



# wwPDB X-ray Structure Validation Summary Report

Feb 26, 2014 – 05:35 PM GMT

PDB ID : 4BPP  
Title : The crystal structure of the eukaryotic 40S ribosomal subunit in complex with eIF1 and eIF1A - Complex 4  
Authors : Weisser, M.; Voigts-Hoffmann, F.; Rabl, J.; Leibundgut, M.; Ban, N.  
Deposited on : 2013-05-27  
Resolution : 3.70 Å(reported)

This is a wwPDB validation summary report for a publicly released PDB entry.

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

A user guide is available at <http://wwpdb.org/ValidationPDFNotes.html>

---

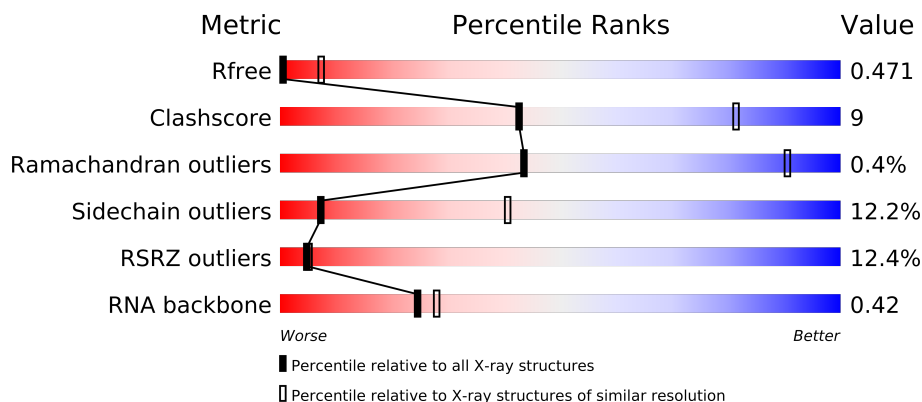
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity	:	4.02b-467
Mogul	:	1.15 2013
Xtriage (Phenix)	:	dev-1323
EDS	:	stable22639
Percentile statistics	:	21963
Refmac	:	5.8.0049
CCP4	:	6.3.0 (Settle)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP)	:	stable22683

# 1 Overall quality at a glance

The reported resolution of this entry is 3.70 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)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	66092	1098 (4.00-3.40)
Clashscore	79885	1009 (3.94-3.46)
Ramachandran outliers	78287	1016 (3.98-3.42)
Sidechain outliers	78261	1014 (3.98-3.42)
RSRZ outliers	66119	1099 (4.00-3.40)
RNA backbone	1838	1008 (4.52-2.80)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	0	211	
2	1	68	
3	2	208	
4	3	197	
5	4	265	
6	5	119	
7	6	81	
8	7	162	
9	8	143	
10	9	189	
11	A	1753	
12	B	241	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
13	C	243	
14	D	181	
15	E	296	
16	F	101	
17	G	200	
18	H	130	
19	I	145	
20	J	120	
21	K	151	
22	L	142	
23	M	155	
24	N	55	
25	O	153	
26	P	149	
27	Q	157	
28	R	343	
29	S	144	
30	T	155	
31	U	126	
32	V	130	
33	W	259	
34	X	80	
35	Y	293	
36	Z	97	

The following table lists non-polymeric compounds that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Geometry	Electron density
38	MG	A	5003	-	X
38	MG	A	5008	-	X
38	MG	A	5019	-	X
38	MG	A	5023	-	X
38	MG	A	5026	-	X
38	MG	A	5034	-	X
38	MG	A	5035	-	X
38	MG	A	5053	-	X
38	MG	A	5070	-	X
38	MG	A	5071	-	X
38	MG	A	5078	-	X

## 2 Entry composition

There are 39 unique types of molecules in this entry. The entry contains 78902 atoms, of which 0 are hydrogen and 0 are deuterium.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, 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 protein called TRANSLATION INITIATION FACTOR EIF-1A FAMILY PROTEIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	0	99	Total	C	N	O	S	0	0	0
			817	517	142	152	6			

There are 19 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
0	-17	MET	-	EXPRESSION TAG	UNP I7MK25
0	-16	GLY	-	EXPRESSION TAG	UNP I7MK25
0	-15	SER	-	EXPRESSION TAG	UNP I7MK25
0	-14	SER	-	EXPRESSION TAG	UNP I7MK25
0	-13	HIS	-	EXPRESSION TAG	UNP I7MK25
0	-12	HIS	-	EXPRESSION TAG	UNP I7MK25
0	-11	HIS	-	EXPRESSION TAG	UNP I7MK25
0	-10	HIS	-	EXPRESSION TAG	UNP I7MK25
0	-9	HIS	-	EXPRESSION TAG	UNP I7MK25
0	-8	HIS	-	EXPRESSION TAG	UNP I7MK25
0	-7	GLU	-	EXPRESSION TAG	UNP I7MK25
0	-6	ASN	-	EXPRESSION TAG	UNP I7MK25
0	-5	LEU	-	EXPRESSION TAG	UNP I7MK25
0	-4	TYR	-	EXPRESSION TAG	UNP I7MK25
0	-3	PHE	-	EXPRESSION TAG	UNP I7MK25
0	-2	GLN	-	EXPRESSION TAG	UNP I7MK25
0	-1	SER	-	EXPRESSION TAG	UNP I7MK25
0	0	ASN	-	EXPRESSION TAG	UNP I7MK25
0	1	ALA	-	EXPRESSION TAG	UNP I7MK25

- Molecule 2 is a protein called 40S RIBOSOMAL PROTEIN RPS28E.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	1	66	Total	C	N	O	S	0	0	0
			511	308	103	96	4			

- Molecule 3 is a protein called 40S RIBOSOMAL PROTEIN S8.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	2	207	Total	C	N	O	S	0	0	0
			1693	1057	336	296	4			

- Molecule 4 is a protein called 40S RIBOSOMAL PROTEIN RPS7E.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	3	196	Total	C	N	O	S	0	0	0
			1629	1048	286	294	1			

- Molecule 5 is a protein called 40S RIBOSOMAL PROTEIN S3A.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	4	221	Total	C	N	O	S	0	0	0
			1775	1121	319	331	4			

- Molecule 6 is a protein called 40S RIBOSOMAL PROTEIN RPS26E.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
6	5	100	Total	C	N	O	S	0	0	0
			812	496	172	138	6			

- Molecule 7 is a protein called 40S RIBOSOMAL PROTEIN S27.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
7	6	80	Total	C	N	O	S	0	0	0
			632	398	110	116	8			

- Molecule 8 is a protein called 40S RIBOSOMAL PROTEIN RPS10E.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
8	7	101	Total	C	N	O	S	0	0	0
			833	546	139	146	2			

- Molecule 9 is a protein called 40S RIBOSOMAL PROTEIN RPS25E.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
9	8	79	Total	C	N	O	S	0	0	0
			615	388	112	113	2			

- Molecule 10 is a protein called 40S RIBOSOMAL PROTEIN RPS31E.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
10	9	93	Total	C	N	O	S	0	0	0
			751	477	143	126	5			

- Molecule 11 is a RNA chain called 18S RRNA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
11	A	1717	Total	C	N	O	P	0	0	0
			36629	16385	6539	11988	1717			

- Molecule 12 is a protein called 40S RIBOSOMAL PROTEIN SA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
12	B	201	Total	C	N	O	S	0	0	0
			1619	1023	285	301	10			

- Molecule 13 is a protein called 40S RIBOSOMAL PROTEIN RPS3E.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
13	C	228	Total	C	N	O	S	0	0	0
			1811	1167	318	318	8			

- Molecule 14 is a protein called 40S RIBOSOMAL PROTEIN RPS9E.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
14	D	180	Total	C	N	O	S	0	0	0
			1478	932	287	254	5			

- Molecule 15 is a protein called 40S RIBOSOMAL PROTEIN RPS2E.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
15	E	229	Total	C	N	O	S	0	0	0
			1818	1171	321	323	3			

- Molecule 16 is a protein called EIF1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
16	F	89	Total	C	N	O	S	0	0	0
			736	465	131	137	3			

- Molecule 17 is a protein called 40S RIBOSOMAL PROTEIN RPS5E.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
17	G	192	Total	C	N	O	S	0	0	0
			1520	961	281	270	8			

- Molecule 18 is a protein called 40S RIBOSOMAL PROTEIN RPS22E.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
18	H	129	Total	C	N	O	S	0	0	0
			1040	671	184	180	5			

- Molecule 19 is a protein called 40S RIBOSOMAL PROTEIN RPS16E.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
19	I	143	Total	C	N	O	S	0	0	0
			1135	715	217	198	5			

- Molecule 20 is a protein called 40S RIBOSOMAL PROTEIN RPS20E.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
20	J	108	Total	C	N	O	S	0	0	0
			859	539	154	160	6			

- Molecule 21 is a protein called 40S RIBOSOMAL PROTEIN RPS14E.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
21	K	140	Total	C	N	O	S	0	0	0
			1063	654	206	197	6			

- Molecule 22 is a protein called 40S RIBOSOMAL PROTEIN S12.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
22	L	140	Total	C	N	O	S	0	0	0
			1086	685	217	179	5			

- Molecule 23 is a protein called 40S RIBOSOMAL PROTEIN RPS18E.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
23	M	153	Total	C	N	O	S	0	0	0
			1231	775	236	215	5			

- Molecule 24 is a protein called 40S RIBOSOMAL PROTEIN RPS29E.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
24	N	54	Total	C	N	O	S	0	0	0
			454	283	92	73	6			

- Molecule 25 is a protein called 40S RIBOSOMAL PROTEIN RPS13E.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
25	O	152	Total	C	N	O	S	0	0	0
			1229	790	233	202	4			

- Molecule 26 is a protein called 40S RIBOSOMAL PROTEIN S24.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
26	P	148	Total	C	N	O	S	0	0	0
			1197	763	221	213				

- Molecule 27 is a protein called 40S RIBOSOMAL PROTEIN RPS11E.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
27	Q	156	Total	C	N	O	S	0	0	0
			1267	813	234	216	4			

- Molecule 28 is a protein called 40S RIBOSOMAL PROTEIN RACK1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
28	R	338	Total	C	N	O	S	0	0	0
			2682	1711	462	501	8			

- Molecule 29 is a protein called 40S RIBOSOMAL PROTEIN RPS15E.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
29	S	128	Total	C	N	O	S	0	0	0
			1010	648	178	180	4			

- Molecule 30 is a protein called 40S RIBOSOMAL PROTEIN RPS19E.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
30	T	154	Total	C	N	O	S	0	0	0
			1242	785	234	221	2			

- Molecule 31 is a protein called 40S RIBOSOMAL PROTEIN RPS12E.



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
31	U	124	Total	C	N	O	S	0	0	0
			952	599	166	182	5			

- Molecule 32 is a protein called 40S RIBOSOMAL PROTEIN RPS17E.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
32	V	119	Total	C	N	O	S	0	0	0
			968	613	180	173	2			

- Molecule 33 is a protein called 40S RIBOSOMAL PROTEIN RPS4E.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
33	W	259	Total	C	N	O	S	0	0	0
			2079	1322	383	370	4			

- Molecule 34 is a protein called 40S RIBOSOMAL PROTEIN RPS30E.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
34	X	74	Total	C	N	O	S	0	0	0
			599	376	124	96	3			

- Molecule 35 is a protein called 40S RIBOSOMAL PROTEIN S6.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
35	Y	228	Total	C	N	O	S	0	0	0
			1826	1157	340	318	11			

- Molecule 36 is a protein called 40S RIBOSOMAL PROTEIN RPS21E.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
36	Z	97	Total	C	N	O	S	0	0	0
			747	458	139	146	4			

- Molecule 37 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
37	9	1	Total	Zn	0	0
			1	1		
37	N	1	Total	Zn	0	0
			1	1		
37	6	1	Total	Zn	0	0
			1	1		

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
37	5	1	Total	Zn	0	0
			1	1		

- Molecule 38 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
38	A	79	Total	Mg	0	0
			79	79		

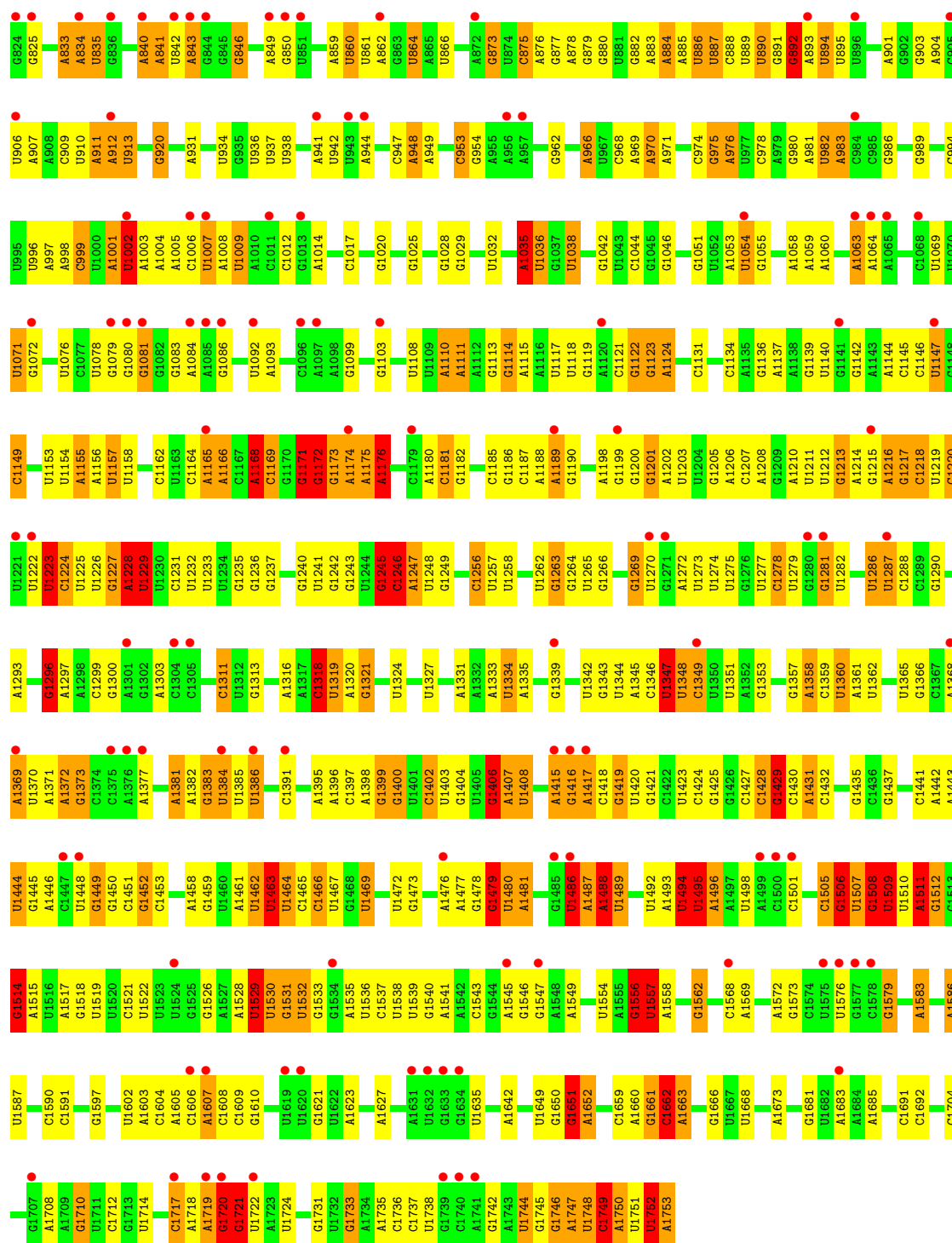
- Molecule 39 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
39	A	474	Total	O	0	0
			474	474		

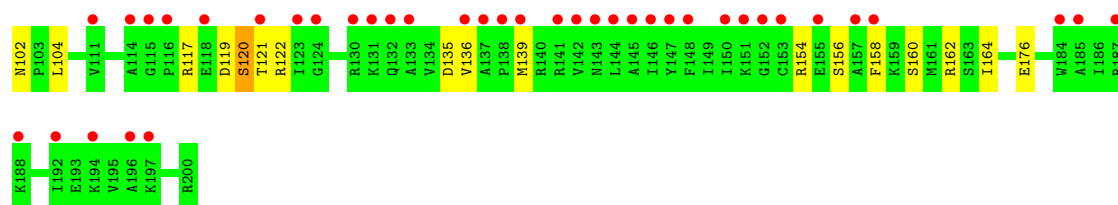












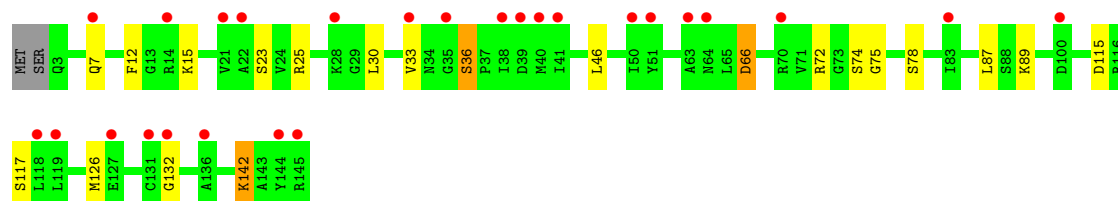
- Molecule 18: 40S RIBOSOMAL PROTEIN RPS22E

Chain H:



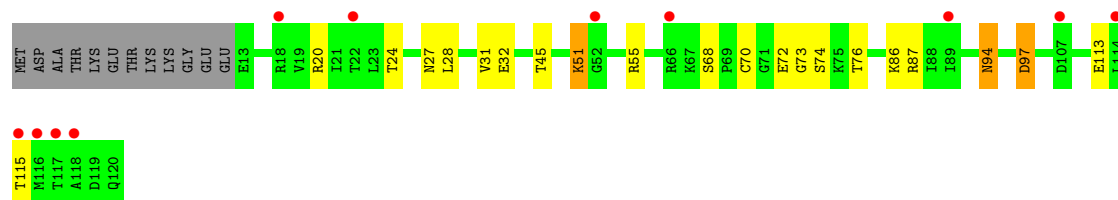
- Molecule 19: 40S RIBOSOMAL PROTEIN RPS16E

Chain I:



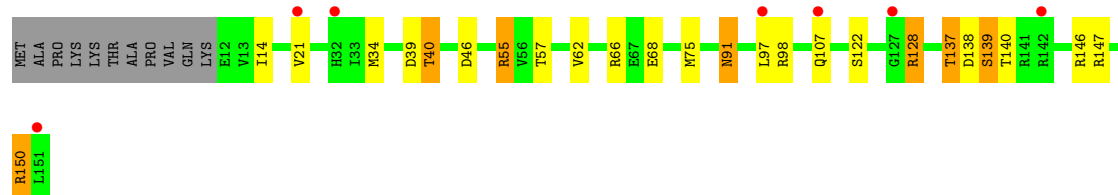
- Molecule 20: 40S RIBOSOMAL PROTEIN RPS20E

Chain J:



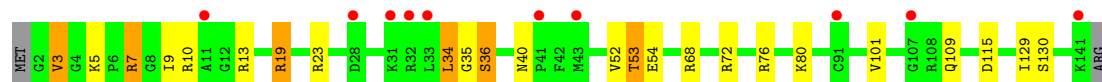
- Molecule 21: 40S RIBOSOMAL PROTEIN RPS14E

Chain K:



- Molecule 22: 40S RIBOSOMAL PROTEIN S12

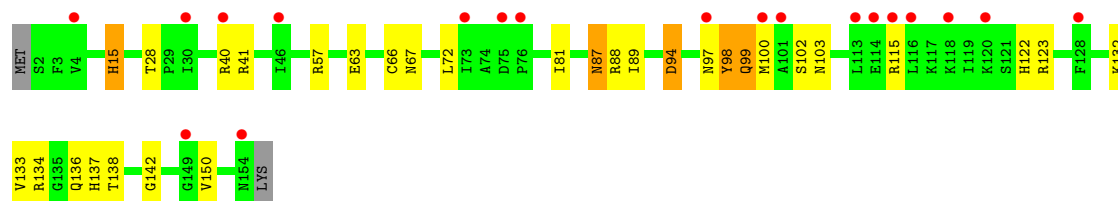
Chain L:



- Molecule 23: 40S RIBOSOMAL PROTEIN RPS18E



Chain M:



- Molecule 24: 40S RIBOSOMAL PROTEIN RPS29E

Chain N:



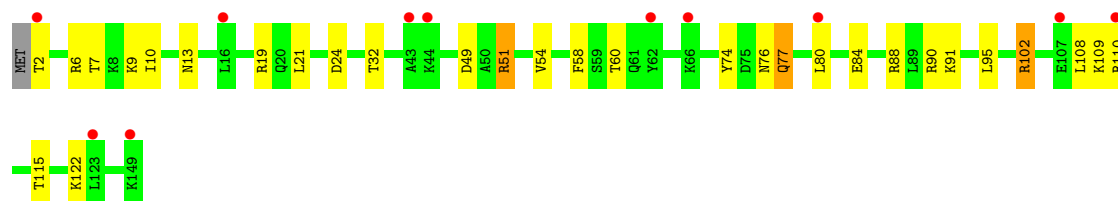
- Molecule 25: 40S RIBOSOMAL PROTEIN RPS13E

Chain O:



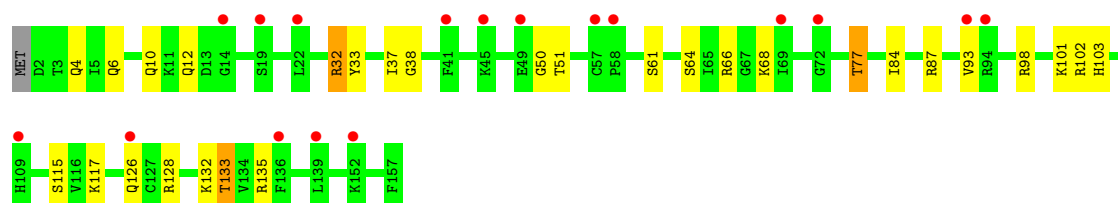
- Molecule 26: 40S RIBOSOMAL PROTEIN S24

Chain P:



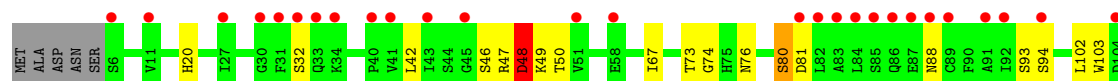
- Molecule 27: 40S RIBOSOMAL PROTEIN RPS11E

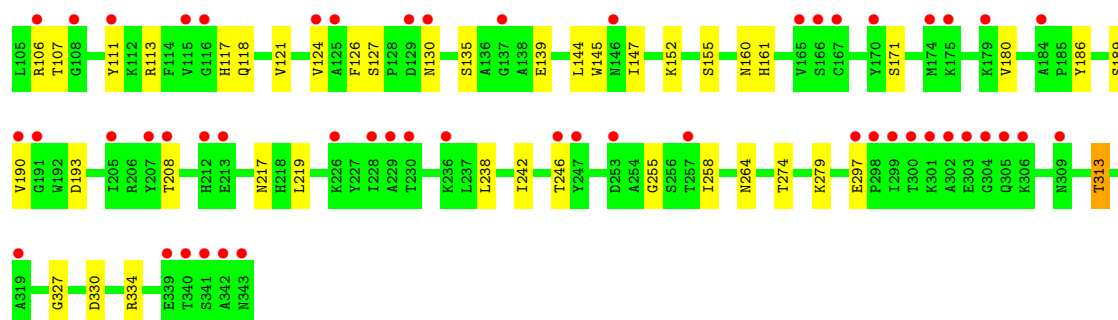
Chain Q:



- Molecule 28: 40S RIBOSOMAL PROTEIN RACK1

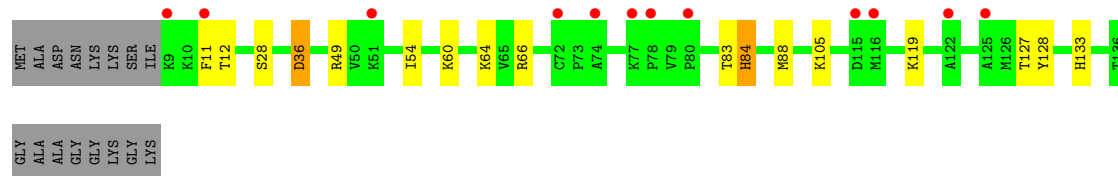
Chain R:





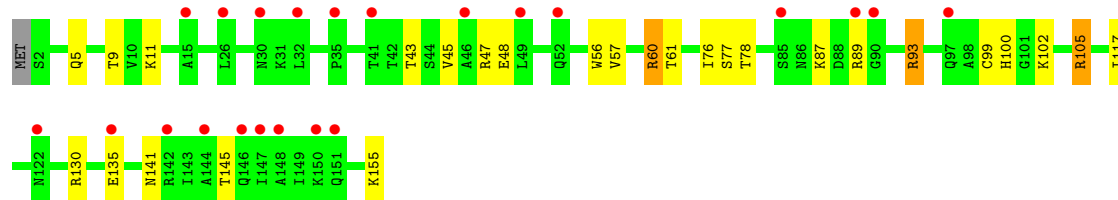
• Molecule 29: 40S RIBOSOMAL PROTEIN RPS15E

Chain S:



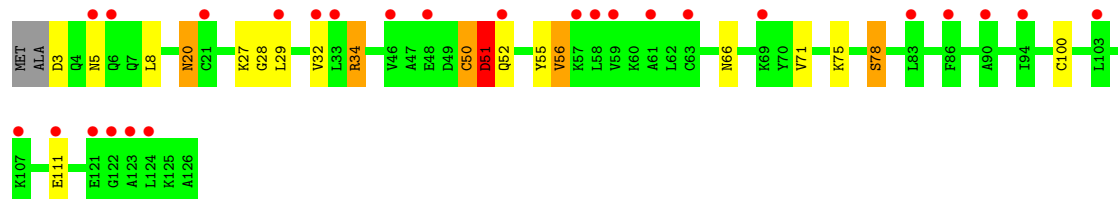
• Molecule 30: 40S RIBOSOMAL PROTEIN RPS19E

Chain T:



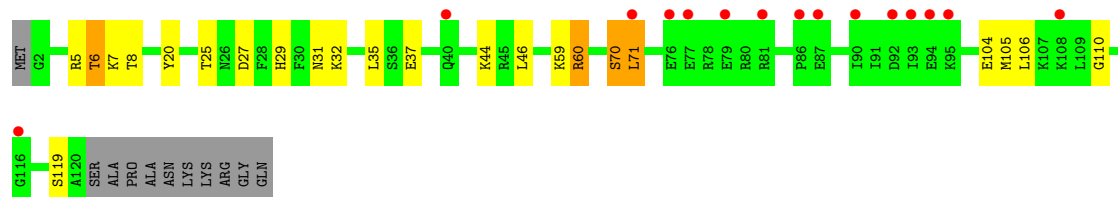
• Molecule 31: 40S RIBOSOMAL PROTEIN RPS12E

Chain U:



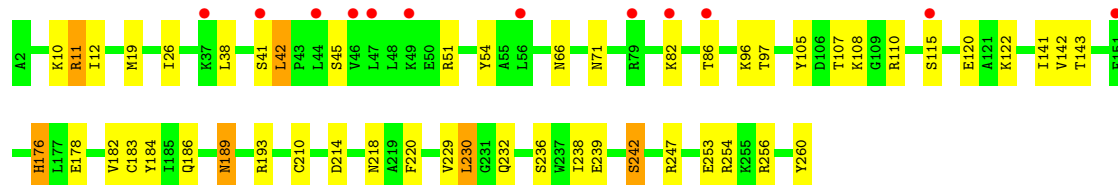
• Molecule 32: 40S RIBOSOMAL PROTEIN RPS17E

Chain V:



• Molecule 33: 40S RIBOSOMAL PROTEIN RPS4E

Chain W: 



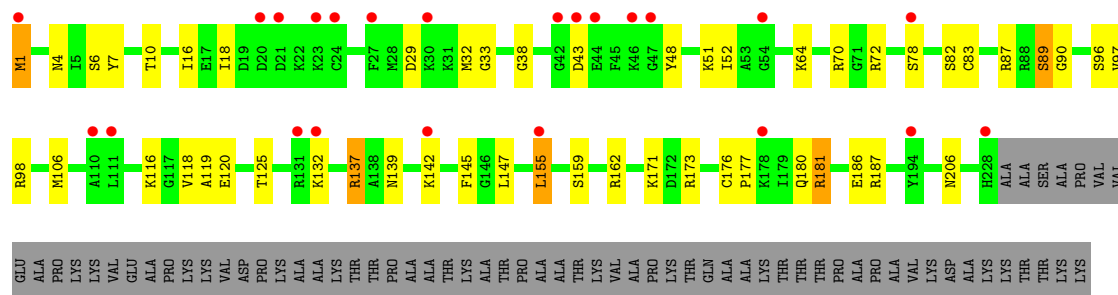
- Molecule 34: 40S RIBOSOMAL PROTEIN RPS30E

Chain X: 



- Molecule 35: 40S RIBOSOMAL PROTEIN S6

Chain Y: 



- Molecule 36: 40S RIBOSOMAL PROTEIN RPS21E

Chain Z: 



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	209.99Å 471.55Å 298.54Å 90.00° 91.02° 90.00°	Depositor
Resolution (Å)	49.75 – 3.70 49.75 – 3.70	Depositor EDS
% Data completeness (in resolution range)	91.0 (49.75-3.70) 86.3 (49.75-3.70)	Depositor EDS
$R_{merge}$	0.15	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.43 (at 3.67Å)	Xtriage
Refinement program	PHENIX (PHENIX.REFINE)	Depositor
R, $R_{free}$	0.202 , 0.229 0.470 , 0.471	Depositor DCC
$R_{free}$ test set	5341 reflections (0.97%)	DCC
Wilson B-factor (Å <sup>2</sup> )	97.6	Xtriage
Anisotropy	0.570	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , 96.3	EDS
Estimated twinning fraction	0.045 for h,-k,-l	Xtriage
L-test for twinning	$\langle  L  \rangle = 0.45$ , $\langle L^2 \rangle = 0.28$	Xtriage
Outliers	0 of 607573 reflections	Xtriage
$F_o, F_c$ correlation	0.59	EDS
Total number of atoms	78902	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	170.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 1.57% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

## 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	0	0.31	0/827	0.57	0/1103
2	1	0.28	0/510	0.65	0/677
3	2	0.32	0/1717	0.61	0/2288
4	3	0.32	0/1656	0.60	0/2223
5	4	0.54	4/1801 (0.2%)	0.70	2/2417 (0.1%)
6	5	0.35	0/823	0.67	0/1100
7	6	0.32	0/640	0.55	0/855
8	7	0.30	0/853	0.56	0/1148
9	8	0.31	0/620	0.62	0/831
10	9	0.31	0/764	0.67	1/1007 (0.1%)
11	A	0.44	1/40993 (0.0%)	1.12	219/63880 (0.3%)
12	B	0.29	0/1652	0.58	0/2240
13	C	0.31	0/1846	0.58	1/2479 (0.0%)
14	D	0.33	0/1501	0.65	0/2003
15	E	0.35	0/1864	0.62	0/2521
16	F	0.30	0/751	0.60	0/1010
17	G	0.31	0/1546	0.62	1/2079 (0.0%)
18	H	0.36	0/1058	0.73	1/1421 (0.1%)
19	I	0.31	0/1151	0.61	0/1540
20	J	0.37	1/868 (0.1%)	0.63	0/1168
21	K	0.33	0/1078	0.70	0/1452
22	L	0.34	0/1103	0.64	0/1471
23	M	0.28	0/1252	0.60	0/1680
24	N	0.34	0/465	0.63	0/619
25	O	0.32	0/1253	0.62	0/1677
26	P	0.31	0/1215	0.60	0/1626
27	Q	0.33	0/1290	0.65	0/1731
28	R	0.29	0/2750	0.60	0/3726
29	S	0.27	0/1028	0.54	0/1374
30	T	0.33	0/1264	0.57	0/1698
31	U	0.30	0/961	0.58	0/1288
32	V	0.30	0/981	0.57	0/1311

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	# Z  >5	RMSZ	# Z  >5
33	W	0.34	0/2119	0.62	0/2849
34	X	0.28	0/612	0.54	0/812
35	Y	0.31	0/1852	0.55	0/2462
36	Z	0.33	0/755	0.61	0/1013
All	All	0.39	6/83419 (0.0%)	0.92	225/120779 (0.2%)

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
4	3	0	1
5	4	0	2
7	6	0	1
9	8	0	1
17	G	0	1
18	H	0	2
20	J	0	1
21	K	0	1
25	O	0	1
29	S	0	2
31	U	0	2
32	V	0	1
33	W	0	1
All	All	0	17

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	4	31	TRP	NE1-CE2	-10.69	1.23	1.37
5	4	31	TRP	CD1-NE1	8.48	1.52	1.38
11	A	1586	A	O3'-P	7.60	1.70	1.61
5	4	31	TRP	CD2-CE2	7.60	1.50	1.41
5	4	31	TRP	CG-CD2	6.46	1.54	1.43

The worst 5 of 225 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
11	A	515	U	N1-C2-O2	12.24	131.37	122.80
11	A	1296	G	N3-C2-N2	-11.41	111.91	119.90
11	A	447	C	C2-N1-C1'	9.34	129.08	118.80

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
11	A	573	A	C6-N1-C2	9.27	124.16	118.60
11	A	493	U	O4'-C1'-N1	9.11	115.48	108.20

There are no chirality outliers.

5 of 17 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
4	3	131	LEU	Peptide
5	4	238	TYR	Peptide
5	4	71	THR	Peptide
7	6	65	THR	Peptide
9	8	99	ASN	Peptide

## 5.2 Close contacts ⓘ

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	0	817	0	0	10	0
2	1	511	0	0	4	0
3	2	1693	0	0	25	0
4	3	1629	0	0	17	0
5	4	1775	0	0	15	0
6	5	812	0	0	11	0
7	6	632	0	0	4	0
8	7	833	0	0	11	0
9	8	615	0	0	4	0
10	9	751	0	0	18	0
11	A	36629	0	0	432	0
12	B	1619	0	0	8	0
13	C	1811	0	0	18	0
14	D	1478	0	0	8	0
15	E	1818	0	0	16	0
16	F	736	0	0	2	0
17	G	1520	0	0	12	0
18	H	1040	0	0	8	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
19	I	1135	0	0	7	0
20	J	859	0	0	10	0
21	K	1063	0	0	13	0
22	L	1086	0	0	11	0
23	M	1231	0	0	16	0
24	N	454	0	0	16	0
25	O	1229	0	0	10	0
26	P	1197	0	0	14	0
27	Q	1267	0	0	12	0
28	R	2682	0	0	21	0
29	S	1010	0	0	6	0
30	T	1242	0	0	16	0
31	U	952	0	0	10	0
32	V	968	0	0	12	0
33	W	2079	0	0	16	0
34	X	599	0	0	4	0
35	Y	1826	0	0	24	0
36	Z	747	0	0	6	0
37	5	1	0	0	0	0
37	6	1	0	0	0	0
37	9	1	0	0	0	0
37	N	1	0	0	0	0
38	A	79	0	0	0	0
39	A	474	0	0	49	0
All	All	78902	0	0	640	0

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 9.

The worst 5 of 640 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
11:A:1377:A:OP2	17:G:54:LYS:NZ	1.65	1.29
11:A:1214:A:OP1	29:S:64:LYS:NZ	1.71	1.23
9:8:81:ARG:NH2	11:A:1505:C:OP2	1.73	1.20
15:E:141:ARG:NH1	36:Z:25:ASP:OD2	1.76	1.15
36:Z:45:SER:O	36:Z:74:ARG:NH1	1.81	1.14

There are no symmetry-related clashes.



## 5.3 Torsion angles

### 5.3.1 Protein backbone ⓘ

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution. 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	
1	0	97/211 (46%)	91 (94%)	6 (6%)	0	100	100
2	1	64/68 (94%)	62 (97%)	2 (3%)	0	100	100
3	2	205/208 (99%)	195 (95%)	10 (5%)	0	100	100
4	3	194/197 (98%)	187 (96%)	7 (4%)	0	100	100
5	4	219/265 (83%)	203 (93%)	13 (6%)	3 (1%)	16	74
6	5	98/119 (82%)	98 (100%)	0	0	100	100
7	6	78/81 (96%)	71 (91%)	6 (8%)	1 (1%)	18	76
8	7	99/162 (61%)	95 (96%)	4 (4%)	0	100	100
9	8	77/143 (54%)	73 (95%)	4 (5%)	0	100	100
10	9	91/189 (48%)	89 (98%)	1 (1%)	1 (1%)	21	79
12	B	199/241 (83%)	197 (99%)	2 (1%)	0	100	100
13	C	226/243 (93%)	214 (95%)	12 (5%)	0	100	100
14	D	178/181 (98%)	174 (98%)	4 (2%)	0	100	100
15	E	227/296 (77%)	212 (93%)	14 (6%)	1 (0%)	43	90
16	F	87/101 (86%)	82 (94%)	5 (6%)	0	100	100
17	G	190/200 (95%)	178 (94%)	10 (5%)	2 (1%)	21	79
18	H	127/130 (98%)	112 (88%)	14 (11%)	1 (1%)	27	83
19	I	141/145 (97%)	132 (94%)	9 (6%)	0	100	100
20	J	106/120 (88%)	104 (98%)	2 (2%)	0	100	100
21	K	138/151 (91%)	133 (96%)	5 (4%)	0	100	100
22	L	138/142 (97%)	125 (91%)	11 (8%)	2 (1%)	16	74
23	M	151/155 (97%)	138 (91%)	11 (7%)	2 (1%)	18	76
24	N	52/55 (94%)	49 (94%)	3 (6%)	0	100	100
25	O	150/153 (98%)	140 (93%)	9 (6%)	1 (1%)	30	85
26	P	146/149 (98%)	135 (92%)	11 (8%)	0	100	100
27	Q	154/157 (98%)	149 (97%)	5 (3%)	0	100	100

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
28	R	336/343 (98%)	307 (91%)	27 (8%)	2 (1%)	33	86
29	S	126/144 (88%)	122 (97%)	3 (2%)	1 (1%)	27	83
30	T	152/155 (98%)	144 (95%)	8 (5%)	0	100	100
31	U	122/126 (97%)	112 (92%)	10 (8%)	0	100	100
32	V	117/130 (90%)	115 (98%)	1 (1%)	1 (1%)	25	82
33	W	257/259 (99%)	247 (96%)	10 (4%)	0	100	100
34	X	72/80 (90%)	68 (94%)	4 (6%)	0	100	100
35	Y	226/293 (77%)	215 (95%)	10 (4%)	1 (0%)	43	90
36	Z	95/97 (98%)	93 (98%)	2 (2%)	0	100	100
All	All	5135/5889 (87%)	4861 (95%)	255 (5%)	19 (0%)	43	90

5 of 19 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
22	L	3	VAL
25	O	152	VAL
28	R	48	ASP
5	4	239	SER
7	6	66	GLY

### 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 X-ray entries followed by that with respect to entries of similar resolution. 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	
1	0	90/192 (47%)	83 (92%)	7 (8%)	18	65
2	1	55/57 (96%)	44 (80%)	11 (20%)	2	13
3	2	184/185 (100%)	158 (86%)	26 (14%)	5	33
4	3	182/183 (100%)	153 (84%)	29 (16%)	4	27
5	4	197/225 (88%)	167 (85%)	30 (15%)	4	29
6	5	90/107 (84%)	76 (84%)	14 (16%)	4	27
7	6	71/72 (99%)	60 (84%)	11 (16%)	4	28
8	7	91/136 (67%)	83 (91%)	8 (9%)	14	60

Continued on next page...

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
9	8	70/109 (64%)	59 (84%)	11 (16%)	4	27
10	9	81/156 (52%)	72 (89%)	9 (11%)	9	46
12	B	180/211 (85%)	164 (91%)	16 (9%)	14	59
13	C	196/210 (93%)	178 (91%)	18 (9%)	13	57
14	D	161/162 (99%)	128 (80%)	33 (20%)	2	12
15	E	193/250 (77%)	168 (87%)	25 (13%)	6	37
16	F	80/92 (87%)	76 (95%)	4 (5%)	34	81
17	G	163/169 (96%)	145 (89%)	18 (11%)	9	47
18	H	116/117 (99%)	97 (84%)	19 (16%)	3	24
19	I	120/122 (98%)	105 (88%)	15 (12%)	7	40
20	J	101/111 (91%)	91 (90%)	10 (10%)	11	53
21	K	112/121 (93%)	97 (87%)	15 (13%)	6	36
22	L	112/114 (98%)	96 (86%)	16 (14%)	5	32
23	M	133/135 (98%)	120 (90%)	13 (10%)	12	53
24	N	48/49 (98%)	41 (85%)	7 (15%)	5	31
25	O	135/136 (99%)	118 (87%)	17 (13%)	7	39
26	P	133/134 (99%)	117 (88%)	16 (12%)	7	41
27	Q	140/141 (99%)	122 (87%)	18 (13%)	6	38
28	R	291/295 (99%)	264 (91%)	27 (9%)	13	56
29	S	108/117 (92%)	97 (90%)	11 (10%)	11	51
30	T	133/134 (99%)	120 (90%)	13 (10%)	12	53
31	U	103/104 (99%)	93 (90%)	10 (10%)	12	54
32	V	107/115 (93%)	98 (92%)	9 (8%)	16	62
33	W	226/226 (100%)	194 (86%)	32 (14%)	5	33
34	X	61/67 (91%)	54 (88%)	7 (12%)	8	44
35	Y	197/244 (81%)	177 (90%)	20 (10%)	11	51
36	Z	82/82 (100%)	74 (90%)	8 (10%)	12	53
All	All	4542/5080 (89%)	3989 (88%)	553 (12%)	7	40

5 of 553 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
16	F	65	ASP

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
20	J	74	SER
33	W	232	GLN
17	G	66	ARG
18	H	65	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

### 5.3.3 RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
11	A	1716/1753 (97%)	647 (37%)	149 (8%)

5 of 647 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
11	A	2	A
11	A	3	C
11	A	4	C
11	A	8	U
11	A	11	A

5 of 149 RNA pucker outliers are listed below:

Mol	Chain	Res	Type
11	A	771	A
11	A	1079	G
11	A	1556	G
11	A	798	G
11	A	890	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 83 ligands modelled in this entry, 83 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.

## 5.7 Other polymers ⓘ

There are no such residues in this entry.

## 5.8 Polymer linkage issues

There are no chain breaks in this entry.

## 6 Fit of model and data ⓘ

### 6.1 Protein, DNA and RNA chains ⓘ

In the following table, the column labelled ‘#RSRZ> 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95<sup>th</sup> percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled ‘Q< 0.9’ lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å <sup>2</sup> )	Q<0.9
1	0	99/211 (46%)	0.86	7 (7%) 16 11	92, 149, 224, 274	0
2	1	66/68 (97%)	0.80	9 (13%) 4 4	203, 261, 324, 343	0
3	2	207/208 (99%)	0.80	16 (7%) 13 10	71, 144, 216, 259	0
4	3	196/197 (99%)	0.77	21 (10%) 6 7	116, 179, 252, 285	0
5	4	221/265 (83%)	0.66	17 (7%) 13 10	103, 155, 208, 245	0
6	5	100/119 (84%)	0.93	11 (11%) 6 6	86, 134, 219, 240	0
7	6	80/81 (98%)	0.80	5 (6%) 19 13	117, 170, 216, 228	0
8	7	101/162 (62%)	1.02	15 (14%) 3 4	130, 187, 230, 255	0
9	8	79/143 (55%)	1.17	17 (21%) 1 2	252, 277, 348, 437	0
10	9	93/189 (49%)	0.97	15 (16%) 2 3	197, 257, 314, 331	0
11	A	1717/1753 (97%)	1.22	245 (14%) 3 4	79, 137, 297, 532	0
12	B	201/241 (83%)	0.59	10 (4%) 28 18	103, 168, 221, 256	0
13	C	228/243 (93%)	0.72	18 (7%) 13 10	89, 156, 238, 280	0
14	D	180/181 (99%)	0.99	15 (8%) 11 9	86, 115, 180, 234	0
15	E	229/296 (77%)	0.93	22 (9%) 8 8	81, 123, 226, 266	0
16	F	89/101 (88%)	1.01	13 (14%) 3 4	93, 146, 216, 259	0
17	G	192/200 (96%)	1.49	58 (30%) 1 2	198, 256, 315, 346	0
18	H	129/130 (99%)	0.73	4 (3%) 47 29	76, 117, 156, 179	0
19	I	143/145 (98%)	1.08	26 (18%) 2 3	152, 218, 272, 312	0
20	J	108/120 (90%)	0.89	11 (10%) 7 7	89, 145, 231, 248	0
21	K	140/151 (92%)	0.93	7 (5%) 28 18	89, 147, 206, 225	0
22	L	140/142 (98%)	0.94	10 (7%) 16 11	77, 105, 148, 213	0
23	M	153/155 (98%)	0.83	19 (12%) 5 5	198, 260, 333, 366	0
24	N	54/55 (98%)	0.74	1 (1%) 64 41	95, 135, 203, 233	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å <sup>2</sup> )	Q<0.9
25	O	152/153 (99%)	1.00	22 (14%) 3 4	91, 153, 215, 269	0
26	P	148/149 (99%)	0.79	11 (7%) 14 11	92, 125, 170, 225	0
27	Q	156/157 (99%)	1.01	17 (10%) 6 7	80, 129, 223, 275	0
28	R	338/343 (98%)	1.23	79 (23%) 1 2	172, 250, 315, 352	0
29	S	128/144 (88%)	0.78	12 (9%) 9 8	231, 263, 309, 342	0
30	T	154/155 (99%)	1.04	22 (14%) 3 4	169, 226, 303, 354	0
31	U	124/126 (98%)	1.15	26 (20%) 1 2	210, 259, 325, 375	0
32	V	119/130 (91%)	0.86	15 (12%) 4 5	135, 188, 258, 292	0
33	W	259/259 (100%)	0.75	12 (4%) 31 20	76, 118, 164, 211	0
34	X	74/80 (92%)	0.99	13 (17%) 2 3	84, 141, 259, 292	0
35	Y	228/293 (77%)	0.85	23 (10%) 7 7	99, 163, 225, 282	0
36	Z	97/97 (100%)	1.13	22 (22%) 1 2	106, 143, 197, 210	0
All	All	6922/7642 (90%)	0.99	866 (12%) 5 5	71, 161, 290, 532	0

The worst 5 of 866 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
14	D	181	ALA	13.8
28	R	125	ALA	11.2
28	R	87	GLU	8.5
14	D	180	SER	8.3
25	O	83	ALA	8.1

## 6.2 Non-standard residues in protein, DNA, RNA chains ⓘ

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

## 6.3 Carbohydrates ⓘ

There are no carbohydrates in this entry.

## 6.4 Ligands ⓘ

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains.

The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSR	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
38	MG	A	5035	1/1	0.30	27.00	196,196,196,196	0
38	MG	A	5053	1/1	1.24	25.66	256,256,256,256	0
38	MG	A	5071	1/1	0.56	22.97	238,238,238,238	0
38	MG	A	5034	1/1	0.62	7.31	220,220,220,220	0
38	MG	A	5008	1/1	0.46	5.51	170,170,170,170	0
38	MG	A	5019	1/1	0.47	5.26	184,184,184,184	0
38	MG	A	5078	1/1	0.36	3.78	241,241,241,241	0
38	MG	A	5026	1/1	0.41	3.07	208,208,208,208	0
38	MG	A	5070	1/1	0.47	2.91	251,251,251,251	0
38	MG	A	5023	1/1	0.35	2.65	127,127,127,127	0
38	MG	A	5038	1/1	0.38	2.00	191,191,191,191	0
38	MG	A	5054	1/1	0.29	1.61	169,169,169,169	0
38	MG	A	5030	1/1	0.26	1.50	229,229,229,229	0
38	MG	A	5048	1/1	0.39	1.48	222,222,222,222	0
38	MG	A	5077	1/1	0.30	1.22	235,235,235,235	0
38	MG	A	5003	1/1	0.52	0.95	175,175,175,175	0
38	MG	A	5029	1/1	0.30	0.76	217,217,217,217	0
38	MG	A	5028	1/1	0.38	0.71	163,163,163,163	0
38	MG	A	5057	1/1	0.35	0.64	178,178,178,178	0
38	MG	A	5059	1/1	0.32	0.64	205,205,205,205	0
38	MG	A	5066	1/1	0.34	0.21	251,251,251,251	0
38	MG	A	5049	1/1	0.33	0.17	207,207,207,207	0
38	MG	A	5056	1/1	0.37	0.13	238,238,238,238	0
38	MG	A	5041	1/1	0.35	-0.15	162,162,162,162	0
38	MG	A	5039	1/1	0.27	-0.31	120,120,120,120	0
38	MG	A	5075	1/1	0.20	-0.31	239,239,239,239	0
38	MG	A	5006	1/1	0.28	-0.65	159,159,159,159	0
38	MG	A	5022	1/1	0.34	-0.76	108,108,108,108	0
38	MG	A	5073	1/1	0.31	-0.78	221,221,221,221	0
38	MG	A	5031	1/1	0.26	-0.86	218,218,218,218	0
37	ZN	6	500	1/1	0.16	-1.05	173,173,173,173	0
38	MG	A	5012	1/1	0.22	-1.10	118,118,118,118	0
38	MG	A	5013	1/1	0.25	-1.22	218,218,218,218	0
38	MG	A	5002	1/1	0.24	-1.26	111,111,111,111	0
38	MG	A	5044	1/1	0.23	-1.27	134,134,134,134	0
37	ZN	N	500	1/1	0.21	-1.30	132,132,132,132	0
38	MG	A	5047	1/1	0.23	-1.39	128,128,128,128	0
38	MG	A	5058	1/1	0.20	-1.40	227,227,227,227	0
38	MG	A	5040	1/1	0.24	-1.43	161,161,161,161	0
37	ZN	9	500	1/1	0.11	-1.45	278,278,278,278	0

*Continued on next page...*



*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSR	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
38	MG	A	5024	1/1	0.21	-1.47	144,144,144,144	0
38	MG	A	5069	1/1	0.16	-1.56	199,199,199,199	0
38	MG	A	5065	1/1	0.31	-1.56	146,146,146,146	0
38	MG	A	5017	1/1	0.25	-1.59	192,192,192,192	0
38	MG	A	5052	1/1	0.21	-1.59	218,218,218,218	0
38	MG	A	5043	1/1	0.24	-1.73	197,197,197,197	0
38	MG	A	5050	1/1	0.21	-1.77	123,123,123,123	0
38	MG	A	5068	1/1	0.22	-1.89	169,169,169,169	0
38	MG	A	5036	1/1	0.26	-1.97	199,199,199,199	0
38	MG	A	5072	1/1	0.27	-2.06	146,146,146,146	0
37	ZN	5	500	1/1	0.14	-2.09	117,117,117,117	0
38	MG	A	5042	1/1	0.22	-2.10	177,177,177,177	0
38	MG	A	5062	1/1	0.13	-2.15	213,213,213,213	0
38	MG	A	5037	1/1	0.23	-2.22	178,178,178,178	0
38	MG	A	5076	1/1	0.20	-2.23	201,201,201,201	0
38	MG	A	5021	1/1	0.21	-2.31	189,189,189,189	0
38	MG	A	5005	1/1	0.28	-2.31	123,123,123,123	0
38	MG	A	5060	1/1	0.20	-2.42	148,148,148,148	0
38	MG	A	5067	1/1	0.12	-2.43	220,220,220,220	0
38	MG	A	5020	1/1	0.24	-2.43	199,199,199,199	0
38	MG	A	5074	1/1	0.24	-2.48	198,198,198,198	0
38	MG	A	5001	1/1	0.16	-2.64	159,159,159,159	0
38	MG	A	5009	1/1	0.19	-2.82	110,110,110,110	0
38	MG	A	5004	1/1	0.19	-2.96	167,167,167,167	0
38	MG	A	5007	1/1	0.23	-3.28	142,142,142,142	0
38	MG	A	5014	1/1	0.30	-3.33	206,206,206,206	0
38	MG	A	5016	1/1	0.15	-3.92	191,191,191,191	0
38	MG	A	5033	1/1	0.18	-3.92	122,122,122,122	0
38	MG	A	5018	1/1	0.28	-3.96	104,104,104,104	0
38	MG	A	5015	1/1	0.24	-4.46	118,118,118,118	0
38	MG	A	5046	1/1	0.20	-4.48	148,148,148,148	0
38	MG	A	5032	1/1	0.21	-4.56	174,174,174,174	0
38	MG	A	5000	1/1	0.15	-4.63	107,107,107,107	0
38	MG	A	5063	1/1	0.17	-4.93	141,141,141,141	0
38	MG	A	5025	1/1	0.16	-5.11	105,105,105,105	0
38	MG	A	5011	1/1	0.23	-5.45	187,187,187,187	0
38	MG	A	5064	1/1	0.12	-5.62	186,186,186,186	0
38	MG	A	5051	1/1	0.15	-5.67	162,162,162,162	0
38	MG	A	5010	1/1	0.13	-7.41	201,201,201,201	0
38	MG	A	5061	1/1	0.16	-13.41	132,132,132,132	0
38	MG	A	5027	1/1	0.14	-40.33	162,162,162,162	0
38	MG	A	5055	1/1	0.31	-	152,152,152,152	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
38	MG	A	5045	1/1	0.12	-	197,197,197,197	0

## 6.5 Other polymers ⓘ

There are no such residues in this entry.