	C	C6-N1-C2	-10.25	116.20	120.30

There are no chirality outliers.

5 of 163 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	f	607	ARG	Sidechain
2	f	610	ARG	Sidechain
2	f	772	TYR	Sidechain
3	g	3	A	Sidechain
3	g	42	G	Sidechain

## 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	5	570	0	598	0	0
2	f	3847	0	3896	0	0
3	g	33012	0	16618	0	0
4	P	649	0	691	0	0
5	r	787	0	828	0	0
6	q	877	0	887	0	0
7	t	955	0	1018	0	0
8	s	884	0	944	0	0
9	w	774	0	827	0	0
10	u	714	0	737	0	0
11	y	649	0	666	0	0
12	1	456	0	478	0	0
13	z	638	0	665	0	0
14	j	1705	0	1732	0	0
15	3	665	0	714	0	0
16	2	426	0	449	0	0
17	h	1625	0	1699	0	0
18	l	1643	0	1710	0	0
19	k	1106	0	1148	0	0
20	n	818	0	808	0	0
21	m	1182	0	1240	0	0
22	p	979	0	1034	0	0
23	o	1022	0	1070	0	0
24	v	1639	0	837	0	0
25	N	126	0	66	0	0
All	All	57748	0	41360	0	0

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). Clashscore could not be calculated for this entry.



There are no clashes within the asymmetric unit.

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	5	69/71 (97%)	64 (93%)	4 (6%)	1 (1%)	12	51
2	f	507/509 (100%)	465 (92%)	33 (6%)	9 (2%)	9	47
4	P	78/80 (98%)	69 (88%)	8 (10%)	1 (1%)	13	54
5	r	96/98 (98%)	80 (83%)	11 (12%)	5 (5%)	2	25
6	q	115/117 (98%)	105 (91%)	9 (8%)	1 (1%)	19	60
7	t	121/123 (98%)	111 (92%)	7 (6%)	3 (2%)	6	40
8	s	112/114 (98%)	106 (95%)	5 (4%)	1 (1%)	19	60
9	w	92/100 (92%)	82 (89%)	10 (11%)	0	100	100
10	u	86/88 (98%)	81 (94%)	5 (6%)	0	100	100
11	y	80/82 (98%)	73 (91%)	4 (5%)	3 (4%)	3	31
12	1	53/55 (96%)	52 (98%)	1 (2%)	0	100	100
13	z	77/79 (98%)	71 (92%)	6 (8%)	0	100	100
14	j	216/218 (99%)	200 (93%)	12 (6%)	4 (2%)	9	46
15	3	83/85 (98%)	82 (99%)	1 (1%)	0	100	100
16	2	49/51 (96%)	37 (76%)	9 (18%)	3 (6%)	1	22
17	h	204/206 (99%)	196 (96%)	6 (3%)	2 (1%)	17	58
18	l	203/205 (99%)	190 (94%)	10 (5%)	3 (2%)	11	51
19	k	148/150 (99%)	136 (92%)	8 (5%)	4 (3%)	5	39
20	n	98/100 (98%)	92 (94%)	3 (3%)	3 (3%)	4	36
21	m	149/151 (99%)	142 (95%)	7 (5%)	0	100	100
22	p	127/129 (98%)	123 (97%)	4 (3%)	0	100	100

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
23	o	125/127 (98%)	114 (91%)	6 (5%)	5 (4%)	3	30
All	All	2888/2938 (98%)	2671 (92%)	169 (6%)	48 (2%)	14	49

5 of 48 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	f	674	LEU
2	f	685	GLU
5	r	57	VAL
7	t	33	CYS
16	2	39	LYS

### 5.3.2 Protein sidechains ⓘ

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all 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	5	62/62 (100%)	60 (97%)	2 (3%)	42	68
2	f	409/409 (100%)	384 (94%)	25 (6%)	20	54
4	P	74/74 (100%)	72 (97%)	2 (3%)	48	72
5	r	86/86 (100%)	85 (99%)	1 (1%)	74	87
6	q	90/90 (100%)	87 (97%)	3 (3%)	41	68
7	t	103/103 (100%)	102 (99%)	1 (1%)	78	89
8	s	92/92 (100%)	88 (96%)	4 (4%)	32	63
9	w	79/83 (95%)	78 (99%)	1 (1%)	71	86
10	u	76/76 (100%)	75 (99%)	1 (1%)	71	86
11	y	65/65 (100%)	64 (98%)	1 (2%)	67	84
12	1	48/48 (100%)	47 (98%)	1 (2%)	56	78
13	z	70/70 (100%)	68 (97%)	2 (3%)	45	70
14	j	180/180 (100%)	176 (98%)	4 (2%)	55	77
15	3	65/65 (100%)	65 (100%)	0	100	100
16	2	44/44 (100%)	43 (98%)	1 (2%)	53	76

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
17	h	170/170 (100%)	168 (99%)	2 (1%)	74	87
18	l	172/172 (100%)	168 (98%)	4 (2%)	53	76
19	k	113/113 (100%)	112 (99%)	1 (1%)	81	90
20	n	87/87 (100%)	87 (100%)	0	100	100
21	m	124/124 (100%)	123 (99%)	1 (1%)	83	91
22	p	104/104 (100%)	100 (96%)	4 (4%)	36	65
23	o	105/105 (100%)	101 (96%)	4 (4%)	36	65
All	All	2418/2422 (100%)	2353 (97%)	65 (3%)	51	72

5 of 65 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
5	r	59	LYS
8	s	65	GLU
22	p	55	LYS
6	q	14	GLN
7	t	35	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

### 5.3.3 RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
24	v	76/77 (98%)	26 (34%)	0
25	N	5/6 (83%)	5 (100%)	2 (40%)
3	g	1538/1539 (99%)	202 (13%)	0
All	All	1619/1622 (99%)	233 (14%)	2 (0%)

5 of 233 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
3	g	4	U
3	g	5	U
3	g	9	G
3	g	31	G
3	g	32	A

All (2) RNA pucker outliers are listed below:

Mol	Chain	Res	Type
25	N	14	A
25	N	16	A

## 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](#)

There are no ligands in this entry.

## 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.