



# wwPDB X-ray Structure Validation Summary Report

Feb 28, 2014 – 09:32 PM GMT

PDB ID : 4KJC  
Title : Control of ribosomal subunit rotation by elongation factor G  
Authors : Pulk, A.; Cate, J.H.D.  
Deposited on : 2013-05-03  
Resolution : 2.90 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>

---

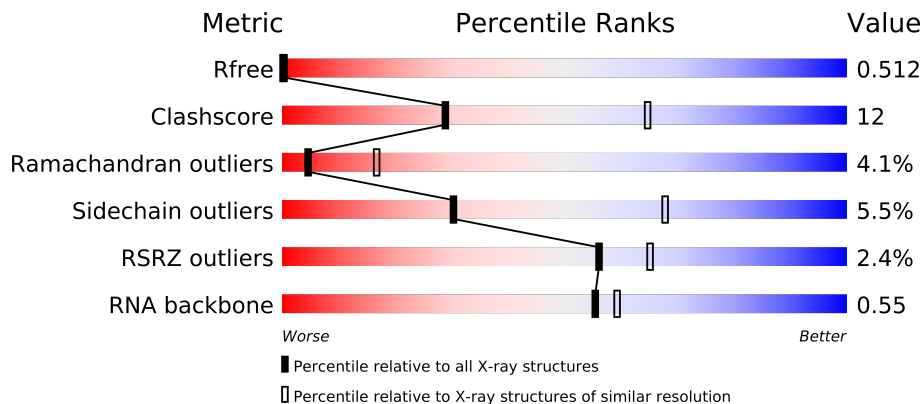
The following versions of software and data (see [references](#)) were used in the production of this report:

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 2.90 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.












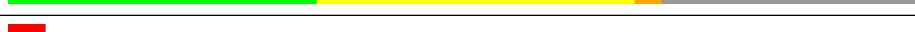

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	66092	1053 (2.90-2.90)
Clashscore	79885	1326 (2.90-2.90)
Ramachandran outliers	78287	1290 (2.90-2.90)
Sidechain outliers	78261	1292 (2.90-2.90)
RSRZ outliers	66119	1054 (2.90-2.90)
RNA backbone	1838	1055 (3.40-2.40)

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	A	1542	
2	B	241	
3	C	233	
4	D	206	
5	E	167	
6	F	135	
7	G	179	
8	H	130	
9	I	130	
10	J	103	
11	K	129	

Continued on next page...

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
12	L	124	
13	M	118	
14	N	101	
15	O	89	
16	P	82	
17	Q	84	
18	R	75	
19	S	92	
20	T	87	
21	U	71	
22	V	704	

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

Mol	Type	Chain	Res	Geometry	Electron density
23	MG	A	1601	-	X
23	MG	A	1604	-	X
23	MG	A	1605	-	X
23	MG	A	1609	-	X
23	MG	A	1613	-	X
23	MG	A	1621	-	X
23	MG	A	1633	-	X
23	MG	A	1634	-	X
23	MG	T	101	-	X

## 2 Entry composition

There are 25 unique types of molecules in this entry. The entry contains 57042 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 RNA chain called 16S rRNA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	1533	Total	C	N	O	P	0	0	0
			32895	14671	6036	10655	1533			

- Molecule 2 is a protein called 30S ribosomal protein S2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	218	Total	C	N	O	S	0	0	0
			1704	1081	305	311	7			

- Molecule 3 is a protein called 30S ribosomal protein S3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	206	Total	C	N	O	S	0	0	0
			1624	1028	305	288	3			

- Molecule 4 is a protein called 30S ribosomal protein S4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	D	205	Total	C	N	O	S	0	0	0
			1643	1026	315	298	4			

- Molecule 5 is a protein called 30S ribosomal protein S5.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	E	150	Total	C	N	O	S	0	0	0
			1105	687	211	201	6			

- Molecule 6 is a protein called 30S ribosomal protein S6.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
6	F	100	Total	C	N	O	S	0	0	0
			817	515	148	148	6			

- Molecule 7 is a protein called 30S ribosomal protein S7.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
7	G	151	Total	C	N	O	S	0	0	0
			1181	735	227	215	4			

- Molecule 8 is a protein called 30S ribosomal protein S8.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
8	H	129	Total	C	N	O	S	0	0	0
			979	616	173	184	6			

- Molecule 9 is a protein called 30S ribosomal protein S9.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
9	I	127	Total	C	N	O	S	0	0	0
			1022	634	206	179	3			

- Molecule 10 is a protein called 30S ribosomal protein S10.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
10	J	98	Total	C	N	O	S	0	0	0
			786	493	150	142	1			

- Molecule 11 is a protein called 30S ribosomal protein S11.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
11	K	117	Total	C	N	O	S	0	0	0
			877	540	174	160	3			

- Molecule 12 is a protein called 30S ribosomal protein S12.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
12	L	123	Total	C	N	O	S	0	0	0
			955	590	196	165	4			

- Molecule 13 is a protein called 30S ribosomal protein S13.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
13	M	114	Total	C	N	O	S	0	0	0
			883	546	178	156	3			

- Molecule 14 is a protein called 30S ribosomal protein S14.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
14	N	96	Total	C	N	O	S	0	0	0
			774	483	160	128	3			

- Molecule 15 is a protein called 30S ribosomal protein S15.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
15	O	88	Total	C	N	O	S	0	0	0
			714	439	144	130	1			

- Molecule 16 is a protein called 30S ribosomal protein S16.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
16	P	82	Total	C	N	O	S	0	0	0
			649	406	128	114	1			

- Molecule 17 is a protein called 30S ribosomal protein S17.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
17	Q	80	Total	C	N	O	S	0	0	0
			648	411	121	113	3			

- Molecule 18 is a protein called 30S ribosomal protein S18.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
18	R	55	Total	C	N	O	0	0	0
			455	288	86	81			

- Molecule 19 is a protein called 30S ribosomal protein S19.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
19	S	79	Total	C	N	O	S	0	0	0
			637	408	120	107	2			

- Molecule 20 is a protein called 30S ribosomal protein S20.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
20	T	85	Total	C	N	O	S	0	0	0
			665	411	137	114	3			

- Molecule 21 is a protein called 30S ribosomal protein S21.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
21	U	51	Total	C	N	O	S	0	0	0
			425	265	86	73	1			

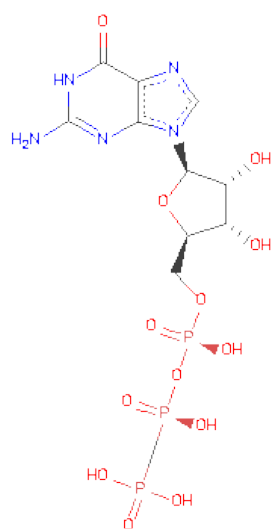
- Molecule 22 is a protein called elongation factor G.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
22	V	689	Total	C	N	O	S	0	0	0
			5319	3345	919	1030	25			

- Molecule 23 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
23	A	40	Total	Mg	0	0
			40	40		
23	T	1	Total	Mg	0	0
			1	1		
23	C	1	Total	Mg	0	0
			1	1		
23	V	1	Total	Mg	0	0
			1	1		
23	E	1	Total	Mg	0	0
			1	1		

- Molecule 24 is PHOSPHOMETHYLPHOSPHONICACID GUANYLATE ESTER (three-letter code: GCP) (formula:  $C_{11}H_{18}N_5O_{13}P_3$ ).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
24	V	1	Total	C	N	O	P	0	0
			32	11	5	13	3		

- Molecule 25 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
25	A	198	Total	O	0	0
			198	198		
25	D	1	Total	O	0	0
			1	1		
25	E	2	Total	O	0	0
			2	2		
25	N	5	Total	O	0	0
			5	5		
25	T	1	Total	O	0	0
			1	1		
25	U	1	Total	O	0	0
			1	1		
25	V	1	Total	O	0	0
			1	1		

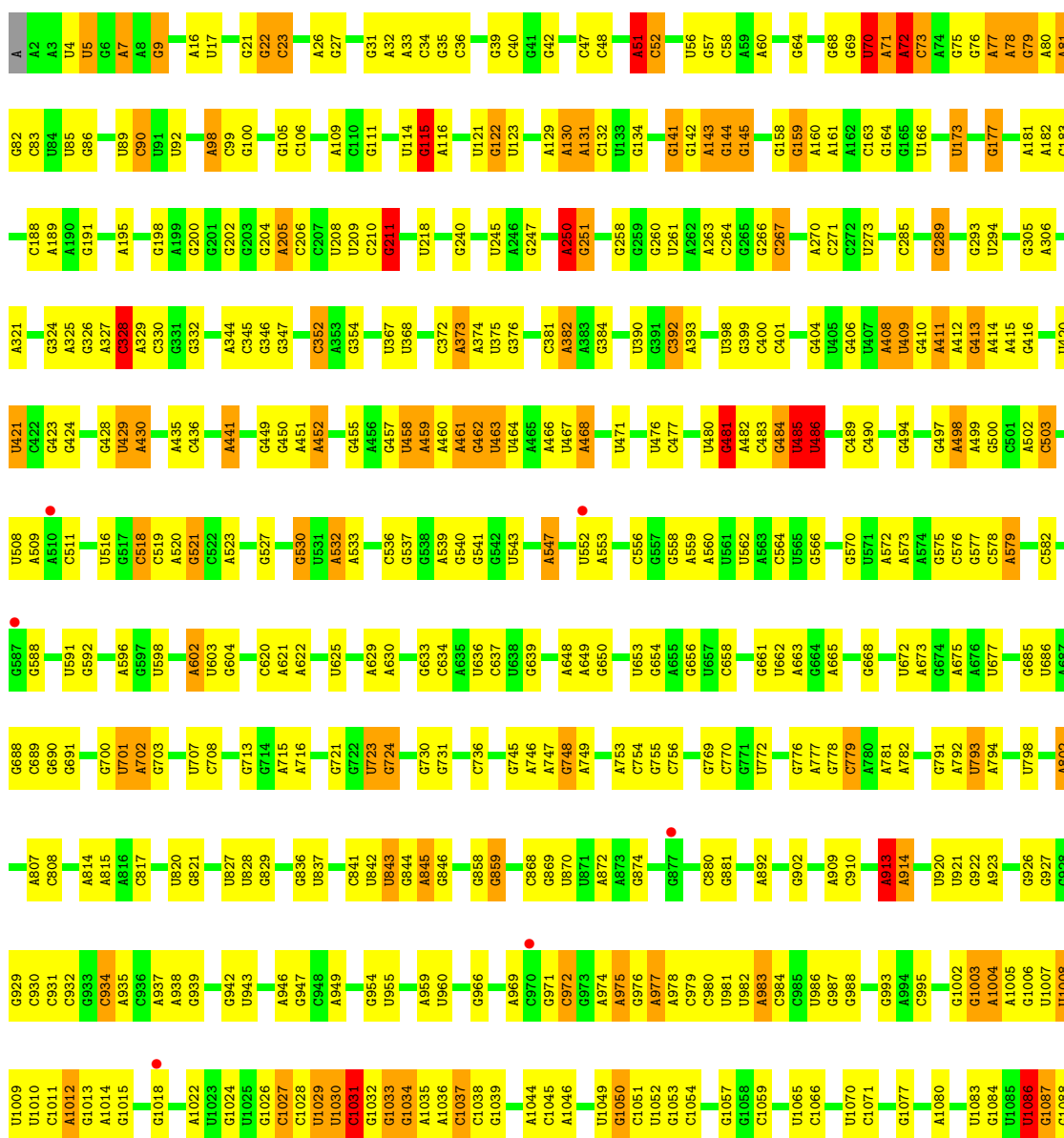


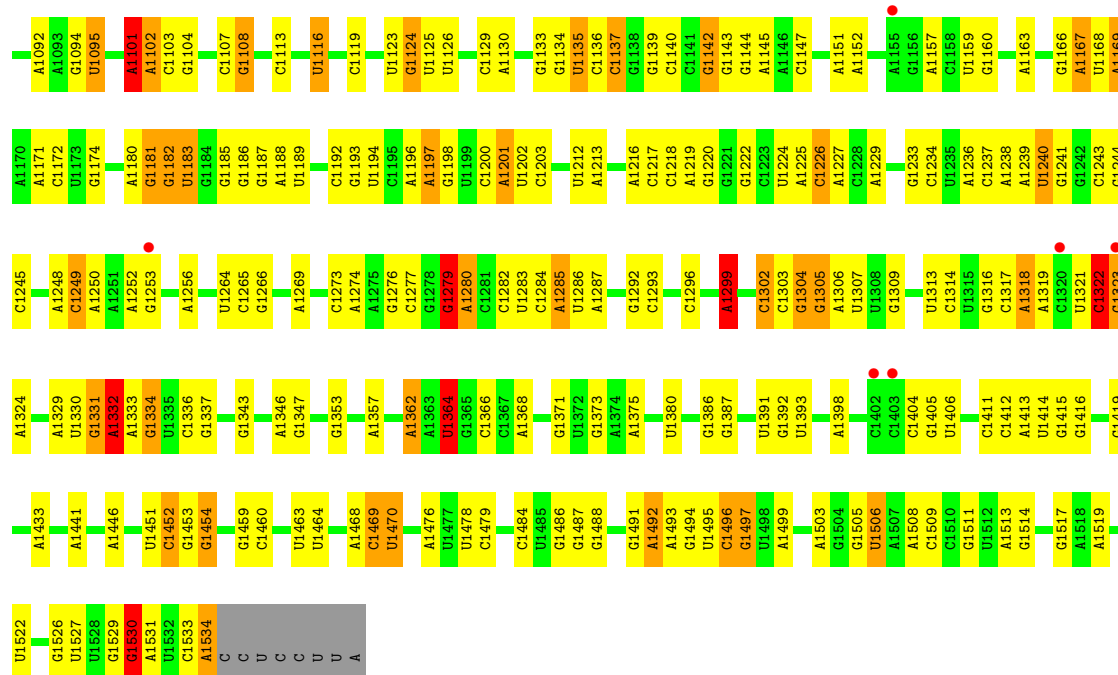
### 3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. 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. A red dot above a residue indicates a poor fit to the electron density ( $\text{RSRZ} > 2$ ). 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: 16S rRNA

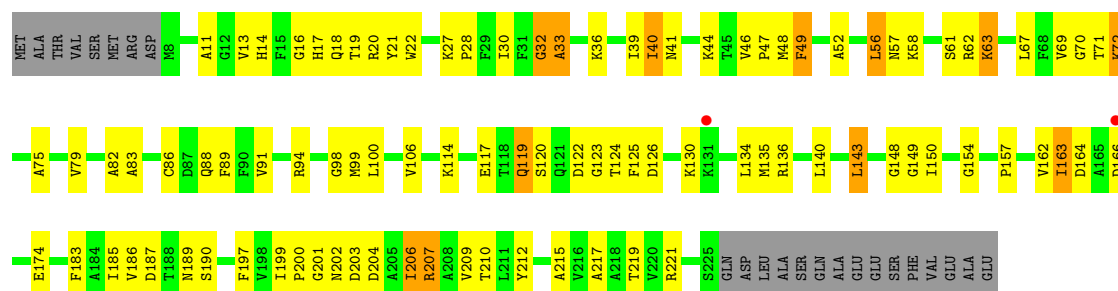
Chain A: 





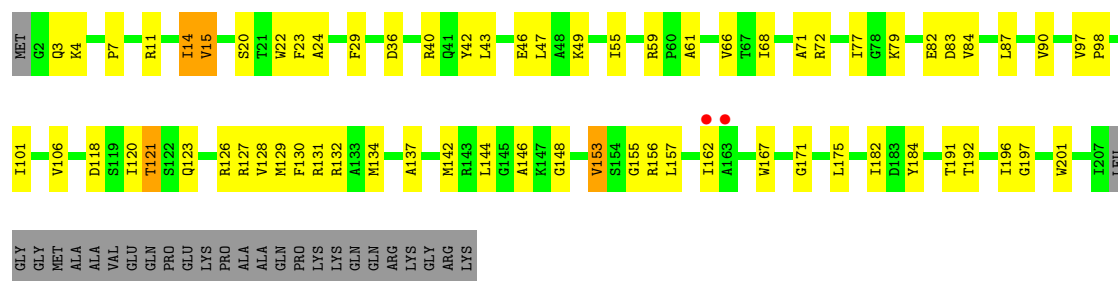
### • Molecule 2: 30S ribosomal protein S2

Chain B:



### • Molecule 3: 30S ribosomal protein S3

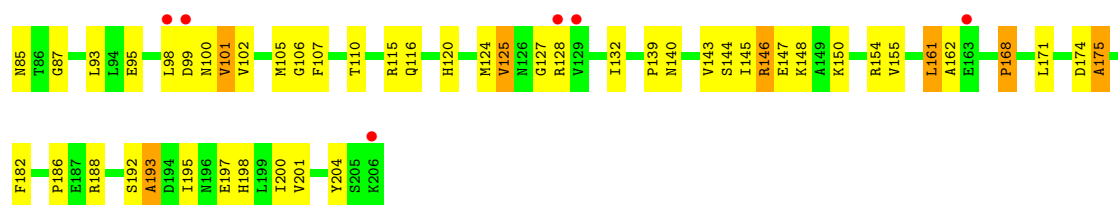
Chain C:



### • Molecule 4: 30S ribosomal protein S4

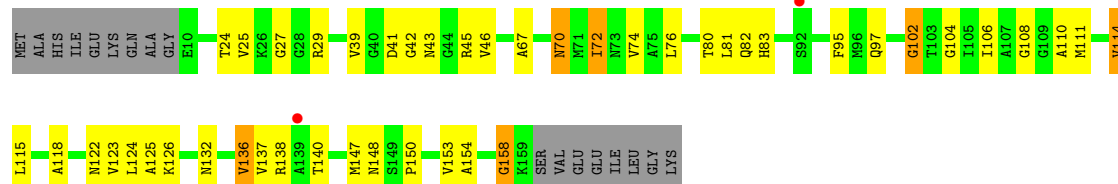
Chain D:





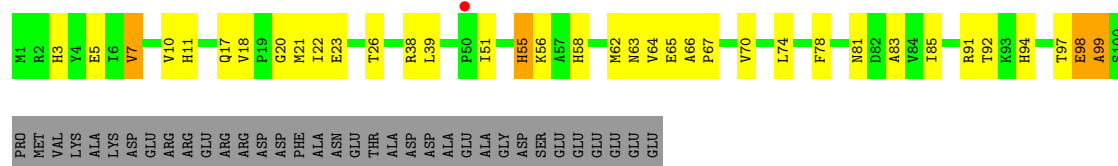
• Molecule 5: 30S ribosomal protein S5

Chain E:



• Molecule 6: 30S ribosomal protein S6

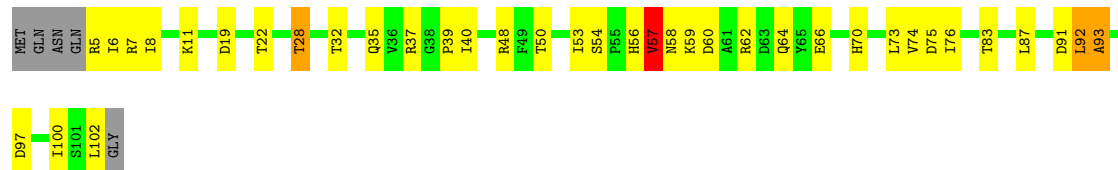
Chain F:





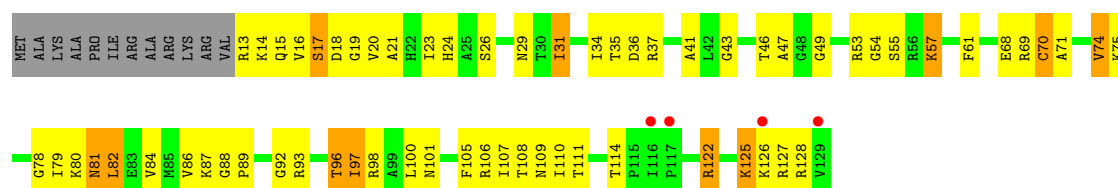
• Molecule 10: 30S ribosomal protein S10

Chain J:



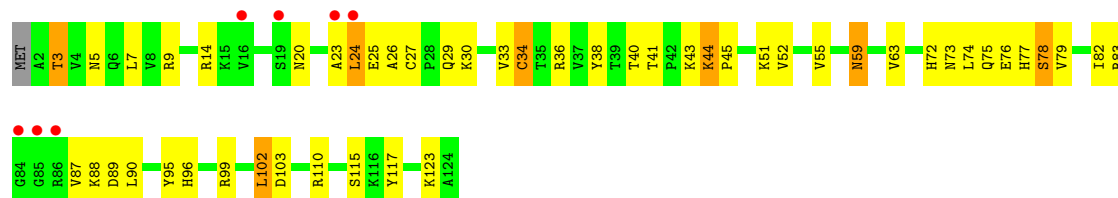
• Molecule 11: 30S ribosomal protein S11

Chain K:



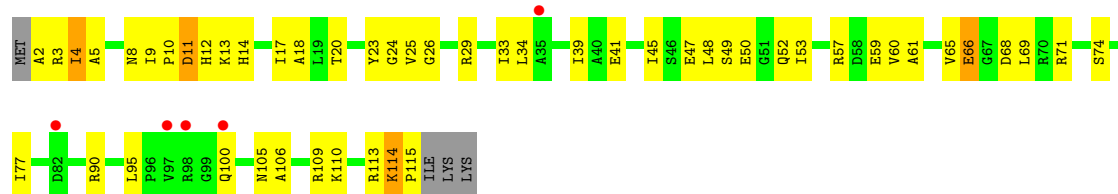
• Molecule 12: 30S ribosomal protein S12

Chain L:



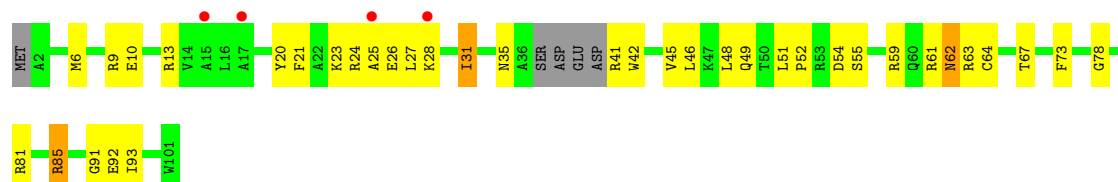
• Molecule 13: 30S ribosomal protein S13

Chain M:



• Molecule 14: 30S ribosomal protein S14

Chain N:



- Molecule 15: 30S ribosomal protein S15

Chain O:



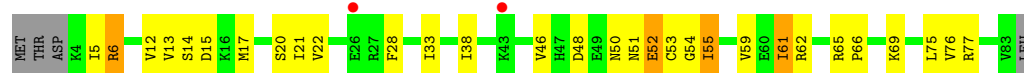
- Molecule 16: 30S ribosomal protein S16

Chain P:



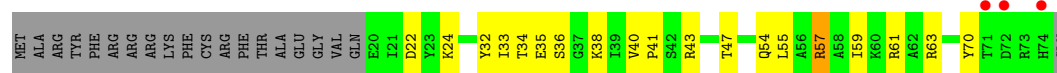
- Molecule 17: 30S ribosomal protein S17

Chain Q:



- Molecule 18: 30S ribosomal protein S18

Chain R:



- Molecule 19: 30S ribosomal protein S19

Chain S:



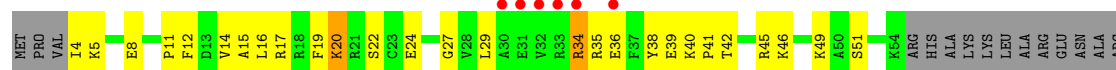
- Molecule 20: 30S ribosomal protein S20

Chain T:

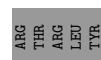


- Molecule 21: 30S ribosomal protein S21

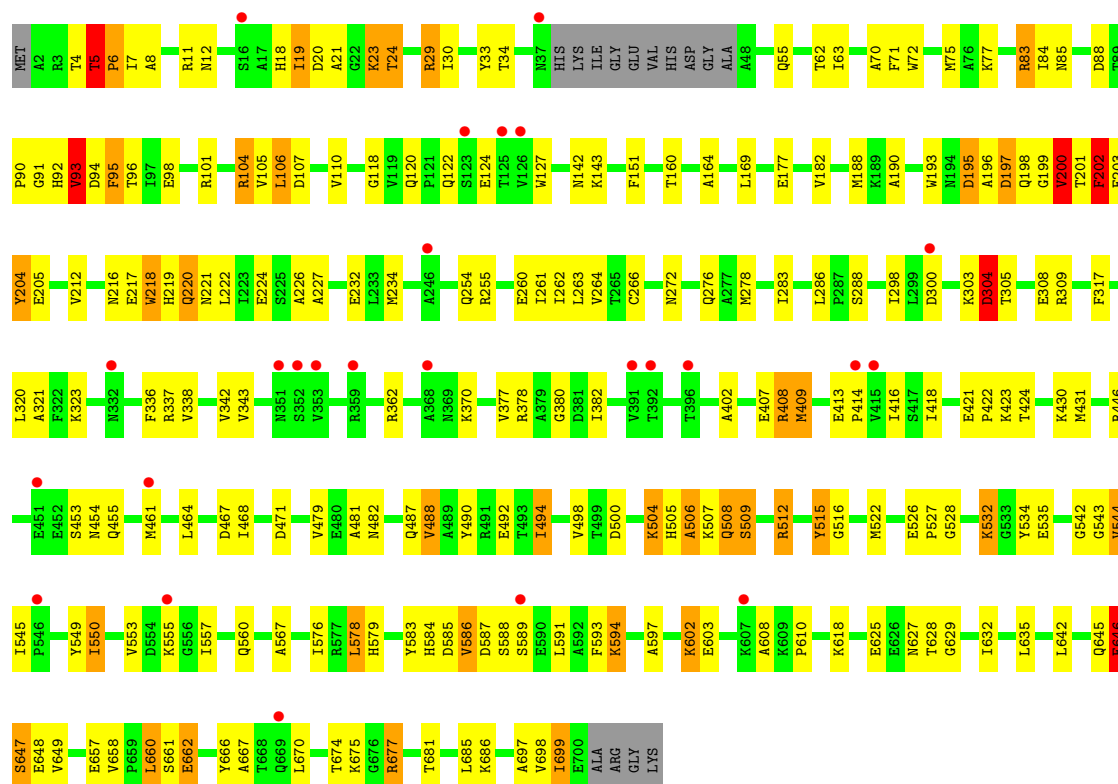
Chain U:



- Molecule 22: elongation factor G



Chain V:



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	361.14Å 360.51Å 429.73Å 90.00° 103.22° 90.00°	Depositor
Resolution (Å)	70.00 – 2.90 69.13 – 2.80	Depositor EDS
% Data completeness (in resolution range)	(Not available) (70.00-2.90) 77.4 (69.13-2.80)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.24 (at 2.81Å)	Xtriage
Refinement program	PHENIX	Depositor
R, $R_{free}$	0.215 , 0.267 0.503 , 0.512	Depositor DCC
$R_{free}$ test set	3890 reflections (0.43%)	DCC
Wilson B-factor (Å <sup>2</sup> )	54.5	Xtriage
Anisotropy	0.324	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.27 , 330.2	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning	$\langle  L  \rangle = 0.45$ , $\langle L^2 \rangle = 0.28$	Xtriage
Outliers	1 of 2019725 reflections (0.000%)	Xtriage
$F_o, F_c$ correlation	0.41	EDS
Total number of atoms	57042	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	45.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 21.58 % of the origin peak, indicating pseudo translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo translational symmetry is equal to 6.8072e-03.*

<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: GCP, 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	0.64	0/36834	1.13	101/57462 (0.2%)
2	B	0.52	0/1735	0.72	0/2338
3	C	0.48	0/1651	0.67	0/2225
4	D	0.52	0/1665	0.73	0/2227
5	E	0.52	0/1118	0.76	0/1504
6	F	0.58	0/835	0.72	0/1128
7	G	0.51	0/1195	0.73	0/1602
8	H	0.45	0/989	0.66	0/1326
9	I	0.56	0/1034	0.76	1/1375 (0.1%)
10	J	0.54	0/796	0.77	0/1077
11	K	0.71	0/893	0.92	2/1205 (0.2%)
12	L	0.50	0/969	0.78	0/1300
13	M	0.60	0/892	0.83	0/1193
14	N	0.48	0/785	0.67	0/1043
15	O	0.50	0/722	0.68	0/964
16	P	0.51	0/659	0.67	0/884
17	Q	0.51	0/657	0.75	0/881
18	R	0.53	0/462	0.77	1/621 (0.2%)
19	S	0.66	0/652	0.79	0/877
20	T	0.49	0/671	0.72	0/888
21	U	0.78	0/430	0.82	0/570
22	V	0.50	0/5418	0.70	1/7329 (0.0%)
All	All	0.60	0/61062	1.01	106/90019 (0.1%)

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
12	L	0	1

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	#Chirality outliers	#Planarity outliers
22	V	0	3
All	All	0	4

There are no bond length outliers.

The worst 5 of 106 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	5	U	C2-N1-C1'	9.96	129.65	117.70
1	A	1087	G	C6-C5-N7	-8.31	125.42	130.40
1	A	481	G	C5-C6-O6	-8.14	123.71	128.60
1	A	1087	G	N7-C8-N9	7.84	117.02	113.10
1	A	1086	U	N3-C2-O2	-7.79	116.75	122.20

There are no chirality outliers.

All (4) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
12	L	23	ALA	Peptide
22	V	218	TRP	Peptide
22	V	304	ASP	Peptide
22	V	588	SER	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	A	32895	0	16553	425	0
2	B	1704	0	1732	72	0
3	C	1624	0	1696	48	0
4	D	1643	0	1707	72	0
5	E	1105	0	1148	37	0
6	F	817	0	808	25	0
7	G	1181	0	1238	33	0
8	H	979	0	1031	28	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
9	I	1022	0	1070	59	0
10	J	786	0	828	34	0
11	K	877	0	887	75	0
12	L	955	0	1016	39	0
13	M	883	0	941	52	0
14	N	774	0	824	29	0
15	O	714	0	734	21	0
16	P	649	0	666	19	0
17	Q	648	0	691	21	0
18	R	455	0	478	15	0
19	S	637	0	665	20	0
20	T	665	0	714	19	0
21	U	425	0	449	36	0
22	V	5319	0	5227	128	0
23	A	40	0	0	0	0
23	C	1	0	0	0	0
23	E	1	0	0	0	0
23	T	1	0	0	0	0
23	V	1	0	0	0	0
24	V	32	0	14	1	0
25	A	198	0	0	33	0
25	D	1	0	0	0	0
25	E	2	0	0	0	0
25	N	5	0	0	0	0
25	T	1	0	0	0	0
25	U	1	0	0	0	0
25	V	1	0	0	1	0
All	All	57042	0	41117	1188	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 12.

The worst 5 of 1188 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
22:V:55:GLN:NE2	22:V:471:ASP:OD2	2.01	0.94
11:K:127:ARG:O	21:U:34:ARG:NH1	2.08	0.86
1:A:946:A:HO2'	1:A:1333:A:HO2'	1.10	0.86
1:A:770:C:N4	25:A:1755:HOH:O	2.09	0.86
2:B:57:ASN:ND2	2:B:219:THR:O	2.11	0.84

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	
2	B	216/241 (90%)	149 (69%)	55 (26%)	12 (6%)	3	8
3	C	204/233 (88%)	183 (90%)	16 (8%)	5 (2%)	9	32
4	D	203/206 (98%)	159 (78%)	35 (17%)	9 (4%)	4	15
5	E	148/167 (89%)	122 (82%)	22 (15%)	4 (3%)	8	30
6	F	98/135 (73%)	75 (76%)	18 (18%)	5 (5%)	3	10
7	G	149/179 (83%)	122 (82%)	25 (17%)	2 (1%)	18	54
8	H	127/130 (98%)	111 (87%)	16 (13%)	0	100	100
9	I	125/130 (96%)	98 (78%)	21 (17%)	6 (5%)	4	12
10	J	96/103 (93%)	73 (76%)	19 (20%)	4 (4%)	4	16
11	K	115/129 (89%)	87 (76%)	24 (21%)	4 (4%)	6	23
12	L	121/124 (98%)	97 (80%)	15 (12%)	9 (7%)	2	4
13	M	112/118 (95%)	91 (81%)	14 (12%)	7 (6%)	2	6
14	N	92/101 (91%)	70 (76%)	20 (22%)	2 (2%)	10	37
15	O	86/89 (97%)	70 (81%)	14 (16%)	2 (2%)	10	36
16	P	80/82 (98%)	61 (76%)	15 (19%)	4 (5%)	3	11
17	Q	78/84 (93%)	57 (73%)	16 (20%)	5 (6%)	2	6
18	R	53/75 (71%)	43 (81%)	10 (19%)	0	100	100
19	S	77/92 (84%)	68 (88%)	9 (12%)	0	100	100
20	T	83/87 (95%)	68 (82%)	14 (17%)	1 (1%)	19	57
21	U	49/71 (69%)	24 (49%)	22 (45%)	3 (6%)	2	7
22	V	685/704 (97%)	556 (81%)	91 (13%)	38 (6%)	3	8
All	All	2997/3280 (91%)	2384 (80%)	491 (16%)	122 (4%)	4	17

5 of 122 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	40	ILE

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
2	B	72	LYS
2	B	119	GLN
3	C	101	ILE
4	D	25	VAL

### 5.3.2 Protein sidechains ⓘ

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all 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	
2	B	180/199 (90%)	170 (94%)	10 (6%)	30	66
3	C	170/190 (90%)	164 (96%)	6 (4%)	48	85
4	D	172/173 (99%)	163 (95%)	9 (5%)	32	71
5	E	113/126 (90%)	106 (94%)	7 (6%)	26	61
6	F	87/116 (75%)	85 (98%)	2 (2%)	63	92
7	G	124/147 (84%)	123 (99%)	1 (1%)	89	97
8	H	104/105 (99%)	98 (94%)	6 (6%)	28	65
9	I	105/107 (98%)	99 (94%)	6 (6%)	29	66
10	J	86/90 (96%)	83 (96%)	3 (4%)	48	85
11	K	90/99 (91%)	82 (91%)	8 (9%)	14	40
12	L	103/104 (99%)	94 (91%)	9 (9%)	15	41
13	M	92/96 (96%)	92 (100%)	0	100	100
14	N	79/84 (94%)	77 (98%)	2 (2%)	60	91
15	O	76/77 (99%)	74 (97%)	2 (3%)	59	90
16	P	65/65 (100%)	63 (97%)	2 (3%)	52	88
17	Q	74/78 (95%)	69 (93%)	5 (7%)	22	55
18	R	48/65 (74%)	48 (100%)	0	100	100
19	S	70/79 (89%)	67 (96%)	3 (4%)	40	78
20	T	65/66 (98%)	60 (92%)	5 (8%)	18	47
21	U	44/61 (72%)	42 (96%)	2 (4%)	38	77
22	V	557/578 (96%)	507 (91%)	50 (9%)	14	39

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
All	All	2504/2705 (93%)	2366 (94%)	138 (6%)	30 68

5 of 138 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
12	L	33	VAL
17	Q	61	ILE
22	V	578	LEU
12	L	78	SER
15	O	87	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 18 such sidechains are listed below:

Mol	Chain	Res	Type
12	L	59	ASN
12	L	96	HIS
22	V	55	GLN
11	K	81	ASN
11	K	109	ASN

### 5.3.3 RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
1	A	1532/1542 (99%)	273 (17%)	18 (1%)

5 of 273 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
1	A	4	U
1	A	5	U
1	A	7	A
1	A	9	G
1	A	32	A

5 of 18 RNA pucker outliers are listed below:

Mol	Chain	Res	Type
1	A	913	A
1	A	1030	U
1	A	1331	G

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	499	A
1	A	701	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 45 ligands modelled in this entry, 44 are monoatomic - leaving 1 for Mogul analysis.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the chemical component dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with  $|Z| > 2$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z  > 2$	Counts	RMSZ	$\# Z  > 2$
24	GCP	V	801	23	34,34,34	2.46	9 (26%)	51,54,54	2.38	9 (17%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the chemical component dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
24	GCP	V	801	23	-	0/20/38/38	0/1/3/3

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
24	V	801	GCP	C2-N2	8.18	1.45	1.32
24	V	801	GCP	C2'-C1'	-5.14	1.46	1.53

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
24	V	801	GCP	PG-C3B	4.71	1.83	1.79
24	V	801	GCP	O4'-C1'	4.57	1.48	1.41
24	V	801	GCP	PB-C3B	4.17	1.83	1.79

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
24	V	801	GCP	C6-C5-N7	-8.47	133.00	134.14
24	V	801	GCP	C4'-O4'-C1'	-6.18	103.03	109.75
24	V	801	GCP	N3-C4-N9	5.19	134.53	126.91
24	V	801	GCP	PB-C3B-PG	-4.89	110.58	117.62
24	V	801	GCP	PA-O3A-PB	-4.52	118.16	131.74

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	A	1533/1542 (99%)	-0.16	12 (0%) 83 89	15, 41, 72, 90	0
2	B	218/241 (90%)	-0.72	2 (0%) 81 88	33, 51, 67, 73	0
3	C	206/233 (88%)	-0.64	2 (0%) 79 86	21, 40, 56, 66	0
4	D	205/206 (99%)	0.04	18 (8%) 10 13	25, 47, 61, 79	0
5	E	150/167 (89%)	-0.46	2 (1%) 74 82	21, 37, 54, 67	0
6	F	100/135 (74%)	-0.60	1 (1%) 79 86	31, 51, 64, 71	0
7	G	151/179 (84%)	-0.46	3 (1%) 62 71	32, 48, 62, 68	0
8	H	129/130 (99%)	-0.55	2 (1%) 68 78	23, 35, 51, 69	0
9	I	127/130 (97%)	-0.39	1 (0%) 83 89	29, 55, 70, 75	0
10	J	98/103 (95%)	-0.58	0 100 100	30, 51, 69, 78	0
11	K	117/129 (90%)	-0.44	4 (3%) 43 51	23, 50, 70, 74	0
12	L	123/124 (99%)	-0.24	7 (5%) 23 28	21, 37, 55, 63	0
13	M	114/118 (96%)	-0.36	5 (4%) 33 40	42, 63, 72, 75	0
14	N	96/101 (95%)	-0.16	4 (4%) 35 41	27, 47, 64, 70	0
15	O	88/89 (98%)	-0.49	2 (2%) 57 66	22, 39, 53, 70	0
16	P	82/82 (100%)	-0.57	1 (1%) 75 83	22, 39, 62, 71	0
17	Q	80/84 (95%)	-0.36	2 (2%) 54 64	27, 41, 51, 57	0
18	R	55/75 (73%)	-0.28	3 (5%) 24 29	26, 40, 55, 64	0
19	S	79/92 (85%)	-0.36	4 (5%) 27 33	43, 58, 70, 74	0
20	T	85/87 (97%)	0.08	7 (8%) 12 15	23, 38, 53, 56	0
21	U	51/71 (71%)	0.24	6 (11%) 5 7	40, 57, 67, 71	0
22	V	689/704 (97%)	-0.32	25 (3%) 41 48	32, 56, 70, 80	0
All	All	4576/4822 (94%)	-0.31	113 (2%) 56 64	15, 46, 69, 90	0

The worst 5 of 113 RSRZ outliers are listed below:



Mol	Chain	Res	Type	RSRZ
4	D	2	ALA	12.0
7	G	83	SER	5.4
4	D	77	LYS	5.1
7	G	81	GLY	4.9
13	M	100	GLN	4.8

## 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
23	MG	A	1601	1/1	0.48	28.43	38,38,38,38	0
23	MG	A	1621	1/1	0.32	19.00	29,29,29,29	0
23	MG	A	1609	1/1	0.21	7.35	29,29,29,29	0
23	MG	A	1605	1/1	0.26	6.63	39,39,39,39	0
23	MG	A	1634	1/1	0.23	5.46	24,24,24,24	0
23	MG	A	1613	1/1	0.16	4.23	41,41,41,41	0
23	MG	T	101	1/1	0.30	4.18	41,41,41,41	0
23	MG	A	1604	1/1	0.28	2.18	27,27,27,27	0
23	MG	A	1633	1/1	0.28	2.05	52,52,52,52	0
23	MG	A	1629	1/1	0.39	1.78	17,17,17,17	0
23	MG	A	1603	1/1	0.13	1.62	38,38,38,38	0
23	MG	A	1627	1/1	0.17	1.54	24,24,24,24	0
23	MG	A	1640	1/1	0.21	1.52	15,15,15,15	0
23	MG	A	1639	1/1	0.18	1.34	20,20,20,20	0
23	MG	A	1628	1/1	0.15	1.04	30,30,30,30	0
23	MG	A	1619	1/1	0.14	0.65	34,34,34,34	0
23	MG	A	1638	1/1	0.15	0.07	26,26,26,26	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
23	MG	A	1622	1/1	0.15	0.00	36,36,36,36	0
23	MG	A	1632	1/1	0.13	-0.19	25,25,25,25	0
23	MG	A	1626	1/1	0.13	-0.37	32,32,32,32	0
23	MG	A	1612	1/1	0.11	-0.38	8,8,8,8	0
24	GCP	V	801	32/32	0.11	-0.62	24,48,54,56	0
23	MG	A	1625	1/1	0.13	-0.86	22,22,22,22	0
23	MG	A	1635	1/1	0.10	-1.00	17,17,17,17	0
23	MG	A	1606	1/1	0.13	-1.00	17,17,17,17	0
23	MG	A	1637	1/1	0.10	-1.08	47,47,47,47	0
23	MG	A	1610	1/1	0.09	-1.27	13,13,13,13	0
23	MG	A	1611	1/1	0.10	-1.33	20,20,20,20	0
23	MG	A	1602	1/1	0.12	-1.48	48,48,48,48	0
23	MG	A	1608	1/1	0.11	-1.88	14,14,14,14	0
23	MG	A	1617	1/1	0.11	-1.94	62,62,62,62	0
23	MG	A	1623	1/1	0.10	-2.05	18,18,18,18	0
23	MG	A	1616	1/1	0.10	-2.45	33,33,33,33	0
23	MG	A	1630	1/1	0.09	-2.46	40,40,40,40	0
23	MG	A	1631	1/1	0.08	-3.42	28,28,28,28	0
23	MG	A	1615	1/1	0.09	-3.43	38,38,38,38	0
23	MG	A	1607	1/1	0.10	-4.29	11,11,11,11	0
23	MG	V	802	1/1	0.05	-4.31	38,38,38,38	0
23	MG	E	201	1/1	0.09	-4.82	39,39,39,39	0
23	MG	A	1636	1/1	0.05	-5.89	21,21,21,21	0
23	MG	A	1618	1/1	0.05	-6.03	19,19,19,19	0
23	MG	A	1614	1/1	0.07	-12.27	37,37,37,37	0
23	MG	A	1624	1/1	0.06	-16.09	14,14,14,14	0
23	MG	C	401	1/1	0.69	-	49,49,49,49	0
23	MG	A	1620	1/1	0.03	-	14,14,14,14	0

## 6.5 Other polymers ⓘ

There are no such residues in this entry.