wwPDB EM Map/Model Validation Report (i)

Apr 10, 2016 – 02:22 PM BST

PDB ID	:	3J7A
EMDB ID:	:	EMD-2660
Title	:	Cryo-EM structure of the Plasmodium falciparum 80S ribosome bound to the anti-protozoan drug emetine, small subunit
Authors	:	Wong, W.; Bai, X.C.; Brown, A.; Fernandez, I.S.; Hanssen, E.; Condron, M.; Tan, Y.H.; Baum, J.; Scheres, S.H.W.
Deposited on	:	2014-06-03
Resolution	:	3.20 Å(reported)

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
A user guide is available at
http://wwpdb.org/validation/2016/EMValidationReportHelp

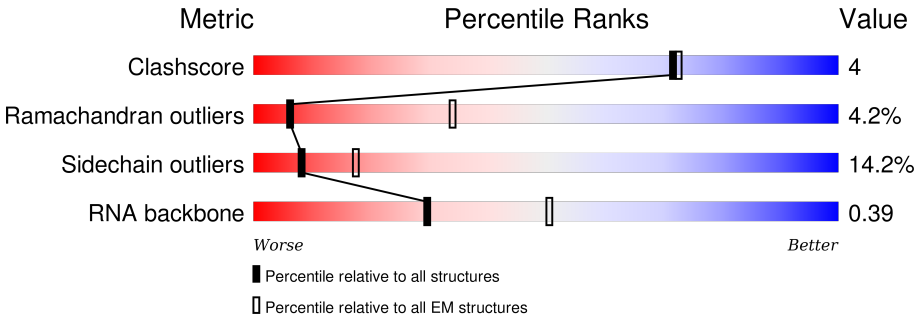
MolProbity	:	4.02b-467
Mogul	:	1.7.1 (RC1), CSD as537be (2016)
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP)	:	trunk27241

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 3.20 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



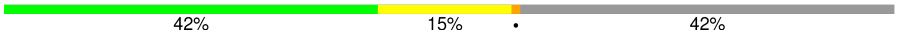
Metric	Whole archive (#Entries)	EM structures (#Entries)
Clashscore	114402	924
Ramachandran outliers	111179	726
Sidechain outliers	111093	686
RNA backbone	3027	244

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 >=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 <=5%

Mol	Chain	Length	Quality of chain
1	A	2092	45% 26% 5% 23%
2	B	262	53% 22% 5% 20%
3	C	263	55% 17% • 26%
4	D	221	56% 14% • 29%
5	E	189	71% 26% • •
6	F	261	71% 23% • •
7	G	272	65% 15% • 18%
8	H	306	50% 14% • 33%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
9	I	195	
10	J	194	
11	K	130	
12	L	218	
13	M	144	
14	N	118	
15	O	137	
16	P	151	
17	Q	145	
18	R	141	
19	S	156	
20	T	54	
21	U	151	
22	V	161	
23	W	137	
24	X	145	
25	Y	170	
26	Z	82	
27	1	133	
28	2	105	
29	3	107	
30	4	82	
31	5	67	
32	6	58	
33	7	74	

2 Entry composition

There are 36 unique types of molecules in this entry. The entry contains 68866 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 18S ribosomal RNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
1	A	1608	34277	15347	6109	11213	1608	0	0

- Molecule 2 is a protein called 40S ribosomal protein eS1.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
2	B	210	1713	1097	301	303	12	0	0

- Molecule 3 is a protein called 40S ribosomal protein uS2.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
3	C	195	1538	990	266	273	9	0	0

- Molecule 4 is a protein called 40S ribosomal protein uS3.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
4	D	157	1228	782	225	214	7	0	0

- Molecule 5 is a protein called 40S ribosomal protein uS4.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
5	E	185	1508	959	287	260	2	0	0

- Molecule 6 is a protein called 40S ribosomal protein eS4.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
6	F	257	2061	1320	377	356	8	0	0

- Molecule 7 is a protein called 40S ribosomal protein uS5.

Mol	Chain	Residues	Atoms					AltConf	Trace
7	G	224	Total	C	N	O	S	0	0
			1757	1132	307	309	9		

- Molecule 8 is a protein called 40S ribosomal protein eS6.

Mol	Chain	Residues	Atoms					AltConf	Trace
8	H	204	Total	C	N	O	S	0	0
			1644	1042	313	283	6		

- Molecule 9 is a protein called 40S ribosomal protein uS7.

Mol	Chain	Residues	Atoms					AltConf	Trace
9	I	180	Total	C	N	O	S	0	0
			1424	893	263	258	10		

- Molecule 10 is a protein called 40S ribosomal protein eS7.

Mol	Chain	Residues	Atoms					AltConf	Trace
10	J	188	Total	C	N	O	S	0	0
			1528	982	264	278	4		

- Molecule 11 is a protein called 40S ribosomal protein uS8.

Mol	Chain	Residues	Atoms					AltConf	Trace
11	K	129	Total	C	N	O	S	0	0
			1037	665	189	178	5		

- Molecule 12 is a protein called 40S ribosomal protein eS8.

Mol	Chain	Residues	Atoms					AltConf	Trace
12	L	172	Total	C	N	O	S	0	0
			1392	878	266	244	4		

- Molecule 13 is a protein called 40S ribosomal protein uS9.

Mol	Chain	Residues	Atoms					AltConf	Trace
13	M	138	Total	C	N	O	S	0	0
			1098	704	200	193	1		

- Molecule 14 is a protein called 40S ribosomal protein uS10.

Mol	Chain	Residues	Atoms					AltConf	Trace
14	N	98	Total	C	N	O	S	0	0
			772	484	135	148	5		

- Molecule 15 is a protein called 40S ribosomal protein eS10.

Mol	Chain	Residues	Atoms					AltConf	Trace
15	O	79	Total	C	N	O	S	0	0
			686	450	116	118	2		

- Molecule 16 is a protein called 40S ribosomal protein uS11.

Mol	Chain	Residues	Atoms					AltConf	Trace
16	P	127	Total	C	N	O	S	0	0
			953	591	184	175	3		

- Molecule 17 is a protein called 40S ribosomal protein uS12.

Mol	Chain	Residues	Atoms					AltConf	Trace
17	Q	144	Total	C	N	O	S	0	0
			1129	712	222	193	2		

- Molecule 18 is a protein called 40S ribosomal protein eS12.

Mol	Chain	Residues	Atoms					AltConf	Trace
18	R	98	Total	C	N	O	S	0	0
			746	474	123	145	4		

- Molecule 19 is a protein called 40S ribosomal protein uS13.

Mol	Chain	Residues	Atoms					AltConf	Trace
19	S	128	Total	C	N	O	S	0	0
			1042	655	204	179	4		

- Molecule 20 is a protein called 40S ribosomal protein uS14.

Mol	Chain	Residues	Atoms					AltConf	Trace
20	T	48	Total	C	N	O	S	0	0
			404	252	85	63	4		

- Molecule 21 is a protein called 40S ribosomal protein uS15.

Mol	Chain	Residues	Atoms					AltConf	Trace
21	U	149	Total	C	N	O	S	0	0
			1202	769	220	210	3		

- Molecule 22 is a protein called 40S ribosomal protein uS17.

Mol	Chain	Residues	Atoms					AltConf	Trace
22	V	146	Total	C	N	O	S	0	0
			1206	772	227	200	7		

- Molecule 23 is a protein called 40S ribosomal protein eS17.

Mol	Chain	Residues	Atoms					AltConf	Trace
23	W	95	Total	C	N	O	S	0	0
			785	498	149	135	3		

- Molecule 24 is a protein called 40S ribosomal protein uS19.

Mol	Chain	Residues	Atoms					AltConf	Trace
24	X	96	Total	C	N	O	S	0	0
			776	497	137	138	4		

- Molecule 25 is a protein called 40S ribosomal protein eS19.

Mol	Chain	Residues	Atoms					AltConf	Trace
25	Y	154	Total	C	N	O	S	0	0
			1266	811	239	214	2		

- Molecule 26 is a protein called 40S ribosomal protein eS21.

Mol	Chain	Residues	Atoms					AltConf	Trace
26	Z	72	Total	C	N	O	S	0	0
			556	346	102	104	4		

- Molecule 27 is a protein called 40S ribosomal protein eS24.

Mol	Chain	Residues	Atoms					AltConf	Trace
27	1	120	Total	C	N	O	S	0	0
			981	629	188	162	2		

- Molecule 28 is a protein called 40S ribosomal protein eS25.

Mol	Chain	Residues	Atoms				AltConf	Trace
28	2	41	Total	C	N	O	0	0
			320	208	56	56		

- Molecule 29 is a protein called 40S ribosomal protein eS26.

Mol	Chain	Residues	Atoms					AltConf	Trace
29	3	95	Total	C	N	O	S	0	0
			781	478	169	128	6		

- Molecule 30 is a protein called 40S ribosomal protein eS27.

Mol	Chain	Residues	Atoms					AltConf	Trace
30	4	76	Total	C	N	O	S	0	0
			586	368	102	107	9		

- Molecule 31 is a protein called 40S ribosomal protein eS28.

Mol	Chain	Residues	Atoms				AltConf	Trace
31	5	58	Total	C	N	O	0	0
			451	282	90	79		

- Molecule 32 is a protein called 40S ribosomal protein eS30.

Mol	Chain	Residues	Atoms				AltConf	Trace
32	6	43	Total	C	N	O	0	0
			345	213	75	57		

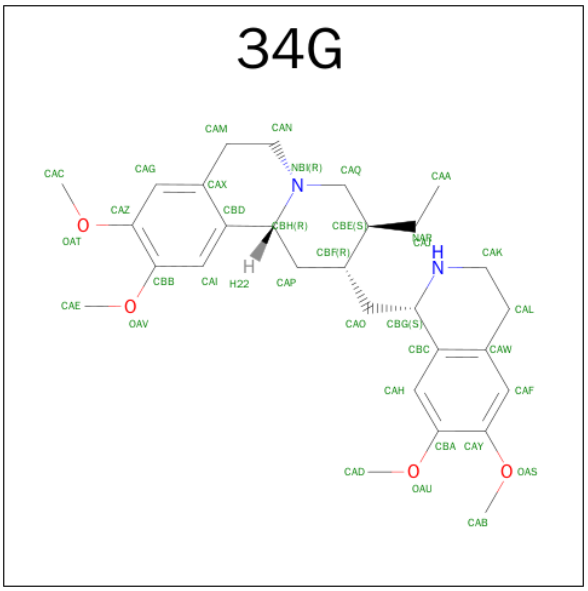
- Molecule 33 is a RNA chain called tRNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
33	7	74	Total	C	N	O	P	0	0
			1571	702	275	521	73		

- Molecule 34 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		AltConf
34	A	67	Total	Mg	0
			67	67	

- Molecule 35 is EMETINE (three-letter code: 34G) (formula: $C_{29}H_{40}N_2O_4$).



Mol	Chain	Residues	Atoms				AltConf
			Total	C	N	O	
35	A	1	35	29	2	4	0

- Molecule 36 is ZINC ION (three-letter code: ZN) (formula: Zn).

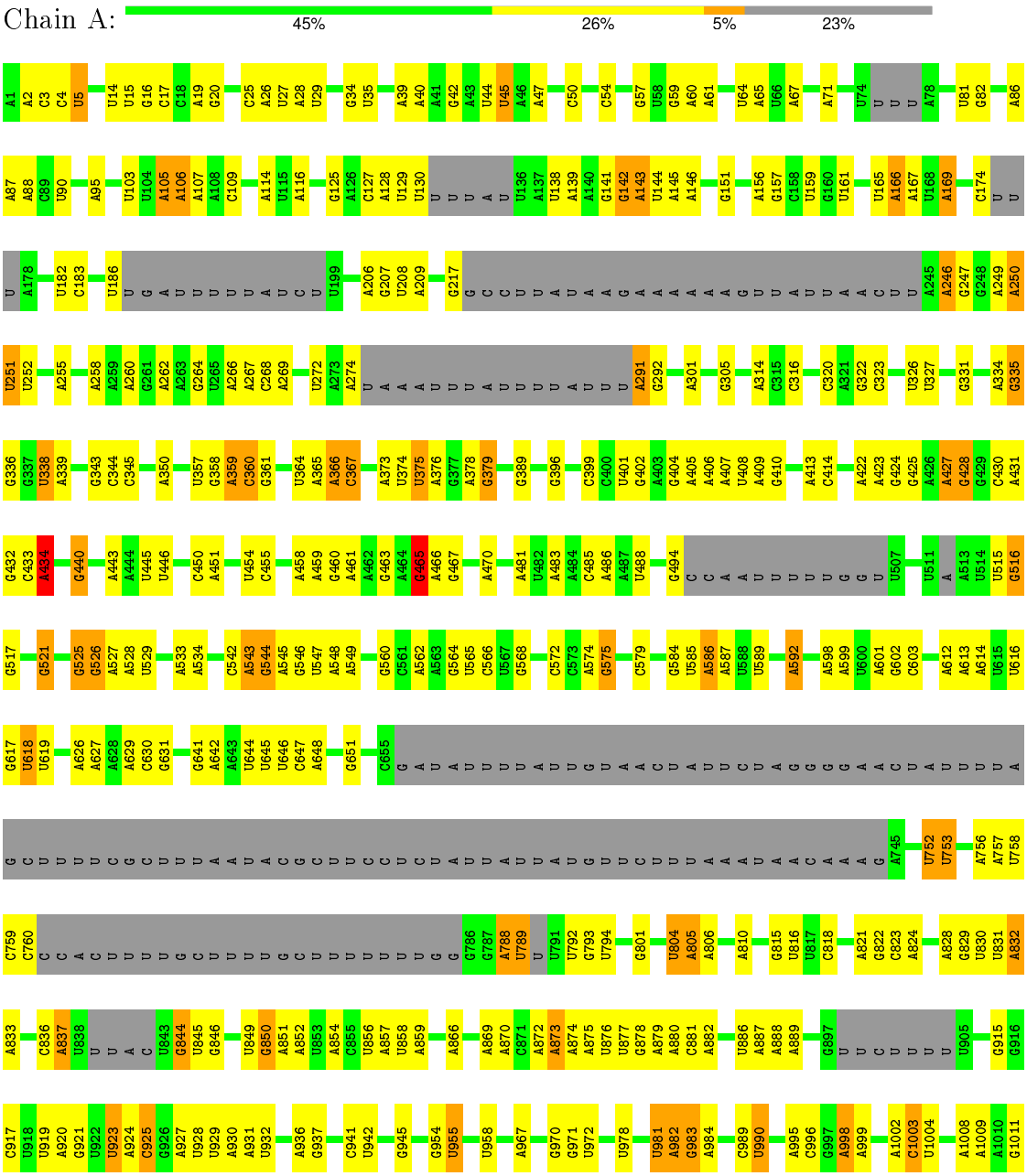
Mol	Chain	Residues	Atoms		AltConf
			Total	Zn	
36	T	1	1	1	0

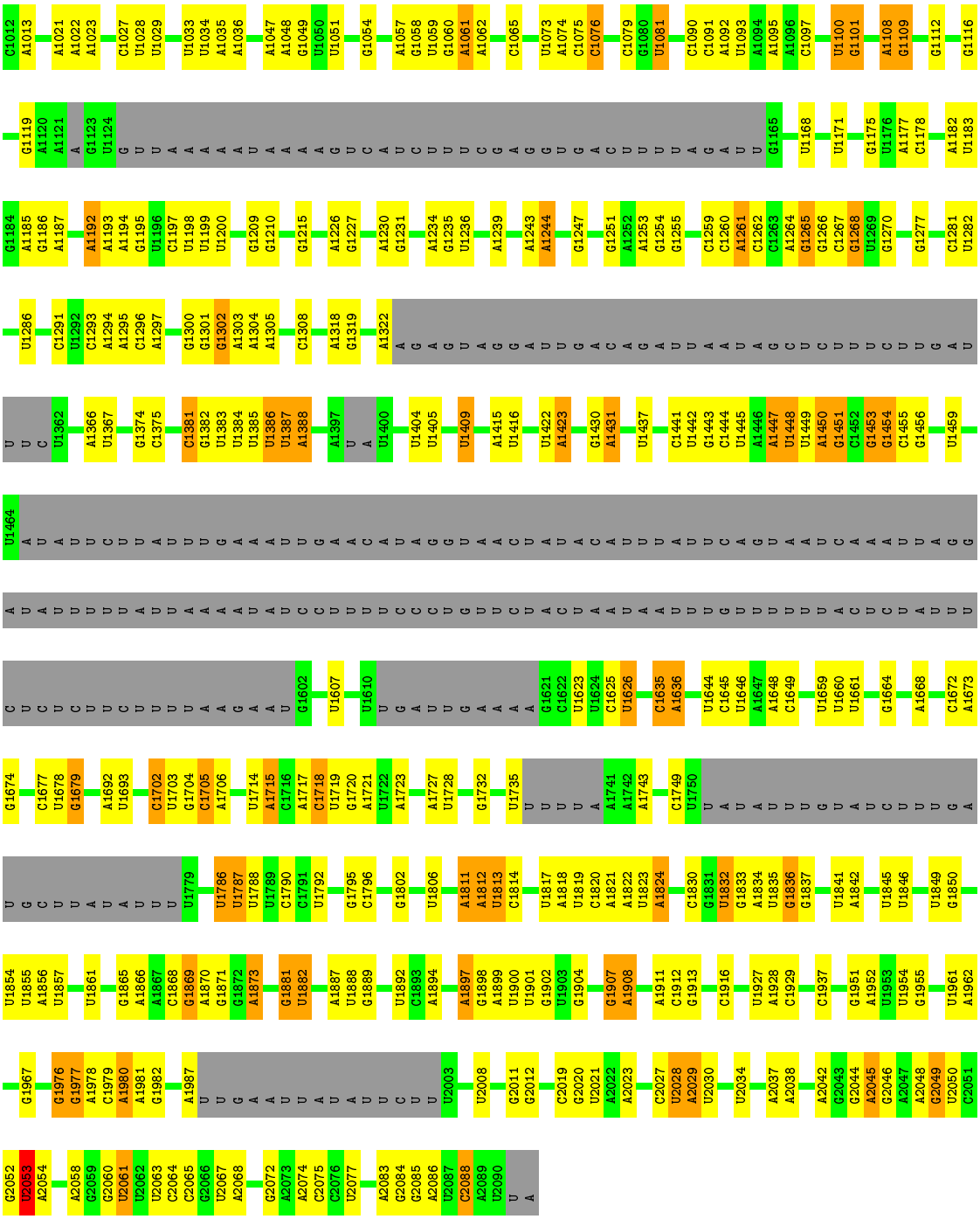
3 Residue-property plots (i)

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: 18S ribosomal RNA

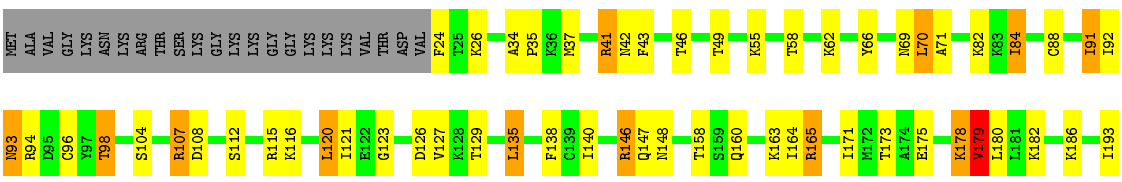
Chain A:





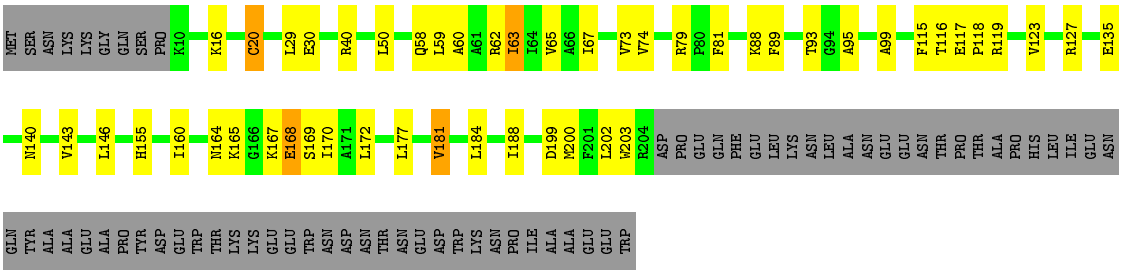
- Molecule 2: 40S ribosomal protein eS1

Chain B:



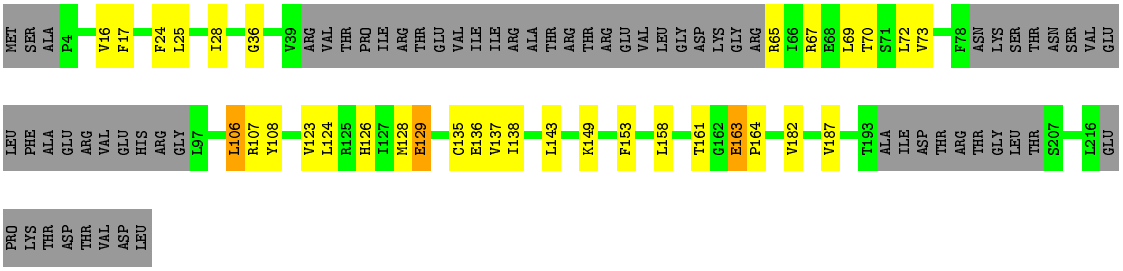
- Molecule 3: 40S ribosomal protein uS2

Chain C:



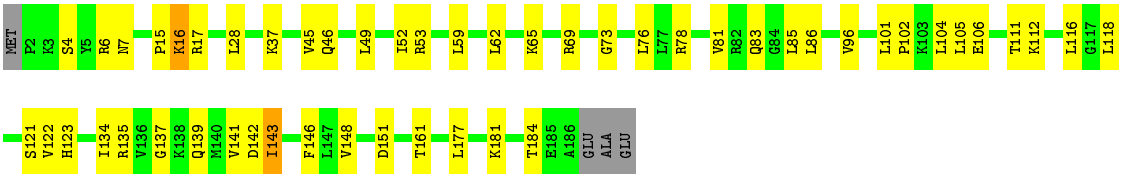
- Molecule 4: 40S ribosomal protein uS3

Chain D:



- Molecule 5: 40S ribosomal protein uS4

Chain E:

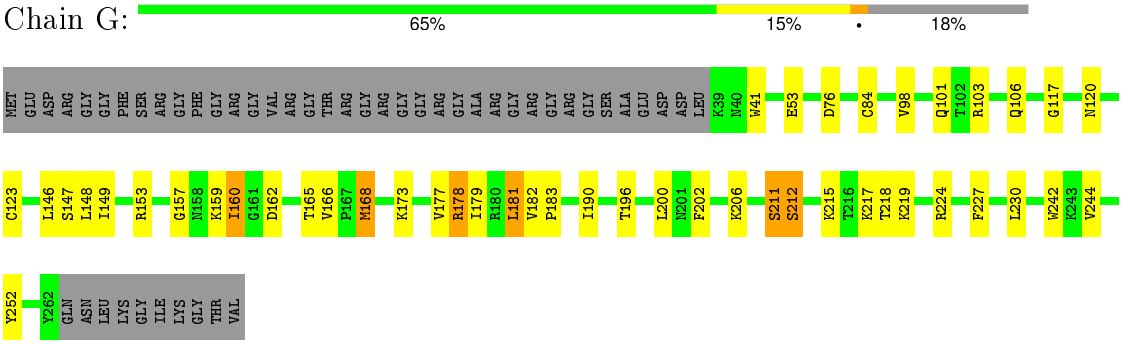


- Molecule 6: 40S ribosomal protein eS4

Chain F:

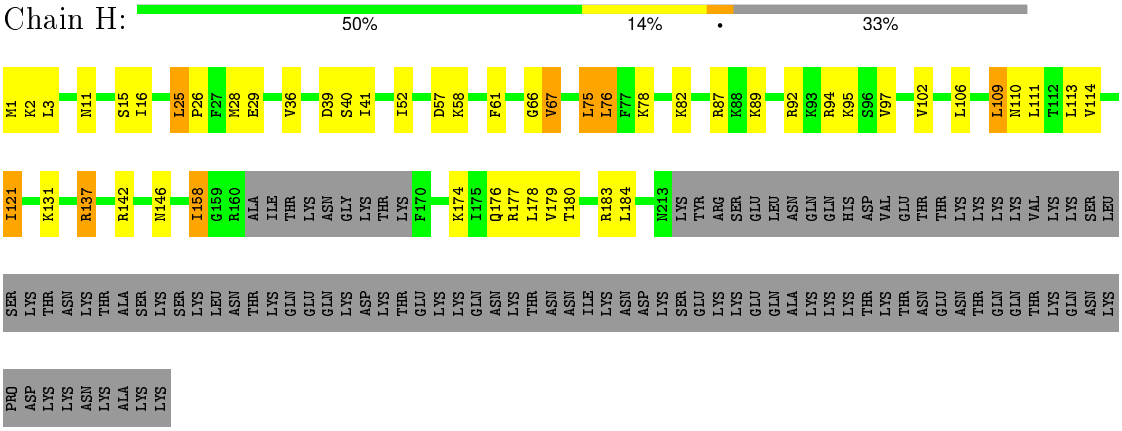
- Molecule 7: 40S ribosomal protein uS5

Chain G:



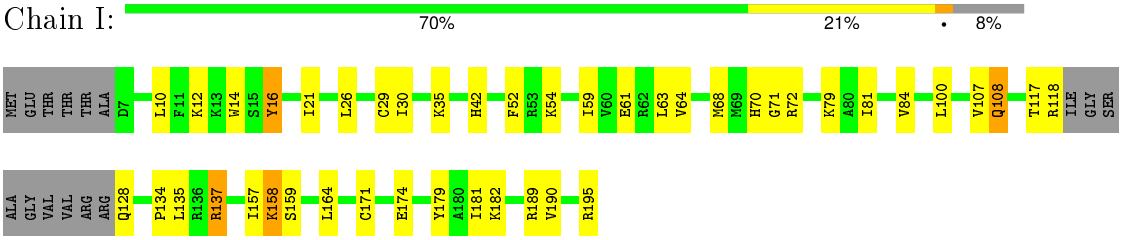
- Molecule 8: 40S ribosomal protein eS6

Chain H:



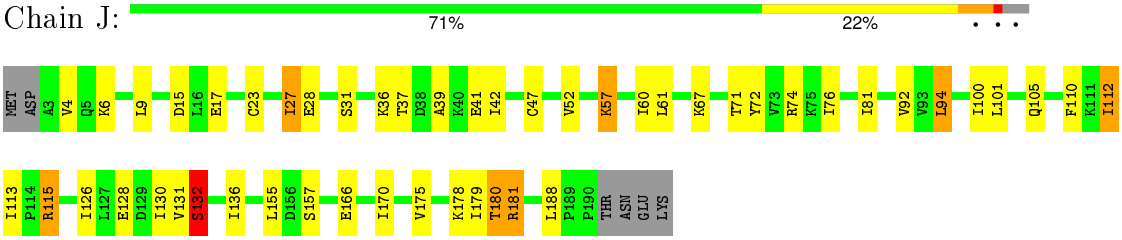
- Molecule 9: 40S ribosomal protein uS7

Chain I:



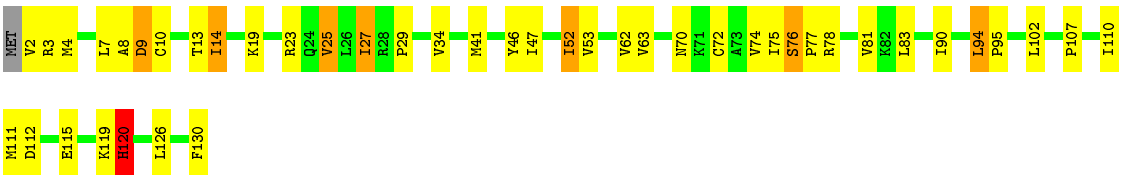
- Molecule 10: 40S ribosomal protein eS7

Chain J:



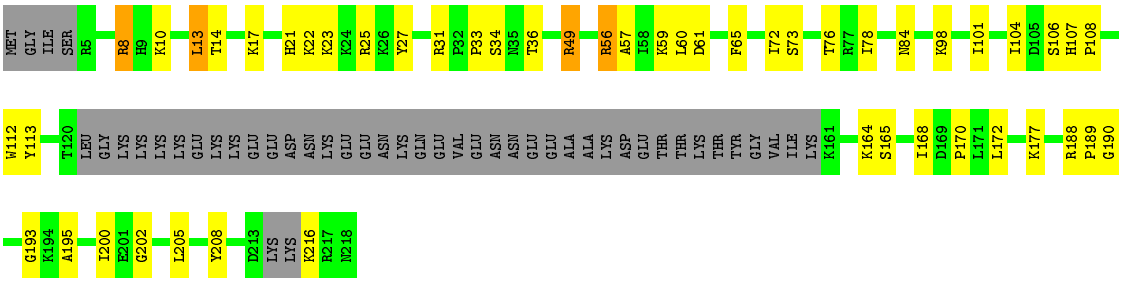
- Molecule 11: 40S ribosomal protein uS8

Chain K:



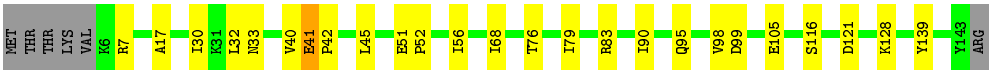
- Molecule 12: 40S ribosomal protein eS8

Chain L:



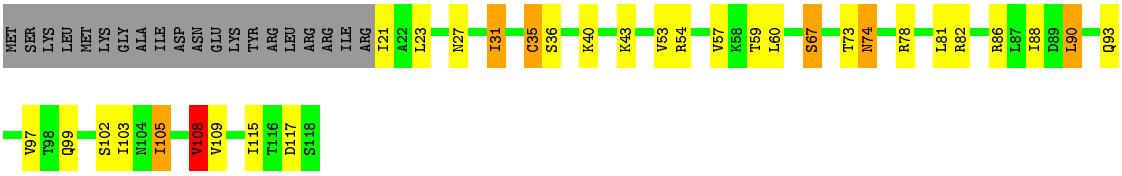
- Molecule 13: 40S ribosomal protein uS9

Chain M:



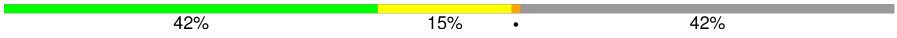
- Molecule 14: 40S ribosomal protein uS10

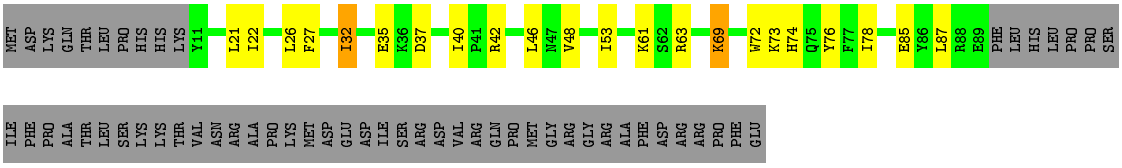
Chain N:



- Molecule 15: 40S ribosomal protein eS10

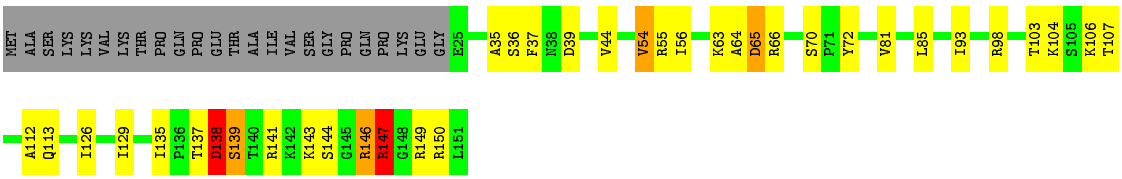
Chain O:





- Molecule 16: 40S ribosomal protein uS11

Chain P:

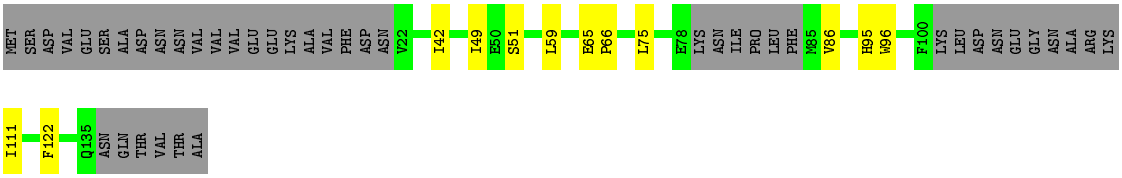


- Molecule 17: 40S ribosomal protein uS12

Chain Q:

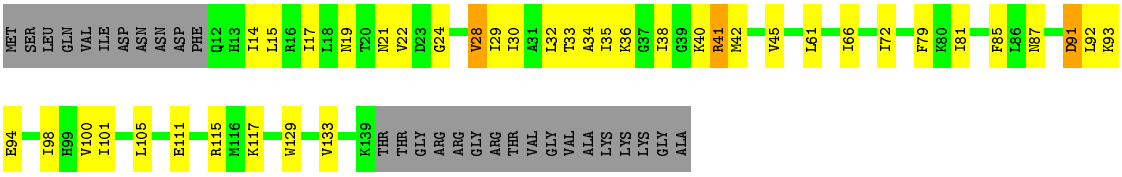
- Molecule 18: 40S ribosomal protein eS12

Chain R:



- Molecule 19: 40S ribosomal protein uS13

Chain S:

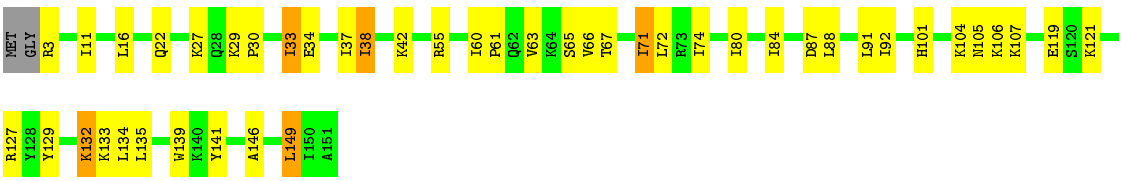


- Molecule 20: 40S ribosomal protein uS14

Chain T:

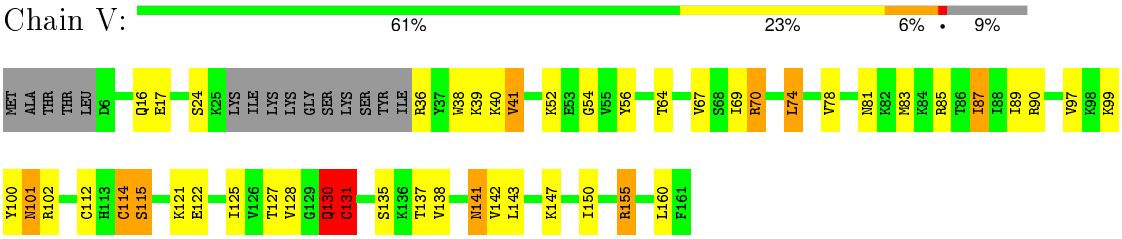
- Molecule 21: 40S ribosomal protein uS15

Chain U:



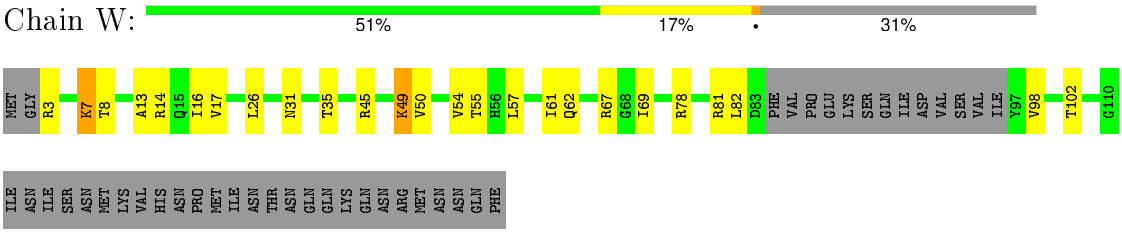
- Molecule 22: 40S ribosomal protein uS17

Chain V:



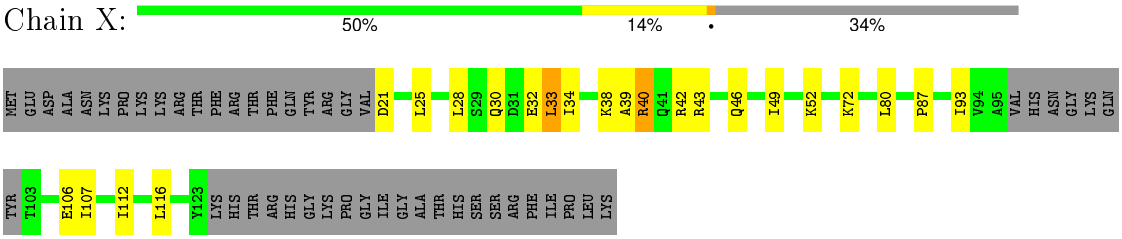
• Molecule 23: 40S ribosomal protein eS17

Chain W:



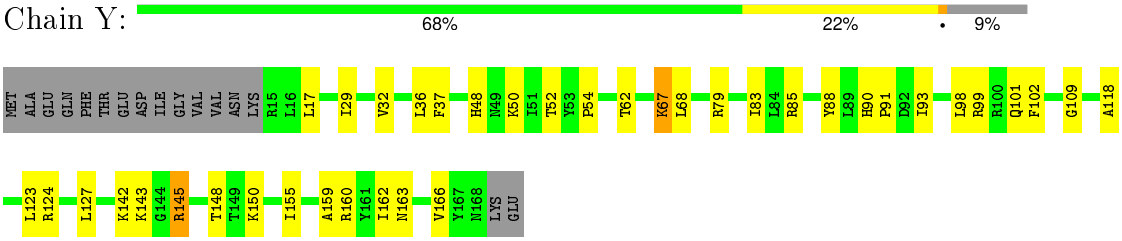
• Molecule 24: 40S ribosomal protein uS19

Chain X:



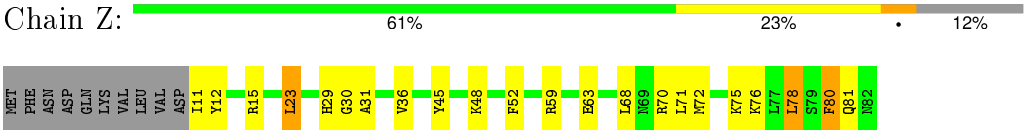
• Molecule 25: 40S ribosomal protein eS19

Chain Y:



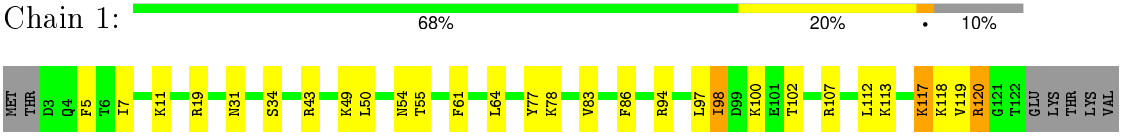
• Molecule 26: 40S ribosomal protein eS21

Chain Z:



• Molecule 27: 40S ribosomal protein eS24

Chain 1:



- Molecule 28: 40S ribosomal protein eS25

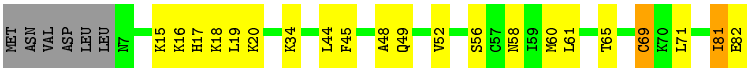
Chain 2:

- Molecule 29: 40S ribosomal protein eS26

Chain 3:

- Molecule 30: 40S ribosomal protein eS27

Chain 4:

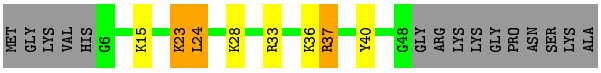


- Molecule 31: 40S ribosomal protein eS28

Chain 5:

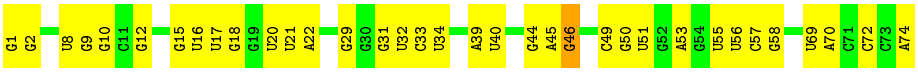
- Molecule 32: 40S ribosomal protein eS30

Chain 6:



- Molecule 33: tRNA

Chain 7:



4 Experimental information (i)

Property	Value	Source
Reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	Depositor
Number of images	105247	Depositor
Resolution determination method	FSC 0.143 gold-standard	Depositor
CTF correction method	Each particle	Depositor
Microscope	FEI POLARA 300	Depositor
Voltage (kV)	300	Depositor
Electron dose ($e^-/Å^2$)	20	Depositor
Minimum defocus (nm)	800	Depositor
Maximum defocus (nm)	3800	Depositor
Magnification	78000	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, 34G, MG

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 5$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	#$\|Z\|>2$	RMSZ	#$\|Z\|>2$
1	A	0.35	3/38345 (0.0%)	0.75	16/59689 (0.0%)
10	J	0.45	0/1544	0.78	0/2064
11	K	0.51	0/1054	0.92	1/1411 (0.1%)
12	L	0.51	0/1416	0.82	1/1890 (0.1%)
13	M	0.45	0/1113	0.71	0/1487
14	N	0.45	0/780	0.81	0/1053
15	O	0.48	0/705	0.73	0/950
16	P	0.47	0/965	0.88	1/1295 (0.1%)
17	Q	0.47	0/1149	0.80	0/1532
18	R	0.47	0/754	0.66	0/1013
19	S	0.48	0/1058	0.82	0/1420
2	B	0.46	0/1737	0.81	0/2321
20	T	0.42	0/411	0.73	0/544
21	U	0.45	0/1223	0.87	0/1634
22	V	0.50	0/1233	0.79	1/1645 (0.1%)
23	W	0.47	0/792	0.86	0/1053
24	X	0.49	0/787	0.81	0/1050
25	Y	0.46	0/1294	0.86	1/1742 (0.1%)
26	Z	0.44	0/564	0.78	0/758
27	1	0.49	0/994	0.86	0/1317
28	2	0.48	0/323	0.67	0/435
29	3	0.48	0/793	0.90	0/1055
3	C	0.43	0/1569	0.79	0/2129
30	4	0.45	0/597	0.73	0/801
31	5	0.40	0/452	0.64	0/599
32	6	0.42	0/348	0.80	0/458
33	7	0.24	0/1754	0.72	0/2732
4	D	0.46	0/1240	0.79	1/1652 (0.1%)
5	E	0.47	0/1532	0.87	0/2048
6	F	0.47	0/2097	0.80	1/2819 (0.0%)
7	G	0.48	0/1799	0.78	0/2429
8	H	0.43	0/1661	0.77	0/2205

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	$\#\|Z\|>2$	RMSZ	$\#\|Z\|>2$
9	I	0.47	0/1443	0.86	0/1936
All	All	0.40	3/73526 (0.0%)	0.77	23/107166 (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
10	J	0	1
11	K	0	3
12	L	0	1
19	S	0	1
2	B	0	1
21	U	0	1
24	X	0	1
27	1	0	1
All	All	0	10

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	788	A	O3'-P	8.01	1.70	1.61
1	A	789	U	C1'-N1	6.93	1.59	1.48
1	A	788	A	C1'-N9	-5.21	1.39	1.46

The worst 5 of 23 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	788	A	P-O3'-C3'	-10.85	106.69	119.70
1	A	788	A	OP2-P-O3'	7.78	122.32	105.20
1	A	2053	U	C2'-C3'-O3'	7.76	126.57	109.50
1	A	850	G	C2'-C3'-O3'	7.72	126.49	109.50
1	A	1381	C	C2'-C3'-O3'	7.23	125.40	109.50

There are no chirality outliers.

5 of 10 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	B	148	ASN	Peptide
10	J	132	SER	Peptide

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Group
11	K	27	ILE	Peptide
11	K	76	SER	Peptide
11	K	94	LEU	Peptide

5.2 Too-close contacts (i)

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	34277	0	17246	210	0
2	B	1713	0	1838	25	0
3	C	1538	0	1600	17	0
4	D	1228	0	1311	14	0
5	E	1508	0	1594	16	0
6	F	2061	0	2200	23	0
7	G	1757	0	1811	12	0
8	H	1644	0	1795	15	0
9	I	1424	0	1471	9	0
10	J	1528	0	1680	20	0
11	K	1037	0	1099	16	0
12	L	1392	0	1447	20	0
13	M	1098	0	1183	7	0
14	N	772	0	813	8	0
15	O	686	0	695	7	0
16	P	953	0	997	20	0
17	Q	1129	0	1196	12	0
18	R	746	0	754	2	0
19	S	1042	0	1095	23	0
20	T	404	0	416	5	0
21	U	1202	0	1299	19	0
22	V	1206	0	1239	16	0
23	W	785	0	858	6	0
24	X	776	0	832	9	0
25	Y	1266	0	1316	12	0
26	Z	556	0	558	7	0
27	1	981	0	1065	9	0
28	2	320	0	338	0	0
29	3	781	0	818	12	0
30	4	586	0	604	5	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
31	5	451	0	485	2	0
32	6	345	0	381	4	0
33	7	1571	0	797	3	0
34	A	67	0	0	0	0
35	A	35	0	40	2	0
36	T	1	0	0	0	0
All	All	68866	0	52871	522	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 522 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)
19:S:33:THR:O	19:S:38:ILE:HG12	1.15	1.32
1:A:759:C:O2	1:A:788:A:C2	1.84	1.30
19:S:35:ILE:O	19:S:38:ILE:CG1	1.76	1.30
1:A:759:C:C2	1:A:788:A:C2	2.30	1.19
19:S:35:ILE:O	19:S:38:ILE:HG13	0.91	1.09

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	208/262 (79%)	168 (81%)	28 (14%)	12 (6%)	2 17
3	C	193/263 (73%)	163 (84%)	22 (11%)	8 (4%)	3 27
4	D	149/221 (67%)	129 (87%)	17 (11%)	3 (2%)	9 48
5	E	183/189 (97%)	159 (87%)	19 (10%)	5 (3%)	6 39
6	F	255/261 (98%)	217 (85%)	31 (12%)	7 (3%)	6 39

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
7	G	222/272 (82%)	195 (88%)	20 (9%)	7 (3%)	5 33
8	H	200/306 (65%)	173 (86%)	19 (10%)	8 (4%)	4 27
9	I	176/195 (90%)	153 (87%)	14 (8%)	9 (5%)	2 20
10	J	186/194 (96%)	160 (86%)	17 (9%)	9 (5%)	3 22
11	K	127/130 (98%)	104 (82%)	16 (13%)	7 (6%)	2 18
12	L	166/218 (76%)	134 (81%)	23 (14%)	9 (5%)	2 19
13	M	136/144 (94%)	116 (85%)	13 (10%)	7 (5%)	2 20
14	N	96/118 (81%)	83 (86%)	8 (8%)	5 (5%)	2 19
15	O	77/137 (56%)	66 (86%)	10 (13%)	1 (1%)	15 59
16	P	125/151 (83%)	103 (82%)	18 (14%)	4 (3%)	5 33
17	Q	142/145 (98%)	128 (90%)	13 (9%)	1 (1%)	26 72
18	R	92/141 (65%)	73 (79%)	13 (14%)	6 (6%)	1 13
19	S	126/156 (81%)	100 (79%)	17 (14%)	9 (7%)	1 10
20	T	46/54 (85%)	43 (94%)	2 (4%)	1 (2%)	8 45
21	U	147/151 (97%)	133 (90%)	10 (7%)	4 (3%)	6 39
22	V	142/161 (88%)	123 (87%)	9 (6%)	10 (7%)	1 10
23	W	91/137 (66%)	80 (88%)	7 (8%)	4 (4%)	3 24
24	X	92/145 (63%)	82 (89%)	6 (6%)	4 (4%)	3 25
25	Y	152/170 (89%)	134 (88%)	12 (8%)	6 (4%)	4 28
26	Z	70/82 (85%)	62 (89%)	3 (4%)	5 (7%)	1 10
27	1	118/133 (89%)	104 (88%)	9 (8%)	5 (4%)	3 26
28	2	35/105 (33%)	30 (86%)	5 (14%)	0	100 100
29	3	93/107 (87%)	80 (86%)	10 (11%)	3 (3%)	5 33
30	4	74/82 (90%)	48 (65%)	21 (28%)	5 (7%)	1 11
31	5	54/67 (81%)	50 (93%)	4 (7%)	0	100 100
32	6	41/58 (71%)	34 (83%)	4 (10%)	3 (7%)	1 9
All	All	4014/4955 (81%)	3427 (85%)	420 (10%)	167 (4%)	6 26

5 of 167 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	84	ILE
2	B	93	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	B	147	GLN
2	B	179	VAL
3	C	20	CYS

5.3.2 Protein sidechains (i)

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	195/238 (82%)	158 (81%)	37 (19%)	2 10
3	C	167/227 (74%)	144 (86%)	23 (14%)	4 21
4	D	132/188 (70%)	119 (90%)	13 (10%)	10 38
5	E	160/167 (96%)	137 (86%)	23 (14%)	4 19
6	F	233/237 (98%)	194 (83%)	39 (17%)	3 13
7	G	191/222 (86%)	160 (84%)	31 (16%)	3 14
8	H	181/279 (65%)	151 (83%)	30 (17%)	3 13
9	I	154/165 (93%)	128 (83%)	26 (17%)	2 13
10	J	177/183 (97%)	156 (88%)	21 (12%)	6 28
11	K	115/116 (99%)	94 (82%)	21 (18%)	2 10
12	L	152/193 (79%)	136 (90%)	16 (10%)	8 35
13	M	116/122 (95%)	110 (95%)	6 (5%)	29 69
14	N	91/109 (84%)	70 (77%)	21 (23%)	1 4
15	O	76/129 (59%)	65 (86%)	11 (14%)	4 19
16	P	99/119 (83%)	83 (84%)	16 (16%)	3 14
17	Q	120/121 (99%)	106 (88%)	14 (12%)	7 30
18	R	83/121 (69%)	81 (98%)	2 (2%)	57 86
19	S	113/136 (83%)	99 (88%)	14 (12%)	6 27
20	T	43/48 (90%)	38 (88%)	5 (12%)	7 30
21	U	132/133 (99%)	110 (83%)	22 (17%)	3 13
22	V	131/144 (91%)	106 (81%)	25 (19%)	2 10

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
23	W	86/127 (68%)	73 (85%)	13 (15%)	3 17
24	X	88/130 (68%)	80 (91%)	8 (9%)	12 42
25	Y	137/151 (91%)	117 (85%)	20 (15%)	4 19
26	Z	60/70 (86%)	50 (83%)	10 (17%)	3 13
27	1	103/115 (90%)	89 (86%)	14 (14%)	5 22
28	2	35/88 (40%)	30 (86%)	5 (14%)	4 19
29	3	87/98 (89%)	76 (87%)	11 (13%)	5 26
30	4	70/76 (92%)	60 (86%)	10 (14%)	4 19
31	5	46/54 (85%)	42 (91%)	4 (9%)	13 45
32	6	36/47 (77%)	33 (92%)	3 (8%)	14 49
All	All	3609/4353 (83%)	3095 (86%)	514 (14%)	8 19

5 of 514 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
10	J	112	ILE
14	N	31	ILE
27	1	98	ILE
10	J	178	LYS
11	K	115	GLU

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

Mol	Chain	Res	Type
11	K	24	GLN
11	K	64	ASN
23	W	31	ASN
9	I	194	ASN
27	1	52	ASN

5.3.3 RNA

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
1	A	1586/2092 (75%)	474 (29%)	71 (4%)
33	7	73/74 (98%)	30 (41%)	3 (4%)
All	All	1659/2166 (76%)	504 (30%)	74 (4%)

5 of 504 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
1	A	2	A
1	A	4	C
1	A	5	U
1	A	17	C
1	A	25	C

5 of 74 RNA pucker outliers are listed below:

Mol	Chain	Res	Type
1	A	923	U
1	A	1209	G
1	A	1976	G
1	A	930	A
1	A	1028	U

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

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates

There are no carbohydrates in this entry.

5.6 Ligand geometry

Of 69 ligands modelled in this entry, 68 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	$\#\|Z\| > 2$
35	34G	A	2168	-	38,39,39	2.84	9 (23%)	51,56,56	1.97	9 (17%)

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
35	34G	A	2168	-	-	0/14/49/49	0/5/5/5

The worst 5 of 9 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
35	A	2168	34G	CBC-CBG	-9.31	1.43	1.52
35	A	2168	34G	CAM-CAX	-7.61	1.38	1.51
35	A	2168	34G	CBD-CBH	-7.50	1.43	1.52
35	A	2168	34G	CAL-CAW	-5.58	1.41	1.51
35	A	2168	34G	CAP-CBH	-5.01	1.48	1.53

The worst 5 of 9 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
35	A	2168	34G	CAW-CBC-CBG	-7.74	113.78	121.78
35	A	2168	34G	CBF-CAP-CBH	-4.98	102.90	111.53
35	A	2168	34G	CAP-CBF-CBE	-3.91	104.18	110.59
35	A	2168	34G	CAE-OAV-CBB	-2.58	113.76	117.53
35	A	2168	34G	CAC-OAT-CAZ	-2.02	114.57	117.53

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
35	A	2168	34G	2	0

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.