



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

Nov 9, 2016 – 10:52 AM EST

PDB ID : 5LMV  
EMDB ID: : EMD-4083  
Title : Structure of bacterial 30S-IF1-IF2-IF3-mRNA-tRNA translation pre-initiation complex(state-III)  
Authors : Hussain, T.; Llacer, J.L.; Wimberly, B.T.; Ramakrishnan, V.  
Deposited on : 2016-08-01  
Resolution : 4.90 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](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)  
EM map analysis : **NOT EXECUTED**  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)  
Validation Pipeline (wwPDB-VP) : rb-20028320

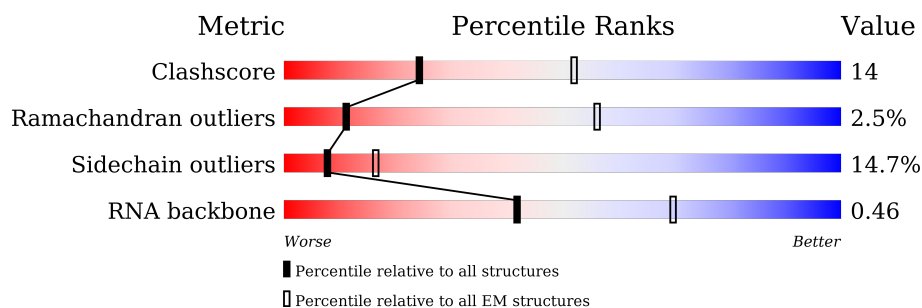
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*ELECTRON MICROSCOPY*

The reported resolution of this entry is 4.90 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.











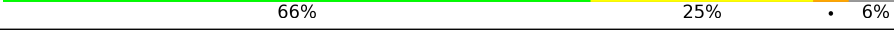

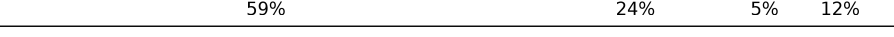


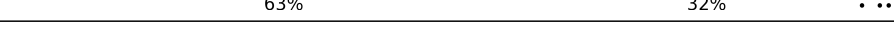




Metric	Whole archive (#Entries)	EM structures (#Entries)
Clashscore	114402	924
Ramachandran outliers	111179	726
Sidechain outliers	111093	686
RNA backbone	3027	244

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	1522	31% 54% 14%
2	B	256	57% 27% 5% 9%
3	C	239	59% 25% 14%
4	D	209	61% 32% 7%
5	E	162	55% 31% 6% 7%
6	F	101	63% 30% 6%
7	G	156	80% 19%
8	H	138	70% 26%

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
9	I	128	
10	J	105	
11	K	129	
12	L	132	
13	M	126	
14	N	61	
15	O	89	
16	P	88	
17	Q	105	
18	R	88	
19	S	93	
20	T	106	
21	V	27	
22	W	72	
23	X	171	
24	Y	42	
25	Z	77	
26	a	571	

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
27	ZN	D	300	-	-	X	-
27	ZN	N	101	-	-	X	-
29	FME	Z	101	-	-	X	-

## 2 Entry composition

There are 29 unique types of molecules in this entry. The entry contains 59470 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 ribosomal RNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	1515	Total	C	N	O	P	0	0
			32548	14490	6022	10523	1513		

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

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

- 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
			1612	1016	314	281	1		

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

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

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

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

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

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

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

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

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

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

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

Mol	Chain	Residues	Atoms				AltConf	Trace
9	I	127	Total	C	N	O	0	0
			1010	639	197	174		

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

Mol	Chain	Residues	Atoms					AltConf	Trace
10	J	98	Total	C	N	O	S	0	0
			792	498	156	137	1		

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

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

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

Mol	Chain	Residues	Atoms					AltConf	Trace
12	L	124	Total	C	N	O	S	0	0
			970	611	195	163	1		

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

Mol	Chain	Residues	Atoms					AltConf	Trace
13	M	118	Total	C	N	O	S	0	0
			937	579	193	163	2		

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

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

- 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
			734	459	147	126	2		

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

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

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

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

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

Mol	Chain	Residues	Atoms				AltConf	Trace
18	R	73	Total	C	N	O	0	0
			598	381	118	99		

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

Mol	Chain	Residues	Atoms					AltConf	Trace
19	S	82	Total	C	N	O	S	0	0
			655	419	120	114	2		

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

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

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

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

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

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

- Molecule 23 is a protein called Translation initiation factor IF-3.

Mol	Chain	Residues	Atoms					AltConf	Trace
23	X	164	Total	C	N	O	S	0	0
			1336	841	245	241	9		

- Molecule 24 is a RNA chain called mRNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
24	Y	20	Total	C	N	O	P	0	0
			439	196	89	134	20		

- Molecule 25 is a RNA chain called tRNAi.

Mol	Chain	Residues	Atoms						AltConf	Trace
25	Z	77	Total	C	N	O	P	S	0	0
			1646	735	297	536	77	1		

- Molecule 26 is a protein called Translation initiation factor IF-2.

Mol	Chain	Residues	Atoms					AltConf	Trace
26	a	502	Total	C	N	O	S	0	0
			3774	2365	679	718	12		

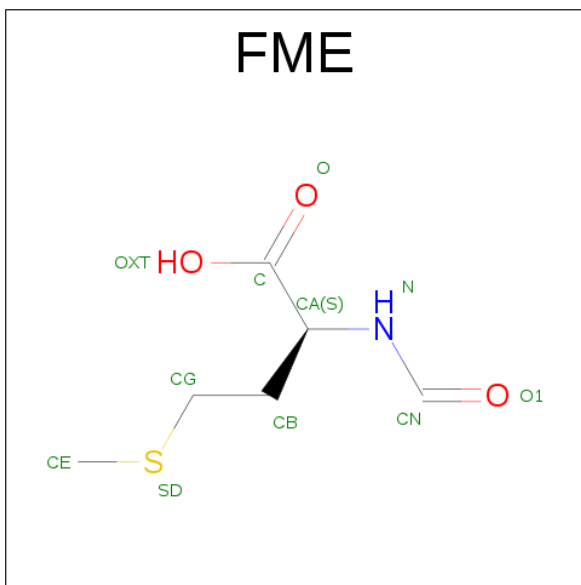
- Molecule 27 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		AltConf
27	D	1	Total	Zn	0
			1	1	
27	N	1	Total	Zn	0
			1	1	

- Molecule 28 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		AltConf
28	W	1	Total	Mg	0
			1	1	
28	Z	1	Total	Mg	0
			1	1	

- Molecule 29 is N-FORMYLMETHIONINE (three-letter code: FME) (formula:  $C_6H_{11}NO_3S$ ).



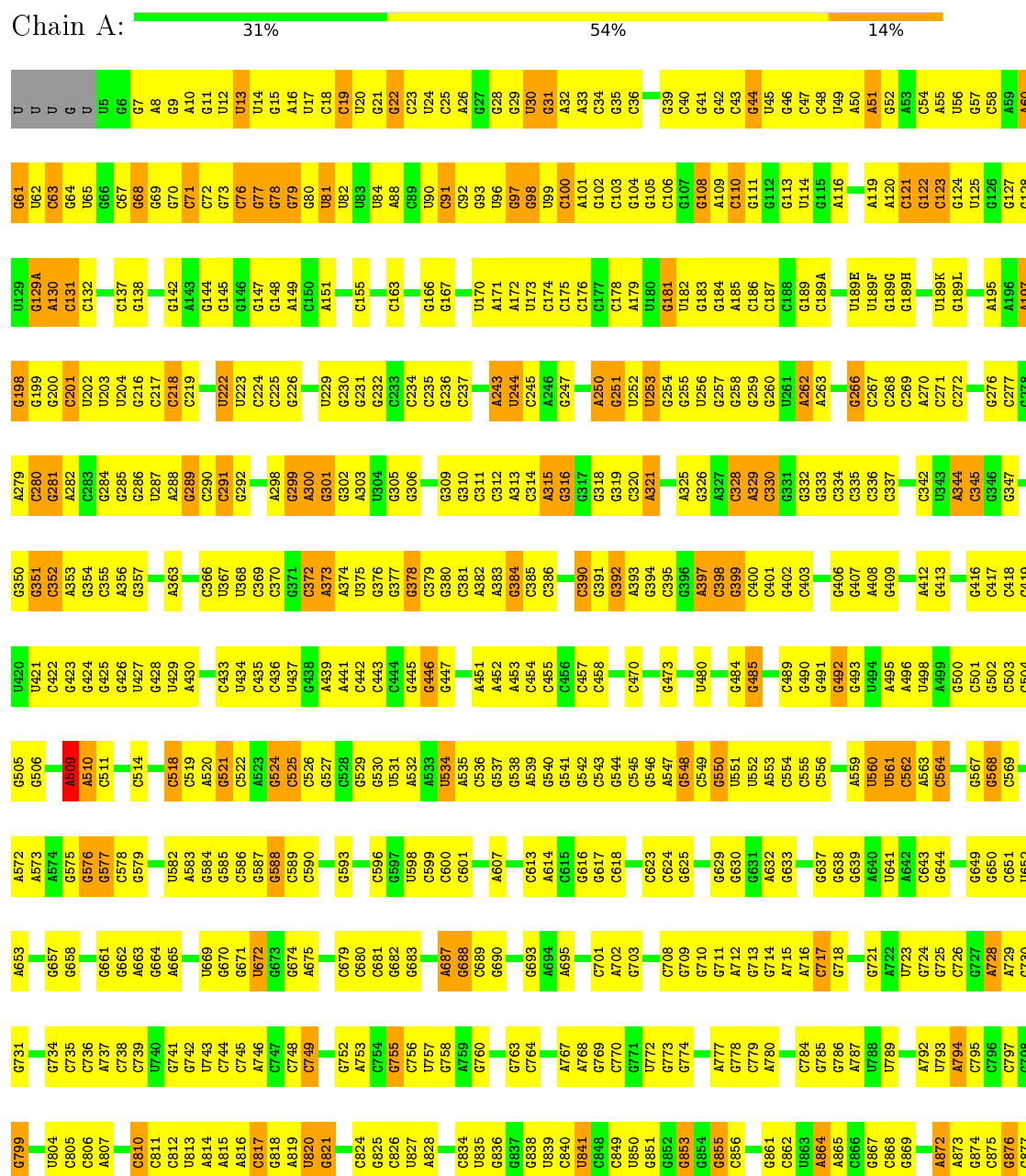
Mol	Chain	Residues	Atoms					AltConf
29	Z	1	Total	C	N	O	S	0
			9	6	1	1	1	

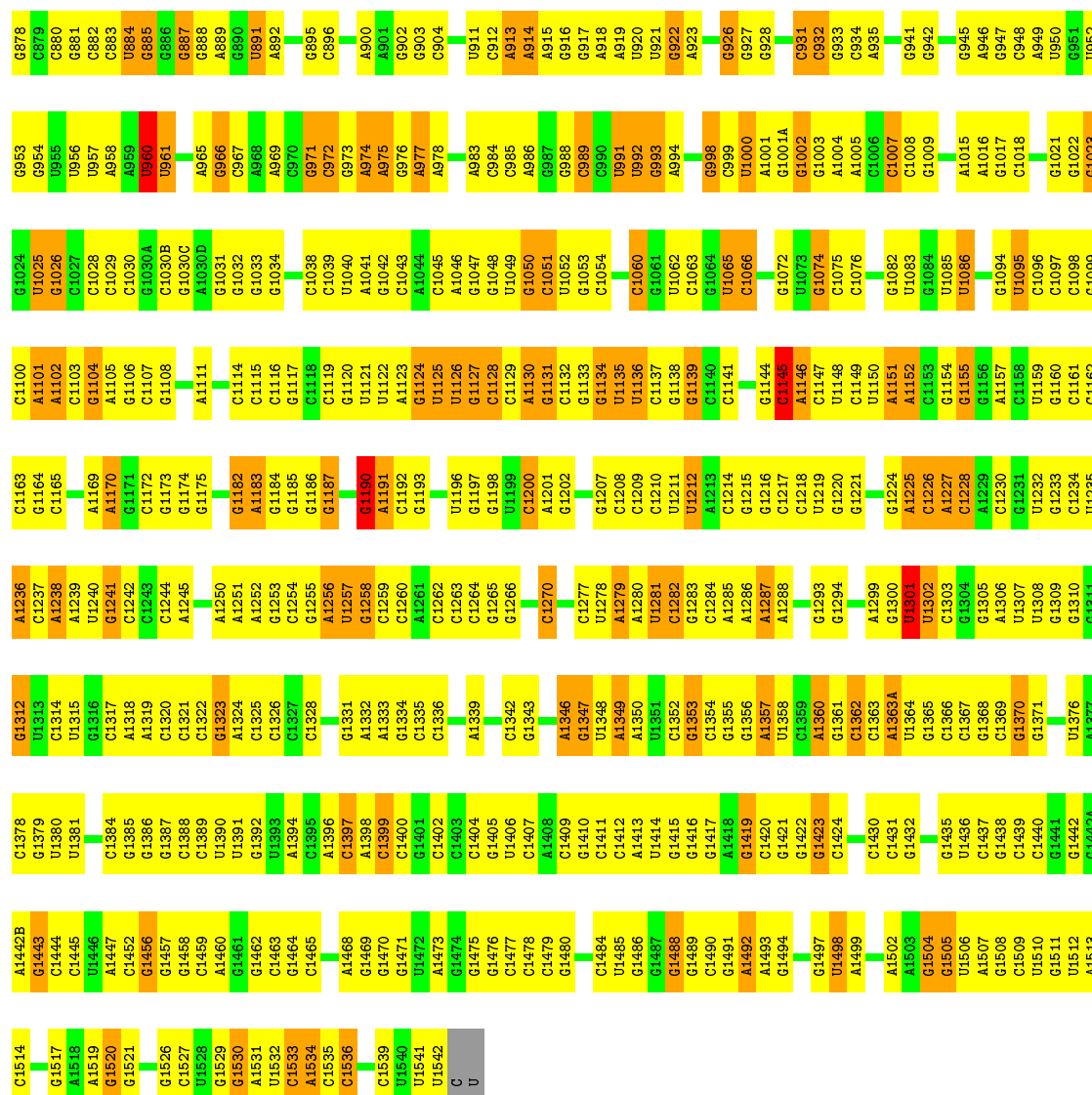


### 3 Residue-property plots

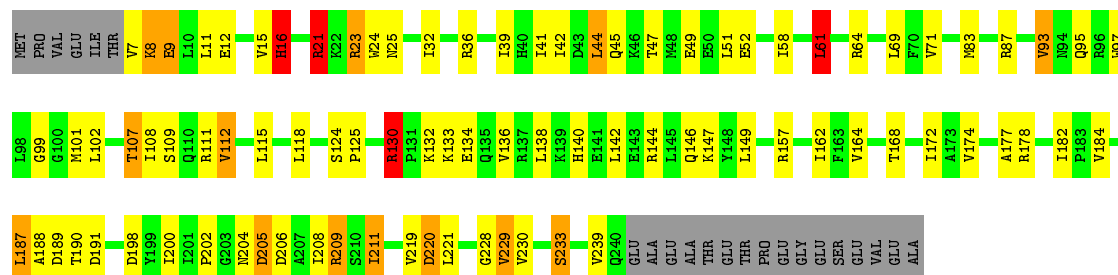
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry. Residues are color-coded according to the number of geometric 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



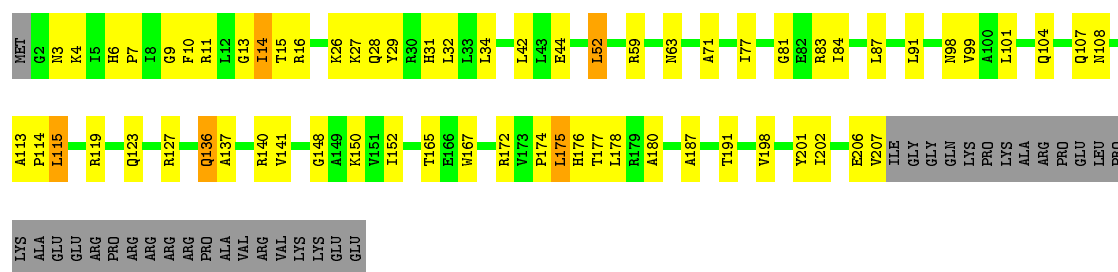


- Molecule 2: 30S ribosomal protein S2



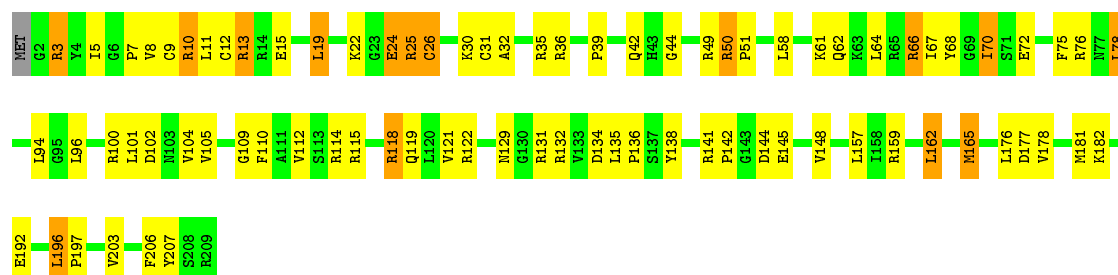
- Molecule 3: 30S ribosomal protein S3





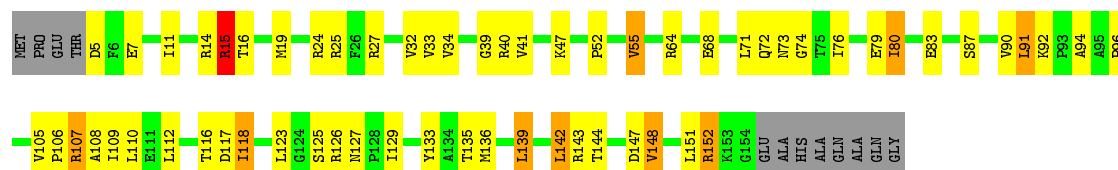
- Molecule 4: 30S ribosomal protein S4

Chain D: 61% 32% 7%



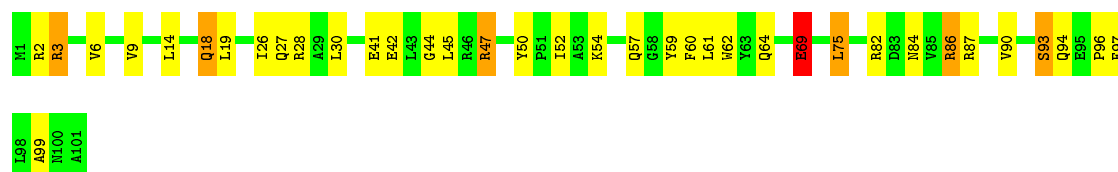
- Molecule 5: 30S ribosomal protein S5

Chain E: 55% 31% 6% 7%



- Molecule 6: 30S ribosomal protein S6

Chain F: 63% 30% 6%



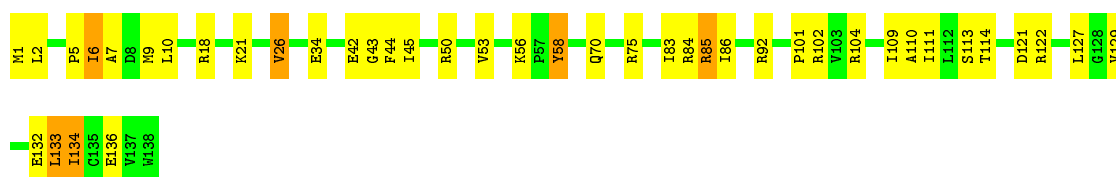
- Molecule 7: 30S ribosomal protein S7

Chain G: 80% 19%



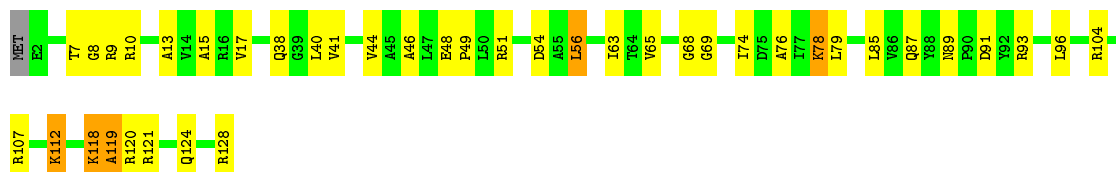
- Molecule 8: 30S ribosomal protein S8

Chain H: 70% 26%



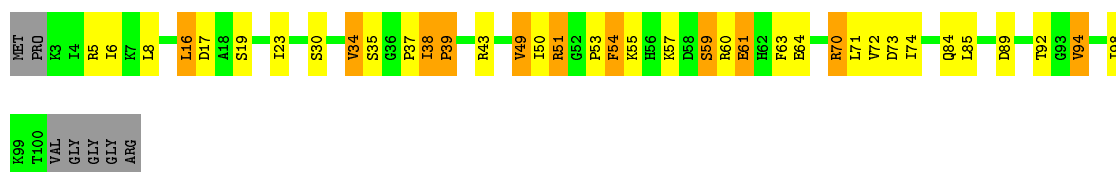
- Molecule 9: 30S ribosomal protein S9

Chain I: 68% 27% 5%



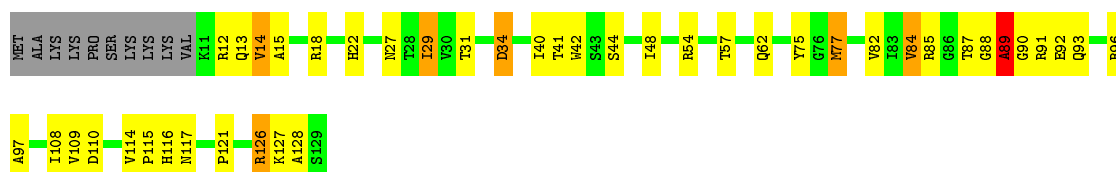
- Molecule 10: 30S ribosomal protein S10

Chain J: 58% 25% 10% 7%



- Molecule 11: 30S ribosomal protein S11

Chain K: 59% 28% 5% 8%



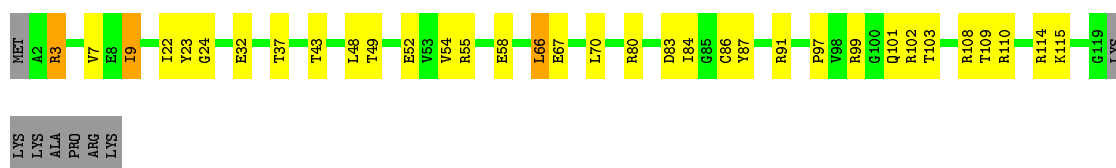
- Molecule 12: 30S ribosomal protein S12

Chain L: 61% 30% 6%



- Molecule 13: 30S ribosomal protein S13

Chain M: 67% 25% 6%



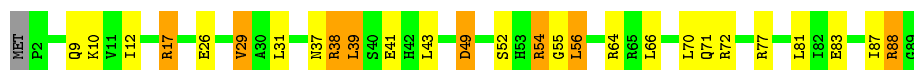
- Molecule 14: 30S ribosomal protein S14 type Z

Chain N: 61% 36% . .



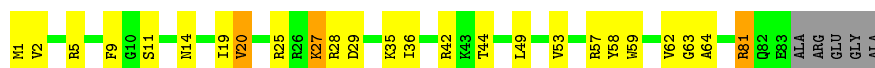
- Molecule 15: 30S ribosomal protein S15

Chain O: 69% 21% 9% .



- Molecule 16: 30S ribosomal protein S16

Chain P: 66% 25% . 6%



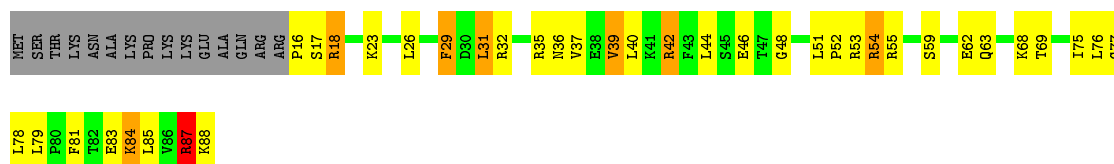
- Molecule 17: 30S ribosomal protein S17

Chain Q: 66% 25% . 6%



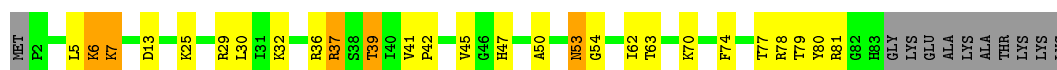
- Molecule 18: 30S ribosomal protein S18

Chain R: 40% 34% 8% . 17%

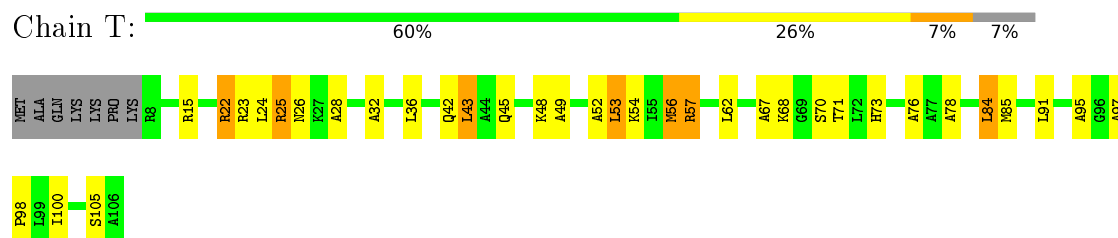


- Molecule 19: 30S ribosomal protein S19

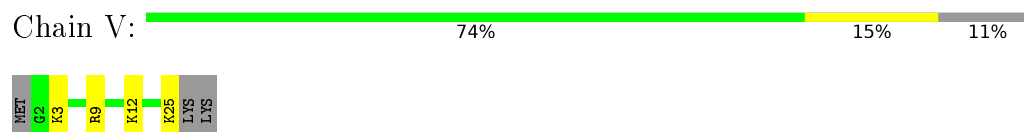
Chain S: 59% 24% 5% 12%



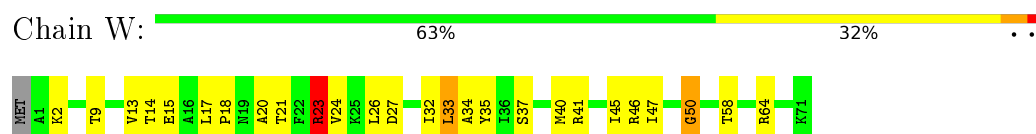
- Molecule 20: 30S ribosomal protein S20



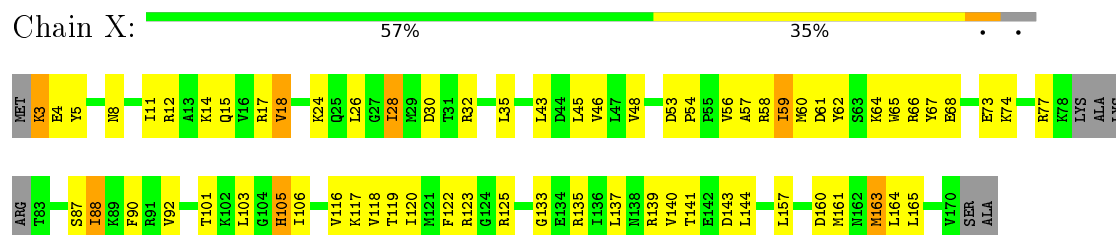
- Molecule 21: 30S ribosomal protein Thx



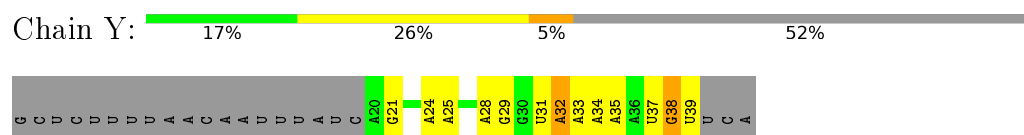
- Molecule 22: Translation initiation factor IF-1



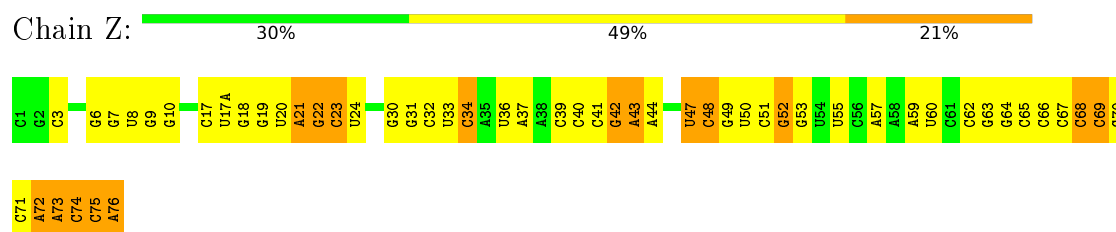
- Molecule 23: Translation initiation factor IF-3



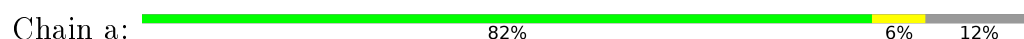
- Molecule 24: mRNA

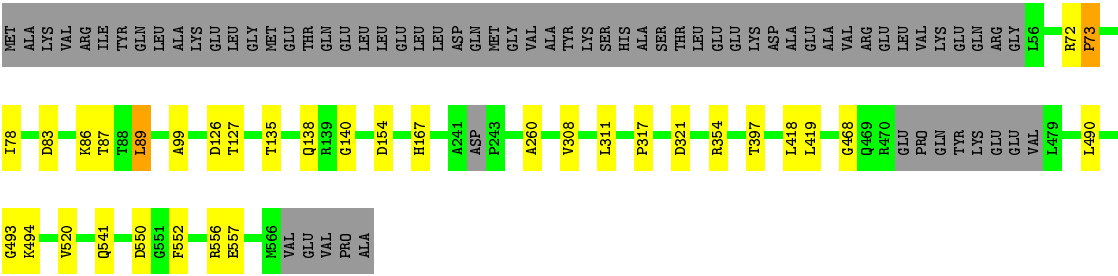


- Molecule 25: tRNAi



- Molecule 26: Translation initiation factor IF-2





## 4 Experimental information

Property	Value	Source
Reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	Depositor
Number of particles used	26324	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	Not provided	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ( $e^-/\text{\AA}^2$ )	Not provided	Depositor
Minimum defocus (nm)	1500	Depositor
Maximum defocus (nm)	3500	Depositor
Magnification	104478	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: OMC, 5MU, ZN, FME, G7M, MG, 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	0.29	1/36426 (0.0%)	0.70	10/56837 (0.0%)
10	J	0.39	0/805	0.74	1/1082 (0.1%)
11	K	0.41	0/900	0.90	2/1213 (0.2%)
12	L	0.35	0/986	0.76	0/1320
13	M	0.41	0/947	0.74	0/1270
14	N	0.40	0/501	0.78	0/664
15	O	0.41	0/745	0.83	0/992
16	P	0.39	0/716	0.73	0/963
17	Q	0.38	0/836	0.75	0/1117
18	R	0.40	0/604	0.85	0/801
19	S	0.42	0/670	0.70	0/903
2	B	0.44	0/1935	0.79	1/2609 (0.0%)
20	T	0.42	0/765	0.87	0/1007
21	V	0.43	0/212	0.69	0/277
22	W	0.48	0/580	1.07	5/782 (0.6%)
23	X	0.49	0/1354	0.74	1/1813 (0.1%)
24	Y	0.36	0/494	0.67	0/770
25	Z	0.34	0/1721	0.71	1/2682 (0.0%)
26	a	0.46	0/3824	0.70	2/5169 (0.0%)
3	C	0.39	0/1636	0.75	2/2205 (0.1%)
4	D	0.41	0/1733	0.80	1/2318 (0.0%)
5	E	0.41	0/1162	0.88	2/1564 (0.1%)
6	F	0.39	0/856	0.78	1/1154 (0.1%)
7	G	0.41	0/1276	0.73	0/1709
8	H	0.38	0/1136	0.76	0/1527
9	I	0.41	0/1029	0.78	1/1379 (0.1%)
All	All	0.35	1/63849 (0.0%)	0.73	30/94127 (0.0%)

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
1	A	0	1
11	K	1	0
23	X	0	1
26	a	0	1
All	All	1	3

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	71	C	O3'-P	-5.32	1.54	1.61

The worst 5 of 30 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
11	K	89	ALA	CB-CA-C	18.64	138.06	110.10
9	I	7	THR	CB-CA-C	-13.26	75.81	111.60
5	E	15	ARG	N-CA-C	-12.58	77.04	111.00
22	W	33	LEU	CB-CA-C	-12.05	87.30	110.20
22	W	23	ARG	N-CA-C	-10.10	83.72	111.00

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
11	K	89	ALA	CA

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	218	C	Sidechain
23	X	53	ASP	Peptide
26	a	73	PRO	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	32548	0	16440	867	0
2	B	1900	0	1951	40	0
3	C	1612	0	1677	41	0
4	D	1703	0	1766	46	0
5	E	1146	0	1207	23	0
6	F	843	0	857	19	0
7	G	1257	0	1296	13	0
8	H	1116	0	1177	16	0
9	I	1010	0	1037	22	0
10	J	792	0	835	30	0
11	K	885	0	904	20	0
12	L	970	0	1057	24	0
13	M	937	0	995	18	0
14	N	492	0	530	18	0
15	O	734	0	771	11	0
16	P	700	0	720	13	0
17	Q	823	0	891	18	0
18	R	598	0	670	26	0
19	S	655	0	672	14	0
20	T	763	0	861	22	0
21	V	208	0	221	2	0
22	W	570	0	599	24	0
23	X	1336	0	1389	45	0
24	Y	439	0	218	10	0
25	Z	1646	0	843	71	0
26	a	3774	0	3747	0	0
27	D	1	0	0	2	0
27	N	1	0	0	2	0
28	W	1	0	0	0	0
28	Z	1	0	0	0	0
29	Z	9	0	10	6	0
All	All	59470	0	43341	1342	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 14.

The worst 5 of 1342 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)
25:Z:76:A:O3'	29:Z:101:FME:C	1.70	1.37
1:A:1358:U:H3	1:A:1363(A):A:N6	1.21	1.33
25:Z:76:A:C3'	29:Z:101:FME:C	2.09	1.28

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:C:28:GLN:O	3:C:32:LEU:HG	1.32	1.27
23:X:3:LYS:CB	23:X:66:ARG:NH2	2.02	1.20

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	B	232/256 (91%)	189 (82%)	32 (14%)	11 (5%)	3	32
3	C	204/239 (85%)	179 (88%)	19 (9%)	6 (3%)	6	44
4	D	206/209 (99%)	185 (90%)	18 (9%)	3 (2%)	13	58
5	E	148/162 (91%)	132 (89%)	13 (9%)	3 (2%)	9	52
6	F	99/101 (98%)	89 (90%)	7 (7%)	3 (3%)	5	44
7	G	153/156 (98%)	139 (91%)	11 (7%)	3 (2%)	9	52
8	H	136/138 (99%)	124 (91%)	10 (7%)	2 (2%)	13	58
9	I	125/128 (98%)	106 (85%)	17 (14%)	2 (2%)	12	57
10	J	96/105 (91%)	81 (84%)	10 (10%)	5 (5%)	2	30
11	K	117/129 (91%)	101 (86%)	12 (10%)	4 (3%)	5	41
12	L	122/132 (92%)	104 (85%)	14 (12%)	4 (3%)	5	42
13	M	116/126 (92%)	106 (91%)	10 (9%)	0	100	100
14	N	58/61 (95%)	48 (83%)	7 (12%)	3 (5%)	2	30
15	O	86/89 (97%)	81 (94%)	5 (6%)	0	100	100
16	P	81/88 (92%)	74 (91%)	7 (9%)	0	100	100
17	Q	97/105 (92%)	88 (91%)	9 (9%)	0	100	100
18	R	71/88 (81%)	66 (93%)	3 (4%)	2 (3%)	6	45
19	S	80/93 (86%)	70 (88%)	6 (8%)	4 (5%)	3	31

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
20	T	97/106 (92%)	84 (87%)	9 (9%)	4 (4%)	3	35
21	V	22/27 (82%)	22 (100%)	0	0	100	100
22	W	69/72 (96%)	58 (84%)	9 (13%)	2 (3%)	6	44
23	X	160/171 (94%)	138 (86%)	20 (12%)	2 (1%)	15	60
26	a	496/571 (87%)	434 (88%)	48 (10%)	14 (3%)	6	45
All	All	3071/3352 (92%)	2698 (88%)	296 (10%)	77 (2%)	11	47

5 of 77 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	24	TRP
6	F	69	GLU
9	I	119	ALA
10	J	34	VAL
10	J	39	PRO

### 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	202/220 (92%)	166 (82%)	36 (18%)	2	17
3	C	160/188 (85%)	143 (89%)	17 (11%)	8	38
4	D	180/181 (99%)	148 (82%)	32 (18%)	2	17
5	E	115/123 (94%)	80 (70%)	35 (30%)	0	3
6	F	90/90 (100%)	76 (84%)	14 (16%)	3	23
7	G	126/127 (99%)	112 (89%)	14 (11%)	8	36
8	H	119/119 (100%)	96 (81%)	23 (19%)	2	14
9	I	98/99 (99%)	85 (87%)	13 (13%)	5	29
10	J	87/92 (95%)	75 (86%)	12 (14%)	4	28
11	K	90/99 (91%)	72 (80%)	18 (20%)	1	12
12	L	104/109 (95%)	89 (86%)	15 (14%)	4	26

Continued on next page...

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
13	M	94/101 (93%)	80 (85%)	14 (15%)	4	25
14	N	49/50 (98%)	42 (86%)	7 (14%)	4	27
15	O	79/80 (99%)	59 (75%)	20 (25%)	1	6
16	P	72/74 (97%)	64 (89%)	8 (11%)	8	36
17	Q	94/97 (97%)	85 (90%)	9 (10%)	10	43
18	R	64/77 (83%)	46 (72%)	18 (28%)	0	4
19	S	71/80 (89%)	60 (84%)	11 (16%)	3	23
20	T	76/82 (93%)	64 (84%)	12 (16%)	3	22
21	V	19/22 (86%)	16 (84%)	3 (16%)	3	22
22	W	62/63 (98%)	55 (89%)	7 (11%)	7	35
23	X	145/150 (97%)	125 (86%)	20 (14%)	4	28
26	a	374/460 (81%)	355 (95%)	19 (5%)	29	67
All	All	2570/2783 (92%)	2193 (85%)	377 (15%)	8	25

5 of 377 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
8	H	134	ILE
11	K	93	GLN
23	X	73	GLU
9	I	78	LYS
10	J	61	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 37 such sidechains are listed below:

Mol	Chain	Res	Type
9	I	89	ASN
11	K	13	GLN
23	X	162	ASN
10	J	56	HIS
10	J	68	HIS

### 5.3.3 RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
1	A	1509/1522 (99%)	407 (26%)	89 (5%)

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
24	Y	19/42 (45%)	7 (36%)	1 (5%)
25	Z	76/77 (98%)	27 (35%)	4 (5%)
All	All	1604/1641 (97%)	441 (27%)	94 (5%)

5 of 441 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
1	A	8	A
1	A	9	G
1	A	13	U
1	A	19	C
1	A	22	G

5 of 94 RNA pucker outliers are listed below:

Mol	Chain	Res	Type
1	A	703	G
1	A	965	A
1	A	1504	G
1	A	748	C
1	A	873	A

## 5.4 Non-standard residues in protein, DNA, RNA chains [i](#)

5 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
25	OMC	Z	32	25	15,22,23	0.72	0	20,31,34	1.53	2 (10%)
25	G7M	Z	46	25	18,26,27	2.72	3 (16%)	21,39,42	2.81	5 (23%)
25	5MU	Z	54	25	13,22,23	0.77	0	16,32,35	3.23	3 (18%)
25	PSU	Z	55	25	15,21,22	1.12	1 (6%)	16,30,33	2.48	5 (31%)
25	4SU	Z	8	25	12,21,22	0.95	0	15,30,33	1.46	3 (20%)

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
25	OMC	Z	32	25	-	0/5/27/28	0/2/2/2
25	G7M	Z	46	25	-	0/3/25/26	0/3/3/3
25	5MU	Z	54	25	-	0/3/25/26	0/2/2/2
25	PSU	Z	55	25	-	0/7/25/26	0/2/2/2
25	4SU	Z	8	25	-	0/3/25/26	0/2/2/2

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
25	Z	55	PSU	C5-C1'	-2.90	1.49	1.52
25	Z	46	G7M	C6-C5	3.71	1.48	1.41
25	Z	46	G7M	C8-N7	7.16	1.46	1.33
25	Z	46	G7M	C8-N9	7.82	1.47	1.33

The worst 5 of 18 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
25	Z	54	5MU	C5-C4-N3	-8.45	118.26	125.35
25	Z	46	G7M	N7-C8-N9	-8.05	96.78	108.67
25	Z	46	G7M	C5-C6-N1	-5.78	115.97	123.52
25	Z	55	PSU	C5-C1'-C2'	-4.41	107.95	115.44
25	Z	8	4SU	C5-C4-N3	-3.45	119.90	123.56

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
25	Z	32	OMC	2	0
25	Z	8	4SU	2	0

## 5.5 Carbohydrates

There are no carbohydrates in this entry.



## 5.6 Ligand geometry

Of 5 ligands modelled in this entry, 4 are monoatomic - leaving 1 for Mogul analysis.

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	$\# Z  > 2$
29	FME	Z	101	-	8,8,10	0.48	0	8,8,11	1.22	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
29	FME	Z	101	-	-	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( $^{\circ}$ )
29	Z	101	FME	CB-CA-N	2.17	113.79	110.20
29	Z	101	FME	CA-N-CN	2.40	127.22	124.12

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 6 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
29	Z	101	FME	6	0

## 5.7 Other polymers

There are no such residues in this entry.

## 5.8 Polymer linkage issues

The following chains have linkage breaks:

Mol	Chain	Number of breaks
1	A	5

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	A	1442(A):G	O3'	1442(B):A	P	5.40
1	A	84:U	O3'	88:A	P	5.02
1	A	93:G	O3'	96:U	P	4.43
1	A	204:U	O3'	216:G	P	4.27
1	A	841:U	O3'	848:C	P	3.83