



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

Jan 5, 2017 – 02:04 PM EST

PDB ID : 5ME1  
EMDB ID: : EMD-3495  
Title : Structure of the 30S Pre-Initiation Complex 2 (30S IC-2) Stalled by GE81112  
Authors : Lopez-Alonso, J.P.; Fabbretti, A.; Kaminishi, T.; Iturrioz, I.; Brandi, L.; Gil  
Carton, D.; Gualerzi, C.; Fucini, P.; Connell, S.  
Deposited on : 2016-11-14  
Resolution : 13.50 Å(reported)  
Based on PDB ID : 2IFE, 4YBB, 3JCN, 1HR0, 1TIF

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

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

---

MolProbity : 4.02b-467  
Mogul : 1.7.1 (RC1), CSD as537be (2016)  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)  
Validation Pipeline (wwPDB-VP) : rb-20028442

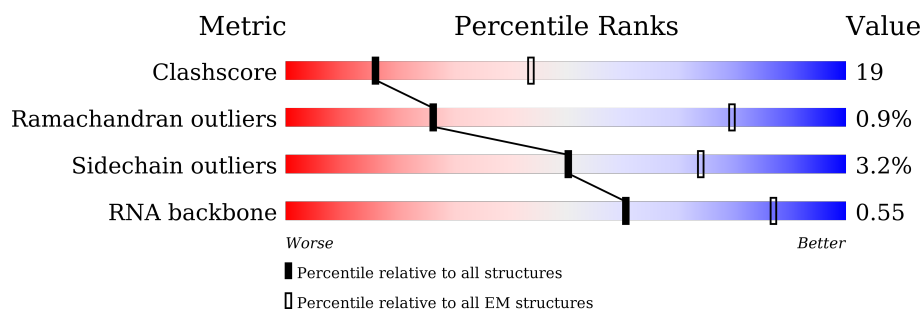
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*ELECTRON MICROSCOPY*

The reported resolution of this entry is 13.50 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.













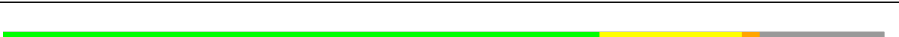


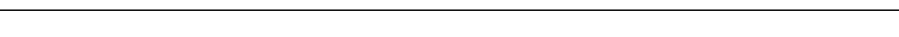
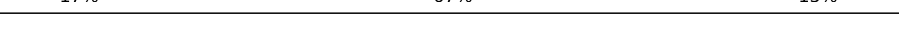
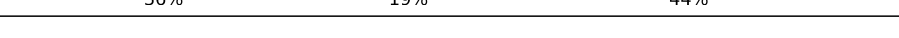

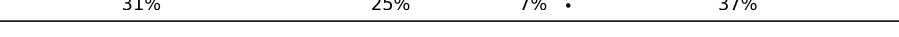
Metric	Whole archive (#Entries)	EM structures (#Entries)
Clashscore	114402	924
Ramachandran outliers	111179	726
Sidechain outliers	111093	686
RNA backbone	3027	244

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	A	1534	70% 24% 6%
2	B	241	71% 20% 7%
3	C	233	73% 15% 12%
4	D	206	87% 12%
5	E	167	66% 25% 7%
6	F	131	69% 11% 19%
7	G	156	73% 22% ..
8	H	130	78% 20% ..

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
9	I	130	
10	J	103	
11	K	129	
12	L	123	
13	M	118	
14	N	101	
15	O	89	
16	P	102	
17	Q	84	
18	R	75	
19	S	92	
20	T	87	
21	U	71	
22	V	72	
23	W	890	
24	Y	171	
25	Z	144	
26	X	77	

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
1	2MG	A	966	-	-	X	-
26	4SU	X	8	-	-	X	-
27	FME	X	101	-	-	X	-

## 2 Entry composition

There are 27 unique types of molecules in this entry. The entry contains 59063 atoms, of which 1 is hydrogen 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 ribosomal RNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	1534	Total	C	N	O	P	0	0
			32930	14694	6041	10661	1534		

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

Mol	Chain	Residues	Atoms					AltConf	Trace
2	B	224	Total	C	N	O	S	0	0
			1753	1109	315	321	8		

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

Mol	Chain	Residues	Atoms					AltConf	Trace
3	C	206	Total	C	N	O	S	0	0
			1624	1028	305	288	3		

- 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	155	Total	C	N	O	S	0	0
			1144	711	216	211	6		

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

Mol	Chain	Residues	Atoms					AltConf	Trace
6	F	106	Total	C	N	O	S	0	0
			862	545	156	154	7		

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

Mol	Chain	Residues	Atoms					AltConf	Trace
7	G	151	Total	C	N	O	S	0	0
			1181	735	227	215	4		

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

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

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

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

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

Mol	Chain	Residues	Atoms					AltConf	Trace
10	J	99	Total	C	N	O	S	0	0
			795	498	152	144	1		

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

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

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

Mol	Chain	Residues	Atoms					AltConf	Trace
12	L	123	Total	C	N	O	S	0	0
			957	591	196	165	5		

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

Mol	Chain	Residues	Atoms					AltConf	Trace
13	M	114	Total	C	N	O	S	0	0
			883	546	178	156	3		

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

Mol	Chain	Residues	Atoms					AltConf	Trace
14	N	100	Total	C	N	O	S	0	0
			805	499	164	139	3		

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

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

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

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

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

Mol	Chain	Residues	Atoms					AltConf	Trace
17	Q	80	Total	C	N	O	S	0	0
			648	411	121	113	3		

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

Mol	Chain	Residues	Atoms				AltConf	Trace
18	R	55	Total	C	N	O	0	0
			455	288	86	81		

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

Mol	Chain	Residues	Atoms					AltConf	Trace
19	S	79	Total	C	N	O	S	0	0
			637	408	120	107	2		

- Molecule 20 is a protein called 30S ribosomal protein S20.

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

- Molecule 21 is a protein called 30S ribosomal protein S21.

Mol	Chain	Residues	Atoms					AltConf	Trace
21	U	56	Total	C	N	O	S	0	0
			465	290	96	78	1		

- Molecule 22 is a protein called Translation initiation factor IF-1.

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

- Molecule 23 is a protein called Translation initiation factor IF-2.

Mol	Chain	Residues	Atoms					AltConf	Trace
23	W	501	Total	C	N	O	S	0	0
			3781	2368	663	735	15		

- Molecule 24 is a protein called Translation initiation factor IF-3.

Mol	Chain	Residues	Atoms					AltConf	Trace
24	Y	76	Total	C	N	O	S	0	0
			623	390	119	112	2		

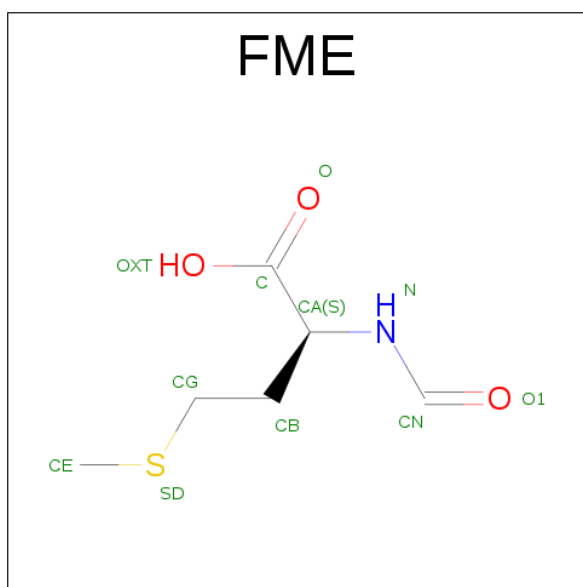
- Molecule 25 is a protein called Translation initiation factor IF-3.

Mol	Chain	Residues	Atoms					AltConf	Trace
25	Z	91	Total	C	N	O	S	0	0
			743	470	135	134	4		

- Molecule 26 is a RNA chain called fMet-tRNA.

Mol	Chain	Residues	Atoms					AltConf	Trace	
26	X	77	Total	C	N	O	P	S	0	0
			1643	733	297	535	77	1		

- Molecule 27 is N-FORMYLMETHIONINE (three-letter code: FME) (formula: C<sub>6</sub>H<sub>11</sub>NO<sub>3</sub>S).



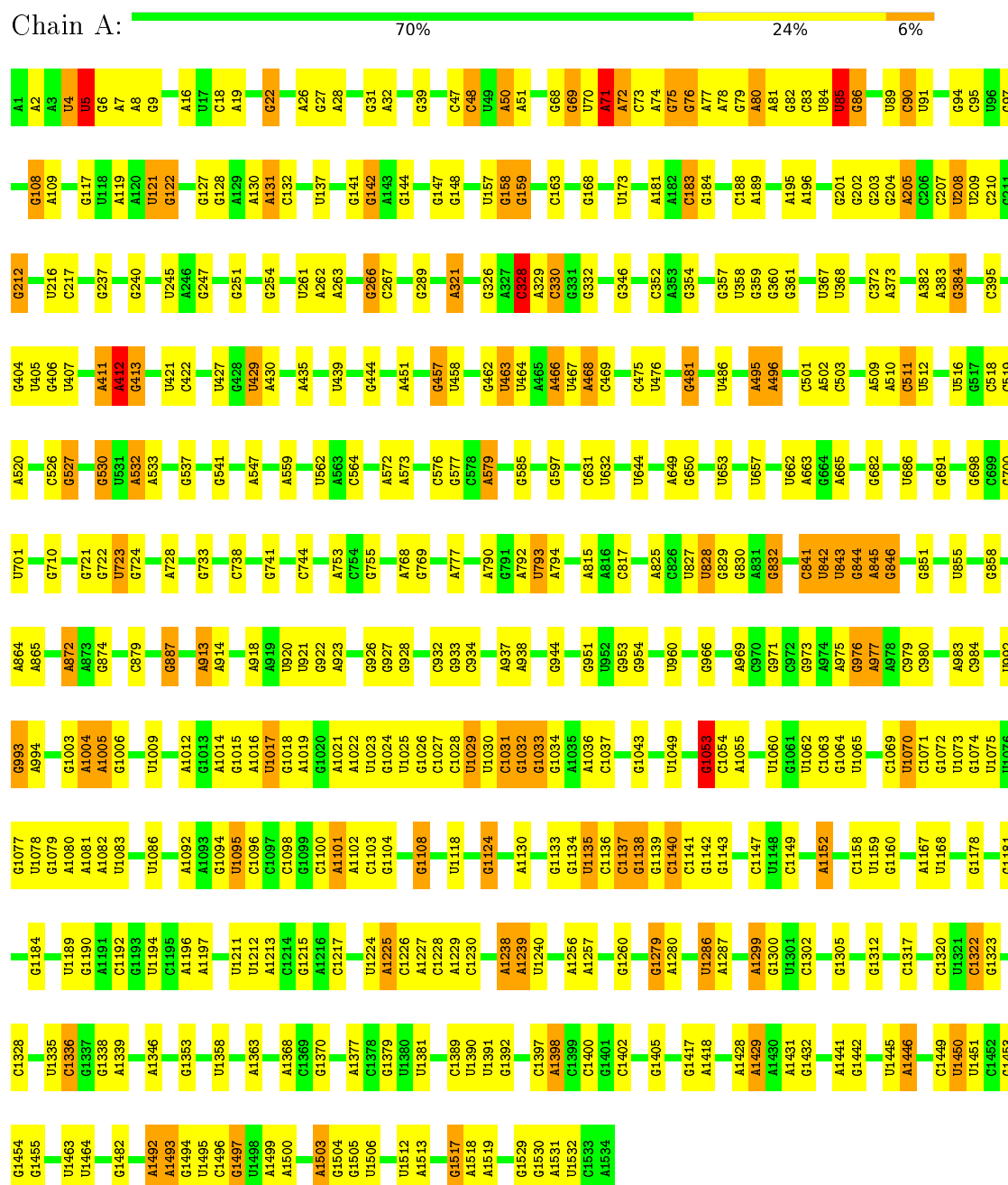
Mol	Chain	Residues	Atoms						AltConf
			Total	C	H	N	O	S	
27	X	1	10	6	1	1	1	1	0



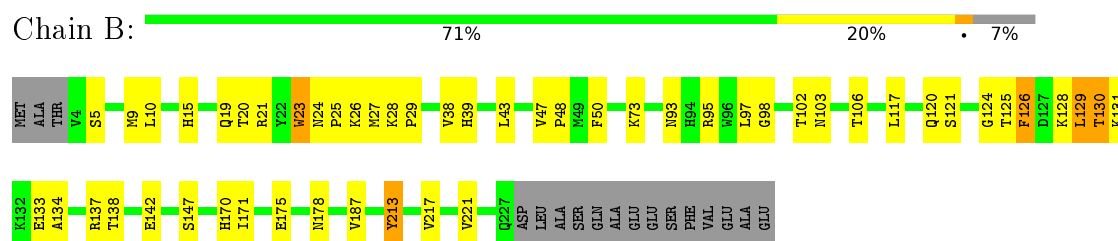
### 3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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.

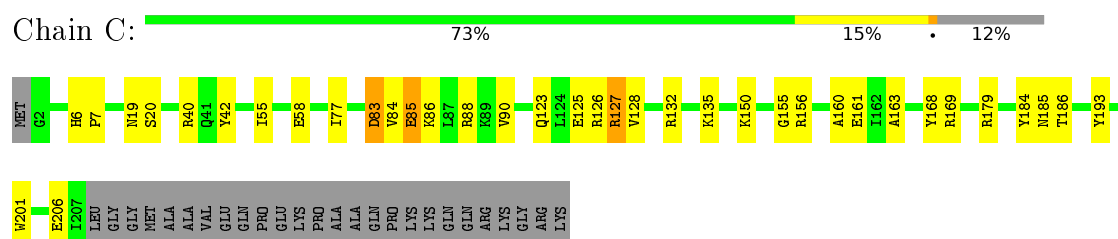
#### • Molecule 1: 16S ribosomal RNA



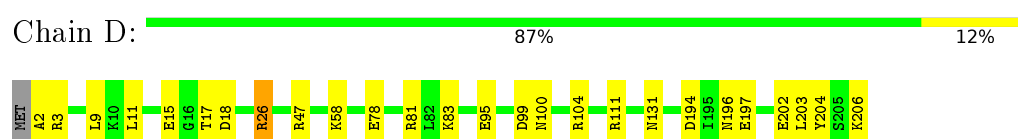
- Molecule 2: 30S ribosomal protein S2



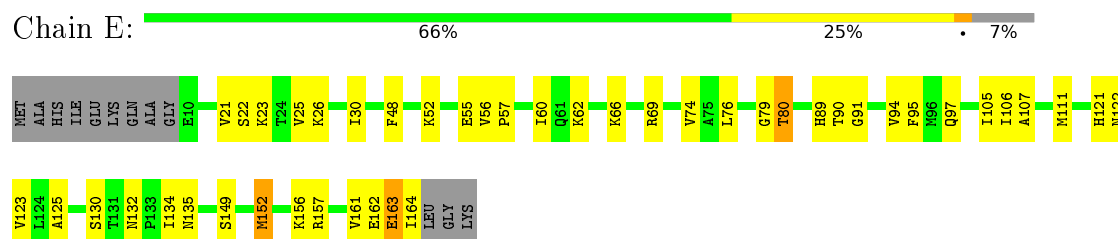
- Molecule 3: 30S ribosomal protein S3



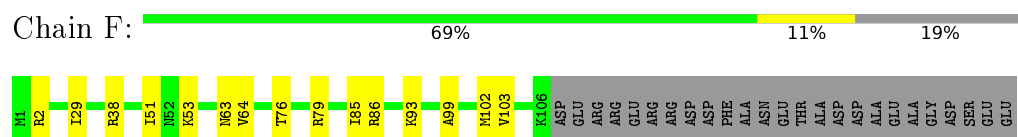
- Molecule 4: 30S ribosomal protein S4



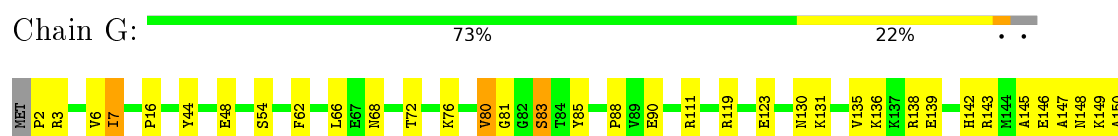
- Molecule 5: 30S ribosomal protein S5

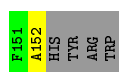


- Molecule 6: 30S ribosomal protein S6



- Molecule 7: 30S ribosomal protein S7





- Molecule 8: 30S ribosomal protein S8

Chain H: 78% 20% ..



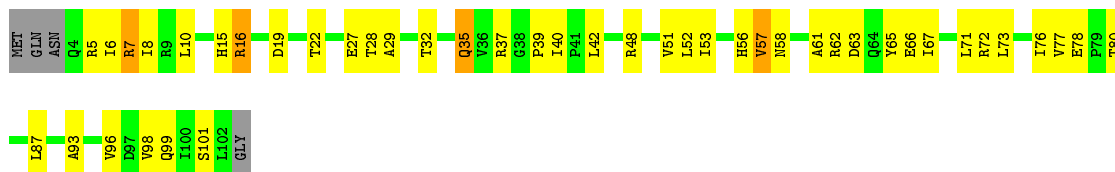
- Molecule 9: 30S ribosomal protein S9

Chain I: 72% 26% .



- Molecule 10: 30S ribosomal protein S10

Chain J: 53% 39% . .



- Molecule 11: 30S ribosomal protein S11

Chain K: 71% 20% 9%



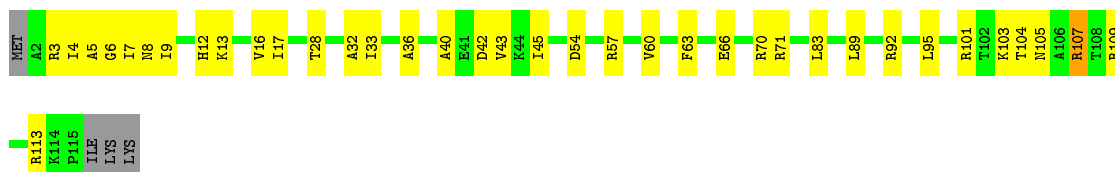
- Molecule 12: 30S ribosomal protein S12

Chain L: 86% 14%

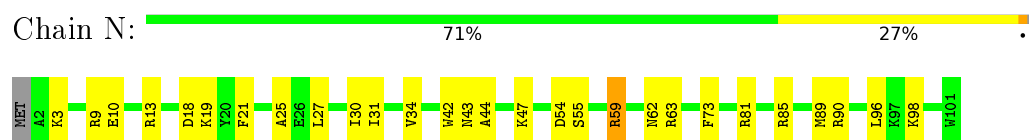


- Molecule 13: 30S ribosomal protein S13

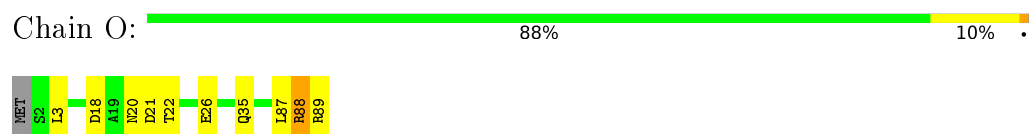
Chain M: 65% 31% . .



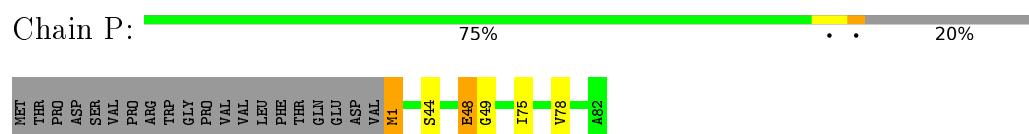
- Molecule 14: 30S ribosomal protein S14



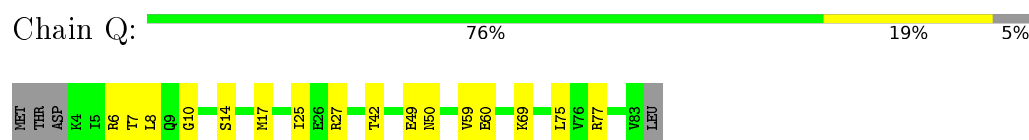
- Molecule 15: 30S ribosomal protein S15



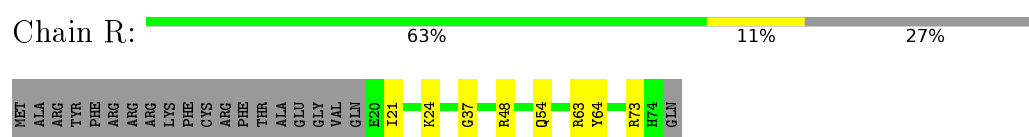
- Molecule 16: 30S ribosomal protein S16



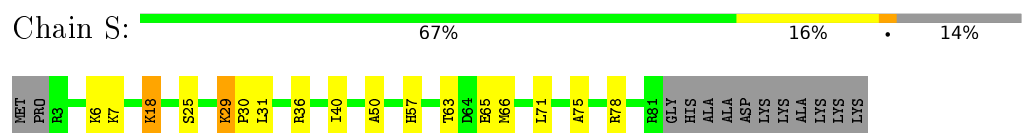
- Molecule 17: 30S ribosomal protein S17



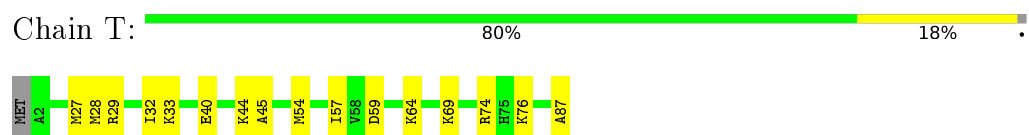
- Molecule 18: 30S ribosomal protein S18



- Molecule 19: 30S ribosomal protein S19

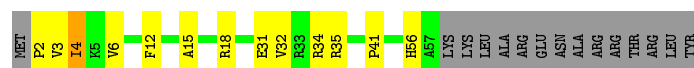


- Molecule 20: 30S ribosomal protein S20



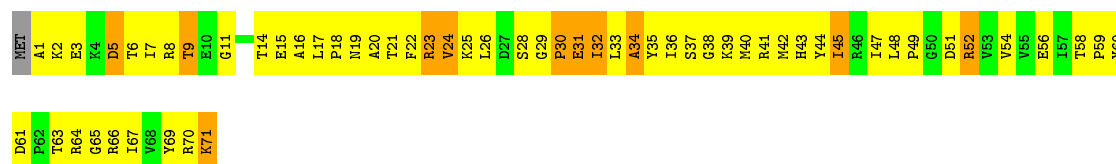
- Molecule 21: 30S ribosomal protein S21

Chain U:  61% 17% • 21%



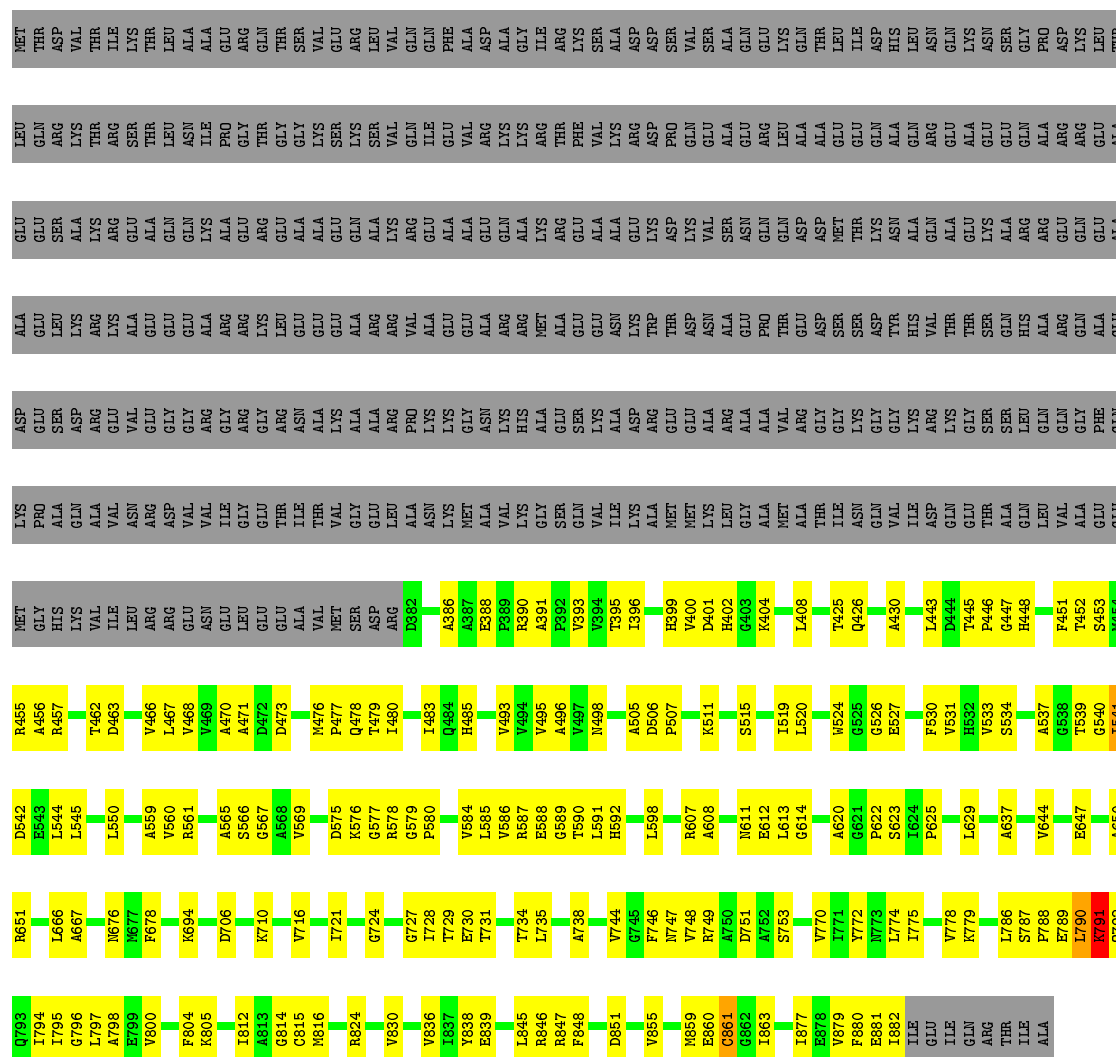
- Molecule 22: Translation initiation factor IF-1

Chain V:  17% 67% 15%



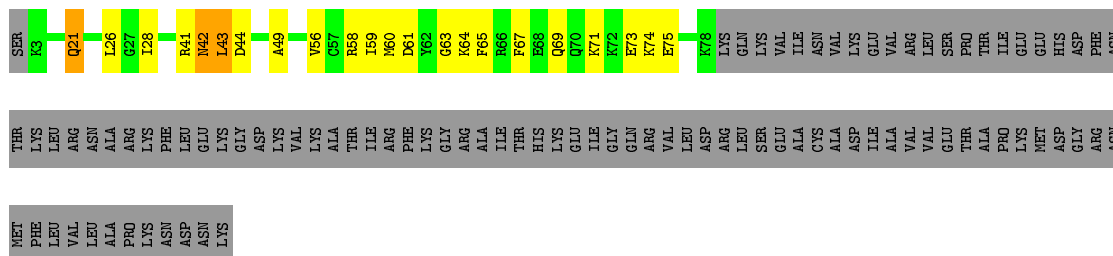
- Molecule 23: Translation initiation factor IF-2

Chain W: 




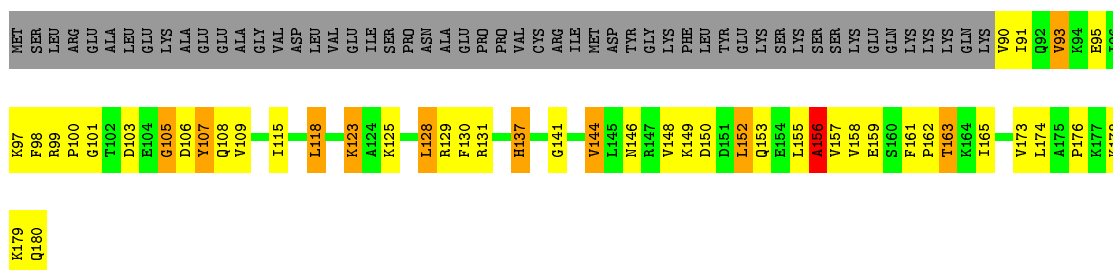
- Molecule 24: Translation initiation factor IF-3

Chain Y:  32% 11% 56%

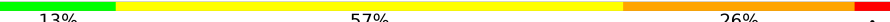


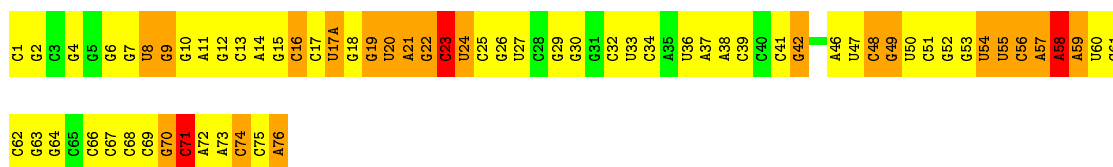
• Molecule 25: Translation initiation factor IF-3

Chain Z:  31% 25% 7% 37%



• Molecule 26: fMet-tRNA

Chain X:  13% 57% 26%



## 4 Experimental information

Property	Value	Source
Reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	Depositor
Number of particles used	26776	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	Not provided	Depositor
Microscope	JEOL 2200FSC	Depositor
Voltage (kV)	200	Depositor
Electron dose ( $e^-/\text{\AA}^2$ )	Not provided	Depositor
Minimum defocus (nm)	Not provided	Depositor
Maximum defocus (nm)	Not provided	Depositor
Magnification	74183	Depositor
Image detector	Not provided	Depositor

## 5 Model quality

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: 5MU, FME, MA6, G7M, D2T, H2U, 2MG, 5MC, UR3, 4OC, 4SU, 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	$\# Z  > 2$	RMSZ	$\# Z  > 2$
1	A	0.58	0/36591	0.98	33/57073 (0.1%)
10	J	0.57	0/805	0.68	0/1089
11	K	0.39	0/893	0.54	0/1205
12	L	0.49	0/960	0.65	0/1286
13	M	0.38	0/892	0.63	0/1193
14	N	0.40	0/817	0.60	0/1088
15	O	0.42	0/722	0.52	0/964
16	P	0.40	0/659	0.59	0/884
17	Q	0.44	0/657	0.61	0/881
18	R	0.41	0/462	0.55	0/621
19	S	0.38	0/652	0.56	0/877
2	B	0.36	0/1784	0.55	0/2403
20	T	0.40	0/676	0.53	0/895
21	U	0.53	0/472	0.57	0/627
22	V	0.39	0/580	0.70	0/782
23	W	0.25	0/3829	0.46	1/5176 (0.0%)
24	Y	0.32	0/629	0.62	0/838
25	Z	1.18	0/751	1.61	13/999 (1.3%)
26	X	0.22	0/1746	0.81	5/2721 (0.2%)
3	C	0.39	0/1651	0.55	0/2225
4	D	0.37	0/1665	0.53	0/2227
5	E	0.48	0/1157	0.61	0/1557
6	F	0.43	0/881	0.56	0/1189
7	G	0.35	0/1195	0.51	0/1602
8	H	0.43	0/989	0.58	0/1326
9	I	0.37	0/1034	0.60	0/1375
All	All	0.52	0/63149	0.86	52/93103 (0.1%)

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.

Mol	Chain	#Chirality outliers	#Planarity outliers
25	Z	0	6

There are no bond length outliers.

The worst 5 of 52 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	117	G	O5'-P-OP2	-9.51	97.14	105.70
25	Z	107	TYR	CB-CG-CD2	-8.11	116.13	121.00
26	X	71	C	N1-C1'-C2'	-7.46	103.80	112.00
25	Z	144	VAL	CA-CB-CG2	-7.42	99.77	110.90
25	Z	156	ALA	CB-CA-C	-6.93	99.71	110.10

There are no chirality outliers.

5 of 6 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
25	Z	103	ASP	Peptide
25	Z	107	TYR	Sidechain
25	Z	137	HIS	Peptide
25	Z	156	ALA	Peptide
25	Z	161	PHE	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	A	32930	0	16565	918	0
2	B	1753	0	1775	134	0
3	C	1624	0	1696	37	0
4	D	1643	0	1707	21	0
5	E	1144	0	1182	118	0
6	F	862	0	864	10	0
7	G	1181	0	1236	110	0
8	H	979	0	1031	18	0
9	I	1022	0	1068	36	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
10	J	795	0	836	39	0
11	K	877	0	887	103	0
12	L	957	0	1016	53	0
13	M	883	0	941	32	0
14	N	805	0	844	27	0
15	O	714	0	734	5	0
16	P	649	0	666	3	0
17	Q	648	0	691	11	0
18	R	455	0	478	7	0
19	S	637	0	665	16	0
20	T	670	0	719	12	0
21	U	465	0	491	8	0
22	V	570	0	592	201	0
23	W	3781	0	3836	318	0
24	Y	623	0	651	52	0
25	Z	743	0	773	185	0
26	X	1643	0	831	380	0
27	X	9	1	10	14	0
All	All	59062	1	42785	1894	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 19.

The worst 5 of 1894 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:A:790:A:C8	26:X:38:A:H4'	1.24	1.71
1:A:790:A:C8	26:X:38:A:C4'	1.77	1.68
12:L:51:LYS:HE3	22:V:60:TYR:CD2	1.28	1.67
1:A:828:U:C2	2:B:25:PRO:CD	1.78	1.66
25:Z:91:ILE:HB	26:X:13:C:C4'	1.23	1.66

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	
2	B	222/241 (92%)	209 (94%)	11 (5%)	2 (1%)	21	67
3	C	204/233 (88%)	193 (95%)	10 (5%)	1 (0%)	34	77
4	D	203/206 (98%)	199 (98%)	4 (2%)	0	100	100
5	E	153/167 (92%)	145 (95%)	8 (5%)	0	100	100
6	F	104/131 (79%)	96 (92%)	8 (8%)	0	100	100
7	G	149/156 (96%)	138 (93%)	10 (7%)	1 (1%)	26	71
8	H	127/130 (98%)	119 (94%)	8 (6%)	0	100	100
9	I	125/130 (96%)	111 (89%)	13 (10%)	1 (1%)	24	69
10	J	97/103 (94%)	88 (91%)	8 (8%)	1 (1%)	19	65
11	K	115/129 (89%)	104 (90%)	11 (10%)	0	100	100
12	L	120/123 (98%)	111 (92%)	9 (8%)	0	100	100
13	M	112/118 (95%)	101 (90%)	8 (7%)	3 (3%)	6	45
14	N	98/101 (97%)	91 (93%)	7 (7%)	0	100	100
15	O	86/89 (97%)	80 (93%)	3 (4%)	3 (4%)	4	39
16	P	80/102 (78%)	73 (91%)	6 (8%)	1 (1%)	15	60
17	Q	78/84 (93%)	73 (94%)	5 (6%)	0	100	100
18	R	53/75 (71%)	51 (96%)	1 (2%)	1 (2%)	10	52
19	S	77/92 (84%)	69 (90%)	7 (9%)	1 (1%)	15	60
20	T	84/87 (97%)	83 (99%)	1 (1%)	0	100	100
21	U	54/71 (76%)	53 (98%)	1 (2%)	0	100	100
22	V	69/72 (96%)	47 (68%)	15 (22%)	7 (10%)	1	14
23	W	499/890 (56%)	480 (96%)	17 (3%)	2 (0%)	39	80
24	Y	74/171 (43%)	73 (99%)	1 (1%)	0	100	100
25	Z	89/144 (62%)	76 (85%)	10 (11%)	3 (3%)	5	40
All	All	3072/3845 (80%)	2863 (93%)	182 (6%)	27 (1%)	26	67

5 of 27 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	130	THR
3	C	127	ARG
9	I	25	ASN

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
10	J	57	VAL
13	M	5	ALA

### 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	B	186/199 (94%)	180 (97%)	6 (3%)	46	76
3	C	170/190 (90%)	164 (96%)	6 (4%)	43	74
4	D	172/173 (99%)	167 (97%)	5 (3%)	50	78
5	E	118/126 (94%)	112 (95%)	6 (5%)	29	66
6	F	92/112 (82%)	91 (99%)	1 (1%)	80	91
7	G	124/129 (96%)	121 (98%)	3 (2%)	57	82
8	H	104/105 (99%)	100 (96%)	4 (4%)	40	73
9	I	105/107 (98%)	102 (97%)	3 (3%)	50	78
10	J	87/90 (97%)	81 (93%)	6 (7%)	19	56
11	K	90/99 (91%)	89 (99%)	1 (1%)	80	91
12	L	102/102 (100%)	101 (99%)	1 (1%)	82	92
13	M	92/96 (96%)	90 (98%)	2 (2%)	60	83
14	N	83/84 (99%)	82 (99%)	1 (1%)	78	90
15	O	76/77 (99%)	74 (97%)	2 (3%)	54	80
16	P	65/84 (77%)	63 (97%)	2 (3%)	47	77
17	Q	74/78 (95%)	72 (97%)	2 (3%)	52	79
18	R	48/65 (74%)	48 (100%)	0	100	100
19	S	70/79 (89%)	68 (97%)	2 (3%)	50	78
20	T	65/66 (98%)	63 (97%)	2 (3%)	47	77
21	U	48/61 (79%)	45 (94%)	3 (6%)	22	59
22	V	62/63 (98%)	54 (87%)	8 (13%)	5	28
23	W	402/713 (56%)	398 (99%)	4 (1%)	82	92

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
24	Y	66/149 (44%)	61 (92%)	5 (8%)	16	53
25	Z	81/128 (63%)	73 (90%)	8 (10%)	10	39
All	All	2582/3175 (81%)	2499 (97%)	83 (3%)	50	76

5 of 83 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
10	J	27	GLU
15	O	89	ARG
25	Z	118	LEU
10	J	35	GLN
13	M	71	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 16 such sidechains are listed below:

Mol	Chain	Res	Type
23	W	402	HIS
23	W	426	GLN
23	W	498	ASN
23	W	399	HIS
24	Y	21	GLN

### 5.3.3 RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
1	A	1528/1534 (99%)	257 (16%)	9 (0%)
26	X	76/77 (98%)	24 (31%)	11 (14%)
All	All	1604/1611 (99%)	281 (17%)	20 (1%)

5 of 281 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
1	A	5	U
1	A	6	G
1	A	9	G
1	A	22	G
1	A	28	A

5 of 20 RNA pucker outliers are listed below:

Mol	Chain	Res	Type
26	X	8	4SU
26	X	19	G
26	X	56	C
1	A	1031	C
1	A	1211	U

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

16 non-standard protein/DNA/RNA residues are 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$
1	2MG	A	1207	1	18,26,27	1.14	2 (11%)	21,38,41	2.37	7 (33%)
1	4OC	A	1402	1	15,23,24	0.64	0	21,32,35	1.91	3 (14%)
1	5MC	A	1407	1	14,22,23	1.49	1 (7%)	17,32,35	1.03	1 (5%)
1	UR3	A	1498	1	13,22,23	0.60	0	18,32,35	0.71	0
1	2MG	A	1516	1	18,26,27	1.04	2 (11%)	21,38,41	2.12	7 (33%)
1	MA6	A	1518	1	18,26,27	0.93	1 (5%)	15,38,41	2.00	2 (13%)
1	MA6	A	1519	1	18,26,27	0.90	1 (5%)	15,38,41	2.54	3 (20%)
1	PSU	A	516	1,22	15,21,22	1.11	1 (6%)	16,30,33	1.64	3 (18%)
1	G7M	A	527	1	18,26,27	1.07	1 (5%)	21,39,42	2.26	5 (23%)
1	2MG	A	966	1,26	18,26,27	1.25	2 (11%)	21,38,41	2.27	6 (28%)
1	5MC	A	967	1	14,22,23	1.22	1 (7%)	17,32,35	1.19	1 (5%)
12	D2T	L	89	12	4,9,10	0.56	0	4,11,13	1.16	0
26	H2U	X	20	26	17,21,22	0.81	0	23,30,33	1.33	5 (21%)
26	5MU	X	54	26	13,22,23	0.77	0	16,32,35	2.62	2 (12%)
26	PSU	X	55	26	15,21,22	1.40	1 (6%)	16,30,33	2.39	5 (31%)
26	4SU	X	8	26	12,21,22	1.13	2 (16%)	15,30,33	1.88	6 (40%)

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
1	2MG	A	1207	1	-	0/5/27/28	0/3/3/3
1	4OC	A	1402	1	-	0/7/29/30	0/2/2/2
1	5MC	A	1407	1	-	0/3/25/26	0/2/2/2
1	UR3	A	1498	1	-	0/3/25/26	0/2/2/2
1	2MG	A	1516	1	-	0/5/27/28	0/3/3/3
1	MA6	A	1518	1	-	0/7/29/30	0/3/3/3
1	MA6	A	1519	1	-	0/7/29/30	0/3/3/3
1	PSU	A	516	1,22	-	0/7/25/26	0/2/2/2
1	G7M	A	527	1	-	0/3/25/26	0/3/3/3
1	2MG	A	966	1,26	-	0/5/27/28	0/3/3/3
1	5MC	A	967	1	-	0/3/25/26	0/2/2/2
12	D2T	L	89	12	-	0/2/12/14	0/0/0/0
26	H2U	X	20	26	-	0/7/38/39	0/2/2/2
26	5MU	X	54	26	-	0/3/25/26	0/2/2/2
26	PSU	X	55	26	-	0/7/25/26	0/2/2/2
26	4SU	X	8	26	-	0/3/25/26	0/2/2/2

The worst 5 of 15 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
26	X	55	PSU	C5-C1'	-4.40	1.48	1.52
26	X	8	4SU	C2-N3	-2.17	1.33	1.38
1	A	1519	MA6	C5-C4	2.52	1.46	1.40
26	X	8	4SU	O4'-C1'	2.55	1.44	1.41
1	A	1516	2MG	C5-C4	2.77	1.46	1.40

The worst 5 of 56 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
26	X	54	5MU	C5-C4-N3	-7.49	119.06	125.35
1	A	1519	MA6	N3-C2-N1	-7.34	123.11	128.87
1	A	1402	4OC	CM4-N4-C4	-6.29	117.57	122.87
1	A	1518	MA6	N3-C2-N1	-5.95	124.20	128.87
1	A	527	G7M	C1'-N9-C4	-4.84	121.40	126.81

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

10 monomers are involved in 51 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	A	1402	4OC	1	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	A	1518	MA6	2	0
1	A	1519	MA6	3	0
1	A	516	PSU	2	0
1	A	527	G7M	2	0
1	A	966	2MG	21	0
26	X	20	H2U	5	0
26	X	54	5MU	5	0
26	X	55	PSU	3	0
26	X	8	4SU	12	0

## 5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

1 ligand 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
27	FME	X	101	26	8,8,10	0.40	0	8,8,11	1.47	2 (25%)

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
27	FME	X	101	26	-	0/7/7/11	0/0/0/0

There are no bond length outliers.

All (2) bond angle outliers are listed below:



Mol	Chain	Res	Type	Atoms	Z	Observed( $^{\circ}$ )	Ideal( $^{\circ}$ )
27	X	101	FME	O1-CN-N	-2.65	120.75	124.80
27	X	101	FME	CA-N-CN	-2.64	120.70	124.12

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 14 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
27	X	101	FME	14	0

## 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
1	A	2

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	A	926:G	O3'	927:G	P	5.84
1	A	1390:U	O3'	1391:U	P	4.62