



## wwPDB EM Map/Model Validation Report ⓘ

Apr 10, 2016 – 02:32 PM BST

PDB ID : 3JAQ  
EMDB ID: : EMD-3049  
Title : Structure of a partial yeast 48S preinitiation complex in open conformation  
Authors : Llacer, J.L.; Hussain, T.; Ramakrishnan, V.  
Deposited on : 2015-06-18  
Resolution : 6.00 Å(reported)  
Based on PDB ID : 1YFG, 2D74, 3CW2, 3J81, 4U1E

This is a wwPDB EM Map/Model Validation Report for a publicly released PDB/EMDB entry.  
For rigid body fitted models, validation errors reported here could stem from errors in the original structure(s) used in the fitting.  
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>

---

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) : trunk27241

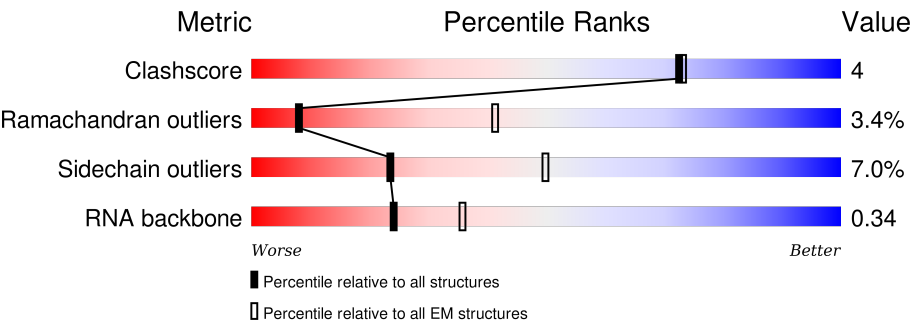


# 1 Overall quality at a glance i

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

The reported resolution of this entry is 6.00 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












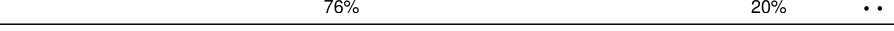







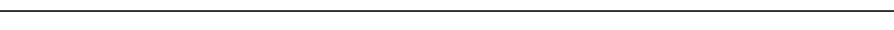

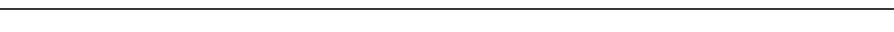
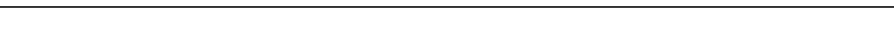


The table below summarises the geometric issues observed across the polymeric chains. The red, orange, yellow and green segments on the bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria. A grey segment represents the fraction of 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	1	75	<div><div>35%</div><div>52%</div><div>13%</div></div>
2	2	1781	<div><div>43%</div><div>46%</div><div>11%</div></div>
3	3	25	<div><div>8%</div><div>92%</div></div>
4	A	254	<div><div>67%</div><div>15%</div><div>18%</div></div>
5	B	255	<div><div>72%</div><div>13%</div><div>•</div><div>13%</div></div>
6	C	259	<div><div>71%</div><div>11%</div><div>•</div><div>16%</div></div>
7	D	237	<div><div>78%</div><div>16%</div><div>6%</div></div>
8	E	261	<div><div>82%</div><div>17%</div><div>•</div></div>

Continued on next page...



Continued from previous page...

Mol	Chain	Length	Quality of chain
9	F	227	
10	G	236	
11	H	190	
12	I	201	
13	J	188	
14	K	106	
15	L	156	
16	M	134	
17	N	151	
18	O	137	
19	P	142	
20	Q	143	
21	R	136	
22	S	146	
23	T	144	
24	U	117	
25	V	87	
26	W	130	
27	X	145	
28	Y	135	
29	Z	108	
30	a	119	
31	b	82	
32	c	67	
33	d	56	

Continued on next page...



*Continued from previous page...*

Mol	Chain	Length	Quality of chain
34	e	63	 75%11%14%
35	f	150	 37%9%54%
36	g	326	 90%8%
37	h	25	 96%
38	i	153	 56%6%38%
39	j	304	 70%11%18%
40	k	527	 71%25%
41	l	285	 43%55%
42	m	108	 77%20%
43	o	92	 100%
44	p	88	 100%
45	q	347	 95%
46	r	34	 100%
47	s	52	 98%



## 2 Entry composition

There are 51 unique types of molecules in this entry. The entry contains 89768 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 Met-tRNAi.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	1	75	Total	C	N	O	P	0	0
			1639	734	298	531	76		

- Molecule 2 is a RNA chain called 18S rRNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	2	1780	Total	C	N	O	P	0	0
			37797	16892	6658	12467	1780		

- Molecule 3 is a RNA chain called mRNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
3	3	2	Total	C	N	O	P	0	0
			42	19	7	14	2		

- Molecule 4 is a protein called uS2.

Mol	Chain	Residues	Atoms					AltConf	Trace
4	A	208	Total	C	N	O	S	0	0
			1626	1040	286	298	2		

- Molecule 5 is a protein called eS1.

Mol	Chain	Residues	Atoms					AltConf	Trace
5	B	222	Total	C	N	O	S	0	0
			1769	1117	324	325	3		

- Molecule 6 is a protein called uS5.

Mol	Chain	Residues	Atoms					AltConf	Trace
6	C	217	Total	C	N	O	S	0	0
			1629	1041	287	297	4		



- Molecule 7 is a protein called uS3.

Mol	Chain	Residues	Atoms					AltConf	Trace
7	D	223	Total	C	N	O	S	0	0
			1744	1108	313	318	5		

- Molecule 8 is a protein called eS4.

Mol	Chain	Residues	Atoms					AltConf	Trace
8	E	260	Total	C	N	O	S	0	0
			2078	1322	393	359	4		

- Molecule 9 is a protein called uS7.

Mol	Chain	Residues	Atoms					AltConf	Trace
9	F	206	Total	C	N	O	S	0	0
			1609	1008	298	300	3		

- Molecule 10 is a protein called eS6.

Mol	Chain	Residues	Atoms					AltConf	Trace
10	G	226	Total	C	N	O	S	0	0
			1812	1134	348	326	4		

- Molecule 11 is a protein called eS7.

Mol	Chain	Residues	Atoms					AltConf	Trace
11	H	184	Total	C	N	O	S	0	0
			1483	950	270	263			

- Molecule 12 is a protein called eS8.

Mol	Chain	Residues	Atoms					AltConf	Trace
12	I	188	Total	C	N	O	S	0	0
			1489	923	300	265	1		

- Molecule 13 is a protein called uS4.

Mol	Chain	Residues	Atoms					AltConf	Trace
13	J	182	Total	C	N	O	S	0	0
			1471	929	287	254	1		

- Molecule 14 is a protein called eS10.



Mol	Chain	Residues	Atoms					AltConf	Trace
14	K	96	Total	C	N	O	S	0	0
			809	533	129	146	1		

- Molecule 15 is a protein called uS17.

Mol	Chain	Residues	Atoms					AltConf	Trace
15	L	155	Total	C	N	O	S	0	0
			1248	798	237	210	3		

- Molecule 16 is a protein called eS12.

Mol	Chain	Residues	Atoms					AltConf	Trace
16	M	117	Total	C	N	O	S	0	0
			885	553	161	171			

- Molecule 17 is a protein called uS15.

Mol	Chain	Residues	Atoms					AltConf	Trace
17	N	150	Total	C	N	O	S	0	0
			1187	756	223	206	2		

- Molecule 18 is a protein called uS11.

Mol	Chain	Residues	Atoms					AltConf	Trace
18	O	127	Total	C	N	O	S	0	0
			942	578	188	173	3		

- Molecule 19 is a protein called uS19.

Mol	Chain	Residues	Atoms					AltConf	Trace
19	P	119	Total	C	N	O	S	0	0
			943	604	171	163	5		

- Molecule 20 is a protein called uS9.

Mol	Chain	Residues	Atoms					AltConf	Trace
20	Q	141	Total	C	N	O	S	0	0
			1105	709	204	192			

- Molecule 21 is a protein called eS17.



Mol	Chain	Residues	Atoms					AltConf	Trace
21	R	111	Total	C	N	O	S	0	0
			892	554	165	170	3		

- Molecule 22 is a protein called uS13.

Mol	Chain	Residues	Atoms					AltConf	Trace
22	S	145	Total	C	N	O	S	0	0
			1193	741	240	210	2		

- Molecule 23 is a protein called eS19.

Mol	Chain	Residues	Atoms					AltConf	Trace
23	T	143	Total	C	N	O	S	0	0
			1110	693	210	207			

- Molecule 24 is a protein called uS10.

Mol	Chain	Residues	Atoms					AltConf	Trace
24	U	106	Total	C	N	O	S	0	0
			845	540	152	152	1		

- Molecule 25 is a protein called eS21.

Mol	Chain	Residues	Atoms					AltConf	Trace
25	V	87	Total	C	N	O	S	0	0
			687	424	126	135	2		

- Molecule 26 is a protein called uS8.

Mol	Chain	Residues	Atoms					AltConf	Trace
26	W	129	Total	C	N	O	S	0	0
			1021	651	187	180	3		

- Molecule 27 is a protein called uS12.

Mol	Chain	Residues	Atoms					AltConf	Trace
27	X	144	Total	C	N	O	S	0	0
			1119	708	218	191	2		

- Molecule 28 is a protein called eS24.



Mol	Chain	Residues	Atoms				AltConf	Trace
28	Y	134	Total	C	N	O		
			1061	665	207	189	0	0

- Molecule 29 is a protein called eS25.

Mol	Chain	Residues	Atoms					AltConf	Trace
29	Z	70	Total	C	N	O	S		
			558	355	104	98	1	0	0

- Molecule 30 is a protein called eS26.

Mol	Chain	Residues	Atoms					AltConf	Trace
30	a	98	Total	C	N	O	S		
			779	480	165	129	5	0	0

- Molecule 31 is a protein called eS27.

Mol	Chain	Residues	Atoms					AltConf	Trace
31	b	81	Total	C	N	O	S		
			609	379	112	113	5	0	0

- Molecule 32 is a protein called eS28.

Mol	Chain	Residues	Atoms					AltConf	Trace
32	c	62	Total	C	N	O	S		
			487	301	97	88	1	0	0

- Molecule 33 is a protein called uS14.

Mol	Chain	Residues	Atoms					AltConf	Trace
33	d	53	Total	C	N	O	S		
			446	280	89	76	1	0	0

- Molecule 34 is a protein called eS30.

Mol	Chain	Residues	Atoms					AltConf	Trace
34	e	54	Total	C	N	O	S		
			433	271	88	73	1	0	0

- Molecule 35 is a protein called eS31.



Mol	Chain	Residues	Atoms					AltConf	Trace
35	f	69	Total	C	N	O	S	0	0
			549	352	102	91	4		

- Molecule 36 is a protein called RACK1.

Mol	Chain	Residues	Atoms					AltConf	Trace
36	g	318	Total	C	N	O	S	0	0
			2466	1561	430	470	5		

- Molecule 37 is a protein called eL41.

Mol	Chain	Residues	Atoms					AltConf	Trace
37	h	25	Total	C	N	O	S	0	0
			233	142	63	27	1		

- Molecule 38 is a protein called eIF1A.

Mol	Chain	Residues	Atoms					AltConf	Trace
38	i	95	Total	C	N	O	S	0	0
			765	475	142	143	5		

- Molecule 39 is a protein called eIF2 alpha.

Mol	Chain	Residues	Atoms					AltConf	Trace
39	j	249	Total	C	N	O	S	0	0
			2006	1283	333	382	8		

- Molecule 40 is a protein called eIF2 gamma.

Mol	Chain	Residues	Atoms					AltConf	Trace
40	k	396	Total	C	N	O	S	0	0
			3034	1932	542	544	16		

- Molecule 41 is a protein called eIF2 beta.

Mol	Chain	Residues	Atoms					AltConf	Trace
41	l	128	Total	C	N	O	S	0	0
			1036	661	186	182	7		

- Molecule 42 is a protein called eIF1.



Mol	Chain	Residues	Atoms					AltConf	Trace
42	m	86	Total	C	N	O	S	0	0
			695	439	128	124	4		

- Molecule 43 is a protein called eIF3a.

Mol	Chain	Residues	Atoms				AltConf	Trace
43	o	92	Total	C	N	O	0	0
			460	276	92	92		

- Molecule 44 is a protein called eIF3c.

Mol	Chain	Residues	Atoms				AltConf	Trace
44	p	88	Total	C	N	O	0	0
			440	264	88	88		

- Molecule 45 is a protein called eIF3i.

Mol	Chain	Residues	Atoms					AltConf	Trace
45	q	342	Total	C	N	O	S	0	0
			2693	1711	443	530	9		

- Molecule 46 is a protein called eIF3b.

Mol	Chain	Residues	Atoms					AltConf	Trace
46	r	34	Total	C	N	O	S	0	0
			300	192	52	54	2		

- Molecule 47 is a protein called eIF3g.

Mol	Chain	Residues	Atoms				AltConf	Trace
47	s	52	Total	C	N	O	0	0
			418	257	82	79		

- Molecule 48 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		AltConf
48	2	79	Total	Mg	0
			79	79	
48	Q	1	Total	Mg	0
			1	1	
48	C	1	Total	Mg	0
			1	1	

*Continued on next page...*



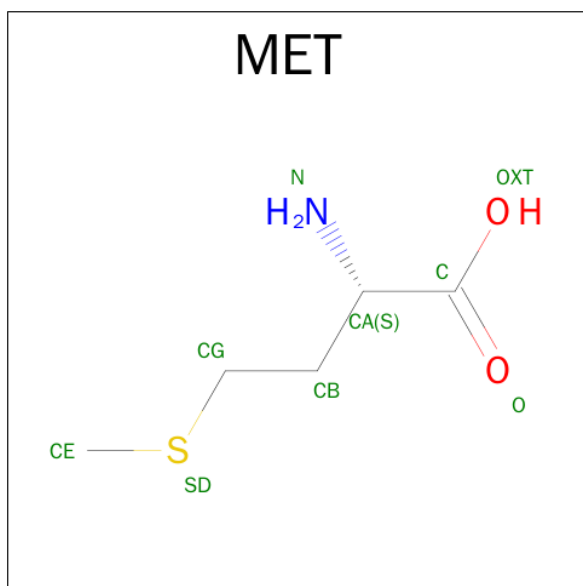
*Continued from previous page...*

Mol	Chain	Residues	Atoms		AltConf
48	k	1	Total	Mg	0
			1	1	

- Molecule 49 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		AltConf
49	b	1	Total	Zn	0
			1	1	
49	a	1	Total	Zn	0
			1	1	
49	l	1	Total	Zn	0
			1	1	
49	f	1	Total	Zn	0
			1	1	

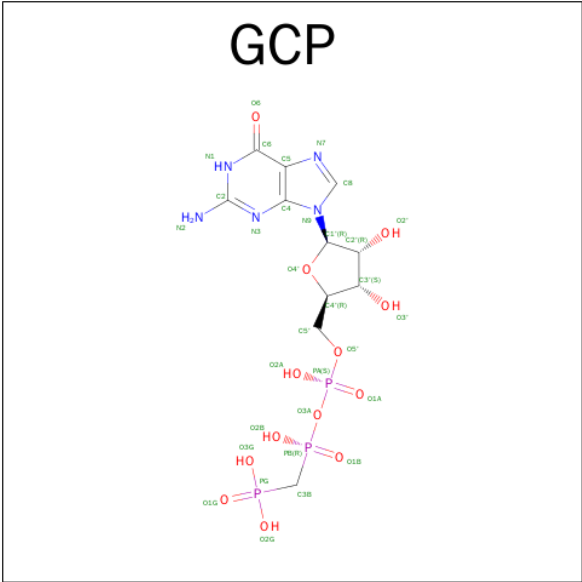
- Molecule 50 is METHIONINE (three-letter code: MET) (formula: C<sub>5</sub>H<sub>11</sub>NO<sub>2</sub>S).



Mol	Chain	Residues	Atoms					AltConf
50	k	1	Total	C	N	O	S	0
			8	5	1	1	1	

- Molecule 51 is PHOSPHOMETHYLPHOSPHONIC ACID GUANYLATE ESTER (three-letter code: GCP) (formula: C<sub>11</sub>H<sub>18</sub>N<sub>5</sub>O<sub>13</sub>P<sub>3</sub>).



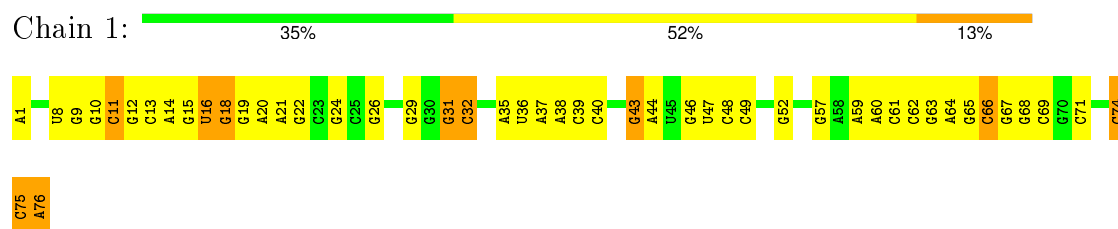




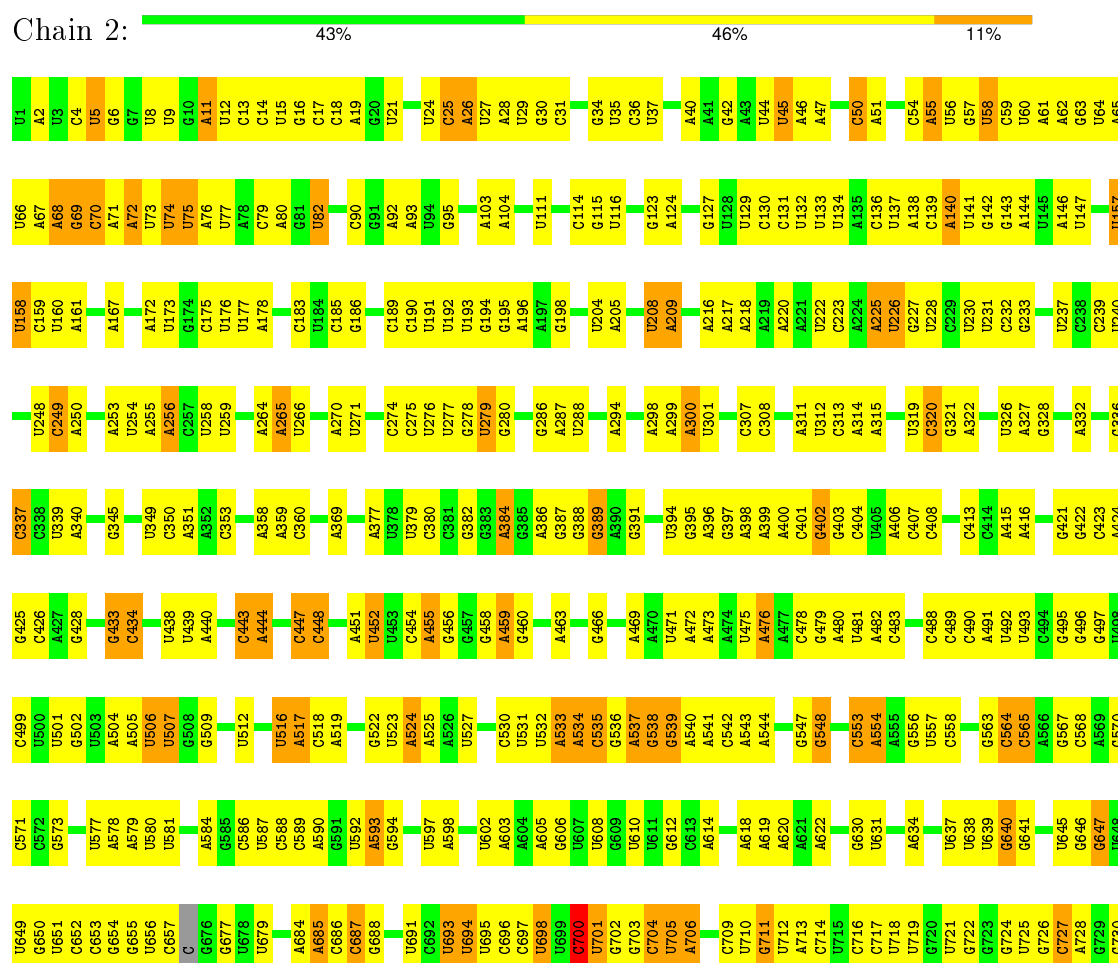
### 3 Residue-property plots

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: Met-tRNAi



#### • Molecule 2: 18S rRNA







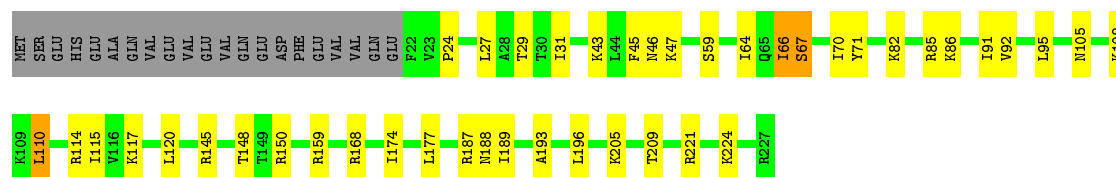







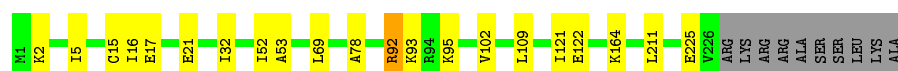
- Molecule 9: uS7

Chain F: 




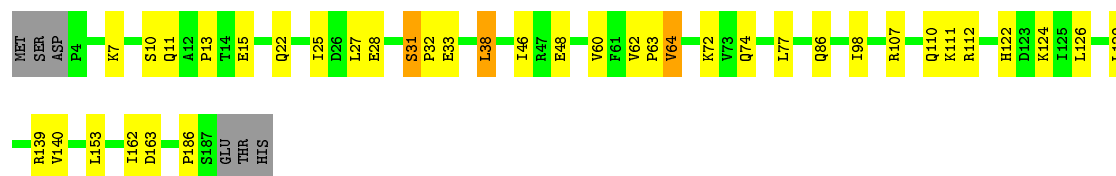
- Molecule 10: eS6

Chain G: 



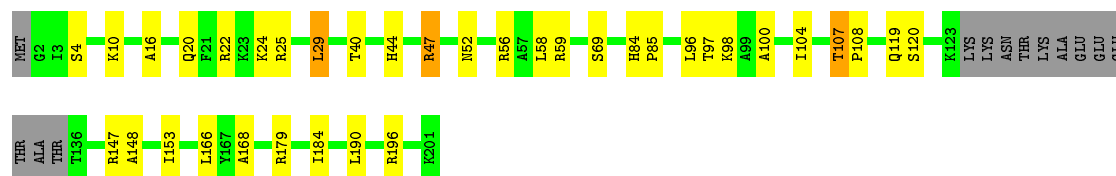
- Molecule 11: eS7

Chain H: 




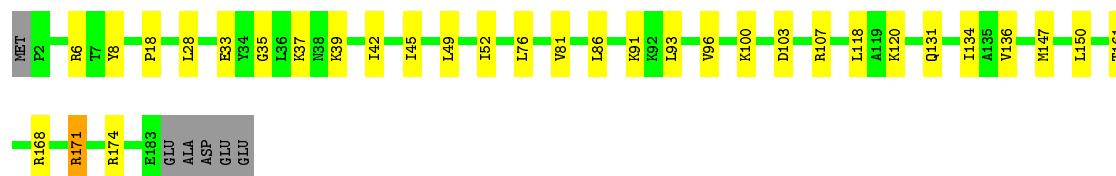
- Molecule 12: eS8

Chain I: 



- Molecule 13: uS4

Chain J: 



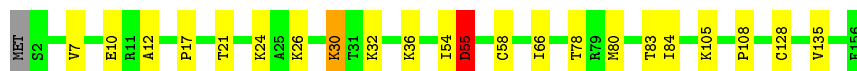
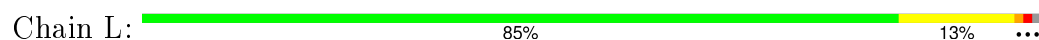
- Molecule 14: eS10

Chain K: 

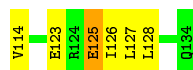
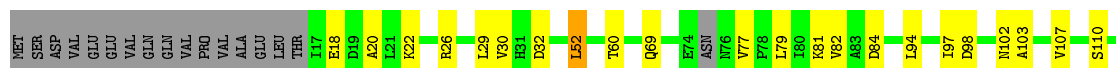




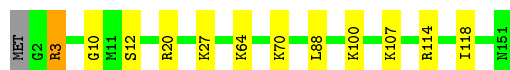
- Molecule 15: uS17



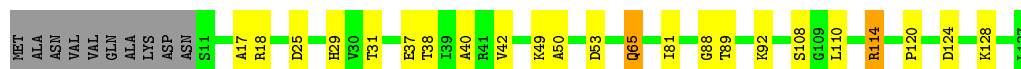
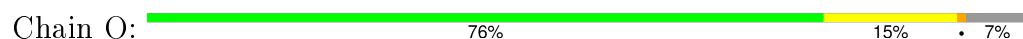
- Molecule 16: eS12



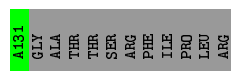
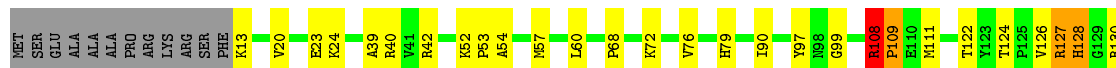
- Molecule 17: uS15



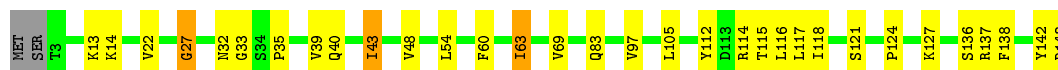
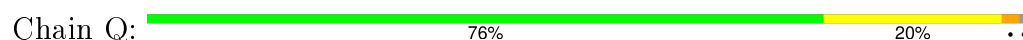
- Molecule 18: uS11



- Molecule 19: uS19

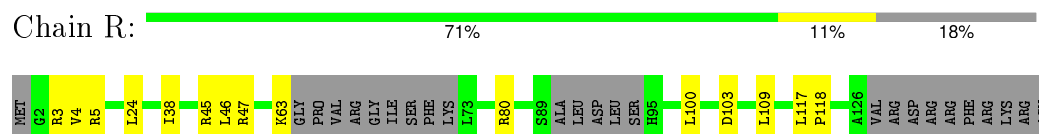


- Molecule 20: uS9

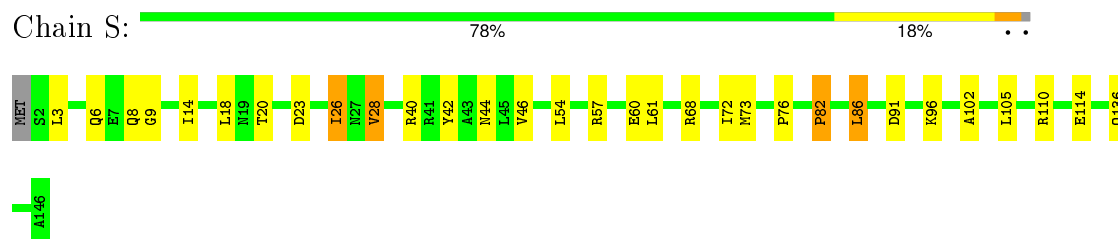


- Molecule 21: eS17

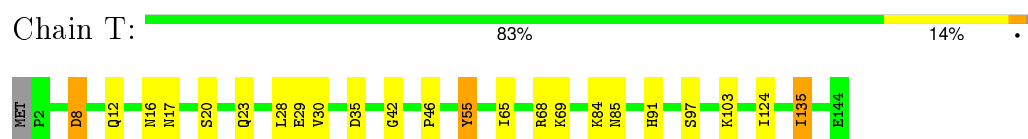




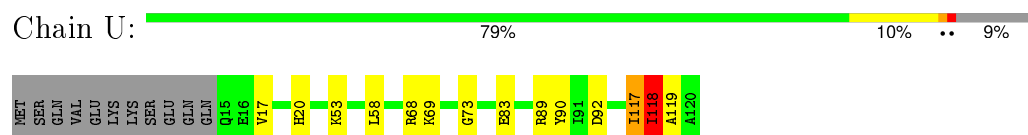
- Molecule 22: uS13



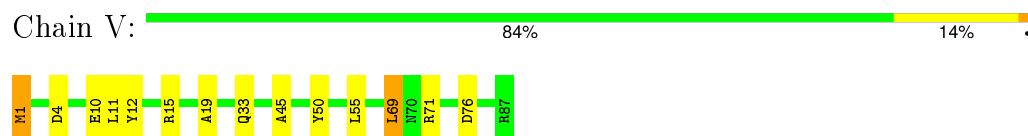
- Molecule 23: eS19



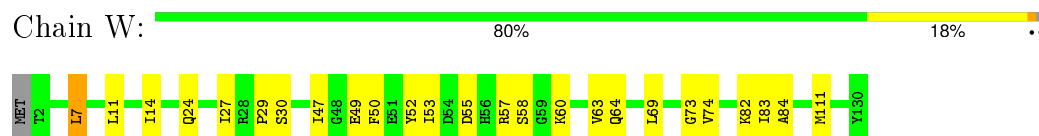
- Molecule 24: uS10



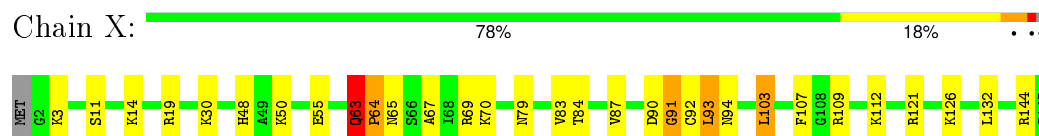
- Molecule 25: eS21



- Molecule 26: uS8




- Molecule 27: uS12



- Molecule 28: eS24

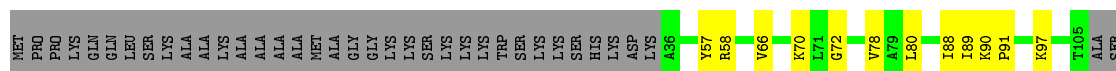


Chain Y:  84% 15% ..



- Molecule 29: eS25

Chain Z:  54% 11% 35%



GLU

- Molecule 30: eS26

Chain a:  74% 8% 18%




- Molecule 31: eS27

Chain b:  93% 6% .



- Molecule 32: eS28

Chain c:  87% 6% 7%



- Molecule 33: uS14

Chain d:  86% 9% 5%




- Molecule 34: eS30

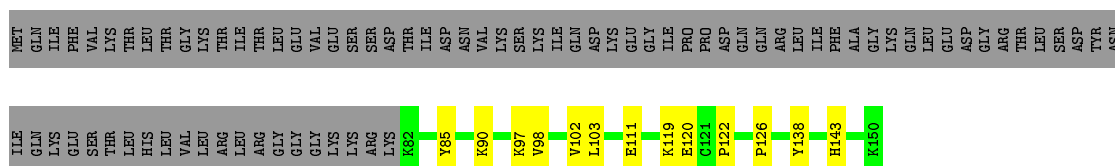
Chain e:  75% 11% 14%



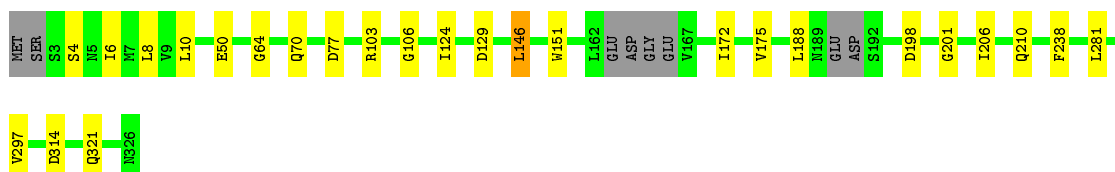
- Molecule 35: eS31

Chain f:  37% 9% 54%





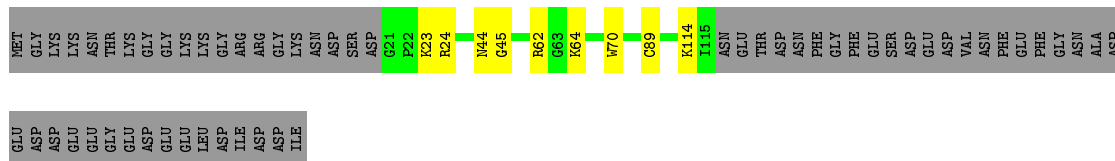
- Molecule 36: RACK1



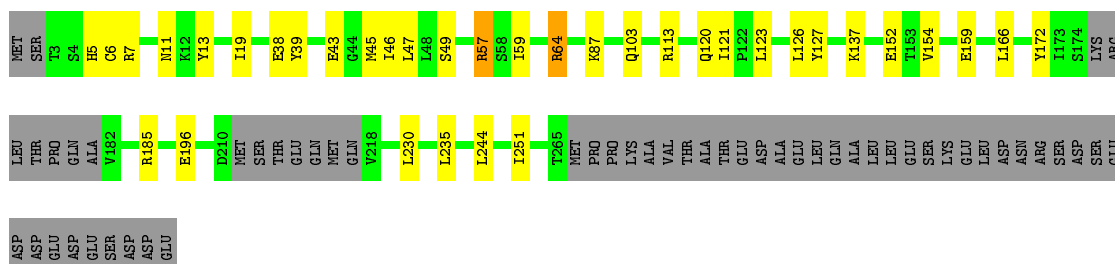
- Molecule 37: eL41



- Molecule 38: eIF1A



- Molecule 39: eIF2 alpha

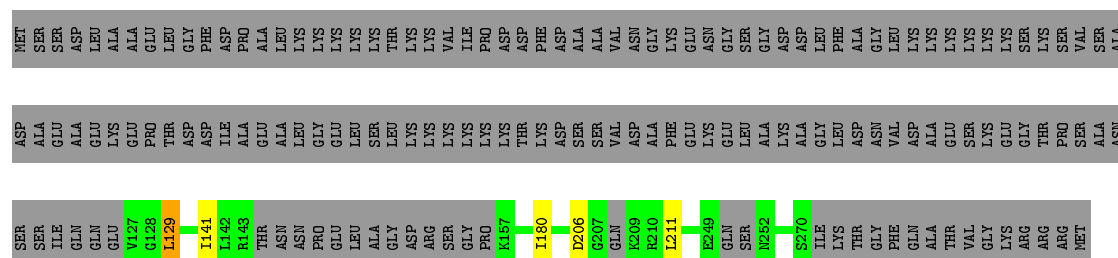


- Molecule 40: eIF2 gamma

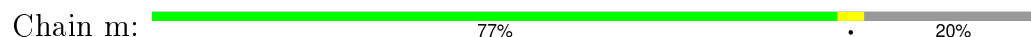




- Molecule 41: eIF2 beta



- Molecule 42: eIF1



- Molecule 43: eIF3a



There are no outlier residues recorded for this chain.

- Molecule 44: eIF3c



There are no outlier residues recorded for this chain.

- Molecule 45: eIF3i



- Molecule 46: eIF3b



There are no outlier residues recorded for this chain.



- Molecule 47: eIF3g

Chain s:  98% .





## 4 Experimental information

Property	Value	Source
Reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	Depositor
Number of images	4547	Depositor
Resolution determination method	FSC 0.143	Depositor
CTF correction method	Each particle	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ( $e^-/\text{\AA}^2$ )	27	Depositor
Minimum defocus (nm)	4000	Depositor
Maximum defocus (nm)	1500	Depositor
Magnification	59000	Depositor
Image detector	FEI FALCON II (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: ZN, H2U, 7MG, MG, GCP, 2MG, 5MC, 1MA, M2G, T6A, 1MG, RIA

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	1	0.45	2/1530 (0.1%)	0.76	4/2380 (0.2%)
10	G	0.38	0/1835	0.60	0/2451
11	H	0.41	0/1507	0.62	0/2028
12	I	0.41	0/1515	0.63	2/2029 (0.1%)
13	J	0.40	0/1495	0.67	2/2001 (0.1%)
14	K	0.46	0/831	0.61	0/1123
15	L	0.42	0/1276	0.58	0/1718
16	M	0.40	0/891	0.65	0/1201
17	N	0.39	0/1210	0.60	0/1628
18	O	0.39	0/953	0.61	0/1279
19	P	0.41	0/962	0.62	0/1294
2	2	0.24	0/42269	0.65	5/65862 (0.0%)
20	Q	0.41	0/1125	0.62	0/1510
21	R	0.39	0/899	0.60	0/1204
22	S	0.40	0/1212	0.66	0/1629
23	T	0.40	0/1129	0.58	0/1520
24	U	0.39	0/857	0.61	0/1158
25	V	0.38	0/696	0.61	0/938
26	W	0.40	0/1039	0.64	1/1399 (0.1%)
27	X	0.42	0/1137	0.66	2/1516 (0.1%)
28	Y	0.40	0/1075	0.59	0/1433
29	Z	0.41	0/567	0.60	0/762
3	3	0.22	0/46	0.64	0/69
30	a	0.38	0/791	0.64	0/1059
31	b	0.39	0/619	0.62	0/837
32	c	0.37	0/489	0.65	0/655
33	d	0.43	0/457	0.58	0/607
34	e	0.40	0/440	0.61	0/586
35	f	0.45	0/562	0.72	1/751 (0.1%)
36	g	0.41	0/2521	0.58	0/3431
37	h	0.35	0/234	0.61	0/300
38	i	0.38	0/775	0.60	0/1034



Mol	Chain	Bond lengths		Bond angles	
		RMSZ	# Z  >2	RMSZ	# Z  >2
39	j	0.41	0/2034	0.63	0/2737
4	A	0.42	0/1666	0.63	1/2279 (0.0%)
40	k	0.39	0/3079	0.60	0/4157
41	l	0.41	0/1051	0.57	0/1402
42	m	0.39	0/703	0.58	0/938
45	q	0.42	0/2757	0.57	0/3733
46	r	0.42	0/306	0.52	0/407
47	s	0.39	0/426	0.54	0/571
5	B	0.40	0/1793	0.61	0/2414
6	C	0.40	0/1659	0.59	0/2252
7	D	0.41	0/1769	0.61	0/2378
8	E	0.40	0/2122	0.62	1/2861 (0.0%)
9	F	0.40	0/1628	0.63	1/2198 (0.0%)
All	All	0.34	2/93937 (0.0%)	0.64	20/135719 (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
27	X	0	1
28	Y	0	1
All	All	0	2

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	1	1	A	OP3-P	-9.98	1.49	1.61
1	1	65	G	P-O5'	-6.66	1.53	1.59

The worst 5 of 20 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
35	f	126	PRO	CA-N-CD	-8.85	99.11	111.50
1	1	65	G	OP1-P-OP2	-8.39	107.01	119.60
2	2	685	A	C2'-C3'-O3'	6.95	124.81	113.70
12	I	29	LEU	CA-CB-CG	6.73	130.78	115.30
1	1	65	G	O5'-P-OP2	6.33	118.29	110.70

There are no chirality outliers.



All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
27	X	63	GLN	Peptide
28	Y	29	HIS	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	1	1639	0	851	18	0
2	2	37797	0	19015	349	0
3	3	42	0	22	0	0
4	A	1626	0	1633	9	0
5	B	1769	0	1829	12	0
6	C	1629	0	1710	9	0
7	D	1744	0	1826	14	0
8	E	2078	0	2157	13	0
9	F	1609	0	1679	14	0
10	G	1812	0	1911	6	0
11	H	1483	0	1579	10	0
12	I	1489	0	1504	14	0
13	J	1471	0	1554	11	0
14	K	809	0	810	10	0
15	L	1248	0	1311	10	0
16	M	885	0	917	9	0
17	N	1187	0	1251	2	0
18	O	942	0	979	6	0
19	P	943	0	989	12	0
20	Q	1105	0	1170	8	0
21	R	892	0	932	3	0
22	S	1193	0	1217	12	0
23	T	1110	0	1124	9	0
24	U	845	0	913	5	0
25	V	687	0	682	4	0
26	W	1021	0	1056	9	0
27	X	1119	0	1198	11	0
28	Y	1061	0	1111	7	0
29	Z	558	0	585	4	0
30	a	779	0	830	0	0

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
31	b	609	0	629	0	0
32	c	487	0	528	0	0
33	d	446	0	436	0	0
34	e	433	0	470	0	0
35	f	549	0	562	0	0
36	g	2466	0	2406	0	0
37	h	233	0	284	0	0
38	i	765	0	769	0	0
39	j	2006	0	2066	0	0
40	k	3034	0	3195	0	0
41	l	1036	0	1080	0	0
42	m	695	0	729	0	0
43	o	460	0	106	0	0
44	p	440	0	102	0	0
45	q	2693	0	2609	0	0
46	r	300	0	297	0	0
47	s	418	0	411	0	0
48	2	79	0	0	0	0
48	C	1	0	0	0	0
48	Q	1	0	0	0	0
48	k	1	0	0	0	0
49	a	1	0	0	0	0
49	b	1	0	0	0	0
49	f	1	0	0	0	0
49	l	1	0	0	0	0
50	k	8	0	8	0	0
51	k	32	0	14	0	0
All	All	89768	0	71046	540	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 4.

The worst 5 of 540 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)
2:2:1330:A:H2'	2:2:1331:C:H5'	1.34	1.08
2:2:534:A:H3'	2:2:535:C:H5''	1.36	1.02
2:2:1330:A:C2'	2:2:1331:C:H5'	1.89	1.02
2:2:1328:A:C2'	2:2:1329:G:H5'	1.90	1.01
2:2:1328:A:H2'	2:2:1329:G:H5'	1.40	1.01



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	
4	A	206/254 (81%)	177 (86%)	23 (11%)	6 (3%)	6	43
5	B	218/255 (86%)	191 (88%)	18 (8%)	9 (4%)	3	35
6	C	215/259 (83%)	193 (90%)	17 (8%)	5 (2%)	8	48
7	D	221/237 (93%)	198 (90%)	17 (8%)	6 (3%)	6	44
8	E	258/261 (99%)	230 (89%)	23 (9%)	5 (2%)	10	52
9	F	204/227 (90%)	170 (83%)	28 (14%)	6 (3%)	6	43
10	G	224/236 (95%)	204 (91%)	17 (8%)	3 (1%)	15	59
11	H	182/190 (96%)	166 (91%)	8 (4%)	8 (4%)	3	33
12	I	184/201 (92%)	163 (89%)	12 (6%)	9 (5%)	3	31
13	J	180/188 (96%)	153 (85%)	19 (11%)	8 (4%)	3	33
14	K	94/106 (89%)	80 (85%)	9 (10%)	5 (5%)	2	29
15	L	153/156 (98%)	133 (87%)	16 (10%)	4 (3%)	7	45
16	M	113/134 (84%)	86 (76%)	16 (14%)	11 (10%)	1	14
17	N	148/151 (98%)	141 (95%)	5 (3%)	2 (1%)	14	58
18	O	125/137 (91%)	101 (81%)	16 (13%)	8 (6%)	2	25
19	P	117/142 (82%)	91 (78%)	18 (15%)	8 (7%)	1	23
20	Q	139/143 (97%)	114 (82%)	13 (9%)	12 (9%)	1	17
21	R	105/136 (77%)	92 (88%)	10 (10%)	3 (3%)	6	43
22	S	143/146 (98%)	115 (80%)	20 (14%)	8 (6%)	2	28
23	T	141/144 (98%)	128 (91%)	13 (9%)	0	100	100
24	U	104/117 (89%)	88 (85%)	11 (11%)	5 (5%)	3	31
25	V	85/87 (98%)	79 (93%)	4 (5%)	2 (2%)	7	47

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
26	W	127/130 (98%)	114 (90%)	9 (7%)	4 (3%)	5	41
27	X	142/145 (98%)	121 (85%)	13 (9%)	8 (6%)	2	28
28	Y	132/135 (98%)	120 (91%)	8 (6%)	4 (3%)	5	42
29	Z	68/108 (63%)	55 (81%)	12 (18%)	1 (2%)	13	57
30	a	96/119 (81%)	83 (86%)	8 (8%)	5 (5%)	2	29
31	b	79/82 (96%)	69 (87%)	8 (10%)	2 (2%)	7	46
32	c	60/67 (90%)	56 (93%)	3 (5%)	1 (2%)	11	55
33	d	51/56 (91%)	37 (72%)	12 (24%)	2 (4%)	4	36
34	e	52/63 (82%)	41 (79%)	9 (17%)	2 (4%)	4	37
35	f	67/150 (45%)	50 (75%)	12 (18%)	5 (8%)	1	20
36	g	312/326 (96%)	267 (86%)	39 (12%)	6 (2%)	10	52
37	h	23/25 (92%)	23 (100%)	0	0	100	100
38	i	93/153 (61%)	82 (88%)	8 (9%)	3 (3%)	5	40
39	j	243/304 (80%)	212 (87%)	24 (10%)	7 (3%)	6	43
40	k	388/527 (74%)	335 (86%)	39 (10%)	14 (4%)	4	38
41	l	120/285 (42%)	106 (88%)	13 (11%)	1 (1%)	24	69
42	m	84/108 (78%)	74 (88%)	8 (10%)	2 (2%)	7	47
45	q	340/347 (98%)	307 (90%)	29 (8%)	4 (1%)	16	61
46	r	32/34 (94%)	30 (94%)	2 (6%)	0	100	100
47	s	50/52 (96%)	44 (88%)	5 (10%)	1 (2%)	9	51
All	All	6118/7123 (86%)	5319 (87%)	594 (10%)	205 (3%)	8	40

5 of 205 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
11	H	31	SER
11	H	64	VAL
11	H	74	GLN
14	K	88	PRO
19	P	126	VAL

### 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	
4	A	174/211 (82%)	159 (91%)	15 (9%)	13	47
5	B	198/228 (87%)	182 (92%)	16 (8%)	15	50
6	C	176/203 (87%)	161 (92%)	15 (8%)	13	48
7	D	185/196 (94%)	171 (92%)	14 (8%)	16	53
8	E	223/224 (100%)	201 (90%)	22 (10%)	10	39
9	F	174/194 (90%)	156 (90%)	18 (10%)	9	37
10	G	192/200 (96%)	184 (96%)	8 (4%)	36	70
11	H	164/170 (96%)	146 (89%)	18 (11%)	8	34
12	I	147/159 (92%)	138 (94%)	9 (6%)	23	60
13	J	153/158 (97%)	144 (94%)	9 (6%)	24	61
14	K	88/96 (92%)	82 (93%)	6 (7%)	20	57
15	L	136/137 (99%)	131 (96%)	5 (4%)	41	73
16	M	93/109 (85%)	90 (97%)	3 (3%)	46	76
17	N	127/128 (99%)	119 (94%)	8 (6%)	22	59
18	O	96/104 (92%)	90 (94%)	6 (6%)	22	59
19	P	101/119 (85%)	89 (88%)	12 (12%)	6	31
20	Q	117/119 (98%)	103 (88%)	14 (12%)	6	31
21	R	102/124 (82%)	93 (91%)	9 (9%)	12	45
22	S	128/129 (99%)	117 (91%)	11 (9%)	13	47
23	T	117/118 (99%)	104 (89%)	13 (11%)	8	34
24	U	96/107 (90%)	92 (96%)	4 (4%)	36	70
25	V	73/73 (100%)	67 (92%)	6 (8%)	14	49
26	W	110/111 (99%)	106 (96%)	4 (4%)	42	74
27	X	119/120 (99%)	109 (92%)	10 (8%)	14	49
28	Y	108/109 (99%)	101 (94%)	7 (6%)	21	58
29	Z	60/88 (68%)	56 (93%)	4 (7%)	20	57
30	a	83/100 (83%)	77 (93%)	6 (7%)	18	55
31	b	71/72 (99%)	68 (96%)	3 (4%)	36	70
32	c	54/59 (92%)	51 (94%)	3 (6%)	26	62

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
33	d	46/48 (96%)	43 (94%)	3 (6%)	21	58
34	e	47/55 (86%)	42 (89%)	5 (11%)	8	36
35	f	58/133 (44%)	51 (88%)	7 (12%)	6	31
36	g	265/272 (97%)	244 (92%)	21 (8%)	15	51
37	h	23/23 (100%)	22 (96%)	1 (4%)	35	70
38	i	81/130 (62%)	75 (93%)	6 (7%)	17	54
39	j	224/274 (82%)	193 (86%)	31 (14%)	4	27
40	k	332/449 (74%)	324 (98%)	8 (2%)	57	82
41	l	119/246 (48%)	114 (96%)	5 (4%)	36	70
42	m	77/96 (80%)	76 (99%)	1 (1%)	76	89
45	q	297/301 (99%)	290 (98%)	7 (2%)	57	82
46	r	33/33 (100%)	33 (100%)	0	100	100
47	s	43/43 (100%)	43 (100%)	0	100	100
All	All	5310/6068 (88%)	4937 (93%)	373 (7%)	23	56

5 of 373 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
18	O	65	GLN
22	S	8	GLN
39	j	159	GLU
19	P	23	GLU
20	Q	69	VAL

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 57 such sidechains are listed below:

Mol	Chain	Res	Type
20	Q	94	GLN
23	T	16	ASN
40	k	465	ASN
20	Q	100	GLN
21	R	29	GLN

### 5.3.3 RNA ⓘ



Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
1	1	73/75 (97%)	38 (52%)	12 (16%)
2	2	1778/1781 (99%)	814 (45%)	118 (6%)
3	3	1/25 (4%)	0	0
All	All	1852/1881 (98%)	852 (46%)	130 (7%)

5 of 852 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
1	1	8	U
1	1	9	1MG
1	1	10	2MG
1	1	11	C
1	1	12	G

5 of 130 RNA pucker outliers are listed below:

Mol	Chain	Res	Type
2	2	619	A
2	2	822	G
2	2	1571	A
2	2	654	G
2	2	724	G

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

11 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	1	10	1	18,26,27	1.43	2 (11%)	21,38,41	2.44	7 (33%)
1	H2U	1	16	1	17,21,22	0.85	1 (5%)	23,30,33	1.41	5 (21%)
1	M2G	1	26	1	18,27,28	1.70	3 (16%)	22,40,43	2.19	6 (27%)
1	T6A	1	37	1	23,34,35	1.17	1 (4%)	26,49,52	2.93	8 (30%)
1	7MG	1	46	1	20,26,27	1.77	3 (15%)	23,39,42	3.16	6 (26%)



Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
1	H2U	1	47	1	17,21,22	0.83	0	23,30,33	1.94	5 (21%)
1	5MC	1	48	1	14,22,23	1.62	2 (14%)	17,32,35	1.18	2 (11%)
1	5MC	1	49	1	14,22,23	1.61	1 (7%)	17,32,35	1.14	1 (5%)
1	1MA	1	58	1	15,25,26	2.05	4 (26%)	15,37,40	1.18	1 (6%)
1	RIA	1	64	1	31,38,39	1.23	2 (6%)	38,57,60	2.08	8 (21%)
1	1MG	1	9	1	17,26,27	1.71	3 (17%)	19,39,42	1.08	2 (10%)

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	1	10	1	-	0/5/27/28	0/3/3/3
1	H2U	1	16	1	-	0/7/38/39	0/2/2/2
1	M2G	1	26	1	-	0/7/29/30	0/3/3/3
1	T6A	1	37	1	-	1/15/41/42	0/3/3/3
1	7MG	1	46	1	-	0/7/37/38	0/3/3/3
1	H2U	1	47	1	-	0/7/38/39	0/2/2/2
1	5MC	1	48	1	-	0/3/25/26	0/2/2/2
1	5MC	1	49	1	-	0/3/25/26	0/2/2/2
1	1MA	1	58	1	-	0/3/25/26	0/3/3/3
1	RIA	1	64	1	-	0/13/51/52	0/4/4/4
1	1MG	1	9	1	-	0/3/25/26	0/3/3/3

The worst 5 of 22 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	1	64	RIA	P'-O3X	-4.10	1.40	1.54
1	1	16	H2U	C2-N3	-2.04	1.34	1.38
1	1	64	RIA	C2-N1	2.04	1.37	1.33
1	1	48	5MC	O4'-C1'	2.17	1.44	1.41
1	1	46	7MG	C1'-N9	2.60	1.50	1.44

The worst 5 of 51 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	1	46	7MG	C5-C4-N3	-8.53	118.05	126.74
1	1	37	T6A	N3-C2-N1	-7.54	122.95	128.87
1	1	64	RIA	O2A-C1'-C2'	-6.69	94.14	107.91
1	1	47	H2U	C5-C6-N1	-5.11	105.17	110.76

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed( $^{\circ}$ )	Ideal( $^{\circ}$ )
1	1	46	7MG	C5-C6-N1	-4.94	116.03	123.39

There are no chirality outliers.

All (1) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
1	1	37	T6A	C5-C6-N6-C10

There are no ring outliers.

1 monomer is involved in 6 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	1	16	H2U	6	0

## 5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

Of 88 ligands modelled in this entry, 86 are monoatomic - leaving 2 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	$\# Z  > 2$
50	MET	k	601	-	5,7,8	0.41	0	4,7,9	0.94	0
51	GCP	k	603	48	29,34,34	1.79	7 (24%)	32,54,54	1.77	5 (15%)

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
50	MET	k	601	-	-	0/4/6/8	0/0/0/0
51	GCP	k	603	48	-	0/15/38/38	0/3/3/3

The worst 5 of 7 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
51	k	603	GCP	PG-O3G	-2.63	1.48	1.54
51	k	603	GCP	PB-O2B	2.14	1.61	1.56
51	k	603	GCP	PG-O2G	2.58	1.61	1.54
51	k	603	GCP	PB-O3A	2.69	1.61	1.58
51	k	603	GCP	C5-C4	3.35	1.48	1.40

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
51	k	603	GCP	C5-C6-N1	-3.99	118.31	123.52
51	k	603	GCP	N3-C2-N1	-3.71	122.51	127.56
51	k	603	GCP	C6-C5-C4	-3.38	116.99	120.86
51	k	603	GCP	O4'-C1'-N9	2.92	113.61	108.11
51	k	603	GCP	C6-N1-C2	5.48	122.31	115.88

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](#)

There are no chain breaks in this entry.