



# wwPDB EM Model Validation Summary Report ⓘ

Apr 15, 2020 – 01:15 PM EDT

PDB ID : 6W2T  
EMDB ID : EMD-21530  
Title : Structure of the Cricket Paralysis Virus 5-UTR IRES (CrPV 5-UTR-IRES)  
bound to the small ribosomal subunit in the closed state (Class 2)  
Authors : Neupane, R.; Pisareva, V.; Rodriguez, C.F.; Pisarev, A.; Fernandez, I.S.  
Deposited on : 2020-03-08  
Resolution : 3.36 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

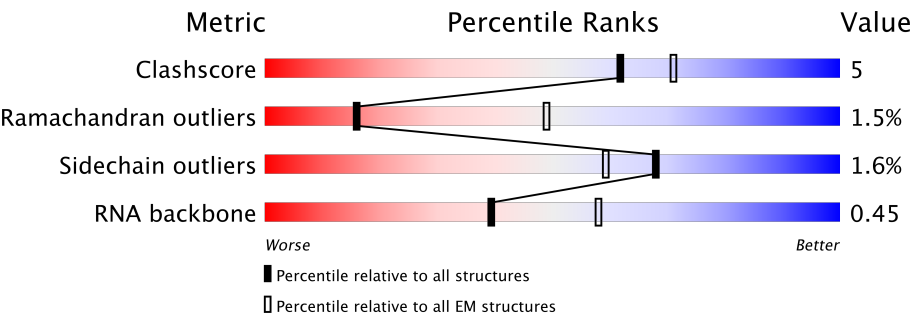


# 1 Overall quality at a glance i

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

The reported resolution of this entry is 3.36 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




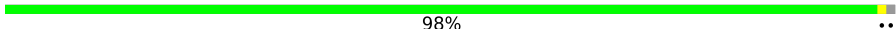


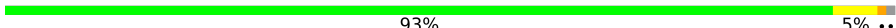



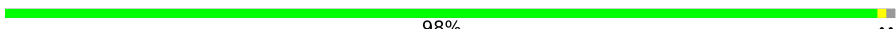














The table below summarises the geometric issues observed across the polymeric chains and their fit to the map. The red, orange, yellow and green segments on the bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria respectively. A grey 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	<div><div>76%</div><div>24%</div></div>
2	B	295	<div><div>65%</div><div>8%</div><div>26%</div></div>
3	C	264	<div><div>76%</div><div>5%</div><div>19%</div></div>
4	D	255	<div><div>80%</div><div>6%</div><div>13%</div></div>
5	F	263	<div><div>95%</div><div>.</div></div>
6	H	249	<div><div>90%</div><div>5%</div><div>5%</div></div>
7	I	194	<div><div>85%</div><div>10%</div><div>5%</div></div>

Continued on next page...






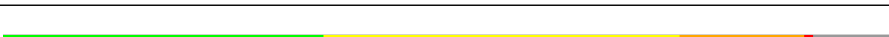
*Continued from previous page...*

Mol	Chain	Length	Quality of chain
8	J	208	 91% 8% .
9	K	194	 87% 9% 5%
10	M	158	 87% . 9%
11	O	151	 98% ..
12	P	151	 84% 6% 10%
13	W	83	 89% 11%
14	X	130	 93% 5% ..
15	Y	143	 89% 9% ..
16	Z	134	 88% . 7%
17	b	115	 83% 5% 12%
18	c	84	 98% ..
19	f	133	 42% . 57%
20	E	281	 73% 7% 19%
21	G	204	 82% 11% 6%
22	L	149	 60% 5% 36%
23	N	132	 77% 12% 11%
24	Q	145	 58% 20% . 21%
25	R	172	 79% . 17%
26	S	135	 90% 8% .
27	T	152	 89% 6% 5%
28	U	145	 89% 8% .
29	V	119	 76% 8% 16%
30	i	125	 58% . 40%
31	d	69	 88% . 10%
32	e	56	 95% . .

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Length	Quality of chain
33	g	156	
34	h	317	
35	3	462	
36	4	364	
37	6	218	
38	7	607	
39	8	374	
40	1	1362	
41	2	913	
42	9	558	
43	5	363	
44	A	377	



## 2 Entry composition

There are 46 unique types of molecules in this entry. The entry contains 109778 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 protein called uS2.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	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 3 is a protein called eS1.

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

- Molecule 4 is a protein called uS5.

Mol	Chain	Residues	Atoms					AltConf	Trace
4	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

*Continued on next page...*



*Continued from previous page...*

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

- Molecule 5 is a protein called eS4.

Mol	Chain	Residues	Atoms					AltConf	Trace
5	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 6 is a protein called eS6.

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

- Molecule 7 is a protein called eS7.

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

- Molecule 8 is a protein called eS8.

Mol	Chain	Residues	Atoms					AltConf	Trace
8	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 9 is a protein called uS4.



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

- Molecule 10 is a protein called uS17.

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

- Molecule 11 is a protein called uS15.

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

- Molecule 12 is a protein called uS11.

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

- Molecule 13 is a protein called eS21.

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

- Molecule 14 is a protein called uS8.

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

- Molecule 15 is a protein called uS12.

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

- Molecule 16 is a protein called eS24.



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

- Molecule 17 is a protein called eS26.

Mol	Chain	Residues	Atoms					AltConf	Trace
17	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 18 is a protein called eS27.

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

- Molecule 19 is a protein called eS30.

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

- Molecule 20 is a protein called uS3.

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

- Molecule 21 is a protein called uS7.

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

- Molecule 22 is a protein called eS10.



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

- Molecule 23 is a protein called eS12.

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

- Molecule 24 is a protein called uS19.

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

- Molecule 25 is a protein called uS9.

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

- Molecule 26 is a protein called eS17.

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

- Molecule 27 is a protein called uS13.

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

- Molecule 28 is a protein called eS19.

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

- Molecule 29 is a protein called uS10.



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

- Molecule 30 is a protein called eS25.

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

- Molecule 31 is a protein called eS28.

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

- Molecule 32 is a protein called eS29.

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

- Molecule 33 is a protein called eS31.

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

- Molecule 34 is a protein called RACK1.

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

- Molecule 35 is a protein called Eukaryotic translation initiation factor 3 subunit E.

Mol	Chain	Residues	Atoms					AltConf	Trace
35	3	419	Total	C	N	O	S	0	0
			3465	2220	586	639	20		

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



Mol	Chain	Residues	Atoms					AltConf	Trace
36	4	272	Total	C	N	O	S	0	0
			2111	1330	359	410	12		

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

Mol	Chain	Residues	Atoms					AltConf	Trace
37	6	215	Total	C	N	O	S	0	0
			1737	1109	285	330	13		

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

Mol	Chain	Residues	Atoms					AltConf	Trace
38	7	372	Total	C	N	O	S	0	0
			3109	2010	519	563	17		

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

Mol	Chain	Residues	Atoms					AltConf	Trace
39	8	365	Total	C	N	O	S	0	0
			2918	1850	493	558	17		

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

Mol	Chain	Residues	Atoms					AltConf	Trace
40	1	534	Total	C	N	O	S	0	0
			4377	2770	778	808	21		

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

Mol	Chain	Residues	Atoms					AltConf	Trace
41	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 42 is a protein called Eukaryotic translation initiation factor 3 subunit D.



Mol	Chain	Residues	Atoms					AltConf	Trace
42	9	356	Total	C	N	O	S	0	0
			2867	1804	500	548	15		

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

Mol	Chain	Residues	Atoms					AltConf	Trace
43	5	324	Total	C	N	O	S	0	0
			2624	1654	452	503	15		

- Molecule 44 is a RNA chain called CrPV 5'-UTR IRES.

Mol	Chain	Residues	Atoms					AltConf	Trace
44	A	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
A	569	U	C	conflict	GB 8895506
A	570	U	C	conflict	GB 8895506
A	571	U	A	conflict	GB 8895506
A	572	U	C	conflict	GB 8895506
A	574	U	C	conflict	GB 8895506
A	575	U	G	conflict	GB 8895506
A	729	G	-	expression tag	GB 8895506
A	730	G	-	expression tag	GB 8895506
A	731	A	-	expression tag	GB 8895506
A	732	U	-	expression tag	GB 8895506
A	733	C	-	expression tag	GB 8895506

- Molecule 45 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		AltConf
45	a	1	Total	Mg	0
			1	1	

- Molecule 46 is ZINC ION (three-letter code: ZN) (formula: Zn).

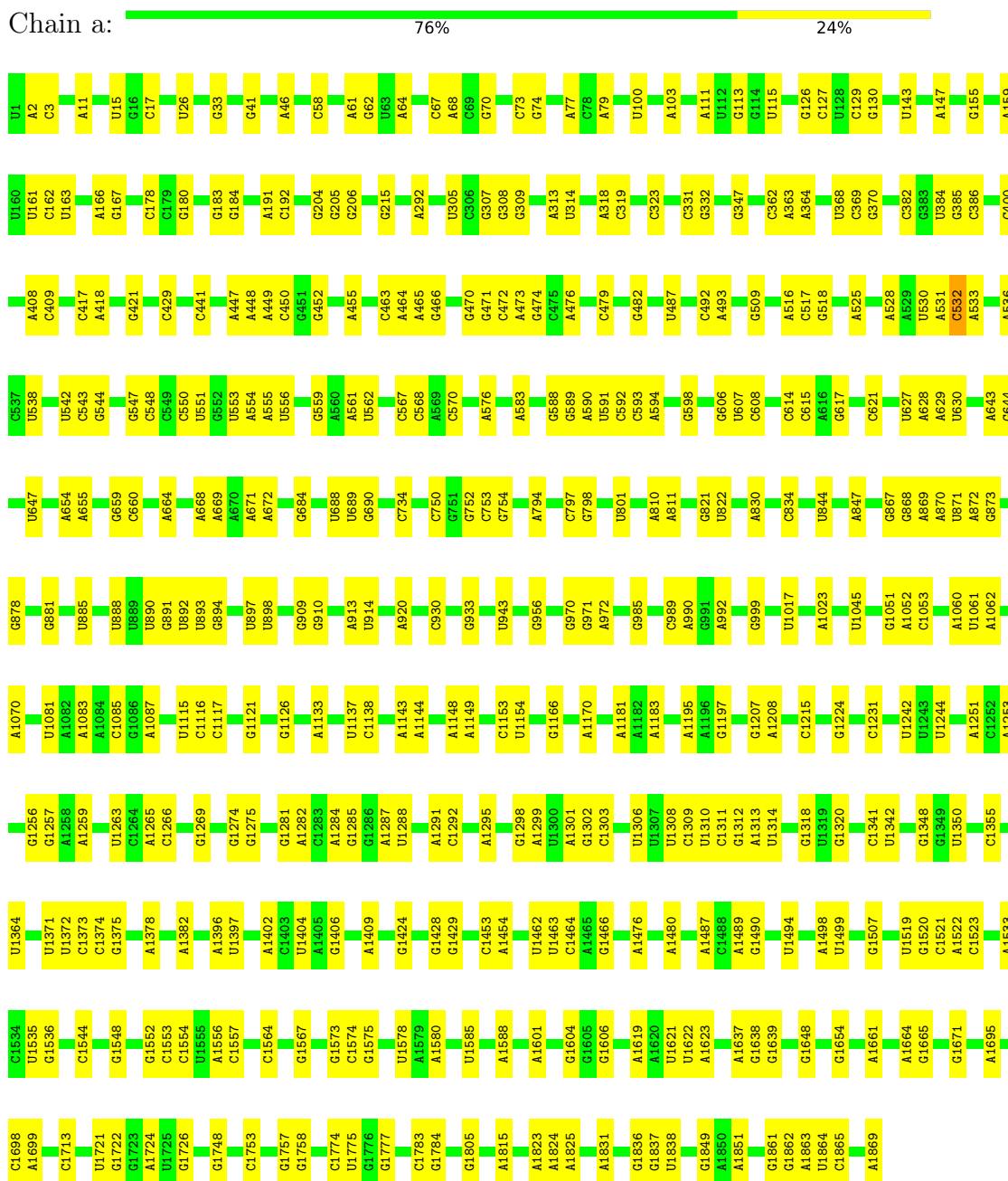
Mol	Chain	Residues	Atoms		AltConf
46	b	1	Total	Zn	0
			1	1	



### 3 Residue-property plots

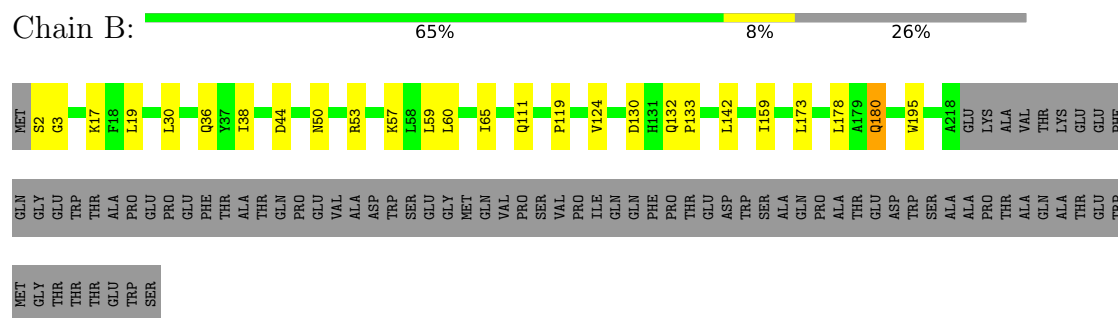
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry. Residues are color-coded according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

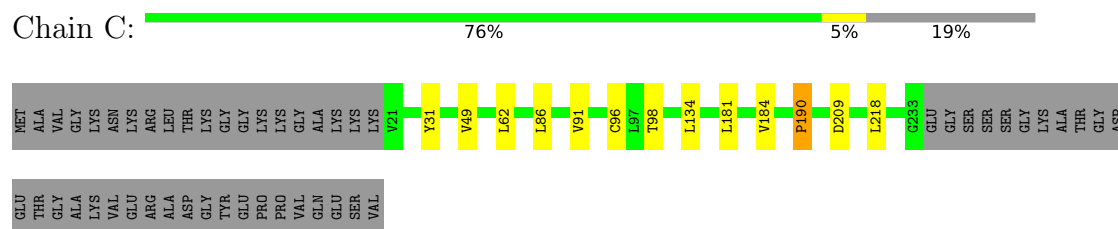




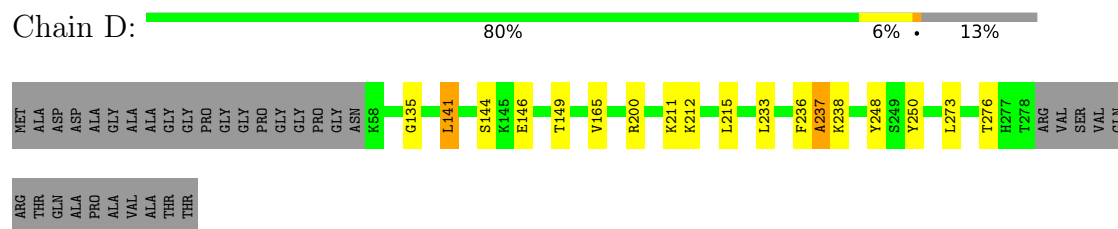
- Molecule 2: uS2



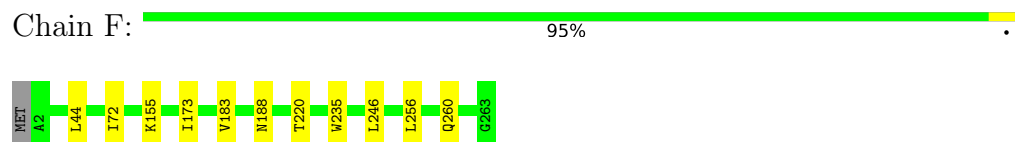
- Molecule 3: eS1



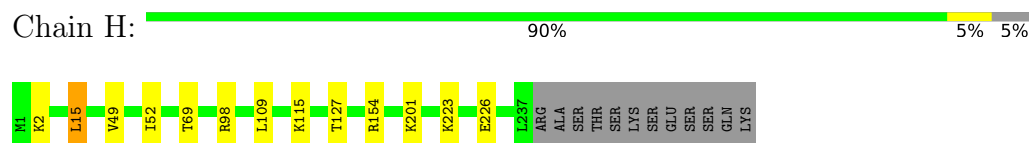
- Molecule 4: uS5



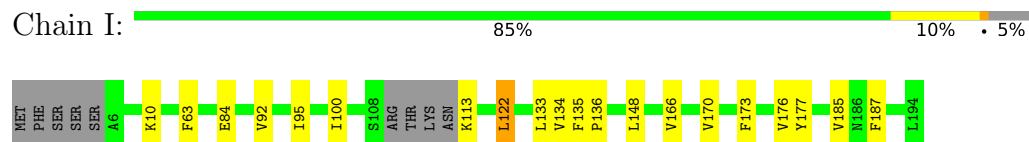
- Molecule 5: eS4



- Molecule 6: eS6




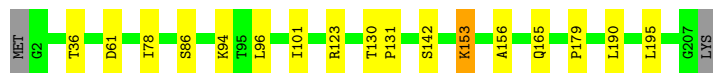
- Molecule 7: eS7






- Molecule 8: eS8

Chain J:  91% 8% .




- Molecule 9: uS4

Chain K:  87% 9% 5%



- Molecule 10: uS17

Chain M:  87% 9% .




- Molecule 11: uS15

Chain O:  98% ..



- Molecule 12: uS11

Chain P:  84% 6% 10%



- Molecule 13: eS21

Chain W:  89% 11%



- Molecule 14: uS8

Chain X:  93% 5% ..



- Molecule 15: uS12

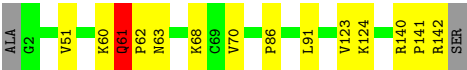


Chain Y: 

89%

9%

..



• Molecule 16: eS24

Chain Z: 

88%

7%

.



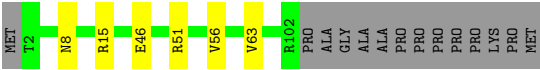
• Molecule 17: eS26

Chain b: 

83%

5%

12%

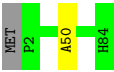


• Molecule 18: eS27

Chain c: 

98%

..



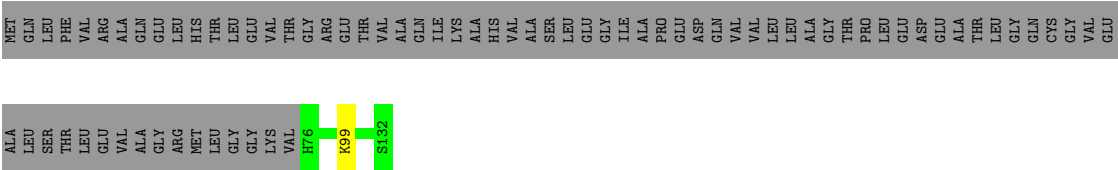
• Molecule 19: eS30

Chain f: 

42%

.

57%



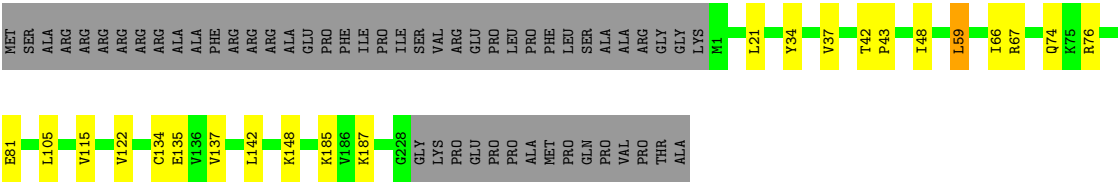
• Molecule 20: uS3

Chain E: 

73%

7%

19%



• Molecule 21: uS7

Chain G: 

82%

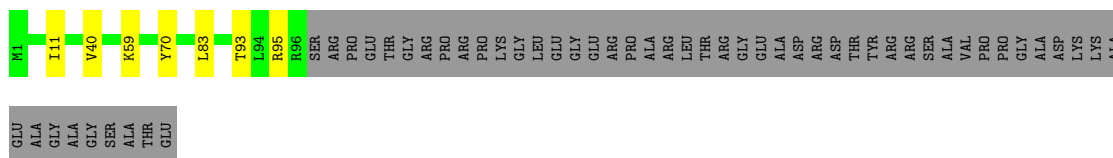
11%

6%

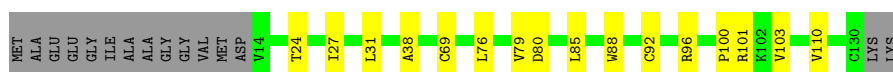
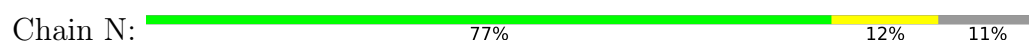




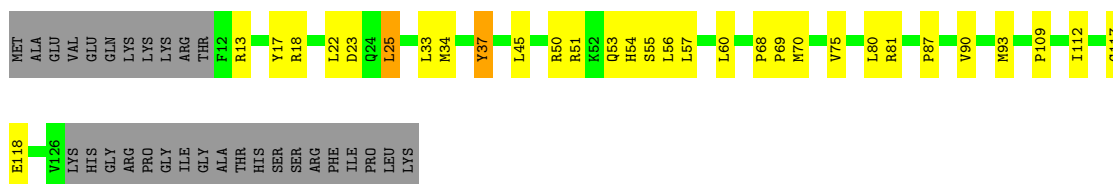
• Molecule 22: eS10



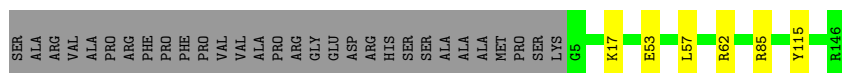
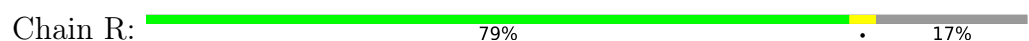
• Molecule 23: eS12



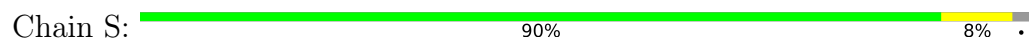
• Molecule 24: uS19



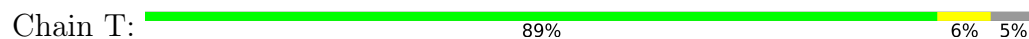
• Molecule 25: uS9



• Molecule 26: eS17




• Molecule 27: uS13






- Molecule 28: eS19

Chain U:  89% 8%



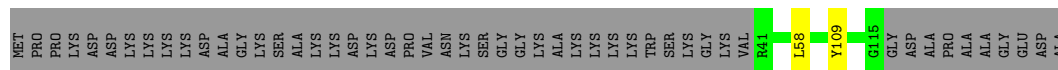
- Molecule 29: uS10

Chain V:  76% 8% 16%




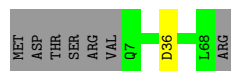
- Molecule 30: eS25

Chain i:  58% 40%



- Molecule 31: eS28

Chain d:  88% 10%



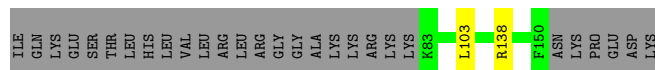
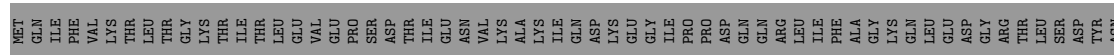
- Molecule 32: eS29

Chain e:  95%



- Molecule 33: eS31

Chain g:  42% 56%



- Molecule 34: RACK1

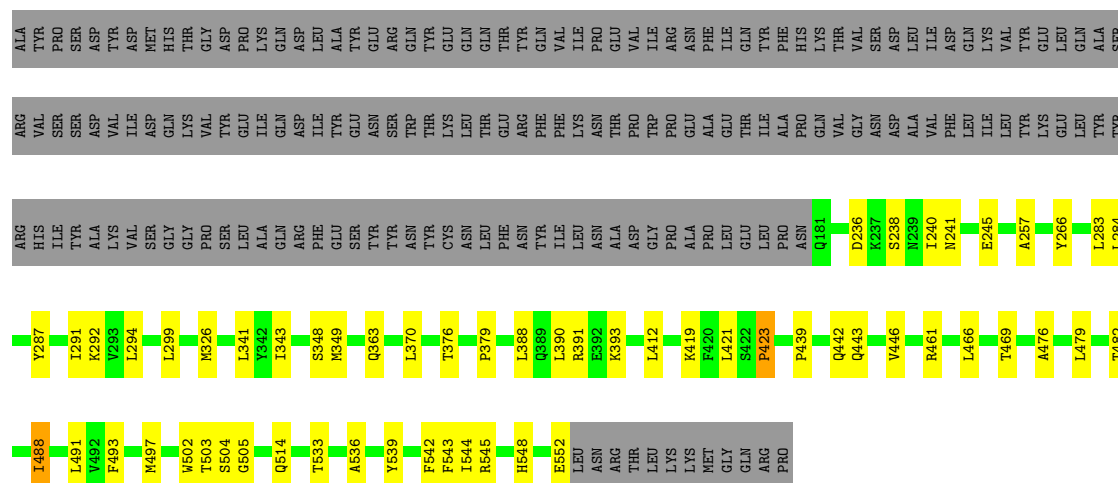
Chain h:  97%





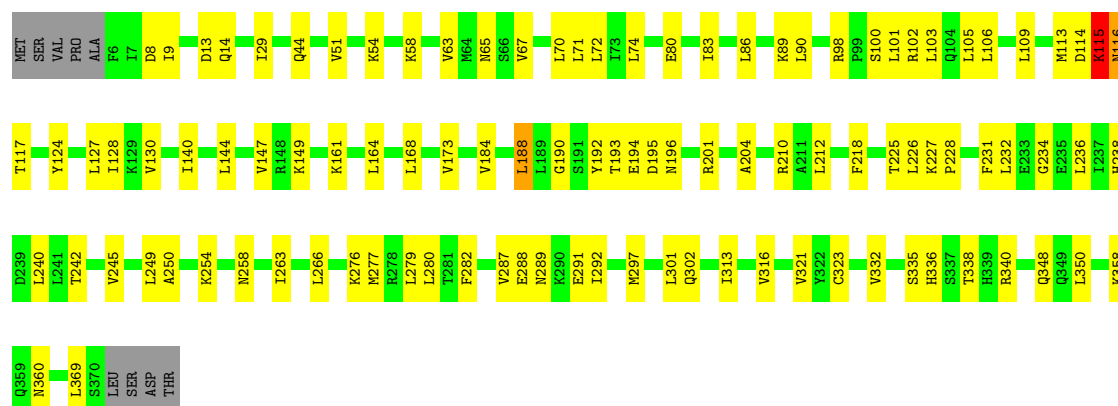






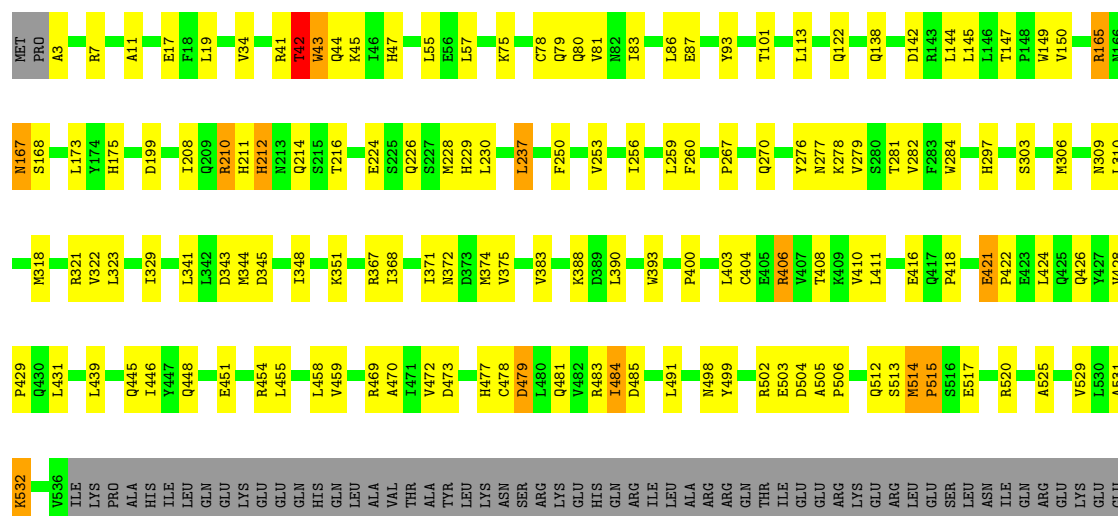
• Molecule 39: Eukaryotic translation initiation factor 3 subunit M

Chain 8:



• Molecule 40: Eukaryotic translation initiation factor 3 subunit A

Chain 1:







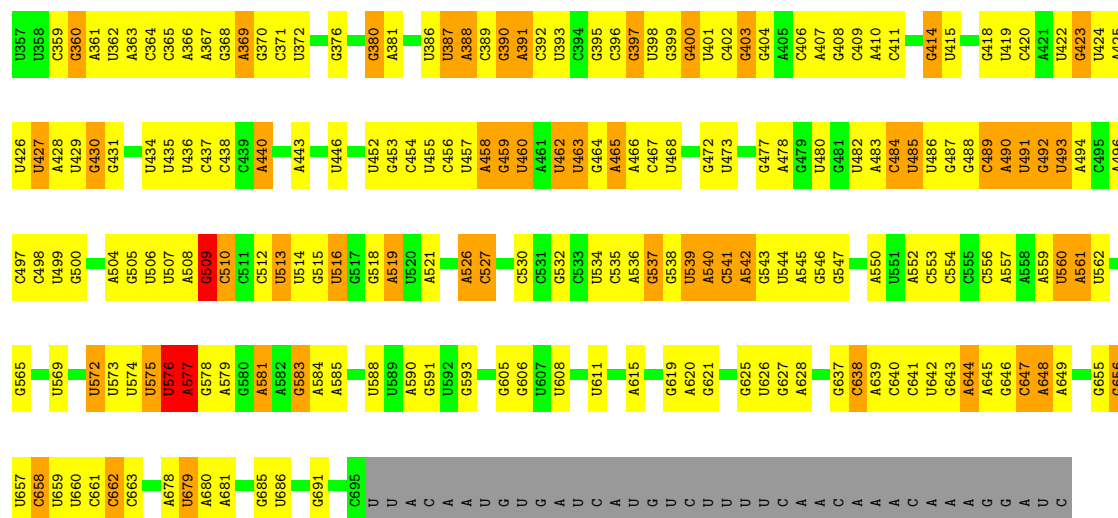






- Molecule 44: CrPV 5'-UTR IRES

Chain A:  36% 40% 14% • 10%





## 4 Experimental information

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, C1	Depositor
Number of particles used	23444	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	0.68	4/63112 (0.0%)
2	B	0.67	0/1743	0.77	0/2370
3	C	0.67	0/1756	0.78	0/2350
4	D	0.67	0/1754	0.78	0/2370
5	F	0.67	0/2114	0.79	0/2843
6	H	0.68	0/1946	0.83	0/2590
7	I	0.68	0/1510	0.78	0/2022
8	J	0.67	0/1715	0.81	0/2287
9	K	0.67	0/1550	0.82	0/2069
10	M	0.66	0/1195	0.79	0/1597
11	O	0.68	0/1226	0.78	0/1649
12	P	0.69	0/1029	0.83	0/1380
13	W	0.70	0/641	0.82	0/858
14	X	0.67	0/1051	0.80	0/1406
15	Y	0.69	0/1116	0.83	0/1490
16	Z	0.67	0/1028	0.80	0/1366
17	b	0.67	0/830	0.83	0/1112
18	c	0.67	0/665	0.82	0/891
19	f	0.68	0/462	0.83	0/607
20	E	0.69	0/1796	0.82	0/2417
21	G	0.70	0/1521	0.80	0/2046
22	L	0.65	0/834	0.78	0/1125
23	N	0.70	0/918	0.82	0/1233
24	Q	0.67	0/974	0.84	0/1301
25	R	0.69	0/1146	0.81	0/1534
26	S	0.69	0/1082	0.81	0/1452
27	T	0.68	0/1208	0.81	0/1618
28	U	0.69	0/1115	0.80	0/1493
29	V	0.68	0/805	0.81	0/1081
30	i	0.69	0/604	0.83	0/810
31	d	0.70	0/490	0.87	0/656
32	e	0.67	0/470	0.78	0/623



Mol	Chain	Bond lengths		Bond angles	
		RMSZ	# Z  >5	RMSZ	# Z  >5
33	g	0.68	0/567	0.81	0/753
34	h	0.69	0/2486	0.81	0/3384
35	3	0.69	0/3538	0.85	2/4786 (0.0%)
36	4	0.71	0/2149	0.85	0/2920
37	6	0.70	0/1772	0.76	1/2396 (0.0%)
38	7	0.69	0/3185	0.78	0/4296
39	8	0.71	0/2963	0.85	3/3998 (0.1%)
40	1	0.68	0/4460	0.81	0/6034
41	2	0.69	0/4522	0.81	1/6102 (0.0%)
42	9	0.68	0/2921	0.79	0/3957
43	5	0.67	0/2675	0.81	0/3609
44	A	0.29	3/8053 (0.0%)	0.76	2/12543 (0.0%)
All	All	0.54	3/116090 (0.0%)	0.76	13/166536 (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
5	F	0	1
14	X	0	1
15	Y	0	1
21	G	0	2
34	h	0	1
35	3	0	12
36	4	0	7
37	6	0	1
39	8	0	3
40	1	0	4
41	2	0	5
42	9	0	2
43	5	0	2
All	All	0	42

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
44	A	509	G	C1'-N9	-6.20	1.38	1.46
44	A	576	U	C1'-N1	5.45	1.56	1.48
44	A	577	A	C1'-N9	-5.24	1.39	1.46



The worst 5 of 13 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed( $^{\circ}$ )	Ideal( $^{\circ}$ )
37	6	131	ILE	CB-CG1-CD1	7.29	134.31	113.90
1	a	688	U	C2'-C3'-O3'	6.84	124.65	113.70
35	3	384	ILE	CB-CG1-CD1	6.39	131.80	113.90
44	A	679	U	C2'-C3'-O3'	6.39	123.92	113.70
35	3	371	ILE	CB-CG1-CD1	6.04	130.80	113.90

There are no chirality outliers.

5 of 42 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
5	F	155	LYS	Peptide
21	G	129	GLY	Peptide
21	G	130	ARG	Peptide
14	X	54	ASP	Peptide
15	Y	61	GLN	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	0	0
2	B	1706	0	1698	17	0
3	C	1729	0	1803	7	0
4	D	1717	0	1812	19	0
5	F	2072	0	2175	5	0
6	H	1923	0	2089	4	0
7	I	1488	0	1582	10	0
8	J	1686	0	1772	8	0
9	K	1525	0	1640	9	0
10	M	1175	0	1249	3	0
11	O	1202	0	1289	0	0
12	P	1016	0	1039	4	0
13	W	634	0	629	6	0
14	X	1034	0	1080	3	0
15	Y	1098	0	1167	5	0
16	Z	1011	0	1083	3	0

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
17	b	816	0	867	0	0
18	c	651	0	672	0	0
19	f	457	0	502	0	0
20	E	1768	0	1866	10	0
21	G	1499	0	1540	8	0
22	L	810	0	836	3	0
23	N	908	0	939	7	0
24	Q	956	0	1002	16	0
25	R	1128	0	1195	2	0
26	S	1068	0	1121	7	0
27	T	1190	0	1249	3	0
28	U	1097	0	1130	5	0
29	V	795	0	862	6	0
30	i	598	0	656	0	0
31	d	488	0	514	0	0
32	e	459	0	452	0	0
33	g	555	0	567	0	0
34	h	2429	0	2386	0	0
35	3	3465	0	3446	77	0
36	4	2111	0	2105	73	0
37	6	1737	0	1706	21	0
38	7	3109	0	3084	44	0
39	8	2918	0	2950	59	0
40	1	4377	0	4433	95	0
41	2	4446	0	4446	94	0
42	9	2867	0	2838	43	0
43	5	2624	0	2592	100	0
44	A	7205	0	3626	61	0
45	a	1	0	0	0	0
46	b	1	0	0	0	0
All	All	109778	0	89993	765	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 5.

The worst 5 of 765 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)
40:1:42:THR:CG2	40:1:81:VAL:HG21	1.44	1.45
41:2:710:ILE:HB	41:2:714:PHE:CE2	1.61	1.35
41:2:438:LEU:CD2	41:2:493:LYS:HD3	1.56	1.35

*Continued on next page...*



*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
41:2:438:LEU:HD21	41:2:493:LYS:CD	1.71	1.19
4:D:141:LEU:HD11	4:D:238:LYS:NZ	1.55	1.19

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	215/295 (73%)	206 (96%)	8 (4%)	1 (0%)	31	67
3	C	211/264 (80%)	187 (89%)	23 (11%)	1 (0%)	31	67
4	D	219/255 (86%)	202 (92%)	16 (7%)	1 (0%)	31	67
5	F	260/263 (99%)	240 (92%)	20 (8%)	0	100	100
6	H	235/249 (94%)	219 (93%)	16 (7%)	0	100	100
7	I	181/194 (93%)	165 (91%)	16 (9%)	0	100	100
8	J	204/208 (98%)	188 (92%)	12 (6%)	4 (2%)	8	37
9	K	183/194 (94%)	167 (91%)	16 (9%)	0	100	100
10	M	139/158 (88%)	128 (92%)	11 (8%)	0	100	100
11	O	147/151 (97%)	138 (94%)	9 (6%)	0	100	100
12	P	134/151 (89%)	120 (90%)	13 (10%)	1 (1%)	24	61
13	W	81/83 (98%)	79 (98%)	2 (2%)	0	100	100
14	X	127/130 (98%)	119 (94%)	7 (6%)	1 (1%)	21	58
15	Y	139/143 (97%)	129 (93%)	6 (4%)	4 (3%)	5	30
16	Z	122/134 (91%)	114 (93%)	8 (7%)	0	100	100
17	b	99/115 (86%)	88 (89%)	7 (7%)	4 (4%)	3	23
18	c	81/84 (96%)	74 (91%)	6 (7%)	1 (1%)	14	49
19	f	55/133 (41%)	53 (96%)	2 (4%)	0	100	100

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
20	E	226/281 (80%)	211 (93%)	14 (6%)	1 (0%)	36	72
21	G	189/204 (93%)	166 (88%)	18 (10%)	5 (3%)	6	32
22	L	94/149 (63%)	86 (92%)	7 (7%)	1 (1%)	16	52
23	N	115/132 (87%)	94 (82%)	19 (16%)	2 (2%)	10	41
24	Q	113/145 (78%)	95 (84%)	15 (13%)	3 (3%)	5	31
25	R	140/172 (81%)	136 (97%)	3 (2%)	1 (1%)	24	61
26	S	130/135 (96%)	110 (85%)	18 (14%)	2 (2%)	11	43
27	T	142/152 (93%)	132 (93%)	7 (5%)	3 (2%)	8	37
28	U	139/145 (96%)	130 (94%)	7 (5%)	2 (1%)	12	45
29	V	98/119 (82%)	88 (90%)	10 (10%)	0	100	100
30	i	73/125 (58%)	70 (96%)	3 (4%)	0	100	100
31	d	60/69 (87%)	54 (90%)	6 (10%)	0	100	100
32	e	53/56 (95%)	48 (91%)	5 (9%)	0	100	100
33	g	66/156 (42%)	52 (79%)	14 (21%)	0	100	100
34	h	310/317 (98%)	278 (90%)	32 (10%)	0	100	100
35	3	417/462 (90%)	315 (76%)	86 (21%)	16 (4%)	3	24
36	4	270/364 (74%)	219 (81%)	45 (17%)	6 (2%)	7	36
37	6	213/218 (98%)	182 (85%)	31 (15%)	0	100	100
38	7	370/607 (61%)	316 (85%)	52 (14%)	2 (0%)	31	67
39	8	363/374 (97%)	283 (78%)	71 (20%)	9 (2%)	6	33
40	1	532/1362 (39%)	429 (81%)	85 (16%)	18 (3%)	4	27
41	2	543/913 (60%)	424 (78%)	104 (19%)	15 (3%)	5	30
42	9	350/558 (63%)	317 (91%)	28 (8%)	5 (1%)	12	45
43	5	322/363 (89%)	255 (79%)	53 (16%)	14 (4%)	3	21
All	All	8160/10782 (76%)	7106 (87%)	931 (11%)	123 (2%)	16	43

5 of 123 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
15	Y	61	GLN
17	b	8	ASN
24	Q	18	ARG
25	R	17	LYS
26	S	99	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	
2	B	180/246 (73%)	178 (99%)	2 (1%)	76	88
3	C	194/231 (84%)	193 (100%)	1 (0%)	90	95
4	D	186/205 (91%)	184 (99%)	2 (1%)	76	88
5	F	223/224 (100%)	222 (100%)	1 (0%)	92	96
6	H	207/218 (95%)	201 (97%)	6 (3%)	45	75
7	I	165/174 (95%)	160 (97%)	5 (3%)	44	74
8	J	178/180 (99%)	177 (99%)	1 (1%)	87	94
9	K	161/168 (96%)	159 (99%)	2 (1%)	74	87
10	M	130/142 (92%)	129 (99%)	1 (1%)	83	92
11	O	130/131 (99%)	129 (99%)	1 (1%)	83	92
12	P	106/119 (89%)	106 (100%)	0	100	100
13	W	68/68 (100%)	67 (98%)	1 (2%)	67	85
14	X	112/113 (99%)	111 (99%)	1 (1%)	81	91
15	Y	113/114 (99%)	112 (99%)	1 (1%)	81	91
16	Z	107/115 (93%)	107 (100%)	0	100	100
17	b	89/99 (90%)	87 (98%)	2 (2%)	55	80
18	c	75/76 (99%)	75 (100%)	0	100	100
19	f	47/106 (44%)	46 (98%)	1 (2%)	56	80
20	E	190/232 (82%)	186 (98%)	4 (2%)	56	80
21	G	158/170 (93%)	158 (100%)	0	100	100
22	L	87/125 (70%)	86 (99%)	1 (1%)	76	88
23	N	99/108 (92%)	98 (99%)	1 (1%)	78	89
24	Q	105/130 (81%)	100 (95%)	5 (5%)	28	63
25	R	117/140 (84%)	116 (99%)	1 (1%)	81	91
26	S	119/121 (98%)	118 (99%)	1 (1%)	83	92
27	T	125/132 (95%)	125 (100%)	0	100	100

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
28	U	111/116 (96%)	111 (100%)	0	100	100
29	V	92/107 (86%)	92 (100%)	0	100	100
30	i	66/103 (64%)	64 (97%)	2 (3%)	44	74
31	d	55/62 (89%)	54 (98%)	1 (2%)	62	82
32	e	48/49 (98%)	46 (96%)	2 (4%)	32	66
33	g	61/140 (44%)	59 (97%)	2 (3%)	41	72
34	h	271/275 (98%)	269 (99%)	2 (1%)	85	93
35	3	384/423 (91%)	369 (96%)	15 (4%)	35	68
36	4	239/282 (85%)	233 (98%)	6 (2%)	50	78
37	6	190/193 (98%)	189 (100%)	1 (0%)	90	95
38	7	342/544 (63%)	340 (99%)	2 (1%)	87	94
39	8	327/335 (98%)	319 (98%)	8 (2%)	52	78
40	1	490/1245 (39%)	473 (96%)	17 (4%)	39	71
41	2	494/812 (61%)	482 (98%)	12 (2%)	52	78
42	9	320/496 (64%)	314 (98%)	6 (2%)	60	82
43	5	293/320 (92%)	291 (99%)	2 (1%)	85	93
All	All	7254/9389 (77%)	7135 (98%)	119 (2%)	68	84

5 of 119 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
35	3	212	THR
36	4	146	LYS
41	2	736	HIS
35	3	222	PHE
35	3	326	VAL

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 62 such sidechains are listed below:

Mol	Chain	Res	Type
36	4	186	ASN
39	8	258	ASN
43	5	283	GLN
38	7	199	GLN
40	1	47	HIS



### 5.3.3 RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
1	a	1681/1697 (99%)	402 (23%)	0
44	A	338/377 (89%)	175 (51%)	23 (6%)
All	All	2019/2074 (97%)	577 (28%)	23 (1%)

5 of 577 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
1	a	2	A
1	a	3	C
1	a	11	A
1	a	15	U
1	a	17	C

5 of 23 RNA pucker outliers are listed below:

Mol	Chain	Res	Type
44	A	518	G
44	A	559	A
44	A	657	U
44	A	526	A
44	A	561	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

The worst 5 of 15 chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	a	834:C	O3'	841:G	P	19.31
1	a	697:G	O3'	729:C	P	19.00
1	a	756:C	O3'	788:G	P	16.00
1	a	323:C	O3'	329:G	P	15.98
1	a	130:G	O3'	141:A	P	15.87