



# wwPDB EM Model Validation Summary Report ⓘ

Apr 15, 2020 – 01:08 PM EDT

PDB ID : 6W2S  
EMDB ID : EMD-21529  
Title : Structure of the Cricket Paralysis Virus 5-UTR IRES (CrPV 5-UTR-IRES)  
bound to the small ribosomal subunit in the open state (Class 1)  
Authors : Neupane, R.; Pisareva, V.; Rodriguez, C.F.; Pisarev, A.; Fernandez, I.S.  
Deposited on : 2020-03-08  
Resolution : 3.00 Å(reported)

This is a wwPDB EM Model Validation Summary Report for a publicly released PDB/EMDB entry.

We welcome your comments at [validation@mail.wwpdb.org](mailto:validation@mail.wwpdb.org)

A user guide is available at

<https://www.wwpdb.org/validation/2017/EMValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

---

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Percentile statistics : 20171227.v01 (using entries in the PDB archive December 27th 2017)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.10.1

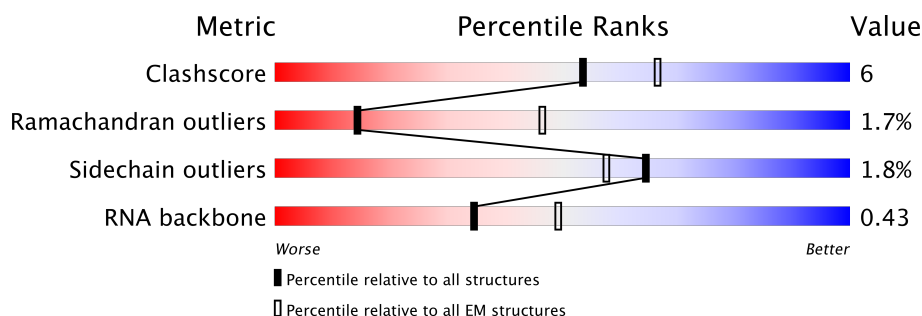
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*ELECTRON MICROSCOPY*

The reported resolution of this entry is 3.00 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.










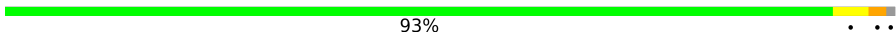



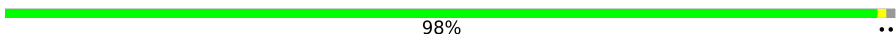













Metric	Whole archive (#Entries)	EM structures (#Entries)
Clashscore	136327	1886
Ramachandran outliers	132723	1663
Sidechain outliers	132532	1531
RNA backbone	3747	458

The table below summarises the geometric issues observed across the polymeric chains and their fit to the map. 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 respectively. 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	1697	69% 27% .
2	O	377	33% 44% 13% . 10%
3	B	295	65% 7% . 26%
4	C	264	76% 5% 19%
5	D	255	82% 5% 13%
6	F	263	94% 5%
7	H	249	87% 7% . 5%

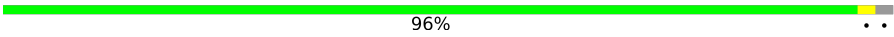

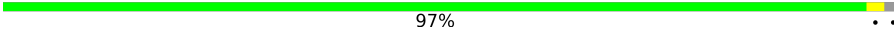

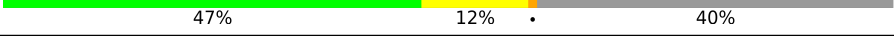

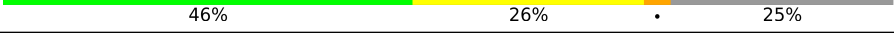

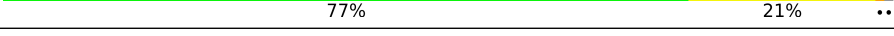
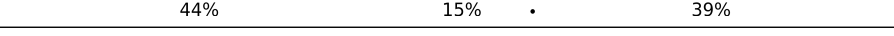
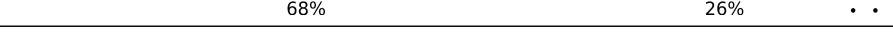
*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
8	I	194	 87% 8% 5%
9	J	208	 90% 9% .
10	K	194	 92% . 5%
11	M	158	 81% 9% 9%
12	O	151	 91% 8% .
13	P	151	 84% 6% 10%
14	W	83	 92% 8%
15	X	130	 93% . .
16	Y	143	 88% 10% ..
17	Z	134	 87% 6% 7%
18	b	115	 83% . 12%
19	c	84	 98% ..
20	f	133	 43% 57%
21	E	281	 75% 6% 19%
22	G	204	 82% 11% 6%
23	L	149	 56% 9% 36%
24	N	132	 68% 20% . 11%
25	Q	145	 61% 17% . 21%
26	R	172	 76% 6% . 17%
27	S	135	 79% 19% ..
28	T	152	 78% 16% 5%
29	U	145	 90% 7% .
30	V	119	 81% . 16%
31	a	125	 60% 40%
32	d	69	 90% 10%

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
33	e	56	 96% ..
34	g	156	 40% . 56%
35	h	317	 97% ..
36	1	1362	 28% 10% . 61%
37	2	913	 47% 12% . 40%
38	3	462	 65% 22% . 9%
39	4	364	 46% 26% . 25%
40	5	363	 55% 30% 5% 11%
41	6	218	 77% 21% ..
42	7	607	 44% 15% . 39%
43	8	374	 68% 26% . .

## 2 Entry composition

There are 45 unique types of molecules in this entry. The entry contains 106911 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 rRNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	1697	Total	C	N	O	P	0	0
			36229	16171	6507	11855	1696		

- Molecule 2 is a RNA chain called CrPV 5'-UTR IRES.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	0	339	Total	C	N	O	P	0	0
			7205	3222	1255	2389	339		

There are 11 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
0	569	U	C	conflict	GB 8895506
0	570	U	C	conflict	GB 8895506
0	571	U	A	conflict	GB 8895506
0	572	U	C	conflict	GB 8895506
0	574	U	C	conflict	GB 8895506
0	575	U	G	conflict	GB 8895506
0	729	G	-	expression tag	GB 8895506
0	730	G	-	expression tag	GB 8895506
0	731	A	-	expression tag	GB 8895506
0	732	U	-	expression tag	GB 8895506
0	733	C	-	expression tag	GB 8895506

- Molecule 3 is a protein called uS2.

Mol	Chain	Residues	Atoms					AltConf	Trace
3	B	217	Total	C	N	O	S	0	0
			1706	1084	296	318	8		

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	114	THR	ALA	conflict	UNP G1TWL4
B	135	THR	MET	conflict	UNP G1TWL4
B	155	ARG	HIS	conflict	UNP G1TWL4
B	162	PRO	LEU	conflict	UNP G1TWL4

- Molecule 4 is a protein called eS1.

Mol	Chain	Residues	Atoms					AltConf	Trace
4	C	213	Total	C	N	O	S	0	0
			1729	1098	309	308	14		

- Molecule 5 is a protein called uS5.

Mol	Chain	Residues	Atoms					AltConf	Trace
5	D	221	Total	C	N	O	S	0	0
			1717	1113	296	299	9		

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	57	ASN	ASP	conflict	UNP G1SWM1
D	97	PHE	CYS	conflict	UNP G1SWM1
D	181	PRO	LEU	conflict	UNP G1SWM1
D	191	VAL	-	insertion	UNP G1SWM1

- Molecule 6 is a protein called eS4.

Mol	Chain	Residues	Atoms					AltConf	Trace
6	F	262	Total	C	N	O	S	0	0
			2072	1323	384	357	8		

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
F	25	GLY	SER	conflict	UNP G1TK17
F	156	VAL	MET	conflict	UNP G1TK17

- Molecule 7 is a protein called eS6.

Mol	Chain	Residues	Atoms					AltConf	Trace
7	H	237	Total	C	N	O	S	0	0
			1923	1200	387	329	7		

- Molecule 8 is a protein called eS7.

Mol	Chain	Residues	Atoms					AltConf	Trace
8	I	185	Total	C	N	O	S	0	0
			1488	952	271	264	1		

- Molecule 9 is a protein called eS8.

Mol	Chain	Residues	Atoms					AltConf	Trace
9	J	206	Total	C	N	O	S	0	0
			1686	1058	332	291	5		

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
J	47	ARG	GLY	conflict	UNP G1TJW1

- Molecule 10 is a protein called uS4.

Mol	Chain	Residues	Atoms					AltConf	Trace
10	K	185	Total	C	N	O	S	0	0
			1525	969	306	248	2		

- Molecule 11 is a protein called uS17.

Mol	Chain	Residues	Atoms					AltConf	Trace
11	M	143	Total	C	N	O	S	0	0
			1175	749	222	198	6		

- Molecule 12 is a protein called uS15.

Mol	Chain	Residues	Atoms					AltConf	Trace
12	O	149	Total	C	N	O	S	0	0
			1202	770	228	203	1		

- Molecule 13 is a protein called uS11.

Mol	Chain	Residues	Atoms					AltConf	Trace
13	P	136	Total	C	N	O	S	0	0
			1016	621	199	190	6		

- Molecule 14 is a protein called eS21.

Mol	Chain	Residues	Atoms					AltConf	Trace
14	W	83	Total	C	N	O	S	0	0
			634	390	116	123	5		

- Molecule 15 is a protein called uS8.

Mol	Chain	Residues	Atoms					AltConf	Trace
15	X	129	Total	C	N	O	S	0	0
			1034	659	193	176	6		

- Molecule 16 is a protein called uS12.

Mol	Chain	Residues	Atoms					AltConf	Trace
16	Y	141	Total	C	N	O	S	0	0
			1098	693	219	183	3		

- Molecule 17 is a protein called eS24.

Mol	Chain	Residues	Atoms					AltConf	Trace
17	Z	124	Total	C	N	O	S	0	0
			1011	640	198	168	5		

- Molecule 18 is a protein called eS26.

Mol	Chain	Residues	Atoms					AltConf	Trace
18	b	101	Total	C	N	O	S	0	0
			816	509	170	132	5		

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
b	28	ARG	CYS	conflict	UNP G1TFE8

- Molecule 19 is a protein called eS27.

Mol	Chain	Residues	Atoms					AltConf	Trace
19	c	83	Total	C	N	O	S	0	0
			651	408	121	115	7		

- Molecule 20 is a protein called eS30.



Mol	Chain	Residues	Atoms					AltConf	Trace
20	f	57	Total	C	N	O	S	0	0
			457	282	101	73	1		

- Molecule 21 is a protein called uS3.

Mol	Chain	Residues	Atoms					AltConf	Trace
21	E	228	Total	C	N	O	S	0	0
			1768	1126	318	316	8		

- Molecule 22 is a protein called uS7.

Mol	Chain	Residues	Atoms					AltConf	Trace
22	G	191	Total	C	N	O	S	0	0
			1499	937	283	272	7		

- Molecule 23 is a protein called eS10.

Mol	Chain	Residues	Atoms					AltConf	Trace
23	L	96	Total	C	N	O	S	0	0
			810	530	143	131	6		

- Molecule 24 is a protein called eS12.

Mol	Chain	Residues	Atoms					AltConf	Trace
24	N	117	Total	C	N	O	S	0	0
			908	570	161	169	8		

- Molecule 25 is a protein called uS19.

Mol	Chain	Residues	Atoms					AltConf	Trace
25	Q	115	Total	C	N	O	S	0	0
			956	610	176	163	7		

- Molecule 26 is a protein called uS9.

Mol	Chain	Residues	Atoms					AltConf	Trace
26	R	142	Total	C	N	O	S	0	0
			1128	717	213	195	3		

- Molecule 27 is a protein called eS17.

Mol	Chain	Residues	Atoms					AltConf	Trace
27	S	132	Total	C	N	O	S	0	0
			1068	670	199	195	4		

- Molecule 28 is a protein called uS13.

Mol	Chain	Residues	Atoms					AltConf	Trace
28	T	144	Total	C	N	O	S	0	0
			1190	746	241	202	1		

- Molecule 29 is a protein called eS19.

Mol	Chain	Residues	Atoms					AltConf	Trace
29	U	141	Total	C	N	O	S	0	0
			1097	688	211	195	3		

- Molecule 30 is a protein called uS10.

Mol	Chain	Residues	Atoms					AltConf	Trace
30	V	100	Total	C	N	O	S	0	0
			795	498	152	141	4		

- Molecule 31 is a protein called eS25.

Mol	Chain	Residues	Atoms					AltConf	Trace
31	a	75	Total	C	N	O	S	0	0
			598	382	111	104	1		

- Molecule 32 is a protein called eS28.

Mol	Chain	Residues	Atoms					AltConf	Trace
32	d	62	Total	C	N	O	S	0	0
			488	297	97	92	2		

- Molecule 33 is a protein called eS29.

Mol	Chain	Residues	Atoms					AltConf	Trace
33	e	55	Total	C	N	O	S	0	0
			459	286	94	74	5		

- Molecule 34 is a protein called eS31.

Mol	Chain	Residues	Atoms					AltConf	Trace
34	g	68	Total	C	N	O	S	0	0
			555	351	103	94	7		

- Molecule 35 is a protein called RACK1.

Mol	Chain	Residues	Atoms					AltConf	Trace
35	h	312	Total	C	N	O	S	0	0
			2429	1531	423	463	12		

- Molecule 36 is a protein called Eukaryotic translation initiation factor 3 subunit A.

Mol	Chain	Residues	Atoms					AltConf	Trace
36	1	534	Total	C	N	O	S	0	0
			4377	2770	778	808	21		

- Molecule 37 is a protein called Eukaryotic translation initiation factor 3 subunit C.

Mol	Chain	Residues	Atoms					AltConf	Trace
37	2	547	Total	C	N	O	S	0	0
			4446	2791	785	837	33		

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
2	671	GLU	VAL	conflict	UNP G1U971

- Molecule 38 is a protein called Eukaryotic translation initiation factor 3 subunit E.

Mol	Chain	Residues	Atoms					AltConf	Trace
38	3	419	Total	C	N	O	S	0	0
			3465	2220	586	639	20		

- Molecule 39 is a protein called Eukaryotic translation initiation factor 3 subunit F.

Mol	Chain	Residues	Atoms					AltConf	Trace
39	4	272	Total	C	N	O	S	0	0
			2111	1330	359	410	12		

- Molecule 40 is a protein called Eukaryotic translation initiation factor 3 subunit H.

Mol	Chain	Residues	Atoms					AltConf	Trace
40	5	324	Total	C	N	O	S	0	0
			2624	1654	452	503	15		

- Molecule 41 is a protein called Eukaryotic translation initiation factor 3 subunit K.

Mol	Chain	Residues	Atoms					AltConf	Trace
41	6	215	Total	C	N	O	S	0	0
			1737	1109	285	330	13		

- Molecule 42 is a protein called Eukaryotic translation initiation factor 3 subunit L.

Mol	Chain	Residues	Atoms					AltConf	Trace
42	7	372	Total	C	N	O	S	0	0
			3109	2010	519	563	17		

- Molecule 43 is a protein called Eukaryotic translation initiation factor 3 subunit M.

Mol	Chain	Residues	Atoms					AltConf	Trace
43	8	365	Total	C	N	O	S	0	0
			2918	1850	493	558	17		

- Molecule 44 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		AltConf
44	A	1	Total	Mg	0
			1	1	

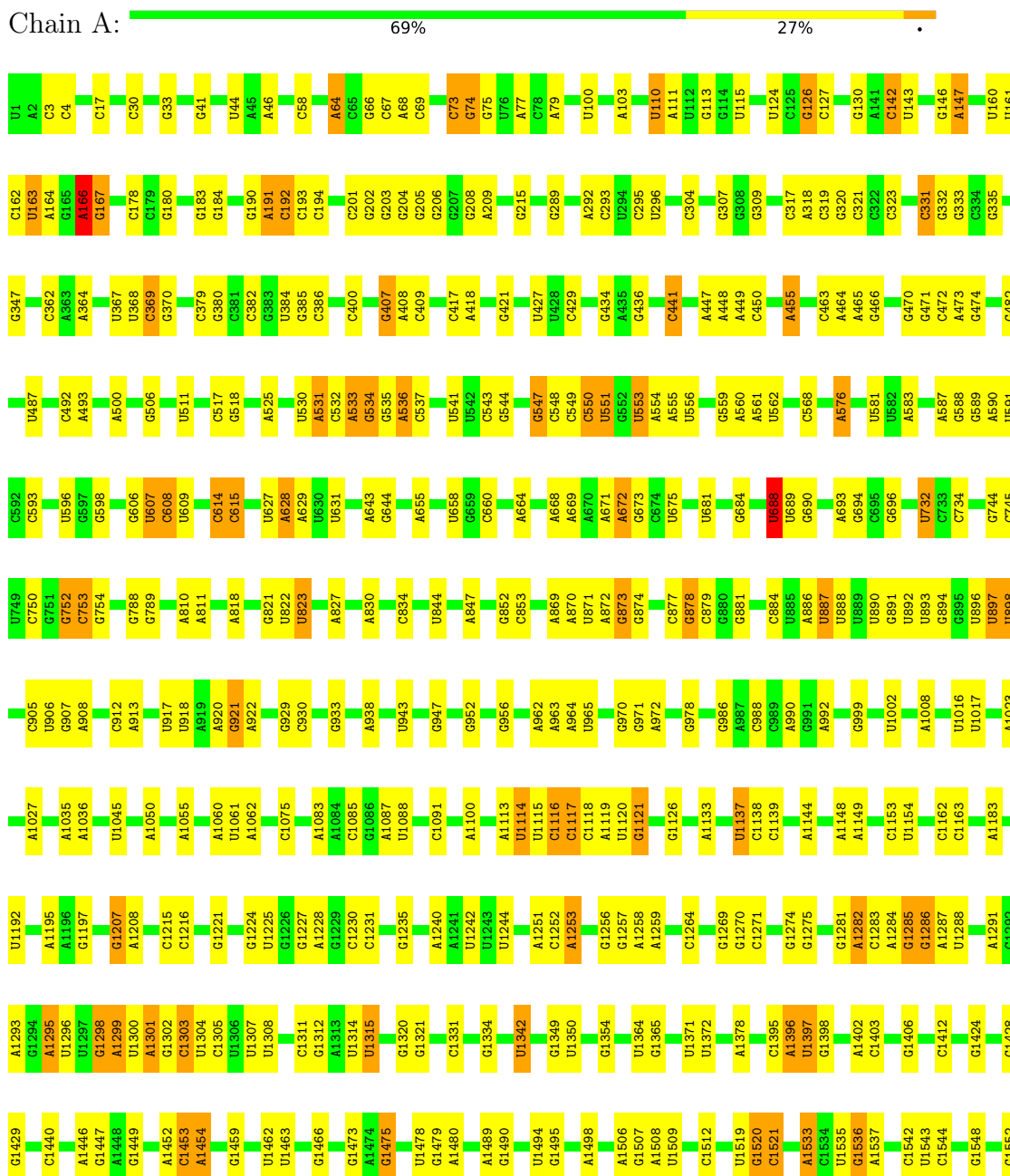
- Molecule 45 is ZINC ION (three-letter code: ZN) (formula: Zn).

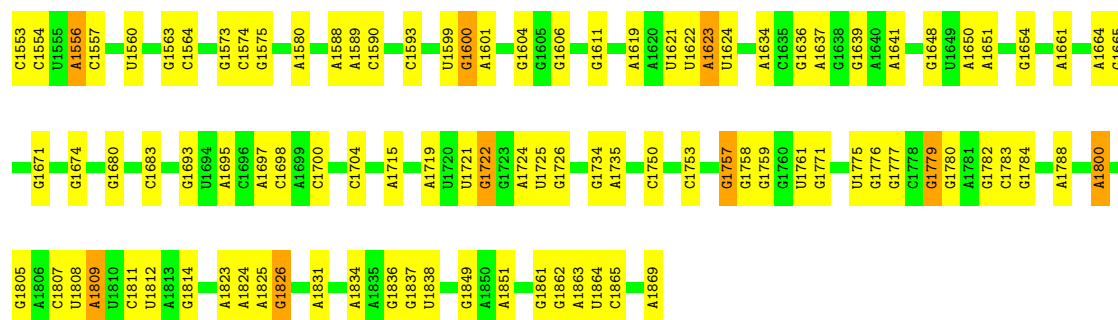
Mol	Chain	Residues	Atoms		AltConf
45	b	1	Total	Zn	0
			1	1	

### 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 the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

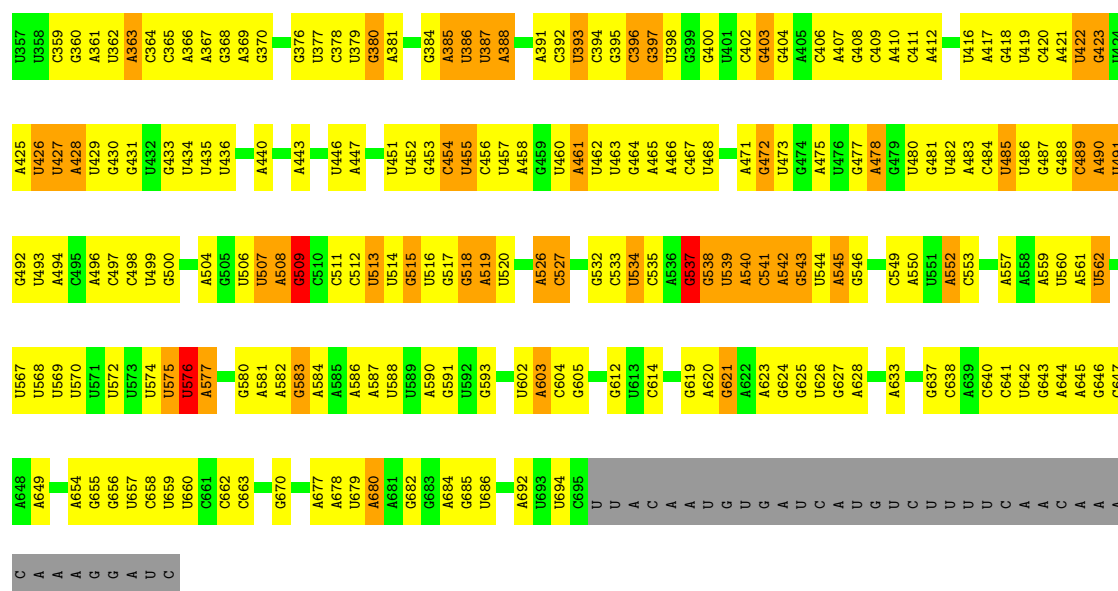
#### • Molecule 1: 18S rRNA





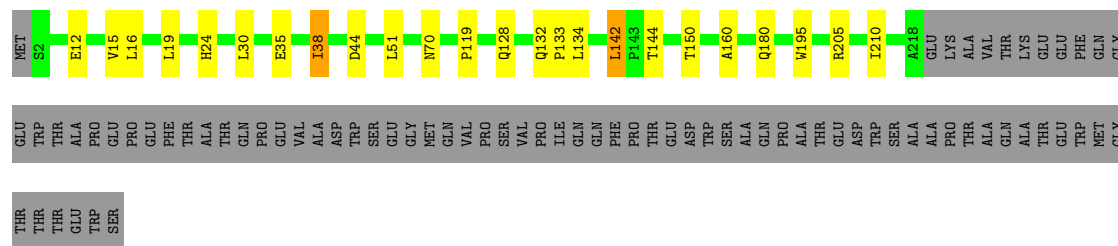
### • Molecule 2: CrPV 5'-UTR IRES

Chain 0: 33% 44% 13% 10%



### • Molecule 3: uS2

Chain B: 65% 7% 26%




### • Molecule 4: eS1

Chain C: 76% 5% 19%



ALA  
LYS  
VAL  
GLU  
ARG  
ALA  
ASP  
GLY  
TYR  
GLU  
PRO  
PRO  
VAL  
GLN  
GLU  
SER  
VAL

• Molecule 5: uS5

Chain D:  82% 5% 13%

MET  
ALA  
ASP  
ASP  
ALA  
GLY  
ALA  
ALA  
ALA  
GLY  
GLY  
PRO  
GLY  
GLY  
PRO  
GLY  
GLY  
PRO  
ASN  
K69  
L66  
I81  
R117  
F127  
L141  
K211  
K212  
L213  
L214  
L215  
A237  
K238  
Y248  
K274  
K278  
ARG  
VAL  
SER  
VAL  
GLN  
ARG  
THR  
GLN  
ALA  
PRO  
ALA  
VAL  
ALA  
THR

• Molecule 6: eS4

Chain F:  94% 5%


MET  
A2  
H17  
L44  
K51  
I72  
R108  
V111  
I129  
F130  
H138  
L139  
R191  
R245  
L256  
K259  
G263

• Molecule 7: eS6

Chain H:  87% 7% 5%


M1  
P8  
V49  
Y84  
R98  
V102  
L106  
L109  
V114  
K115  
K116  
T125  
R159  
N163  
K164  
P169  
R170  
T171  
K172  
A173  
Q186  
T199  
K200  
K201  
L237  
ARG  
ALA  
SER  
THR  
SER  
LYS  
SER  
GLU  
SER  
SER  
GLN  
LYS

• Molecule 8: eS7

Chain I:  87% 8% 5%

MET  
PHE  
SER  
SER  
A6  
F19  
A38  
R41  
E50  
T60  
L69  
E82  
S88  
V92  
A96  
T100  
S108  
ARG  
THR  
LYS  
ASN  
K113  
L122  
R143  
H157  
E188  
L194

• Molecule 9: eS8

Chain J:  90% 9%


MET  
G2  
R5  
I33  
R56  
A57  
T76  
V81  
K94  
V101  
V102  
P108  
R123  
T130  
P131  
R141  
I145  
R178  
G183  
L190  
G207  
LYS

• Molecule 10: uS4

Chain K:  92% 5%

MET  
P2  
V46  
L50  
R80  
F105  
L106  
P170  
R175  
G186  
ALA  
GLY  
ASP  
ASP  
GLU  
GLU  
GLU  
ASP

• Molecule 11: uS17

Chain M:  81% 9% 9%


MET  
ALA  
D3  
Y10  
Q18  
V23  
LEU  
GLY  
GLY  
THR  
GLY  
LYS  
GLU  
K32  
L33  
P34  
L42  
P47  
R69  
I72  
N80  
R101  
C116  
V120  
V126  
T127  
V128  
V145  
K153  
GLN  
PHE  
GLN  
LYS  
PHE

- Molecule 12: uS15

Chain O:  91% 8%



- Molecule 13: uS11

Chain P:  84% 6% 10%



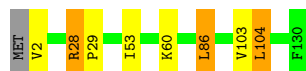
- Molecule 14: eS21

Chain W:  92% 8%




- Molecule 15: uS8

Chain X:  93% ..




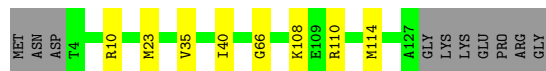
- Molecule 16: uS12

Chain Y:  88% 10% ..




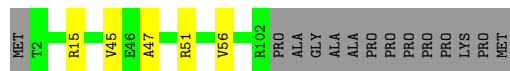
- Molecule 17: eS24

Chain Z:  87% 6% 7%



- Molecule 18: eS26

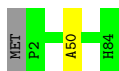
Chain b:  83% 12%



- Molecule 19: eS27

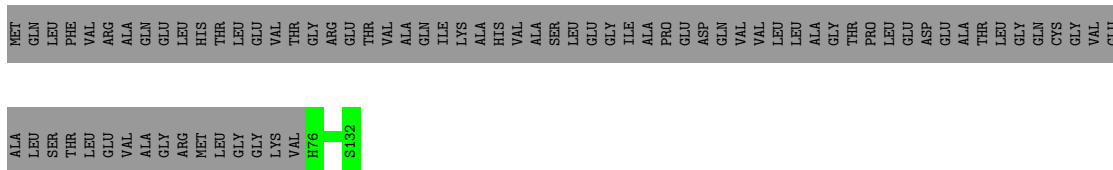


Chain c:  98% .



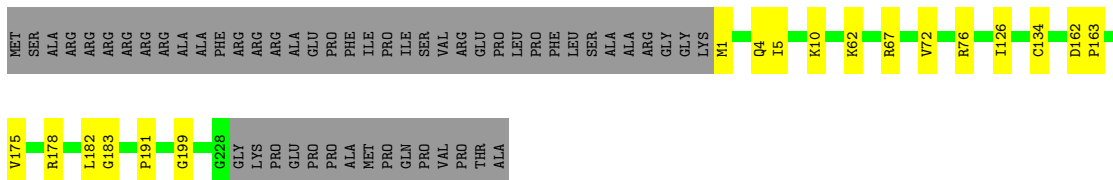
- Molecule 20: eS30

Chain f:  43% 57%

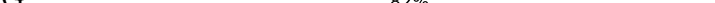


- Molecule 21: uS3

Chain E:  75% 6% 19%



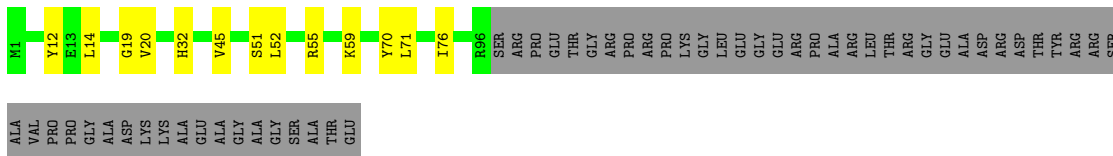
- Molecule 22: uS7

Chain G:  82% 11% 6%



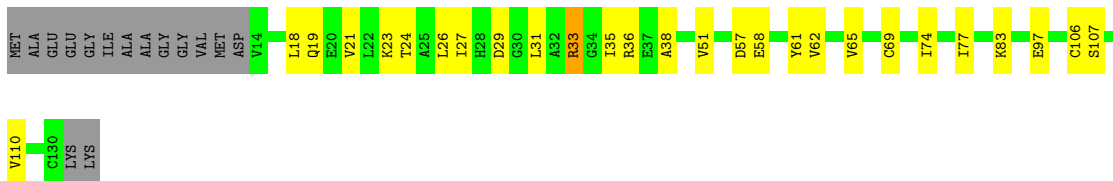
- Molecule 23: eS10

Chain L:  56% 9% 36%



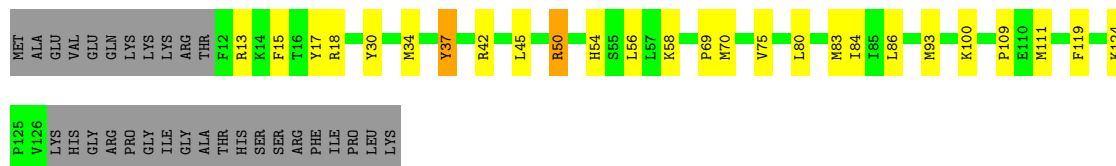
- Molecule 24: eS12

Chain N:  68% 20% • 11%




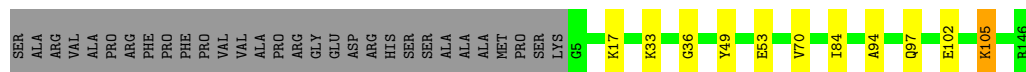
- Molecule 25: uS19

Chain Q:  61% 17% 21%




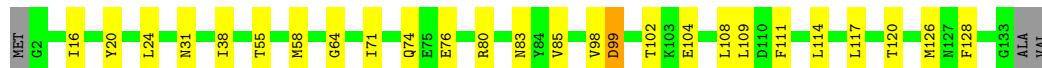
- Molecule 26: uS9

Chain R:  76% 6% 17%




- Molecule 27: eS17

Chain S:  79% 19% 2%




- Molecule 28: uS13

Chain T:  78% 16% 5%




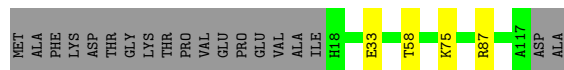
- Molecule 29: eS19

Chain U:  90% 7% 3%



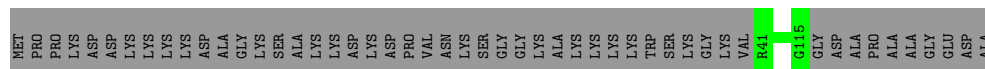
- Molecule 30: uS10

Chain V:  81% 16% 3%

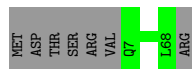


- Molecule 31: eS25

Chain a:  60% 40%



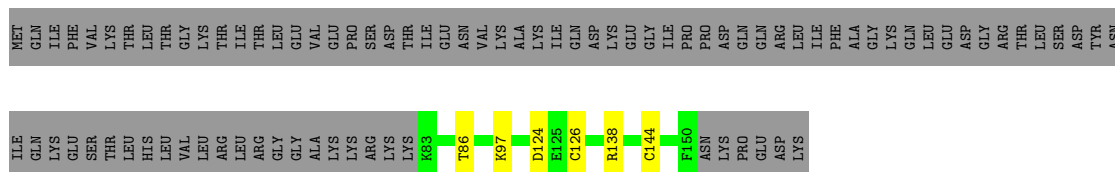
## • Molecule 32: eS28

Chain d:  90% 10%

## • Molecule 33: eS29

Chain e:  96% ..

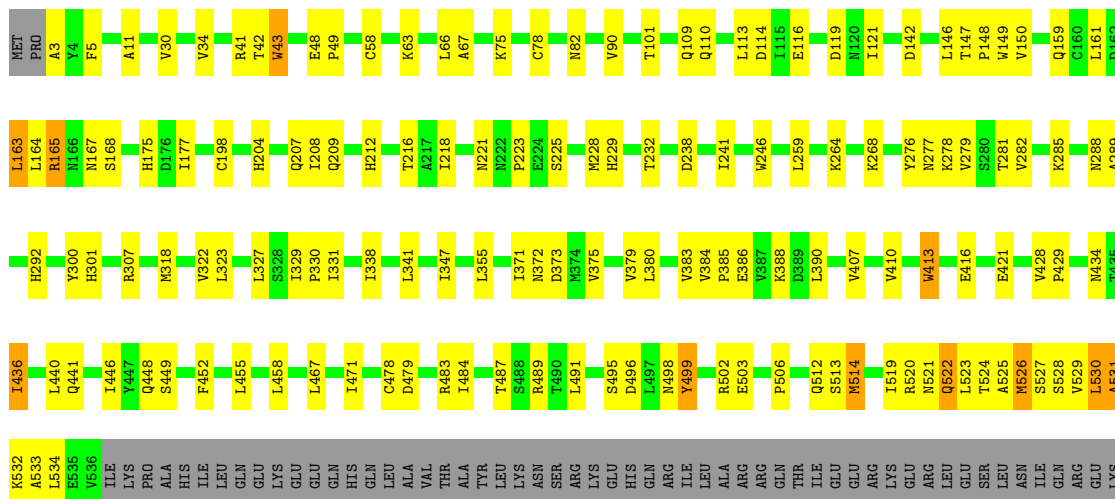
## • Molecule 34: eS31

Chain g:  40% . 56%

## • Molecule 35: RACK1

Chain h:  97% ..

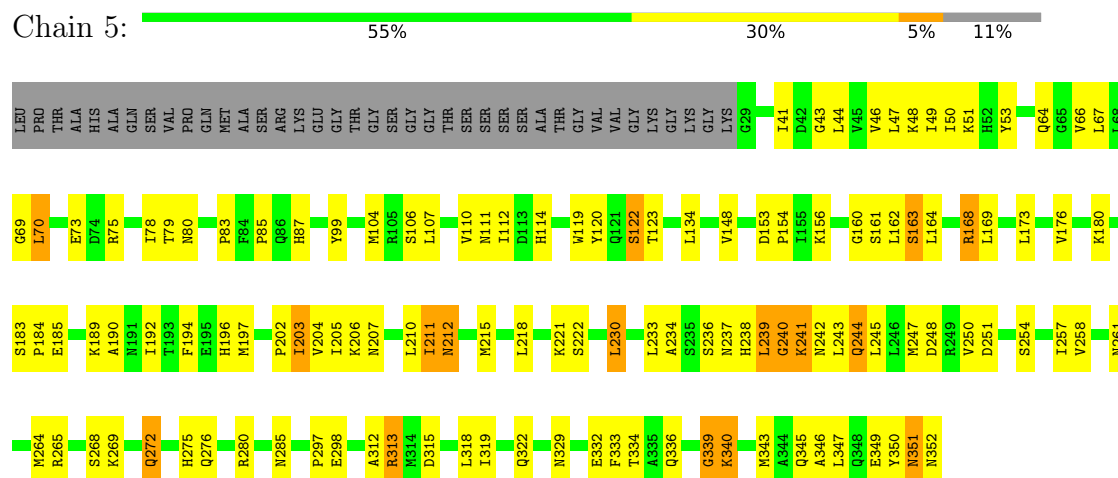
## • Molecule 36: Eukaryotic translation initiation factor 3 subunit A

Chain 1:  28% 10% . 61%

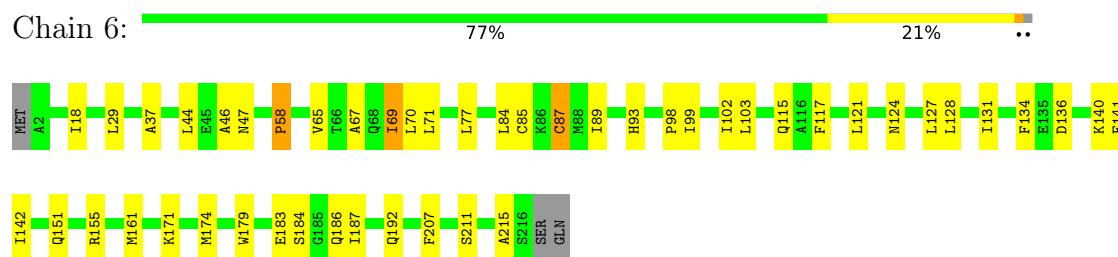




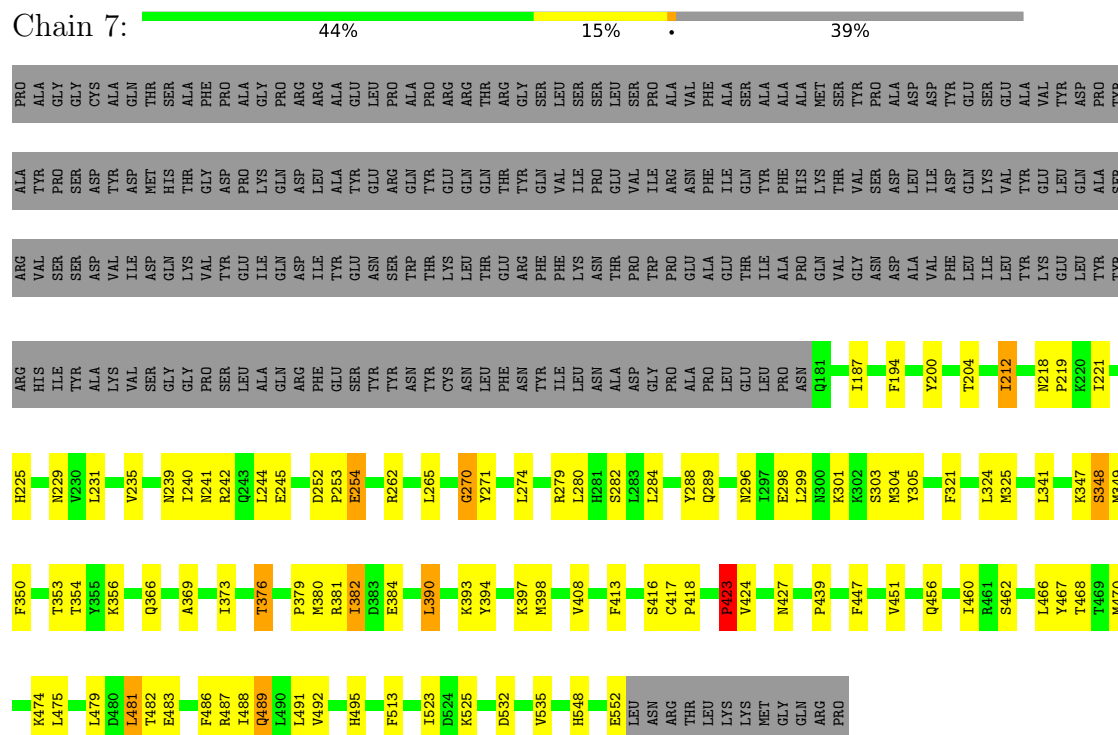
- Molecule 40: Eukaryotic translation initiation factor 3 subunit H



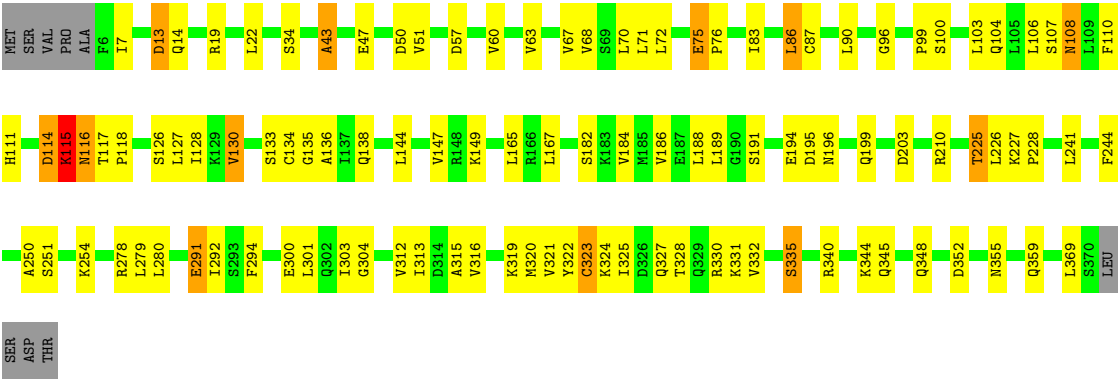
- Molecule 41: Eukaryotic translation initiation factor 3 subunit K



- Molecule 42: Eukaryotic translation initiation factor 3 subunit L



● Molecule 43: Eukaryotic translation initiation factor 3 subunit M



## 4 Experimental information

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, C1	Depositor
Number of particles used	36100	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	PHASE FLIPPING AND AMPLITUDE CORRECTION	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ( $e^-/\text{\AA}^2$ )	56.90	Depositor
Minimum defocus (nm)	Not provided	Depositor
Maximum defocus (nm)	Not provided	Depositor
Magnification	Not provided	Depositor
Image detector	GATAN K2 SUMMIT (4k x 4k)	Depositor



## 5 Model quality ⓘ

### 5.1 Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: ZN, 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	$\# Z  > 5$	RMSZ	$\# Z  > 5$
1	A	0.21	0/40505	0.69	5/63112 (0.0%)
2	O	0.30	2/8053 (0.0%)	0.77	2/12543 (0.0%)
3	B	0.67	0/1743	0.78	0/2370
4	C	0.67	0/1756	0.78	0/2350
5	D	0.67	0/1754	0.78	0/2370
6	F	0.67	0/2114	0.79	0/2843
7	H	0.68	0/1946	0.83	0/2590
8	I	0.68	0/1510	0.79	0/2022
9	J	0.67	0/1715	0.82	0/2287
10	K	0.67	0/1550	0.80	0/2069
11	M	0.66	0/1195	0.80	0/1597
12	O	0.67	0/1226	0.79	0/1649
13	P	0.69	0/1029	0.82	0/1380
14	W	0.70	0/641	0.82	0/858
15	X	0.67	0/1051	0.79	0/1406
16	Y	0.68	0/1116	0.81	0/1490
17	Z	0.67	0/1028	0.80	0/1366
18	b	0.66	0/830	0.82	0/1112
19	c	0.67	0/665	0.81	0/891
20	f	0.68	0/462	0.86	0/607
21	E	0.69	0/1796	0.81	0/2417
22	G	0.70	0/1521	0.81	0/2046
23	L	0.65	0/834	0.77	0/1125
24	N	0.70	0/918	0.81	0/1233
25	Q	0.67	0/974	0.85	0/1301
26	R	0.69	0/1146	0.81	0/1534
27	S	0.69	0/1082	0.83	0/1452
28	T	0.68	0/1208	0.82	0/1618
29	U	0.69	0/1115	0.79	0/1493
30	V	0.69	0/805	0.81	0/1081
31	a	0.69	0/604	0.81	0/810
32	d	0.70	0/490	0.85	0/656

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	# Z  >5	RMSZ	# Z  >5
33	e	0.67	0/470	0.79	0/623
34	g	0.68	0/567	0.84	0/753
35	h	0.69	0/2486	0.80	0/3384
36	1	0.68	0/4460	0.81	0/6034
37	2	0.69	0/4522	0.82	1/6102 (0.0%)
38	3	0.69	0/3538	0.83	0/4786
39	4	0.71	1/2149 (0.0%)	0.89	0/2920
40	5	0.67	0/2674	0.88	1/3606 (0.0%)
41	6	0.73	1/1772 (0.1%)	0.79	1/2396 (0.0%)
42	7	0.71	2/3185 (0.1%)	0.86	3/4296 (0.1%)
43	8	0.71	0/2963	0.85	1/3998 (0.0%)
All	All	0.54	6/113168 (0.0%)	0.77	14/162576 (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
16	Y	0	1
18	b	0	1
22	G	0	1
25	Q	0	1
27	S	0	1
36	1	0	4
37	2	0	4
38	3	0	6
39	4	0	3
40	5	0	6
42	7	0	3
43	8	0	4
All	All	0	35

The worst 5 of 6 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
41	6	69	ILE	CG1-CD1	-7.53	0.98	1.50
42	7	187	ILE	CG1-CD1	-6.45	1.05	1.50
39	4	351	ILE	CG1-CD1	6.23	1.93	1.50
2	0	509	G	C1'-N9	-5.98	1.38	1.46
2	0	576	U	C1'-N1	5.62	1.57	1.48

The worst 5 of 14 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
42	7	382	ILE	CB-CG1-CD1	9.61	140.82	113.90
42	7	240	ILE	CB-CG1-CD1	8.06	136.47	113.90
42	7	187	ILE	CB-CG1-CD1	7.27	134.25	113.90
1	A	688	U	C2'-C3'-O3'	6.76	124.51	113.70
1	A	166	A	C2'-C3'-O3'	6.62	124.30	113.70

There are no chirality outliers.

5 of 35 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
22	G	129	GLY	Peptide
25	Q	54	HIS	Peptide
27	S	64	GLY	Peptide
16	Y	61	GLN	Peptide
18	b	45	VAL	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	36229	0	18304	152	0
2	0	7205	0	3626	54	0
3	B	1706	0	1698	18	0
4	C	1729	0	1803	5	0
5	D	1717	0	1812	17	0
6	F	2072	0	2175	9	0
7	H	1923	0	2089	13	0
8	I	1488	0	1582	7	0
9	J	1686	0	1772	12	0
10	K	1525	0	1640	4	0
11	M	1175	0	1249	7	0
12	O	1202	0	1289	5	0
13	P	1016	0	1039	5	0
14	W	634	0	629	4	0
15	X	1034	0	1080	5	0
16	Y	1098	0	1167	7	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
17	Z	1011	0	1083	5	0
18	b	816	0	867	0	0
19	c	651	0	672	0	0
20	f	457	0	502	0	0
21	E	1768	0	1866	11	0
22	G	1499	0	1540	10	0
23	L	810	0	836	6	0
24	N	908	0	939	13	0
25	Q	956	0	1002	13	0
26	R	1128	0	1195	6	0
27	S	1068	0	1121	17	0
28	T	1190	0	1249	16	0
29	U	1097	0	1130	5	0
30	V	795	0	862	3	0
31	a	598	0	656	0	0
32	d	488	0	514	0	0
33	e	459	0	452	0	0
34	g	555	0	567	0	0
35	h	2429	0	2386	0	0
36	1	4377	0	4433	151	0
37	2	4446	0	4444	118	0
38	3	3465	0	3446	69	0
39	4	2111	0	2105	169	0
40	5	2624	0	2591	249	0
41	6	1737	0	1706	33	0
42	7	3109	0	3084	65	0
43	8	2918	0	2950	66	0
44	A	1	0	0	0	0
45	b	1	0	0	0	0
All	All	106911	0	87152	1080	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 6.

The worst 5 of 1080 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)
5:D:127:PHE:CD1	5:D:141:LEU:HD21	1.26	1.61
5:D:127:PHE:CD1	5:D:141:LEU:CD2	1.95	1.50
36:1:523:LEU:HA	36:1:526:MET:SD	1.71	1.30
37:2:710:ILE:O	37:2:714:PHE:HD2	1.09	1.29

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
36:1:526:MET:HG2	40:5:236:SER:O	1.26	1.28

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	
3	B	215/295 (73%)	201 (94%)	13 (6%)	1 (0%)	31	71
4	C	211/264 (80%)	189 (90%)	21 (10%)	1 (0%)	31	71
5	D	219/255 (86%)	201 (92%)	17 (8%)	1 (0%)	31	71
6	F	260/263 (99%)	248 (95%)	12 (5%)	0	100	100
7	H	235/249 (94%)	216 (92%)	18 (8%)	1 (0%)	36	76
8	I	181/194 (93%)	160 (88%)	20 (11%)	1 (1%)	27	67
9	J	204/208 (98%)	188 (92%)	14 (7%)	2 (1%)	17	56
10	K	183/194 (94%)	170 (93%)	13 (7%)	0	100	100
11	M	139/158 (88%)	130 (94%)	9 (6%)	0	100	100
12	O	147/151 (97%)	137 (93%)	9 (6%)	1 (1%)	24	64
13	P	134/151 (89%)	120 (90%)	14 (10%)	0	100	100
14	W	81/83 (98%)	79 (98%)	2 (2%)	0	100	100
15	X	127/130 (98%)	119 (94%)	7 (6%)	1 (1%)	21	61
16	Y	139/143 (97%)	128 (92%)	7 (5%)	4 (3%)	5	26
17	Z	122/134 (91%)	115 (94%)	7 (6%)	0	100	100
18	b	99/115 (86%)	87 (88%)	10 (10%)	2 (2%)	8	37
19	c	81/84 (96%)	75 (93%)	5 (6%)	1 (1%)	14	51
20	f	55/133 (41%)	51 (93%)	4 (7%)	0	100	100
21	E	226/281 (80%)	209 (92%)	17 (8%)	0	100	100

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
22	G	189/204 (93%)	165 (87%)	21 (11%)	3 (2%)	11	43
23	L	94/149 (63%)	83 (88%)	10 (11%)	1 (1%)	16	53
24	N	115/132 (87%)	96 (84%)	16 (14%)	3 (3%)	6	30
25	Q	113/145 (78%)	92 (81%)	17 (15%)	4 (4%)	4	22
26	R	140/172 (81%)	130 (93%)	9 (6%)	1 (1%)	24	64
27	S	130/135 (96%)	111 (85%)	17 (13%)	2 (2%)	11	45
28	T	142/152 (93%)	132 (93%)	8 (6%)	2 (1%)	12	47
29	U	139/145 (96%)	126 (91%)	12 (9%)	1 (1%)	24	64
30	V	98/119 (82%)	92 (94%)	6 (6%)	0	100	100
31	a	73/125 (58%)	67 (92%)	6 (8%)	0	100	100
32	d	60/69 (87%)	59 (98%)	1 (2%)	0	100	100
33	e	53/56 (95%)	49 (92%)	3 (6%)	1 (2%)	9	39
34	g	66/156 (42%)	54 (82%)	9 (14%)	3 (4%)	3	16
35	h	310/317 (98%)	274 (88%)	34 (11%)	2 (1%)	27	67
36	1	532/1362 (39%)	430 (81%)	85 (16%)	17 (3%)	4	24
37	2	543/913 (60%)	428 (79%)	105 (19%)	10 (2%)	9	40
38	3	417/462 (90%)	333 (80%)	67 (16%)	17 (4%)	3	18
39	4	270/364 (74%)	217 (80%)	45 (17%)	8 (3%)	5	26
40	5	320/363 (88%)	255 (80%)	54 (17%)	11 (3%)	4	22
41	6	213/218 (98%)	179 (84%)	31 (15%)	3 (1%)	12	47
42	7	370/607 (61%)	301 (81%)	57 (15%)	12 (3%)	4	24
43	8	363/374 (97%)	290 (80%)	61 (17%)	12 (3%)	4	23
All	All	7808/10224 (76%)	6786 (87%)	893 (11%)	129 (2%)	14	42

5 of 129 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
9	J	94	LYS
15	X	28	ARG
16	Y	61	GLN
22	G	134	VAL
24	N	57	ASP

### 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	
3	B	180/246 (73%)	177 (98%)	3 (2%)	63	88
4	C	194/231 (84%)	191 (98%)	3 (2%)	67	89
5	D	186/205 (91%)	183 (98%)	3 (2%)	65	89
6	F	223/224 (100%)	222 (100%)	1 (0%)	92	97
7	H	207/218 (95%)	203 (98%)	4 (2%)	60	87
8	I	165/174 (95%)	162 (98%)	3 (2%)	62	87
9	J	178/180 (99%)	178 (100%)	0	100	100
10	K	161/168 (96%)	161 (100%)	0	100	100
11	M	130/142 (92%)	128 (98%)	2 (2%)	67	89
12	O	130/131 (99%)	128 (98%)	2 (2%)	67	89
13	P	106/119 (89%)	105 (99%)	1 (1%)	81	93
14	W	68/68 (100%)	67 (98%)	1 (2%)	67	89
15	X	112/113 (99%)	109 (97%)	3 (3%)	48	81
16	Y	113/114 (99%)	112 (99%)	1 (1%)	81	93
17	Z	107/115 (93%)	106 (99%)	1 (1%)	81	93
18	b	89/99 (90%)	87 (98%)	2 (2%)	55	84
19	c	75/76 (99%)	75 (100%)	0	100	100
20	f	47/106 (44%)	47 (100%)	0	100	100
21	E	190/232 (82%)	189 (100%)	1 (0%)	90	96
22	G	158/170 (93%)	156 (99%)	2 (1%)	71	91
23	L	87/125 (70%)	86 (99%)	1 (1%)	76	92
24	N	99/108 (92%)	96 (97%)	3 (3%)	44	79
25	Q	105/130 (81%)	101 (96%)	4 (4%)	36	73
26	R	117/140 (84%)	116 (99%)	1 (1%)	81	93
27	S	119/121 (98%)	119 (100%)	0	100	100
28	T	125/132 (95%)	123 (98%)	2 (2%)	65	89

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
29	U	111/116 (96%)	109 (98%)	2 (2%)	62	87
30	V	92/107 (86%)	92 (100%)	0	100	100
31	a	66/103 (64%)	66 (100%)	0	100	100
32	d	55/62 (89%)	55 (100%)	0	100	100
33	e	48/49 (98%)	48 (100%)	0	100	100
34	g	61/140 (44%)	58 (95%)	3 (5%)	27	66
35	h	271/275 (98%)	268 (99%)	3 (1%)	76	92
36	1	490/1245 (39%)	476 (97%)	14 (3%)	45	80
37	2	494/812 (61%)	477 (97%)	17 (3%)	40	76
38	3	384/423 (91%)	375 (98%)	9 (2%)	53	84
39	4	239/282 (85%)	233 (98%)	6 (2%)	50	82
40	5	293/320 (92%)	283 (97%)	10 (3%)	40	76
41	6	190/193 (98%)	188 (99%)	2 (1%)	76	92
42	7	342/544 (63%)	335 (98%)	7 (2%)	58	86
43	8	327/335 (98%)	319 (98%)	8 (2%)	52	83
All	All	6934/8893 (78%)	6809 (98%)	125 (2%)	64	87

5 of 125 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
36	1	238	ASP
37	2	462	THR
42	7	481	LEU
36	1	268	LYS
36	1	373	ASP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 85 such sidechains are listed below:

Mol	Chain	Res	Type
36	1	82	ASN
37	2	386	ASN
42	7	289	GLN
36	1	167	ASN
36	1	434	ASN



### 5.3.3 RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
1	A	1682/1697 (99%)	406 (24%)	26 (1%)
2	0	338/377 (89%)	186 (55%)	20 (5%)
All	All	2020/2074 (97%)	592 (29%)	46 (2%)

5 of 592 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
1	A	3	C
1	A	4	C
1	A	17	C
1	A	33	G
1	A	41	G

5 of 46 RNA pucker outliers are listed below:

Mol	Chain	Res	Type
1	A	1137	U
1	A	1824	A
2	0	581	A
1	A	1295	A
1	A	1664	A

## 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 2 ligands modelled in this entry, 2 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

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

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	15
40	5	1

The worst 5 of 16 chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	A	697:G	O3'	729:C	P	18.10
1	A	834:C	O3'	841:G	P	17.97
1	A	130:G	O3'	141:A	P	17.73
1	A	756:C	O3'	788:G	P	15.69
1	A	323:C	O3'	329:G	P	15.54