



# wwPDB X-ray Structure Validation Summary Report

Mar 1, 2014 – 04:40 AM GMT

PDB ID : 3DF3

Title : Crystal structure of the bacterial ribosome from Escherichia coli in complex with hygromycin B. This file contains the 30S subunit of the second 70S ribosome, with hygromycin B bound. The entire crystal structure contains two 70S ribosomes.

Authors : Borovinskaya, M.A.; Shoji, S.; Fredrick, K.; Cate, J.H.D.

Deposited on : 2008-06-11

Resolution : 3.50 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>

---

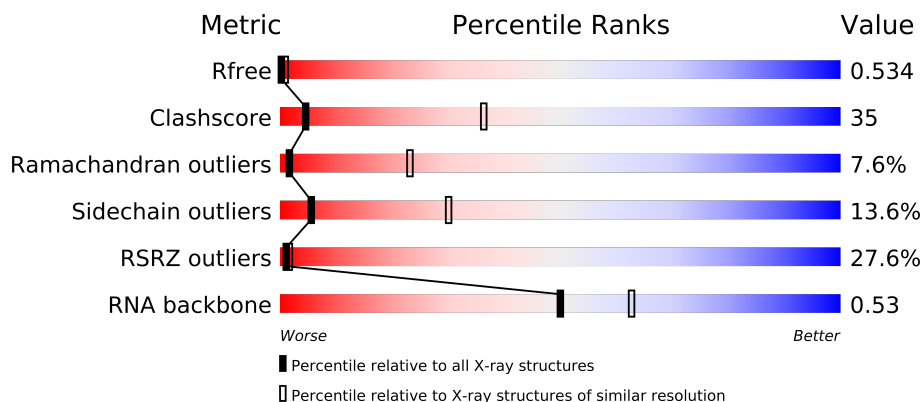
The following versions of software and data (see [references](#)) were used in the production of this report:

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 3.50 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.



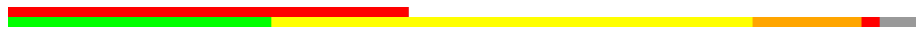



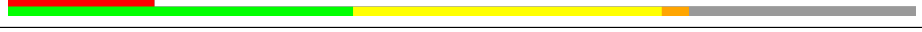
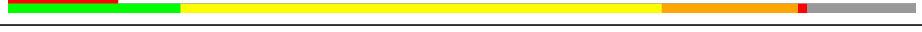



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	66092	1243 (3.70-3.30)
Clashscore	79885	1039 (3.66-3.34)
Ramachandran outliers	78287	1000 (3.66-3.34)
Sidechain outliers	78261	1000 (3.66-3.34)
RSRZ outliers	66119	1243 (3.70-3.30)
RNA backbone	1838	1007 (4.22-2.76)

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	C	232	
3	D	205	
4	E	166	
5	F	135	
6	G	178	
7	H	129	
8	I	129	
9	J	103	
10	K	128	
11	L	123	
12	M	117	

Continued on next page...

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
13	N	100	
14	O	89	
15	P	82	
16	Q	83	
17	R	74	
18	S	91	
19	T	86	
20	B	240	
21	U	71	

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

Mol	Type	Chain	Res	Geometry	Electron density
22	MG	A	2042	-	X
22	MG	A	2062	-	X
22	MG	A	2168	-	X
22	MG	A	2182	-	X
22	MG	A	2346	-	X
22	MG	A	2360	-	X

## 2 Entry composition

There are 24 unique types of molecules in this entry. The entry contains 51794 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 ribosomal RNA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	1530	Total	C	N	O	P	0	0	0
			32831	14642	6024	10635	1530			

- Molecule 2 is a protein called 30S ribosomal protein S3.

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

- Molecule 3 is a protein called 30S ribosomal protein S4.

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

- Molecule 4 is a protein called 30S ribosomal protein S5.

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

- Molecule 5 is a protein called 30S ribosomal protein S6.

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

- Molecule 6 is a protein called 30S ribosomal protein S7.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
6	G	152	Total	C	N	O	S	0	0	0
			1196	745	230	217	4			

- Molecule 7 is a protein called 30S ribosomal protein S8.

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

- Molecule 8 is a protein called 30S ribosomal protein S9.

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

- Molecule 9 is a protein called 30S ribosomal protein S10.

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

- Molecule 10 is a protein called 30S ribosomal protein S11.

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

- Molecule 11 is a protein called 30S ribosomal protein S12.

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

- Molecule 12 is a protein called 30S ribosomal protein S13.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
12	M	113	Total	C	N	O	S	0	0	0
			876	541	177	155	3			

- Molecule 13 is a protein called 30S ribosomal protein S14.

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

- Molecule 14 is a protein called 30S ribosomal protein S15.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
14	O	88	Total	C	N	O	S	0	0	0
			715	440	146	128	1			

- Molecule 15 is a protein called 30S ribosomal protein S16.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
15	P	80	Total	C	N	O	S	0	0	0
			638	400	126	111	1			

- Molecule 16 is a protein called 30S ribosomal protein S17.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
16	Q	81	Total	C	N	O	S	0	0	0
			656	417	122	114	3			

- Molecule 17 is a protein called 30S ribosomal protein S18.

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

- Molecule 18 is a protein called 30S ribosomal protein S19.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
18	S	80	Total	C	N	O	S	0	0	0
			644	413	121	108	2			

- Molecule 19 is a protein called 30S ribosomal protein S20.

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

- Molecule 20 is a protein called 30S ribosomal protein S2.

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

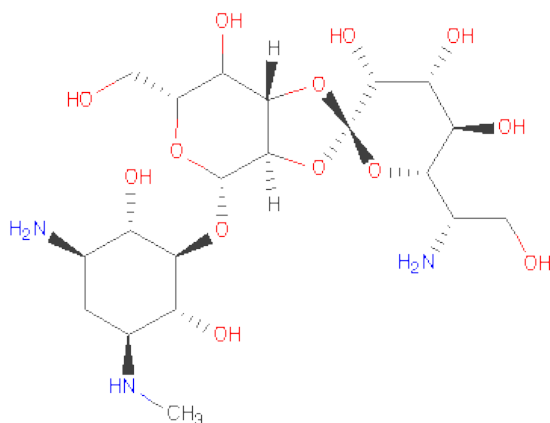
- Molecule 21 is a protein called 30S ribosomal protein S21.

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

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
22	A	61	Total	Mg	0	0
			61	61		
22	E	1	Total	Mg	0	0
			1	1		

- Molecule 23 is HYGROMYCIN B (three-letter code: HYG) (formula: C<sub>20</sub>H<sub>37</sub>N<sub>3</sub>O<sub>13</sub>).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
23	A	1	Total	C	N	O	0	0
			36	20	3	13		

- Molecule 24 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
24	A	299	Total	O	0	0
			299	299		
24	E	5	Total	O	0	0
			5	5		
24	L	1	Total	O	0	0
			1	1		

*Continued on next page...*

*Continued from previous page...*

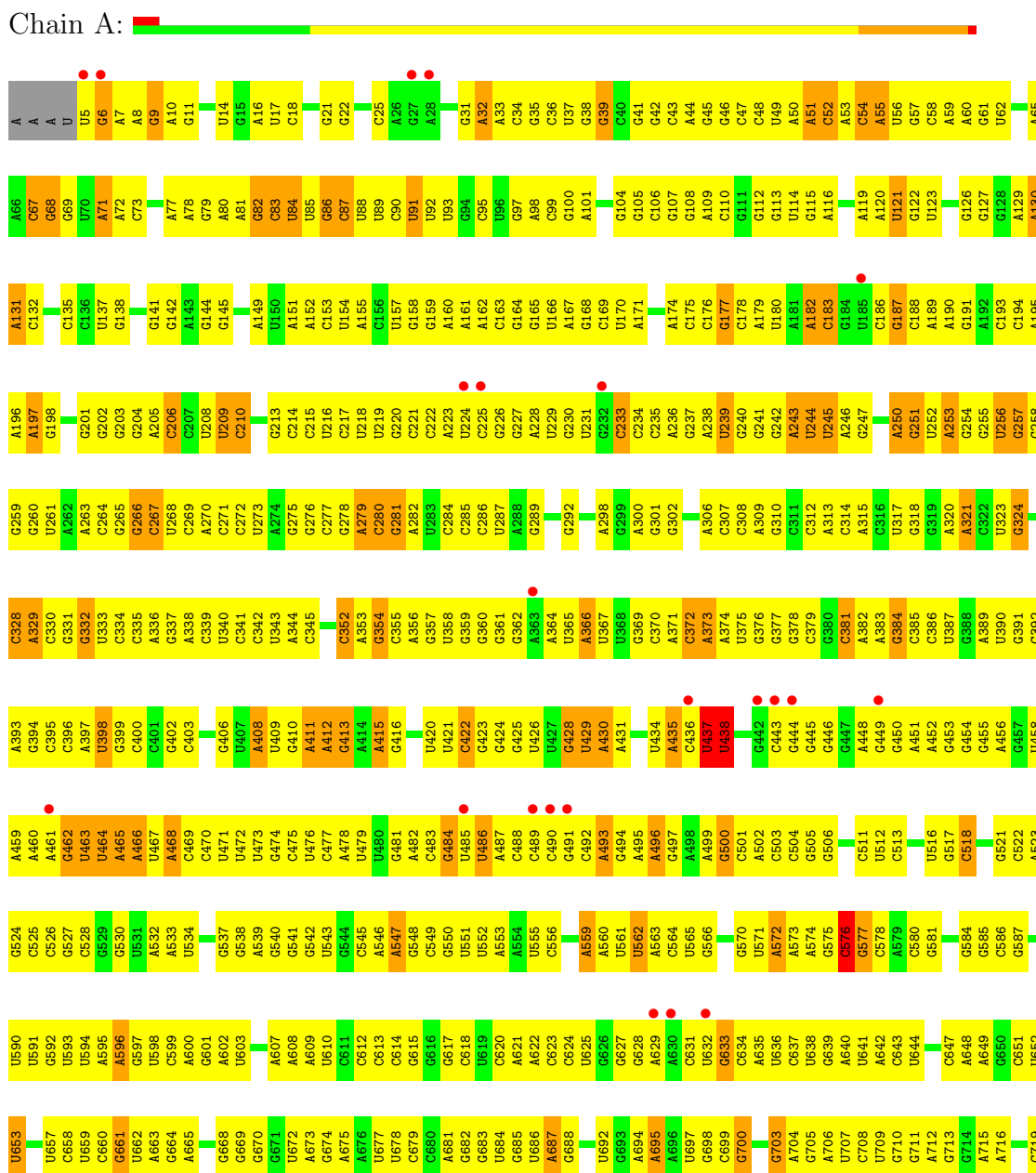
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
24	N	4	Total	O	0	0
			4	4		

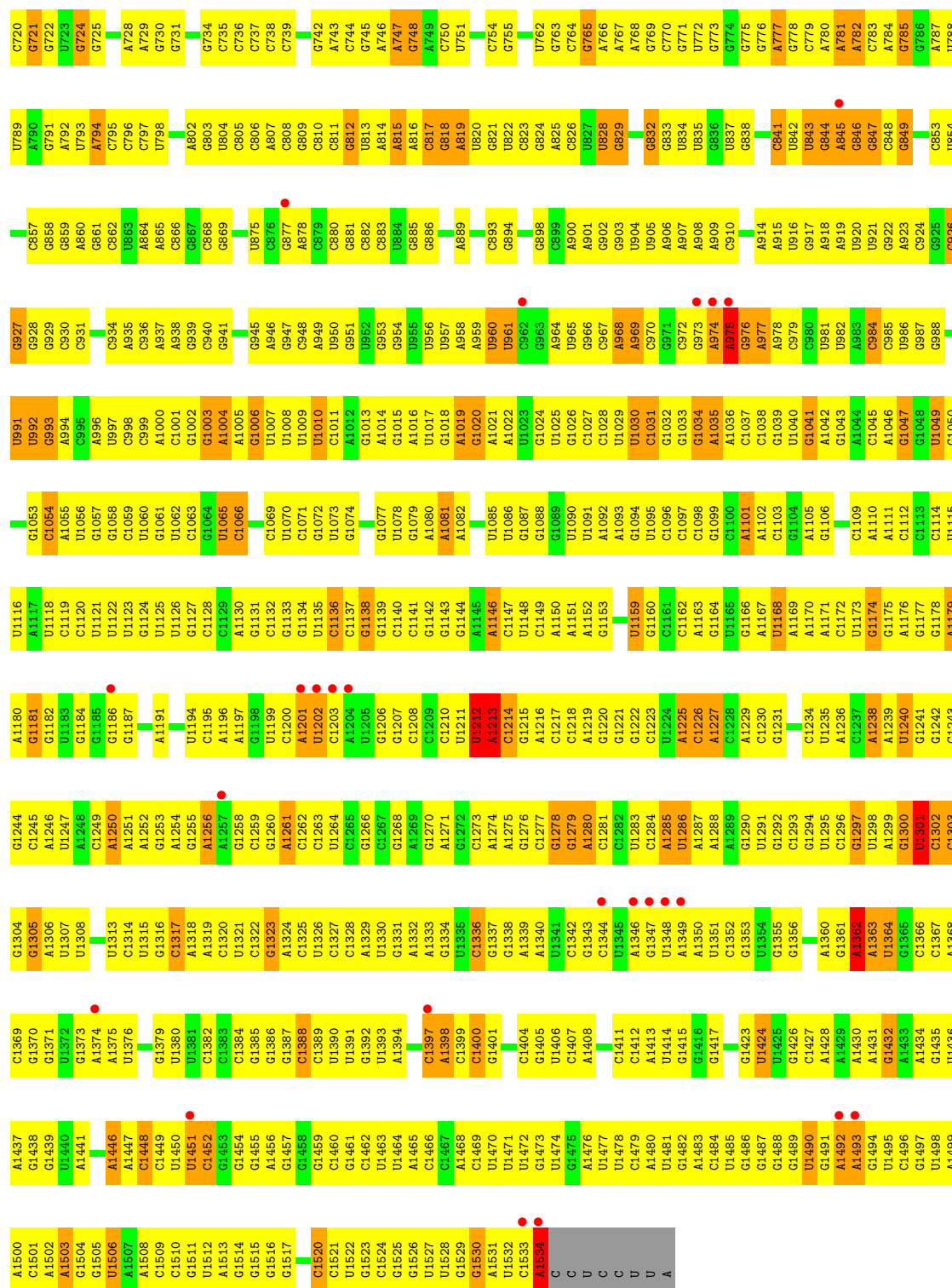


### 3 Residue-property plots

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 ( $RSRZ > 2$ ). 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.

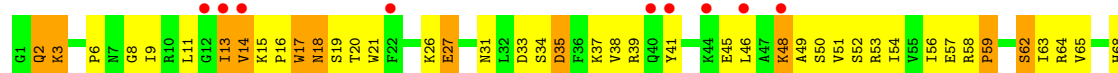
#### • Molecule 1: 16S ribosomal RNA

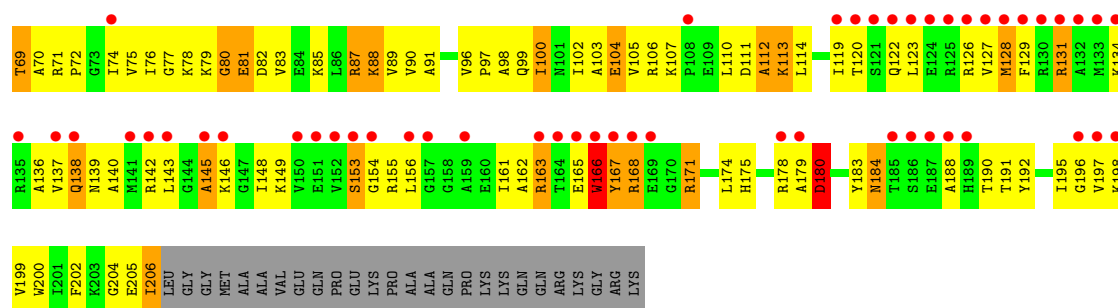




• Molecule 2: 30S ribosomal protein S3

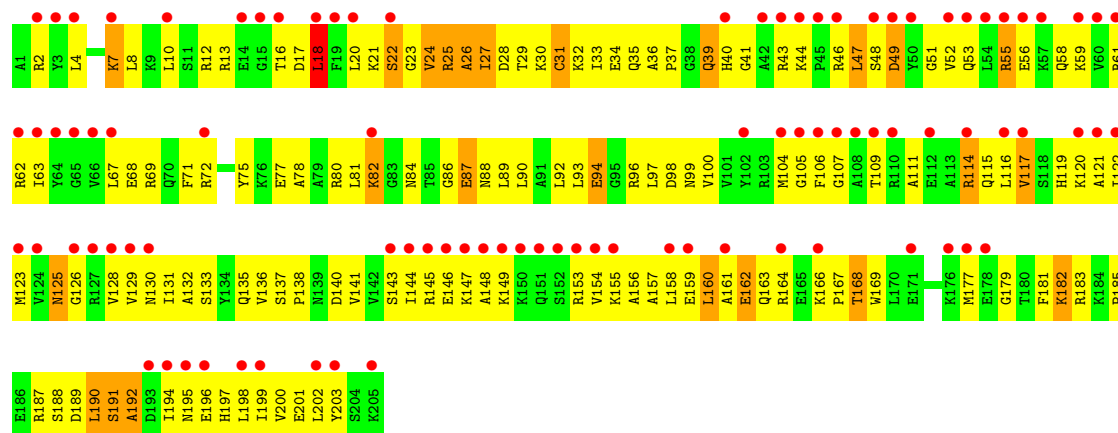
Chain C:





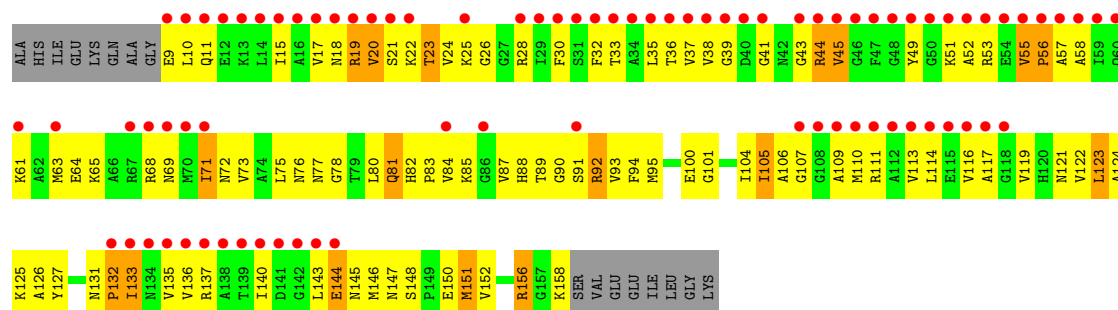
• Molecule 3: 30S ribosomal protein S4

Chain D:



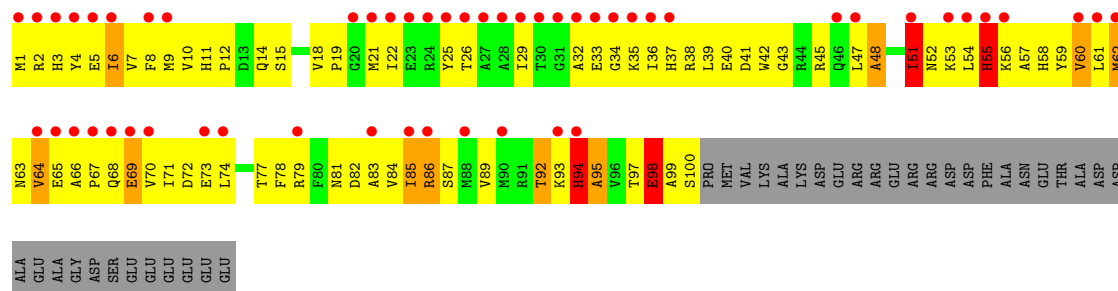
• Molecule 4: 30S ribosomal protein S5

Chain E:

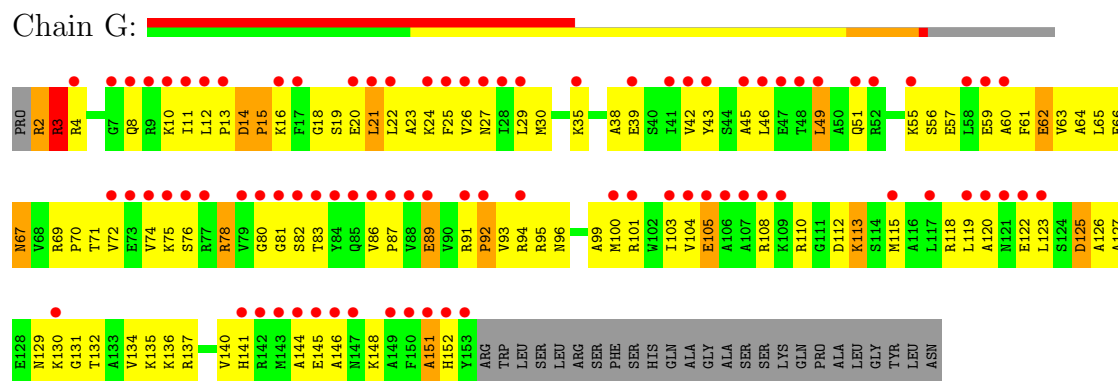


• Molecule 5: 30S ribosomal protein S6

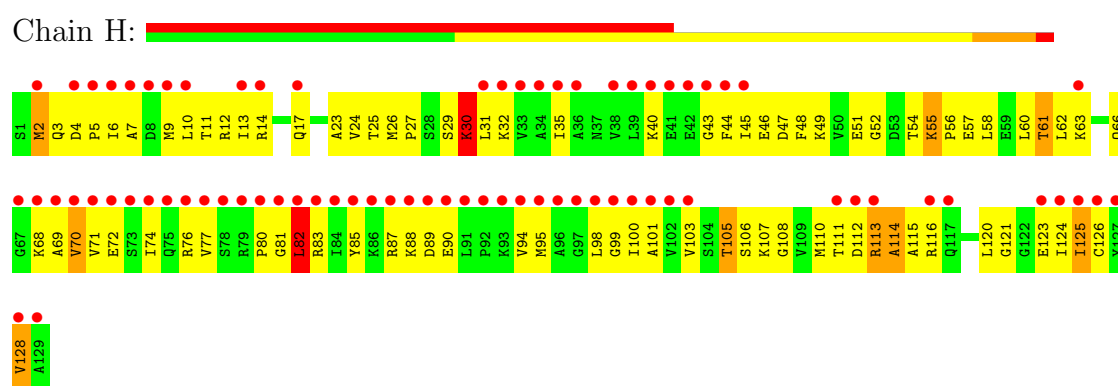
Chain F:



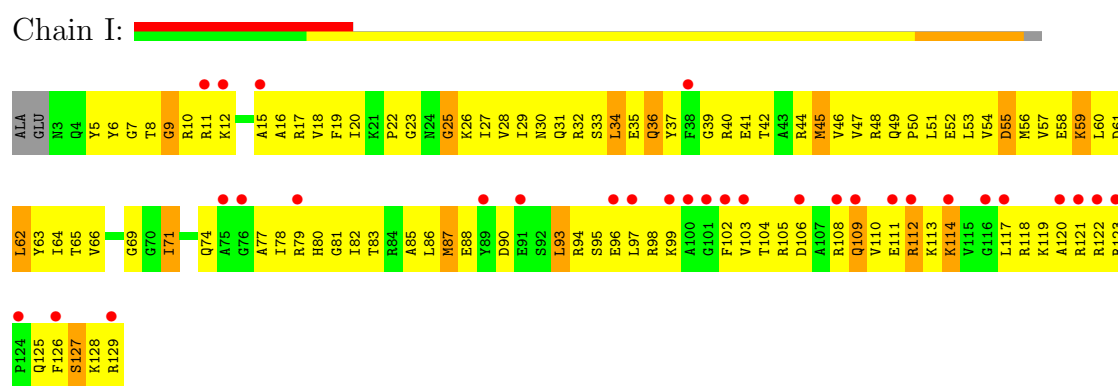
- Molecule 6: 30S ribosomal protein S7



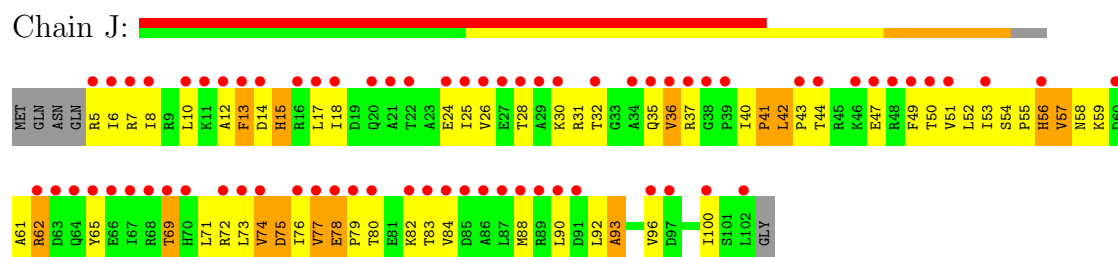
- Molecule 7: 30S ribosomal protein S8



- Molecule 8: 30S ribosomal protein S9

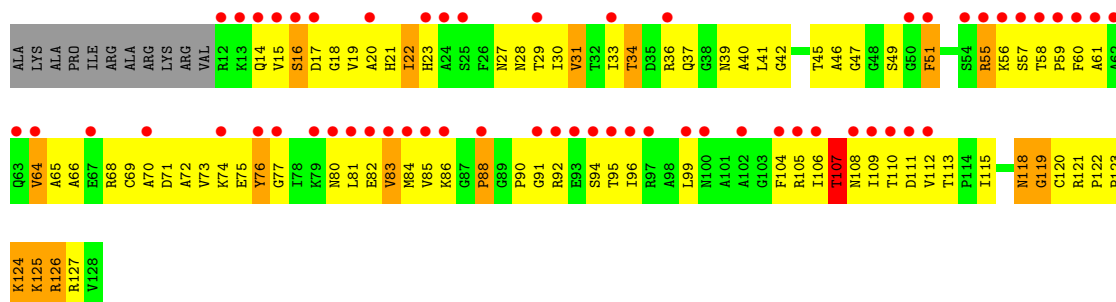


- Molecule 9: 30S ribosomal protein S10



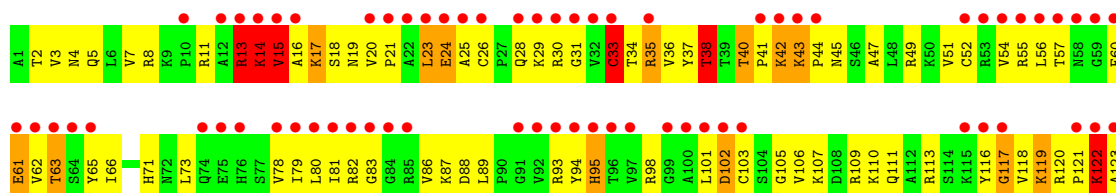
- Molecule 10: 30S ribosomal protein S11

Chain K: 



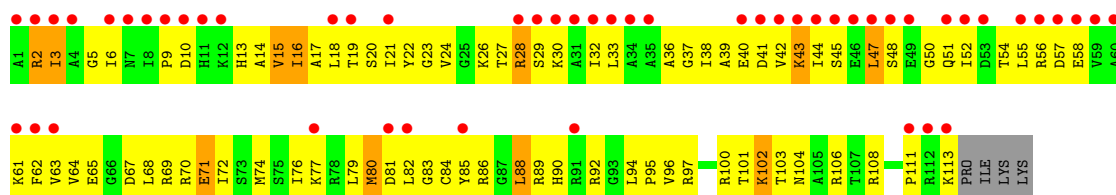
- Molecule 11: 30S ribosomal protein S12

Chain L: 



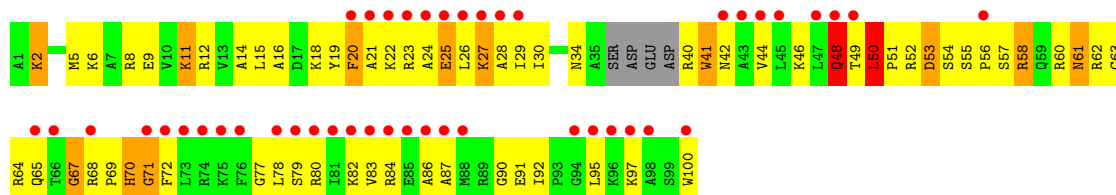
- Molecule 12: 30S ribosomal protein S13

Chain M: 



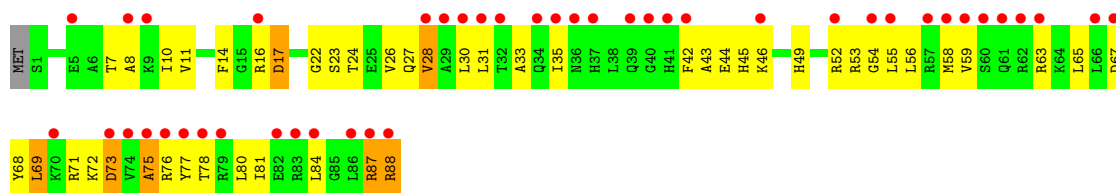
- Molecule 13: 30S ribosomal protein S14

Chain N: 



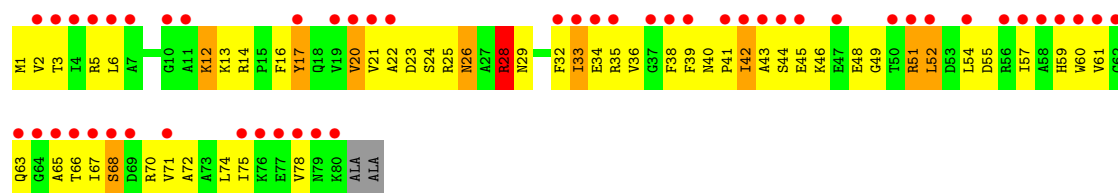
- Molecule 14: 30S ribosomal protein S15

Chain O: 



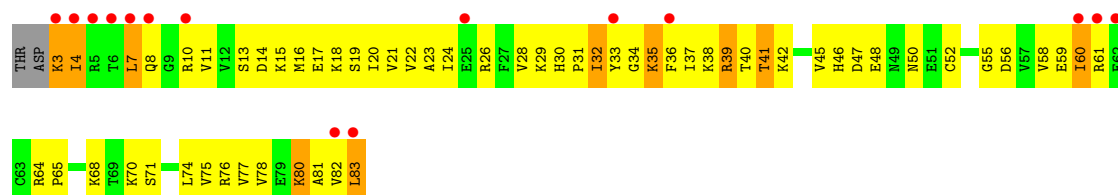
- Molecule 15: 30S ribosomal protein S16

Chain P:



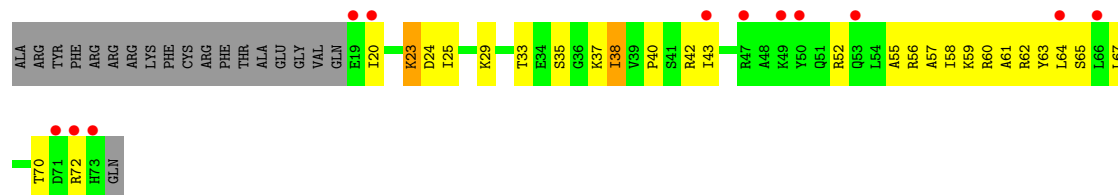
- Molecule 16: 30S ribosomal protein S17

Chain Q:



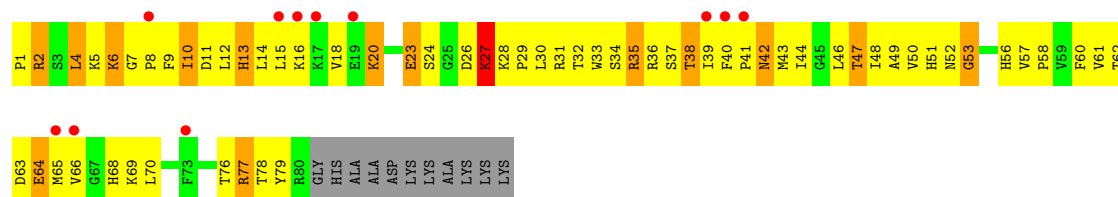
- Molecule 17: 30S ribosomal protein S18

Chain R:



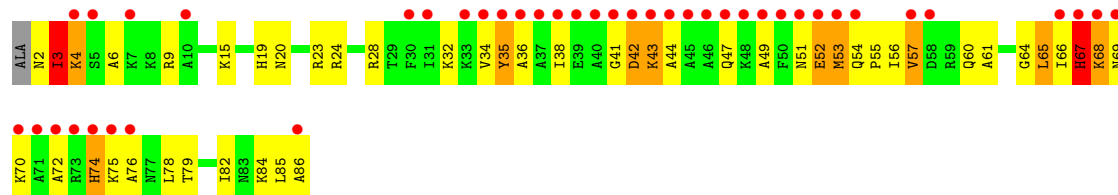
- Molecule 18: 30S ribosomal protein S19

Chain S:



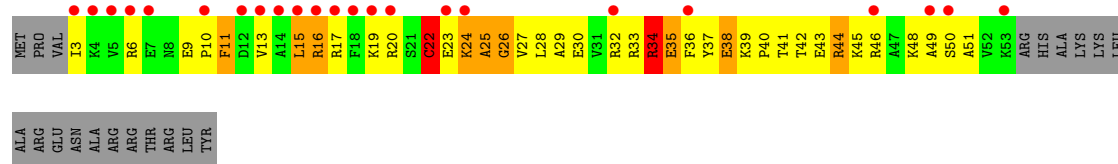
- Molecule 19: 30S ribosomal protein S20

Chain T:



- Molecule 20: 30S ribosomal protein S2

Chain B:



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	208.70Å 379.50Å 739.30Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	70.00 – 3.50 137.77 – 3.50	Depositor EDS
% Data completeness (in resolution range)	62.1 (70.00-3.50) 62.3 (137.77-3.50)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.07	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.88 (at 3.49Å)	Xtriage
Refinement program	CNS	Depositor
R, $R_{free}$	0.269 , 0.318 0.526 , 0.534	Depositor DCC
$R_{free}$ test set	22206 reflections (4.89%)	DCC
Wilson B-factor (Å <sup>2</sup> )	117.9	Xtriage
Anisotropy	0.294	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.22 , 20.3	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Outliers	0 of 454411 reflections	Xtriage
$F_o, F_c$ correlation	0.58	EDS
Total number of atoms	51794	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	71.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.15% 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: MG, HYG

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	0.26	1/36762 (0.0%)	0.77	17/57350 (0.0%)
2	C	0.23	0/1651	0.44	0/2225
3	D	0.23	0/1665	0.43	0/2227
4	E	0.23	0/1118	0.45	0/1504
5	F	0.24	0/835	0.44	0/1128
6	G	0.23	0/1211	0.44	0/1624
7	H	0.23	0/989	0.45	0/1326
8	I	0.24	0/1034	0.44	0/1375
9	J	0.22	0/796	0.47	0/1077
10	K	0.24	0/893	0.46	0/1205
11	L	0.22	0/969	0.47	0/1300
12	M	0.21	0/884	0.45	0/1181
13	N	0.24	0/785	0.43	0/1043
14	O	0.22	0/723	0.44	0/966
15	P	0.25	0/648	0.44	0/870
16	Q	0.24	0/665	0.48	0/892
17	R	0.23	0/462	0.44	0/621
18	S	0.25	0/660	0.46	0/888
19	T	0.24	0/671	0.41	0/888
20	B	0.25	0/1735	0.44	0/2338
21	U	0.26	0/430	0.46	0/570
All	All	0.25	1/55586 (0.0%)	0.68	17/82598 (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
1	A	0	13

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	495	A	N3-C4	-5.48	1.31	1.34

The worst 5 of 17 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	1213	A	O5'-P-OP1	-26.18	79.28	110.70
1	A	1213	A	O5'-P-OP2	17.64	131.86	110.70
1	A	1212	U	OP1-P-O3'	15.45	139.18	105.20
1	A	1049	U	O5'-P-OP1	-7.79	98.69	105.70
1	A	975	A	C5'-C4'-C3'	-7.25	104.41	116.00

There are no chirality outliers.

5 of 13 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	187	G	Sidechain
1	A	281	G	Sidechain
1	A	324	G	Sidechain
1	A	437	U	Sidechain
1	A	82	G	Sidechain

## 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	32831	0	16521	1375	0
2	C	1624	0	1699	136	0
3	D	1643	0	1710	156	0
4	E	1105	0	1148	115	0
5	F	817	0	808	88	0
6	G	1196	0	1246	110	0
7	H	979	0	1034	93	0
8	I	1022	0	1070	151	0
9	J	786	0	828	82	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
10	K	877	0	887	101	0
11	L	955	0	1019	94	0
12	M	876	0	937	107	0
13	N	774	0	827	90	0
14	O	715	0	742	40	0
15	P	638	0	656	55	0
16	Q	656	0	702	73	0
17	R	455	0	478	34	0
18	S	644	0	675	98	0
19	T	665	0	714	49	0
20	B	1704	0	1732	208	0
21	U	425	0	449	68	0
22	A	61	0	0	0	0
22	E	1	0	0	0	0
23	A	36	0	37	1	0
24	A	299	0	0	0	0
24	E	5	0	0	0	0
24	L	1	0	0	0	0
24	N	4	0	0	0	0
All	All	51794	0	35919	3069	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 35.

The worst 5 of 3069 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:1533:C:H2'	1:A:1534:A:H3'	1.34	1.08
1:A:1221:G:H4'	18:S:76:THR:HG21	1.41	1.02
21:U:16:ARG:HA	21:U:16:ARG:HE	1.23	1.02
8:I:20:ILE:HA	8:I:62:LEU:HD12	1.37	1.01
20:B:202:ASN:HD22	20:B:204:ASP:H	1.06	1.00

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	C	204/232 (88%)	148 (72%)	40 (20%)	16 (8%)	1	22
3	D	203/205 (99%)	134 (66%)	54 (27%)	15 (7%)	2	24
4	E	148/166 (89%)	117 (79%)	25 (17%)	6 (4%)	4	44
5	F	98/135 (73%)	67 (68%)	21 (21%)	10 (10%)	1	14
6	G	150/178 (84%)	124 (83%)	20 (13%)	6 (4%)	5	44
7	H	127/129 (98%)	90 (71%)	28 (22%)	9 (7%)	2	25
8	I	125/129 (97%)	88 (70%)	29 (23%)	8 (6%)	2	29
9	J	96/103 (93%)	68 (71%)	18 (19%)	10 (10%)	1	14
10	K	115/128 (90%)	80 (70%)	27 (24%)	8 (7%)	2	26
11	L	121/123 (98%)	79 (65%)	28 (23%)	14 (12%)	1	12
12	M	111/117 (95%)	69 (62%)	38 (34%)	4 (4%)	5	49
13	N	92/100 (92%)	56 (61%)	26 (28%)	10 (11%)	1	13
14	O	86/89 (97%)	65 (76%)	16 (19%)	5 (6%)	3	32
15	P	78/82 (95%)	53 (68%)	15 (19%)	10 (13%)	0	10
16	Q	79/83 (95%)	56 (71%)	17 (22%)	6 (8%)	2	23
17	R	53/74 (72%)	41 (77%)	10 (19%)	2 (4%)	5	46
18	S	78/91 (86%)	58 (74%)	18 (23%)	2 (3%)	8	57
19	T	83/86 (96%)	59 (71%)	19 (23%)	5 (6%)	2	31
20	B	216/240 (90%)	148 (68%)	46 (21%)	22 (10%)	1	14
21	U	49/71 (69%)	28 (57%)	14 (29%)	7 (14%)	0	7
All	All	2312/2561 (90%)	1628 (70%)	509 (22%)	175 (8%)	2	23

5 of 175 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	C	112	ALA
2	C	180	ASP
2	C	205	GLU
4	E	20	VAL
5	F	41	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 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	C	170/189 (90%)	145 (85%)	25 (15%)	4	26
3	D	172/172 (100%)	149 (87%)	23 (13%)	6	31
4	E	113/125 (90%)	98 (87%)	15 (13%)	6	31
5	F	87/116 (75%)	75 (86%)	12 (14%)	5	29
6	G	125/146 (86%)	108 (86%)	17 (14%)	5	30
7	H	104/104 (100%)	96 (92%)	8 (8%)	18	64
8	I	105/106 (99%)	93 (89%)	12 (11%)	8	40
9	J	86/90 (96%)	80 (93%)	6 (7%)	21	68
10	K	90/98 (92%)	77 (86%)	13 (14%)	5	27
11	L	103/103 (100%)	84 (82%)	19 (18%)	2	14
12	M	91/95 (96%)	82 (90%)	9 (10%)	11	49
13	N	79/83 (95%)	67 (85%)	12 (15%)	4	25
14	O	76/77 (99%)	70 (92%)	6 (8%)	18	62
15	P	65/65 (100%)	57 (88%)	8 (12%)	7	35
16	Q	75/77 (97%)	65 (87%)	10 (13%)	6	31
17	R	48/64 (75%)	46 (96%)	2 (4%)	40	84
18	S	71/78 (91%)	53 (75%)	18 (25%)	1	5
19	T	65/65 (100%)	53 (82%)	12 (18%)	2	13
20	B	180/198 (91%)	150 (83%)	30 (17%)	3	19
21	U	44/61 (72%)	35 (80%)	9 (20%)	2	10
All	All	1949/2112 (92%)	1683 (86%)	266 (14%)	5	30

5 of 266 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
10	K	34	THR
11	L	119	LYS
20	B	121	GLN
10	K	64	VAL
11	L	18	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 69 such sidechains are listed below:

Mol	Chain	Res	Type
9	J	56	HIS
11	L	28	GLN
20	B	41	ASN
9	J	70	HIS
10	K	39	ASN

### 5.3.3 RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
1	A	1529/1542 (99%)	240 (15%)	27 (1%)

5 of 240 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
1	A	6	G
1	A	9	G
1	A	14	U
1	A	32	A
1	A	39	G

5 of 27 RNA pucker outliers are listed below:

Mol	Chain	Res	Type
1	A	960	U
1	A	1049	U
1	A	1397	C
1	A	975	A
1	A	279	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 63 ligands modelled in this entry, 62 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	$\# Z  > 2$
23	HYG	A	3001	-	39,39,39	1.50	7 (17%)	60,60,60	1.45	6 (10%)

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
23	HYG	A	3001	-	-	1/14/87/87	0/2/4/4

The worst 5 of 7 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
23	A	3001	HYG	C3-C2	5.00	1.56	1.52
23	A	3001	HYG	C27-C33	2.65	1.56	1.52
23	A	3001	HYG	O22-C17	-2.31	1.38	1.43
23	A	3001	HYG	O28-C23	2.18	1.44	1.40
23	A	3001	HYG	C16-C15	2.10	1.57	1.53

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
23	A	3001	HYG	C23-O28-C27	4.38	115.84	111.82
23	A	3001	HYG	O22-C17-C16	4.35	122.21	111.24
23	A	3001	HYG	C10-N9-C4	3.77	115.93	113.82
23	A	3001	HYG	O8-C1-C2	-3.36	101.78	109.66
23	A	3001	HYG	O35-C34-C33	-2.92	103.94	111.25

There are no chirality outliers.

All (1) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
23	A	3001	HYG	C26-C27-C33-N36

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	1530/1542 (99%)	-0.09	46 (3%) 48 22	8, 57, 142, 180	0
2	C	206/232 (88%)	1.52	60 (29%) 1 1	15, 81, 138, 180	0
3	D	205/205 (100%)	2.18	91 (44%) 1 1	5, 63, 135, 180	0
4	E	150/166 (90%)	2.85	82 (54%) 0 1	5, 62, 135, 175	0
5	F	100/135 (74%)	2.36	53 (53%) 0 1	14, 78, 126, 166	0
6	G	152/178 (85%)	2.48	84 (55%) 0 1	27, 98, 156, 177	0
7	H	129/129 (100%)	2.87	75 (58%) 0 1	5, 53, 117, 153	0
8	I	127/129 (98%)	1.28	31 (24%) 1 2	32, 103, 162, 180	0
9	J	98/103 (95%)	3.43	71 (72%) 0 0	42, 107, 156, 180	0
10	K	117/128 (91%)	2.05	58 (49%) 1 1	5, 57, 112, 179	0
11	L	123/123 (100%)	2.52	67 (54%) 0 1	5, 44, 109, 165	0
12	M	113/117 (96%)	2.33	52 (46%) 1 1	32, 108, 156, 180	0
13	N	96/100 (96%)	2.11	44 (45%) 1 1	38, 99, 137, 171	0
14	O	88/89 (98%)	2.10	44 (50%) 0 1	8, 54, 111, 165	0
15	P	80/82 (97%)	2.72	51 (63%) 0 1	5, 51, 143, 164	0
16	Q	81/83 (97%)	0.83	15 (18%) 2 2	5, 51, 121, 157	0
17	R	55/74 (74%)	1.24	12 (21%) 1 2	13, 69, 132, 149	0
18	S	80/91 (87%)	0.49	11 (13%) 4 3	49, 113, 171, 180	0
19	T	85/86 (98%)	2.26	42 (49%) 1 1	14, 58, 121, 177	0
20	B	218/240 (90%)	1.98	76 (34%) 1 1	26, 106, 160, 180	0
21	U	51/71 (71%)	2.10	23 (45%) 1 1	19, 85, 151, 180	0
All	All	3884/4103 (94%)	1.25	1088 (28%) 1 1	5, 71, 147, 180	0

The worst 5 of 1088 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
4	E	12	GLU	11.4
4	E	54	GLU	11.4
6	G	152	HIS	11.1
7	H	126	CYS	10.7
2	C	133	MET	10.6

## 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
22	MG	A	2168	1/1	0.41	15.71	50,50,50,50	1
22	MG	A	2346	1/1	0.21	10.87	106,106,106,106	0
22	MG	A	2042	1/1	0.22	3.08	93,93,93,93	0
22	MG	A	2360	1/1	0.29	2.97	80,80,80,80	0
22	MG	A	2182	1/1	0.20	2.29	23,23,23,23	1
22	MG	A	2236	1/1	0.26	1.75	128,128,128,128	0
22	MG	A	2062	1/1	0.69	1.63	132,132,132,132	0
22	MG	A	2196	1/1	0.25	1.43	28,28,28,28	0
22	MG	A	2283	1/1	0.19	1.35	121,121,121,121	0
22	MG	A	2302	1/1	0.29	1.01	8,8,8,8	0
22	MG	A	2049	1/1	0.24	0.87	67,67,67,67	0
22	MG	A	2023	1/1	0.19	0.84	24,24,24,24	0
22	MG	A	2121	1/1	0.15	0.79	125,125,125,125	0
22	MG	A	2324	1/1	0.14	0.68	104,104,104,104	0
22	MG	A	2148	1/1	0.29	0.53	50,50,50,50	0
22	MG	A	2340	1/1	0.14	0.31	99,99,99,99	0
22	MG	A	2245	1/1	0.20	0.26	48,48,48,48	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
22	MG	A	2273	1/1	0.27	0.13	48,48,48,48	0
22	MG	A	2213	1/1	0.22	-0.02	6,6,6,6	0
22	MG	A	2366	1/1	0.27	-0.03	23,23,23,23	0
22	MG	A	2068	1/1	0.19	-0.07	93,93,93,93	0
23	HYG	A	3001	36/36	0.20	-0.10	45,45,45,45	0
22	MG	A	2334	1/1	0.18	-0.14	5,5,5,5	0
22	MG	A	2056	1/1	0.17	-0.24	33,33,33,33	0
22	MG	E	2155	1/1	0.24	-0.28	127,127,127,127	0
22	MG	A	2036	1/1	0.07	-0.72	46,46,46,46	0
22	MG	A	2141	1/1	0.21	-0.77	29,29,29,29	0
22	MG	A	2097	1/1	0.19	-0.77	5,5,5,5	0
22	MG	A	2084	1/1	0.13	-0.81	149,149,149,149	0
22	MG	A	2261	1/1	0.09	-0.84	19,19,19,19	0
22	MG	A	2091	1/1	0.07	-0.88	53,53,53,53	0
22	MG	A	2107	1/1	0.10	-0.88	73,73,73,73	0
22	MG	A	2114	1/1	0.06	-0.97	73,73,73,73	0
22	MG	A	2295	1/1	0.05	-1.05	74,74,74,74	0
22	MG	A	2202	1/1	0.12	-1.10	33,33,33,33	0
22	MG	A	2225	1/1	0.08	-1.31	101,101,101,101	0
22	MG	A	2256	1/1	0.09	-1.32	58,58,58,58	0
22	MG	A	2231	1/1	0.09	-1.33	94,94,94,94	0
22	MG	A	2330	1/1	0.11	-1.37	11,11,11,11	0
22	MG	A	2208	1/1	0.10	-1.43	56,56,56,56	0
22	MG	A	2250	1/1	0.07	-1.54	46,46,46,46	0
22	MG	A	2077	1/1	0.08	-1.64	47,47,47,47	0
22	MG	A	2007	1/1	0.07	-1.68	5,5,5,5	0
22	MG	A	2353	1/1	0.10	-1.80	94,94,94,94	0
22	MG	A	2319	1/1	0.07	-1.85	30,30,30,30	0
22	MG	A	2011	1/1	0.05	-2.12	35,35,35,35	0
22	MG	A	2101	1/1	0.09	-2.15	6,6,6,6	0
22	MG	A	2306	1/1	0.08	-2.19	39,39,39,39	0
22	MG	A	2073	1/1	0.08	-2.22	73,73,73,73	0
22	MG	A	2267	1/1	0.10	-2.33	59,59,59,59	0
22	MG	A	2016	1/1	0.08	-2.38	8,8,8,8	0
22	MG	A	2218	1/1	0.09	-2.57	89,89,89,89	0
22	MG	A	2189	1/1	0.13	-2.60	7,7,7,7	0
22	MG	A	2175	1/1	0.12	-2.78	75,75,75,75	0
22	MG	A	2288	1/1	0.05	-3.10	58,58,58,58	0
22	MG	A	2278	1/1	0.08	-3.53	57,57,57,57	0
22	MG	A	2001	1/1	0.06	-3.85	5,5,5,5	0
22	MG	A	2029	1/1	0.07	-4.31	95,95,95,95	0
22	MG	A	2128	1/1	0.09	-5.62	63,63,63,63	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
22	MG	A	2239	1/1	0.04	-5.69	16,16,16,16	0
22	MG	A	2312	1/1	0.20	-	73,73,73,73	0
22	MG	A	2161	1/1	0.38	-	26,26,26,26	1
22	MG	A	2134	1/1	0.13	-	137,137,137,137	0

## 6.5 Other polymers

There are no such residues in this entry.