wwPDB X-ray Structure Validation Summary Report (i)

Jan 17, 2022 – 10:17 AM JST

PDB ID	:	7DUH
Title	:	Crystal structure of the Thermus thermophilus (HB8) 30S ribosomal subunit with mRNA and cognate transfer RNA anticodon stem-loop and sisomicin derivative N1"AC bound
Authors	:	DeMirci, H.; Destan, E.
Deposited on	:	2021-01-09
Resolution	:	3.75 Å(reported)

This is a wwPDB X-ray Structure Validation Summary Report for a publicly released PDB entry.

We welcome your comments at *validation@mail.wwpdb.org*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

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

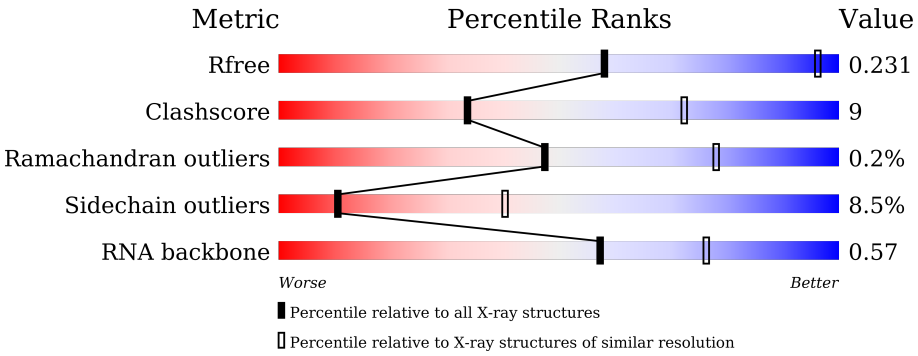
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.25
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.25

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 3.75 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.



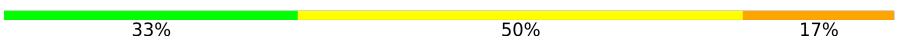
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	1039 (3.94-3.58)
Clashscore	141614	1051 (3.92-3.60)
Ramachandran outliers	138981	1015 (3.92-3.60)
Sidechain outliers	138945	1011 (3.92-3.60)
RNA backbone	3102	1035 (4.52-3.00)

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 of the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions <=5%

Mol	Chain	Length	Quality of chain
1	A	1522	55% 36% 8%
2	B	256	62% 26% 9%
3	C	239	64% 19% 14%
4	D	209	66% 32%
5	E	162	73% 17% 7%
6	F	101	82% 15%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
7	G	156	
8	H	138	
9	I	128	
10	J	105	
11	K	129	
12	L	135	
13	M	126	
14	N	61	
15	O	89	
16	P	88	
17	Q	105	
18	R	88	
19	S	93	
20	T	106	
21	U	27	
22	Y	6	
23	W	15	

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
24	HJO	A	1601	X	-	-	-

2 Entry composition (i)

There are 27 unique types of molecules in this entry. The entry contains 52555 atoms, of which 0 are hydrogens and 0 are deuteriums.

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 30S Ribosomal RNA rRNA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
1	A	1512	32507	14477	6012	10506	1512	0	0	0

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	234	1900	1213	341	341	5	0	0	0

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	C	206	1612	1016	314	281	1	0	0	0

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
4	D	208	1703	1066	339	291	7	0	0	0

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
5	E	150	1146	724	217	201	4	0	0	0

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
6	F	101	843	531	155	154	3	0	0	0

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
7	G	155	Total	C	N	O	S	0	0	0
			1257	781	252	218	6			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
8	H	138	Total	C	N	O	S	0	0	0
			1116	705	215	193	3			

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

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
9	I	127	Total	C	N	O	0	0	0
			1010	639	197	174			

- 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
			792	498	156	137	1			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
11	K	116	Total	C	N	O	S	0	0	0
			864	537	164	160	3			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
12	L	124	Total	C	N	O	S	0	0	0
			972	612	195	163	2			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
13	M	118	Total	C	N	O	S	0	0	0
			937	579	193	163	2			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
14	N	60	Total 492	C 312	N 104	O 72	S 4	0	0	0

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
15	O	87	Total 729	C 457	N 146	O 124	S 2	0	0	0

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
16	P	83	Total 700	C 443	N 139	O 117	S 1	0	0	0

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
17	Q	99	Total 823	C 528	N 152	O 141	S 2	0	0	0

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

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
18	R	70	Total 574	C 367	N 112	O 95	0	0	0

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
19	S	80	Total 647	C 414	N 119	O 112	S 2	0	0	0

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
20	T	99	Total 763	C 470	N 162	O 129	S 2	0	0	0

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

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
21	U	24	Total	C	N	O	0	0	0
			208	128	50	30			

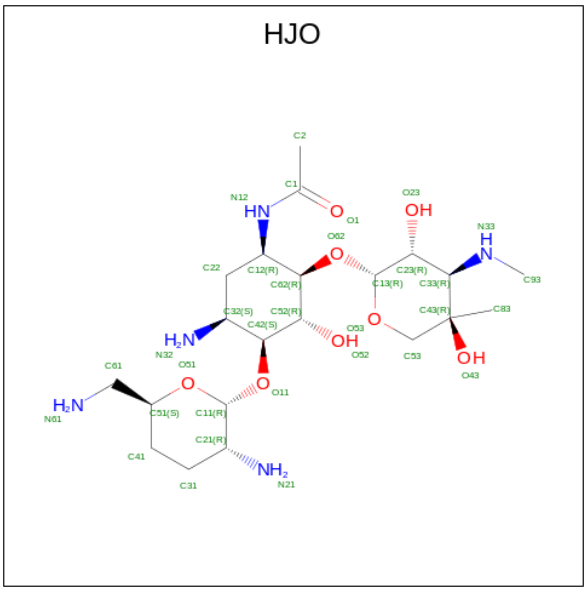
- Molecule 22 is a RNA chain called RNA (5'-R(*UP*UP*UP*UP*UP*U)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
22	Y	6	Total	C	N	O	P	0	0	0
			117	54	12	46	5			

- Molecule 23 is a RNA chain called RNA (5'-R(*GP*GP*GP*AP*UP*UP*GP*AP*AP*AP *AP*UP*CP*CP*C)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
23	W	15	Total	C	N	O	P	0	0	0
			319	144	60	101	14			

- Molecule 24 is N-[(1R,2R,3R,4S,5S)-4-[(2R,3R,6S)-6-(aminomethyl)-3-azanyl-oxan-2-yl]oxy-5-azanyl-2-[(2R,3R,4R)-5-methyl-4-(methylamino)-3,5-bis(oxidanyl)oxan-2-yl]oxy-3-oxidanyl-cyclohexyl]ethanamide (three-letter code: HJO) (formula: $C_{21}H_{41}N_5O_8$) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
24	A	1	Total	C	N	O	0	0
			34	21	5	8		

- Molecule 25 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
25	A	232	Total Mg 232 232	0	0
25	B	3	Total Mg 3 3	0	0
25	C	3	Total Mg 3 3	0	0
25	D	4	Total Mg 4 4	0	0
25	E	1	Total Mg 1 1	0	0
25	F	1	Total Mg 1 1	0	0
25	H	1	Total Mg 1 1	0	0
25	P	2	Total Mg 2 2	0	0
25	Q	1	Total Mg 1 1	0	0
25	S	1	Total Mg 1 1	0	0

- Molecule 26 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
26	D	1	Total Zn 1 1	0	0
26	N	1	Total Zn 1 1	0	0

- Molecule 27 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
27	A	223	Total O 223 223	0	0
27	C	1	Total O 1 1	0	0
27	D	3	Total O 3 3	0	0
27	E	5	Total O 5 5	0	0
27	L	2	Total O 2 2	0	0
27	N	1	Total O 1 1	0	0

Continued on next page...

Continued from previous page...

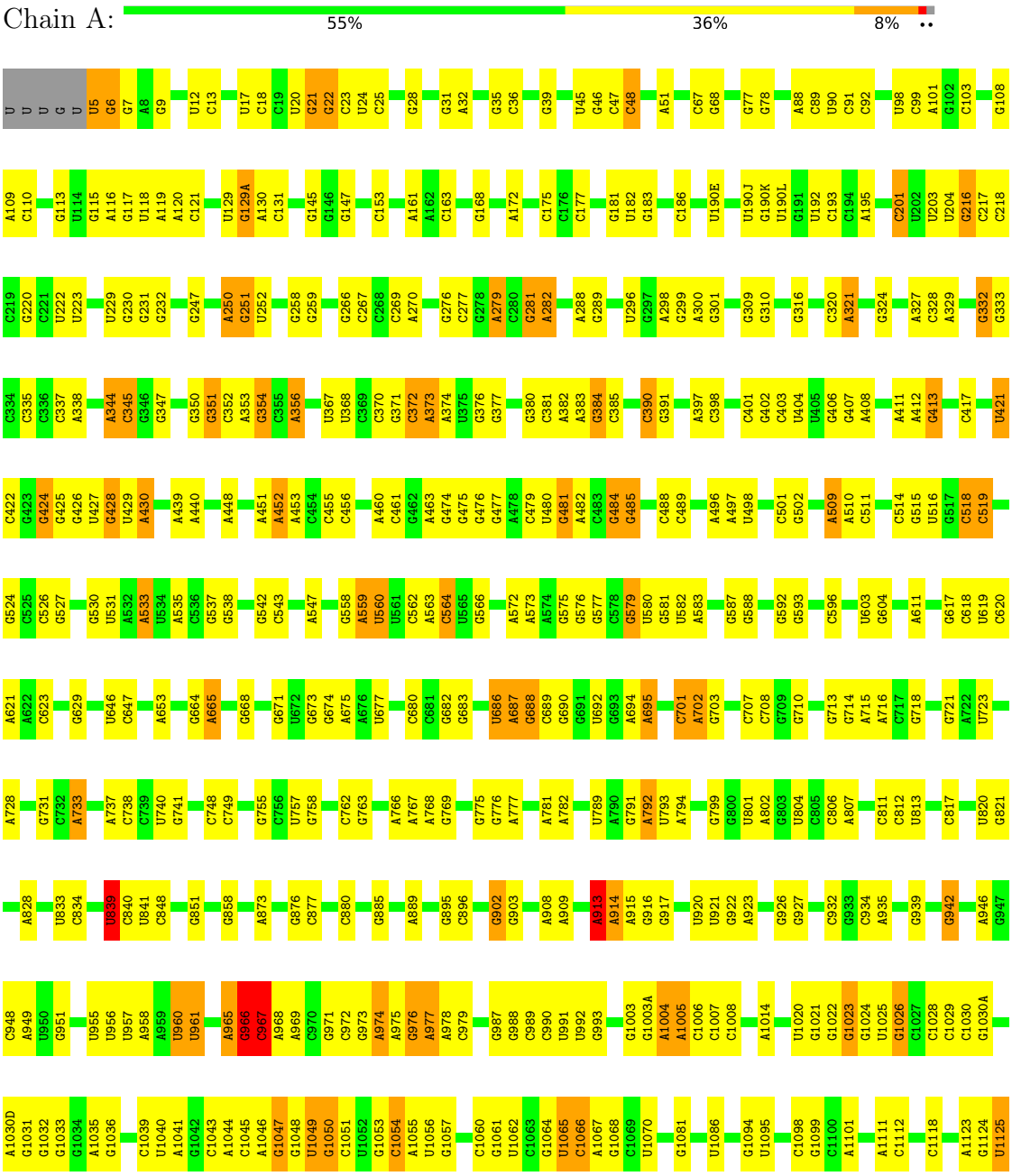
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
27	Q	2	Total O 2 2	0	0
27	T	2	Total O 2 2	0	0

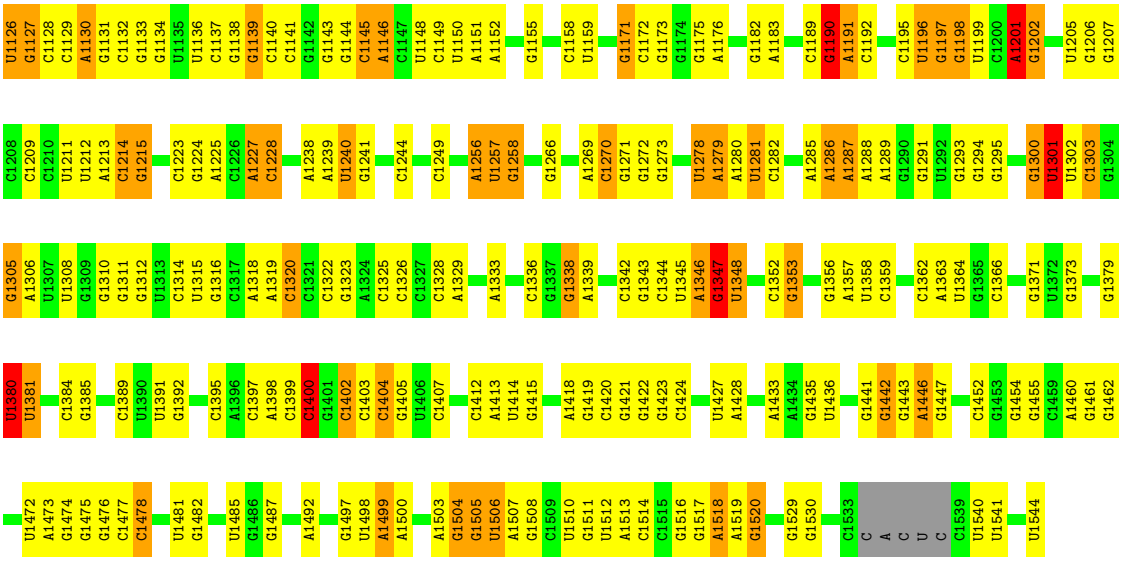
3 Residue-property plots (i)

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

- Molecule 1: 30S Ribosomal RNA rRNA

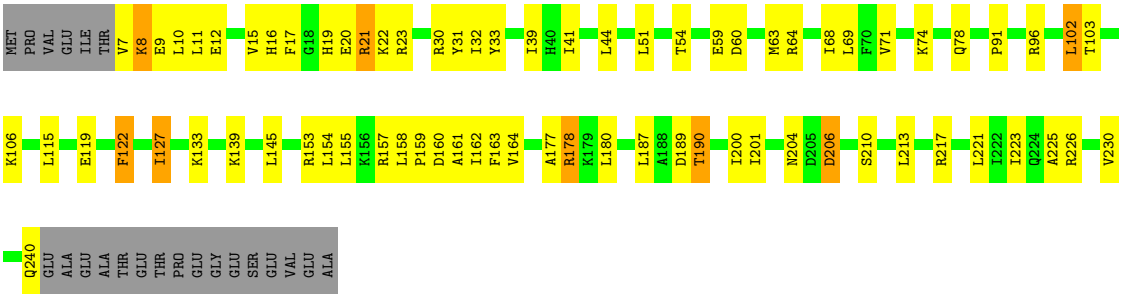
Chain A:





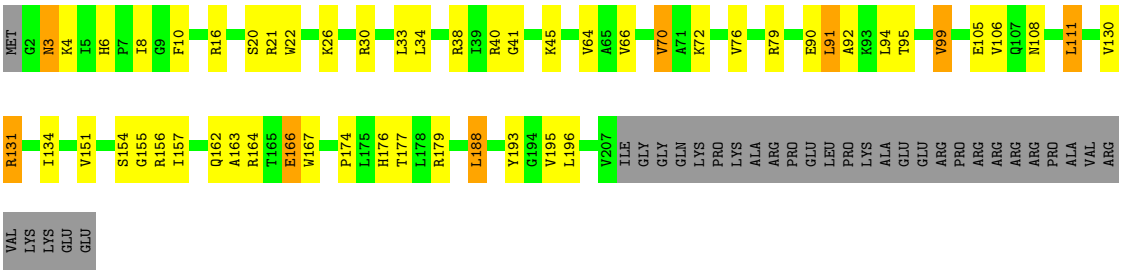
• Molecule 2: 30S ribosomal protein S2

Chain B:



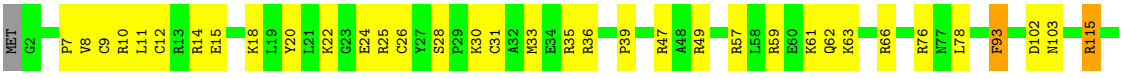
• Molecule 3: 30S ribosomal protein S3

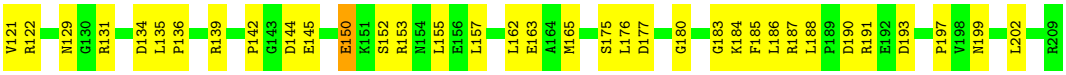
Chain C:



• Molecule 4: 30S ribosomal protein S4

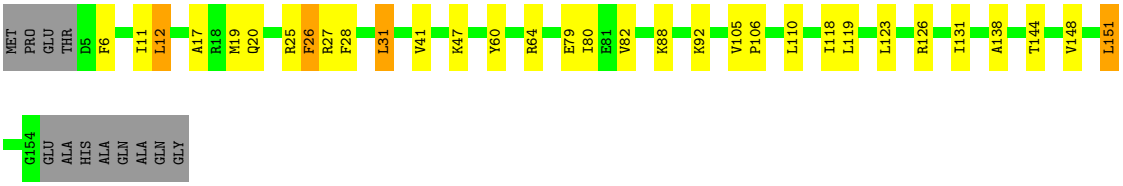
Chain D:





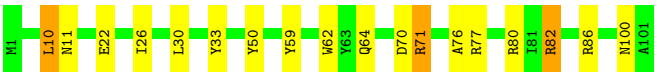
• Molecule 5: 30S ribosomal protein S5

Chain E:



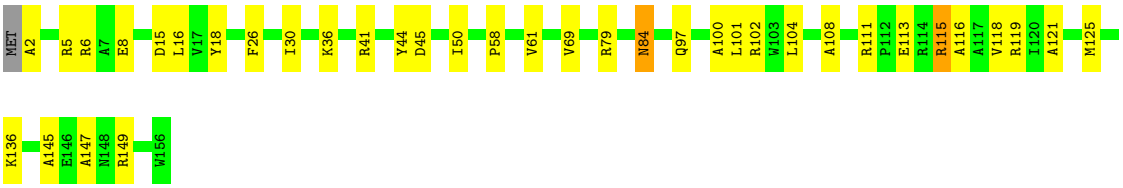
• Molecule 6: 30S ribosomal protein S6

Chain F:



• Molecule 7: 30S ribosomal protein S7

Chain G:

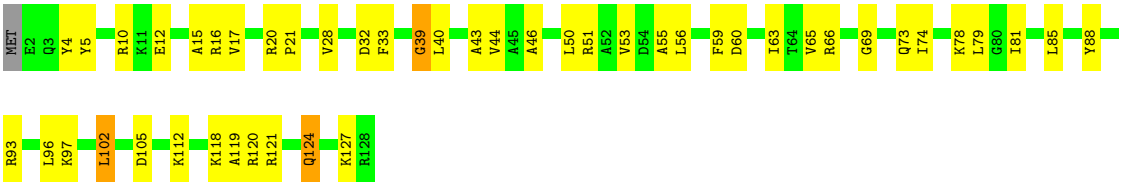


• Molecule 8: 30S ribosomal protein S8

Chain H:

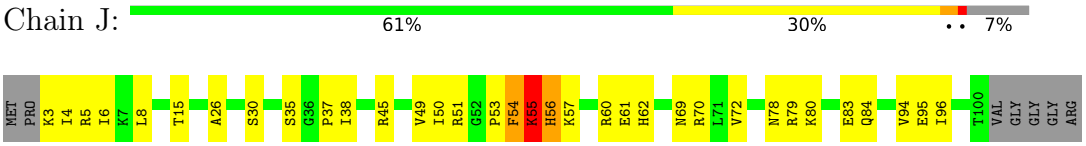
• Molecule 9: 30S ribosomal protein S9

Chain I:



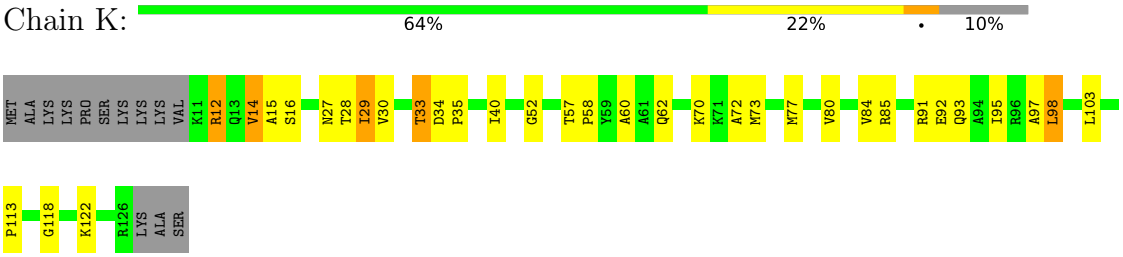
• Molecule 10: 30S ribosomal protein S10

Chain J:



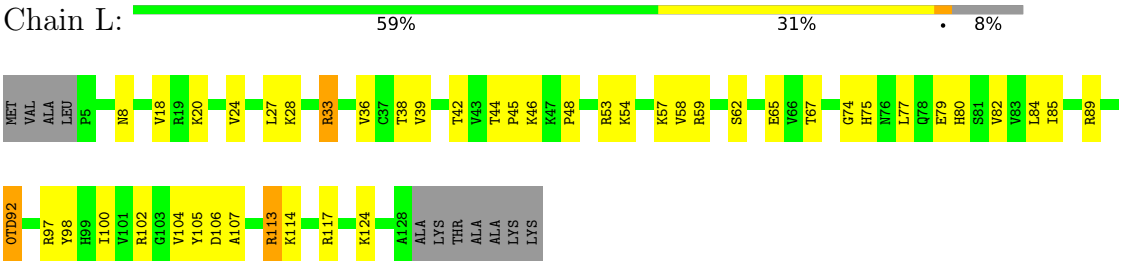
• Molecule 11: 30S ribosomal protein S11

Chain K:



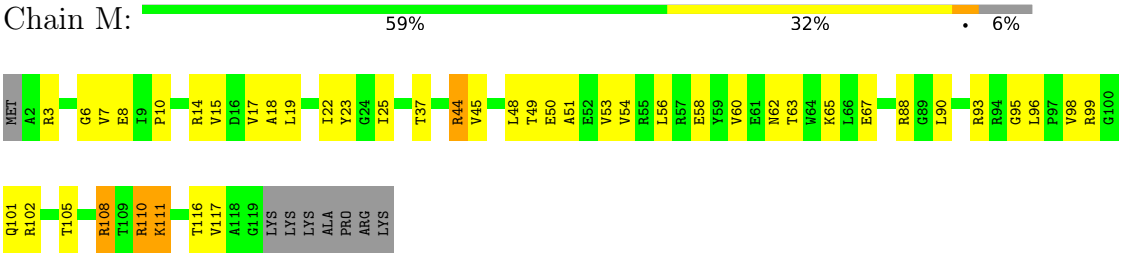
• Molecule 12: 30S ribosomal protein S12

Chain L:



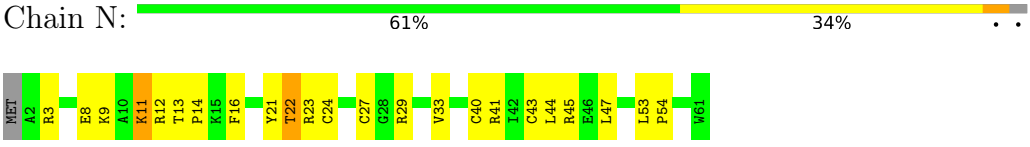
• Molecule 13: 30S ribosomal protein S13

Chain M:



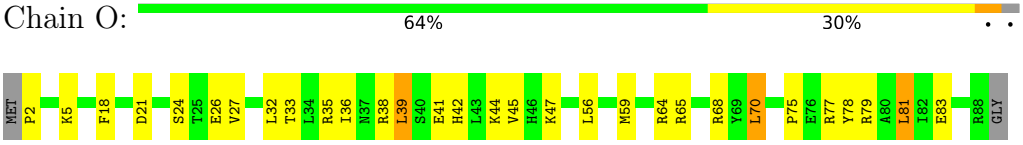
• Molecule 14: 30S ribosomal protein S14 type Z

Chain N:



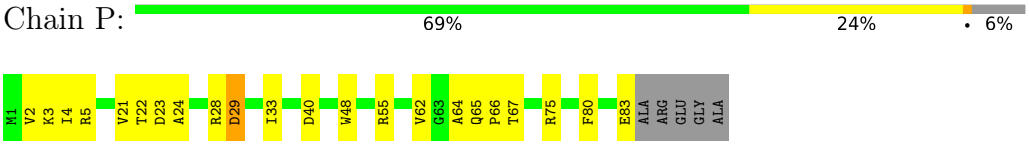
• Molecule 15: 30S ribosomal protein S15

Chain O:



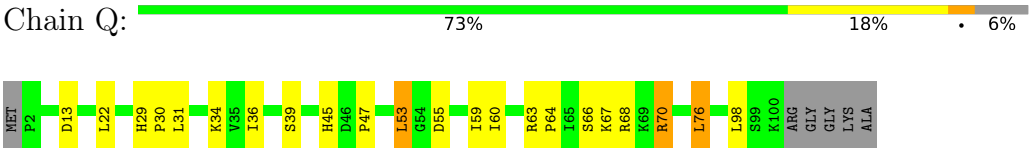
• Molecule 16: 30S ribosomal protein S16

Chain P:



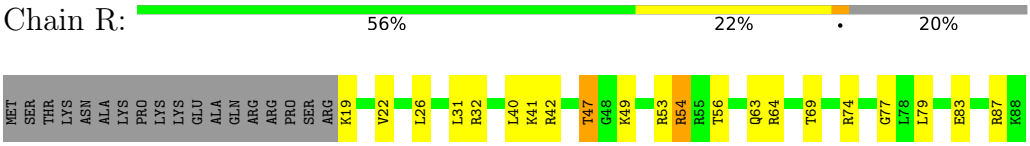
• Molecule 17: 30S ribosomal protein S17

Chain Q:



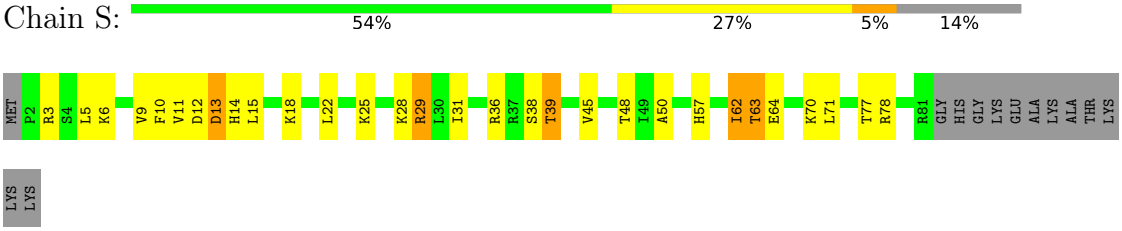
• Molecule 18: 30S ribosomal protein S18

Chain R:



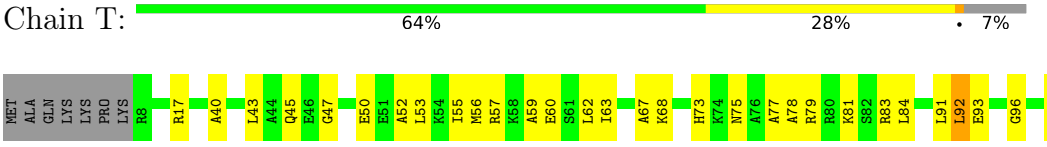
• Molecule 19: 30S ribosomal protein S19

Chain S:



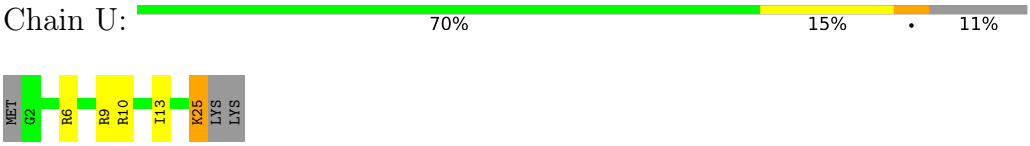
• Molecule 20: 30S ribosomal protein S20

Chain T:



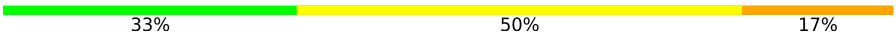
• Molecule 21: 30S ribosomal protein Thx

Chain U:



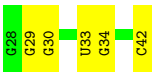
• Molecule 22: RNA (5'-R(*UP*UP*UP*UP*UP*U)-3')

Chain Y:



• Molecule 23: RNA (5'-R(*GP*GP*GP*AP*UP*UP*GP*AP*AP*AP*AP*UP*CP*CP*C)-3')

Chain W:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 41 21 2	Depositor
Cell constants a, b, c, α, β, γ	400.83Å 400.83Å 174.67Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	39.83 – 3.75 39.83 – 2.92	Depositor EDS
% Data completeness (in resolution range)	91.3 (39.83-3.75) 63.8 (39.83-2.92)	Depositor EDS
R_{merge}	1.64	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	-0.70 (at 2.90Å)	Xtriage
Refinement program	PHENIX DEV-3318	Depositor
R, R_{free}	0.188 , 0.230 0.188 , 0.231	Depositor DCC
R_{free} test set	2000 reflections (0.70%)	wwPDB-VP
Wilson B-factor (Å^2)	37.1	Xtriage
Anisotropy	0.019	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.14 , -10.0	EDS
L-test for twinning[2]	$< \|L\| > = 0.45$, $< L^2 > = 0.28$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.87	EDS
Total number of atoms	52555	wwPDB-VP
Average B, all atoms (Å^2)	161.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 1.22% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: 5MC, PSU, M2G, HJO, MG, 0TD, MA6, UR3, 4OC, ZN, 2MG, 7MG

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	0.25	0/36040	0.77	13/56243 (0.0%)
2	B	0.26	0/1935	0.42	0/2609
3	C	0.26	0/1636	0.46	0/2205
4	D	0.27	0/1733	0.41	0/2318
5	E	0.27	0/1162	0.46	0/1564
6	F	0.24	0/856	0.41	0/1154
7	G	0.25	0/1276	0.38	0/1709
8	H	0.27	0/1136	0.46	0/1527
9	I	0.26	0/1029	0.48	1/1379 (0.1%)
10	J	0.26	0/805	0.49	0/1082
11	K	0.27	0/879	0.44	0/1187
12	L	0.26	0/977	0.48	0/1306
13	M	0.25	0/947	0.45	0/1270
14	N	0.27	0/501	0.44	0/664
15	O	0.24	0/740	0.40	0/987
16	P	0.26	0/716	0.46	0/963
17	Q	0.27	0/836	0.46	0/1117
18	R	0.26	0/579	0.43	0/768
19	S	0.25	0/661	0.46	0/890
20	T	0.26	0/765	0.40	0/1007
21	U	0.24	0/212	0.41	0/277
22	Y	0.27	0/128	0.88	0/196
23	W	0.26	0/357	0.79	0/555
All	All	0.25	0/55906	0.69	14/82977 (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
3	C	0	1
8	H	0	1
10	J	0	2
13	M	0	2
All	All	0	6

There are no bond length outliers.

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

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	839	U	C2-N1-C1'	7.92	127.20	117.70
1	A	839	U	N1-C2-O2	7.03	127.72	122.80
1	A	839	U	N3-C2-O2	-6.45	117.68	122.20
1	A	1347	G	P-O3'-C3'	5.57	126.38	119.70
1	A	1301	U	P-O3'-C3'	5.55	126.36	119.70

There are no chirality outliers.

5 of 6 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
3	C	166	GLU	Peptide
8	H	71	GLY	Peptide
10	J	54	PHE	Peptide
10	J	55	LYS	Peptide
13	M	6	GLY	Peptide

5.2 Too-close contacts (i)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry-related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	32507	0	16434	405	0
2	B	1900	0	1951	38	0
3	C	1612	0	1676	30	0
4	D	1703	0	1763	40	0
5	E	1146	0	1207	21	0
6	F	843	0	857	12	0
7	G	1257	0	1296	21	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
8	H	1116	0	1176	16	0
9	I	1010	0	1037	30	0
10	J	792	0	835	25	0
11	K	864	0	881	21	0
12	L	972	0	1058	27	0
13	M	937	0	995	26	0
14	N	492	0	529	24	0
15	O	729	0	768	16	0
16	P	700	0	720	11	0
17	Q	823	0	893	12	0
18	R	574	0	644	14	0
19	S	647	0	673	23	0
20	T	763	0	861	14	0
21	U	208	0	221	4	0
22	Y	117	0	62	3	0
23	W	319	0	164	3	0
24	A	34	0	0	1	0
25	A	232	0	0	0	0
25	B	3	0	0	0	0
25	C	3	0	0	0	0
25	D	4	0	0	0	0
25	E	1	0	0	0	0
25	F	1	0	0	0	0
25	H	1	0	0	0	0
25	P	2	0	0	0	0
25	Q	1	0	0	0	0
25	S	1	0	0	0	0
26	D	1	0	0	0	0
26	N	1	0	0	0	0
27	A	223	0	0	3	0
27	C	1	0	0	0	0
27	D	3	0	0	0	0
27	E	5	0	0	0	0
27	L	2	0	0	0	0
27	N	1	0	0	0	0
27	Q	2	0	0	0	0
27	T	2	0	0	1	0
All	All	52555	0	36701	742	0

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 9.

The worst 5 of 742 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:481:G:HO2'	1:A:482:A:H8	0.97	0.94
1:A:1214:C:H3'	1:A:1215:G:H8	1.37	0.87
1:A:456:C:H42	1:A:476:G:H1	1.28	0.81
20:T:100:ILE:HG22	20:T:102:GLY:H	1.45	0.80
17:Q:22:LEU:HD11	17:Q:39:SER:HB3	1.64	0.79

There are no symmetry-related clashes.

5.3 Torsion angles (i)

5.3.1 Protein backbone (i)

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all 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	232/256 (91%)	218 (94%)	14 (6%)	0	100	100
3	C	204/239 (85%)	179 (88%)	25 (12%)	0	100	100
4	D	206/209 (99%)	198 (96%)	8 (4%)	0	100	100
5	E	148/162 (91%)	142 (96%)	6 (4%)	0	100	100
6	F	99/101 (98%)	96 (97%)	3 (3%)	0	100	100
7	G	153/156 (98%)	144 (94%)	9 (6%)	0	100	100
8	H	136/138 (99%)	128 (94%)	6 (4%)	2 (2%)	10	45
9	I	125/128 (98%)	111 (89%)	14 (11%)	0	100	100
10	J	96/105 (91%)	81 (84%)	12 (12%)	3 (3%)	4	33
11	K	114/129 (88%)	108 (95%)	6 (5%)	0	100	100
12	L	121/135 (90%)	114 (94%)	7 (6%)	0	100	100
13	M	116/126 (92%)	101 (87%)	15 (13%)	0	100	100
14	N	58/61 (95%)	49 (84%)	9 (16%)	0	100	100
15	O	85/89 (96%)	84 (99%)	1 (1%)	0	100	100
16	P	81/88 (92%)	77 (95%)	4 (5%)	0	100	100
17	Q	97/105 (92%)	94 (97%)	3 (3%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
18	R	68/88 (77%)	64 (94%)	4 (6%)	0	100	100
19	S	78/93 (84%)	70 (90%)	8 (10%)	0	100	100
20	T	97/106 (92%)	87 (90%)	10 (10%)	0	100	100
21	U	22/27 (82%)	22 (100%)	0	0	100	100
All	All	2336/2541 (92%)	2167 (93%)	164 (7%)	5 (0%)	47	78

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
10	J	55	LYS
10	J	56	HIS
8	H	71	GLY
10	J	54	PHE
8	H	72	PRO

5.3.2 Protein sidechains (i)

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	202/220 (92%)	180 (89%)	22 (11%)	6	29
3	C	160/188 (85%)	141 (88%)	19 (12%)	5	26
4	D	180/181 (99%)	165 (92%)	15 (8%)	11	41
5	E	115/123 (94%)	108 (94%)	7 (6%)	18	51
6	F	90/90 (100%)	87 (97%)	3 (3%)	38	64
7	G	126/127 (99%)	118 (94%)	8 (6%)	18	49
8	H	119/119 (100%)	111 (93%)	8 (7%)	16	48
9	I	98/99 (99%)	90 (92%)	8 (8%)	11	41
10	J	87/92 (95%)	82 (94%)	5 (6%)	20	53
11	K	88/99 (89%)	82 (93%)	6 (7%)	16	47
12	L	103/110 (94%)	93 (90%)	10 (10%)	8	34

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
13	M	94/101 (93%)	84 (89%)	10 (11%)	6 30
14	N	49/50 (98%)	47 (96%)	2 (4%)	30 59
15	O	79/80 (99%)	70 (89%)	9 (11%)	5 28
16	P	72/74 (97%)	67 (93%)	5 (7%)	15 47
17	Q	94/97 (97%)	87 (93%)	7 (7%)	13 44
18	R	61/77 (79%)	56 (92%)	5 (8%)	11 41
19	S	71/80 (89%)	63 (89%)	8 (11%)	6 28
20	T	76/82 (93%)	66 (87%)	10 (13%)	4 23
21	U	19/22 (86%)	17 (90%)	2 (10%)	7 30
All	All	1983/2111 (94%)	1814 (92%)	169 (8%)	10 40

5 of 169 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
13	M	48	LEU
17	Q	76	LEU
13	M	110	ARG
15	O	81	LEU
19	S	15	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
2	B	19	HIS
2	B	204	ASN
8	H	78	GLN
9	I	124	GLN
10	J	84	GLN

5.3.3 RNA (i)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
1	A	1508/1522 (99%)	243 (16%)	45 (2%)
22	Y	5/6 (83%)	2 (40%)	0
23	W	14/15 (93%)	2 (14%)	0
All	All	1527/1543 (98%)	247 (16%)	45 (2%)

5 of 247 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
1	A	6	G
1	A	9	G
1	A	22	G
1	A	32	A
1	A	39	G

5 of 45 RNA pucker outliers are listed below:

Mol	Chain	Res	Type
1	A	1049	U
1	A	1256	A
1	A	1065	U
1	A	1182	G
1	A	1285	A

5.4 Non-standard residues in protein, DNA, RNA chains (i)

15 non-standard protein/DNA/RNA residues are modelled in this entry.

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	$\#\|Z\| > 2$
1	PSU	A	1540	1	17,21,22	1.04	1 (5%)	20,30,33	3.23	6 (30%)
1	MA6	A	1518	1	19,26,27	1.08	1 (5%)	18,38,41	0.69	0
1	7MG	A	527	1	22,26,27	2.20	6 (27%)	28,39,42	1.78	7 (25%)
1	M2G	A	966	1	20,27,28	2.00	4 (20%)	22,40,43	2.42	5 (22%)
1	5MC	A	1400	1	15,22,23	0.89	0	19,32,35	1.09	1 (5%)
1	5MC	A	1404	1	15,22,23	0.89	0	19,32,35	1.07	2 (10%)
1	5MC	A	1407	1	15,22,23	0.90	0	19,32,35	1.05	2 (10%)
1	2MG	A	1207	1	19,26,27	2.27	4 (21%)	21,38,41	1.98	2 (9%)
1	UR3	A	1498	1	14,22,23	1.05	2 (14%)	15,32,35	1.05	1 (6%)
1	PSU	A	1541	1,25	17,21,22	1.06	2 (11%)	20,30,33	3.19	6 (30%)
1	PSU	A	516	1,25	17,21,22	1.04	2 (11%)	20,30,33	3.13	6 (30%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\|>2$	Counts	RMSZ	$\#\|Z\|>2$
1	MA6	A	1519	1	19,26,27	1.23	2 (10%)	18,38,41	0.60	0
1	5MC	A	967	1	15,22,23	0.94	0	19,32,35	1.09	2 (10%)
12	0TD	L	92	12	4,9,10	0.80	0	3,11,13	1.86	1 (33%)
1	4OC	A	1402	1	16,23,24	0.99	1 (6%)	17,32,35	0.74	0

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
1	PSU	A	1540	1	-	0/7/25/26	0/2/2/2
1	MA6	A	1518	1	-	0/7/29/30	0/3/3/3
1	7MG	A	527	1	-	0/7/37/38	0/3/3/3
1	M2G	A	966	1	-	3/7/29/30	0/3/3/3
1	5MC	A	1400	1	-	2/5/25/26	0/2/2/2
1	5MC	A	1404	1	-	0/5/25/26	0/2/2/2
1	5MC	A	1407	1	-	0/5/25/26	0/2/2/2
1	2MG	A	1207	1	-	0/5/27/28	0/3/3/3
1	UR3	A	1498	1	-	0/5/25/26	0/2/2/2
1	PSU	A	1541	1,25	-	0/7/25/26	0/2/2/2
1	PSU	A	516	1,25	-	0/7/25/26	0/2/2/2
1	MA6	A	1519	1	-	4/7/29/30	0/3/3/3
1	5MC	A	967	1	-	2/5/25/26	0/2/2/2
12	0TD	L	92	12	-	2/3/12/14	-
1	4OC	A	1402	1	-	3/9/29/30	0/2/2/2

The worst 5 of 25 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	1207	2MG	C2-N2	6.86	1.39	1.34
1	A	1207	2MG	C6-N1	5.98	1.43	1.33
1	A	527	7MG	C4-N3	5.89	1.41	1.34
1	A	966	M2G	C6-N1	5.88	1.43	1.33
1	A	527	7MG	C2-N2	5.06	1.44	1.33

The worst 5 of 41 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	1541	PSU	N3-C2-N1	-11.07	119.63	128.43

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	1540	PSU	N3-C2-N1	-10.93	119.74	128.43
1	A	516	PSU	N3-C2-N1	-10.73	119.90	128.43
1	A	966	M2G	C5-C6-N1	-8.09	112.37	123.43
1	A	1207	2MG	C5-C6-N1	-7.59	113.05	123.43

There are no chirality outliers.

5 of 16 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
12	L	92	0TD	O-C-CA-CB
12	L	92	0TD	CG-CB-SB-CSB
1	A	966	M2G	C4'-C5'-O5'-P
1	A	967	5MC	O4'-C4'-C5'-O5'
1	A	1400	5MC	O4'-C4'-C5'-O5'

There are no ring outliers.

7 monomers are involved in 12 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	A	1518	MA6	2	0
1	A	966	M2G	2	0
1	A	1400	5MC	1	0
1	A	1404	5MC	2	0
1	A	967	5MC	2	0
12	L	92	0TD	3	0
1	A	1402	4OC	1	0

5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

Of 252 ligands modelled in this entry, 251 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	HJO	A	1601	-	32,36,36	1.68	7 (21%)	39,53,53	2.66	15 (38%)

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	HJO	A	1601	-	1/1/13/15	10/15/69/69	0/3/3/3

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
24	A	1601	HJO	O53-C53	5.01	1.50	1.43
24	A	1601	HJO	C1-N12	3.89	1.47	1.34
24	A	1601	HJO	O51-C51	2.86	1.49	1.44
24	A	1601	HJO	O53-C13	2.71	1.47	1.41
24	A	1601	HJO	C23-C33	-2.26	1.48	1.53

The worst 5 of 15 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
24	A	1601	HJO	O43-C43-C83	-10.23	85.71	108.13
24	A	1601	HJO	C11-O51-C51	-6.66	105.76	113.13
24	A	1601	HJO	C53-O53-C13	4.19	118.27	111.53
24	A	1601	HJO	O62-C13-C23	-3.98	97.78	108.10
24	A	1601	HJO	O11-C11-C21	3.20	113.59	108.23

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
24	A	1601	HJO	C43

5 of 10 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
24	A	1601	HJO	C21-C11-O11-C42
24	A	1601	HJO	C23-C33-N33-C93

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms
24	A	1601	HJO	C41-C51-C61-N61
24	A	1601	HJO	O51-C51-C61-N61
24	A	1601	HJO	O51-C11-O11-C42

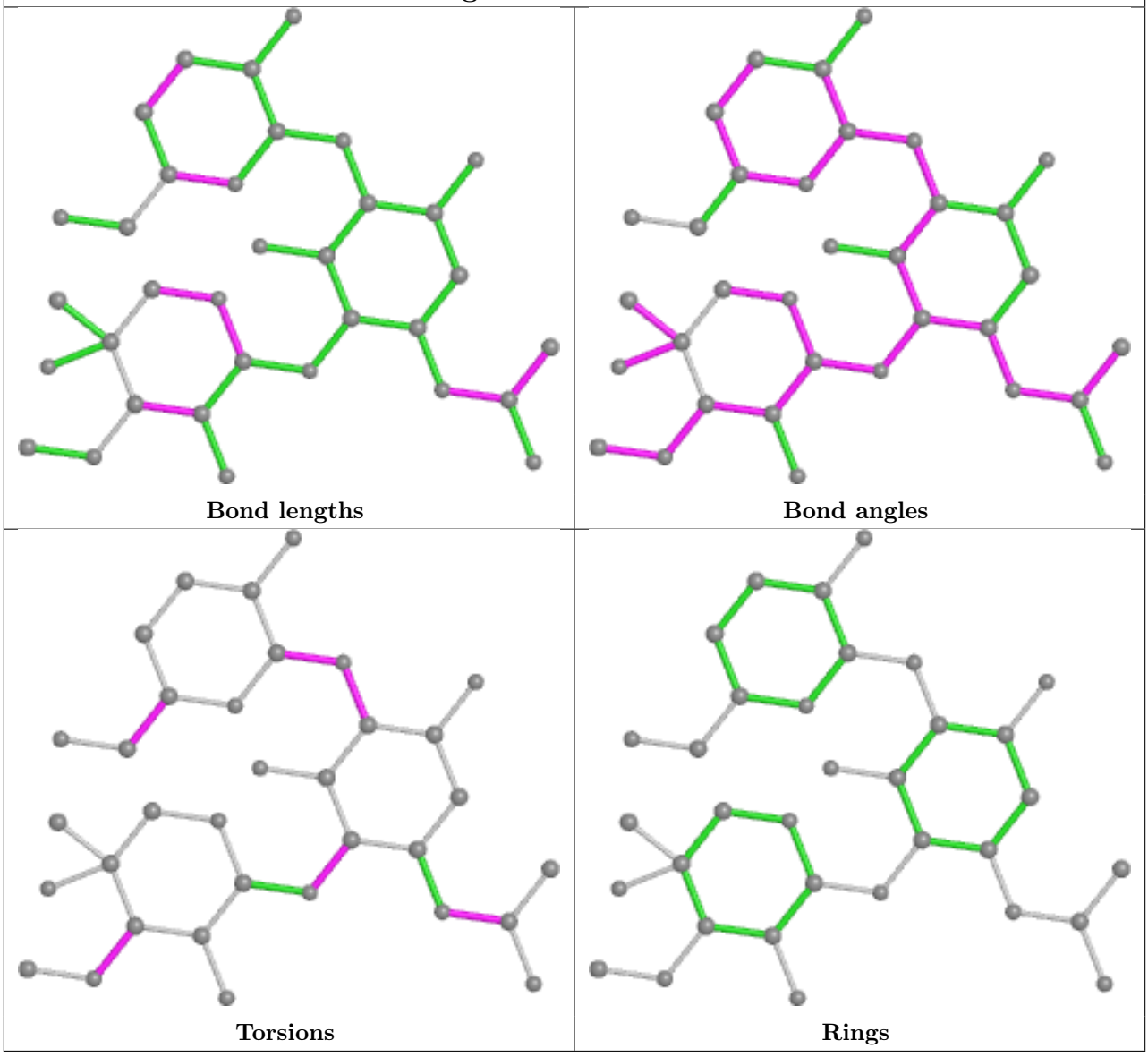
There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
24	A	1601	HJO	1	0

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less then 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.

Ligand HJO A 1601



5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

There are no chain breaks in this entry.

6 Fit of model and data

6.1 Protein, DNA and RNA chains

Unable to reproduce the depositors R factor - this section is therefore empty.

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

Unable to reproduce the depositors R factor - this section is therefore empty.

6.3 Carbohydrates

Unable to reproduce the depositors R factor - this section is therefore empty.

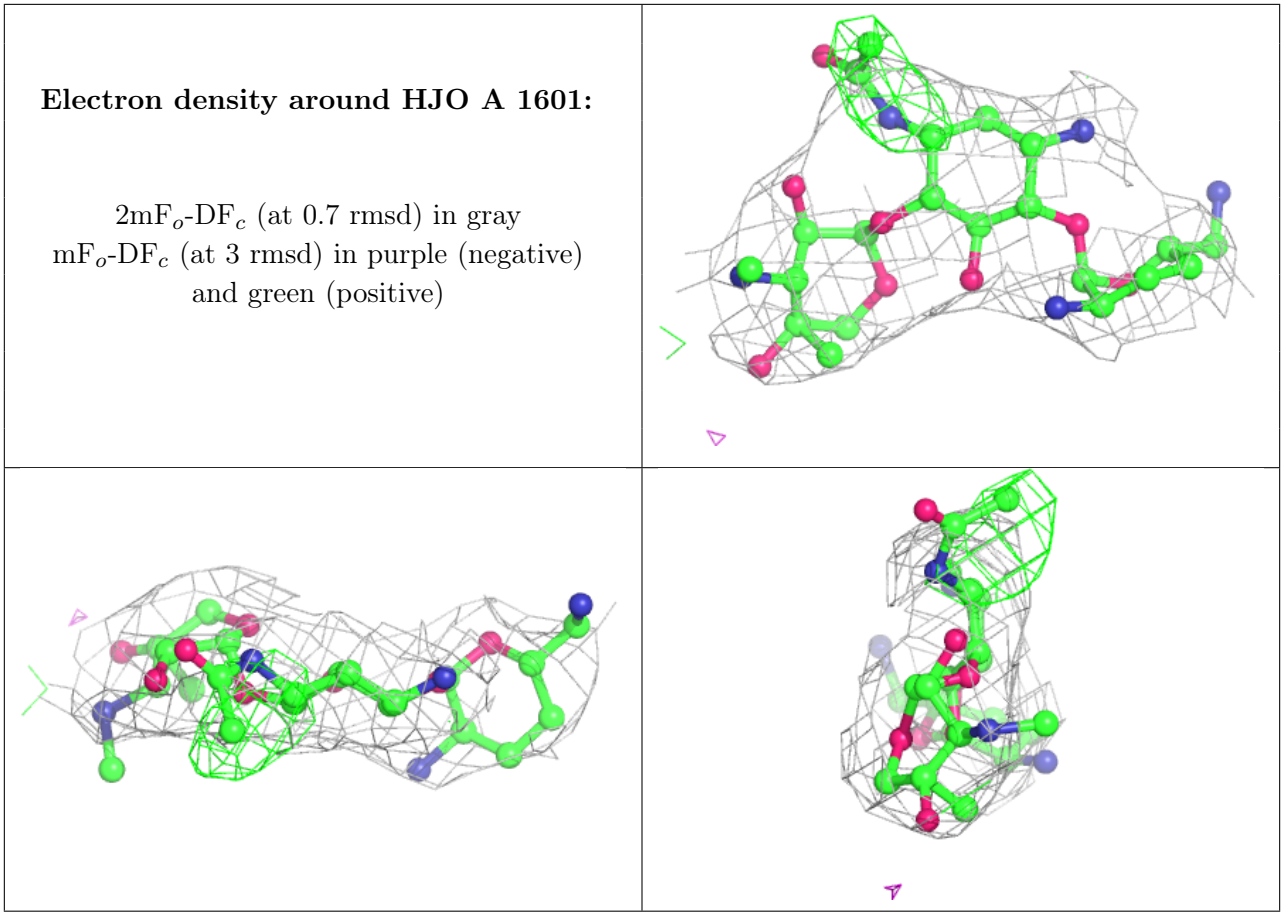
6.4 Ligands

Unable to reproduce the depositors R factor - this section is therefore empty.

The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation to approximate a three-dimensional view.

Electron density around HJO A 1601:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)



6.5 Other polymers

Unable to reproduce the depositors R factor - this section is therefore empty.