WORLDWIDE wwPDB PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report (i)

Jan 3, 2022 – 02:17 PM JST

PDB ID	:	7DUK
Title	:	Crystal structure of the Thermus thermophilus (HB8) 30S ribosomal subunit with mRNA and cognate transfer RNA anticodon stem-loop and sisomicin derivative N1,3"MS bound
Authors	:	DeMirci, H.; Destan, E.
Deposited on	:	2021-01-09
Resolution	:	3.60 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:

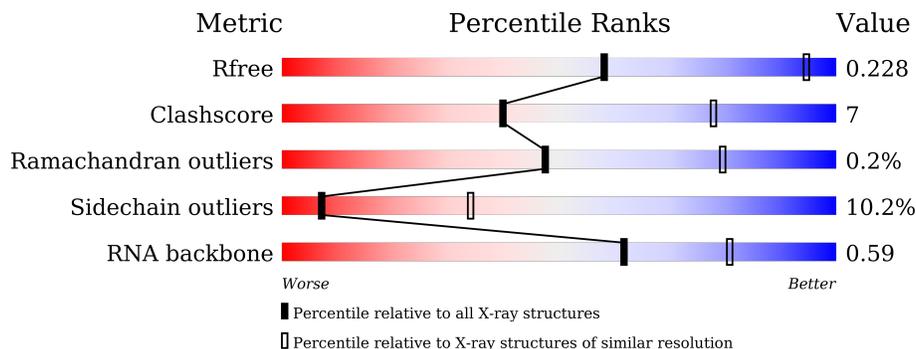|		
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.25
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh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

1   Overall quality at a glance ⓘ

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

The reported resolution of this entry is 3.60 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	Percentile Ranks	Value
Rfree		0.228
Clashscore		7
Ramachandran outliers		0.2%
Sidechain outliers		10.2%
RNA backbone		0.59

Worse   *Better*

▮ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R$_{free}$	130704	1257 (3.70-3.50)
Clashscore	141614	1353 (3.70-3.50)
Ramachandran outliers	138981	1307 (3.70-3.50)
Sidechain outliers	138945	1307 (3.70-3.50)
RNA backbone	3102	1017 (4.20-3.00)

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	61% 31% 7% ••
2	B	256	64% 25% • 9%
3	C	239	62% 21% • 14%
4	D	209	69% 27% •
5	E	162	67% 22% • 7%
6	F	101	83% 15% •

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
7	G	156	76% 22%
8	H	138	77% 17% 7%
9	I	128	69% 26% 5%
10	J	105	61% 29% 7%
11	K	129	64% 22% 10%
12	L	135	62% 28% 8%
13	M	126	67% 24% 6%
14	N	61	54% 41%
15	O	89	73% 19% 6%
16	P	88	78% 14% 6%
17	Q	105	69% 22% 6%
18	R	88	58% 19% 20%
19	S	93	59% 22% 5% 14%
20	T	106	70% 21% 7%
21	U	27	56% 33% 11%
22	Y	6	50% 33% 17%
23	W	15	67% 33%

2 Entry composition ⓘ

There are 27 unique types of molecules in this entry. The entry contains 52689 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
1	A	1512	Total 32644	C 14540	N 6040	O 10546	P 1518	0	6	0

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

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

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

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

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

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

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

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

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
6	F	101	Total 843	C 531	N 155	O 154	S 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 1257	C 781	N 252	O 218	S 6	0	0	0

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

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

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

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

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
10	J	98	Total 792	C 498	N 156	O 137	S 1	0	0	0

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
20	T	99	763	470	162	129	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 208	C 128	N 50	O 30	0	0	0

- 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 117	C 54	N 12	O 46	P 5	0	0	0

- 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 319	C 144	N 60	O 101	P 14	0	0	0

- Molecule 24 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
24	A	234	Total 234	Mg 234	0	0
24	B	2	Total 2	Mg 2	0	0
24	C	3	Total 3	Mg 3	0	0
24	D	3	Total 3	Mg 3	0	0
24	E	1	Total 1	Mg 1	0	0
24	F	1	Total 1	Mg 1	0	0
24	J	1	Total 1	Mg 1	0	0
24	P	2	Total 2	Mg 2	0	0
24	Q	1	Total 1	Mg 1	0	0
24	S	1	Total 1	Mg 1	0	0

- Molecule 25 is (1S,2S,3R,4S,6R)-4,6-diamino-3-{[(2S,3R)-3-amino-6-(aminomethyl)-3,4-dihydro-2H-pyran-2-yl]oxy}-2-hydroxycyclohexyl 3-deoxy-4-C-methyl-3-(methylamino)-beta-L-arabinopyranoside (three-letter code: SIS) (formula: $C_{19}H_{37}N_5O_7$) (labeled as "Ligand of

Interest" by depositor).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
25	A	1	Total 31	C 19	N 5	O 7	0	0

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

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

- Molecule 27 is water.

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

Continued on next page...

*Continued from previous page...*

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

3 Residue-property plots ⓘ

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.
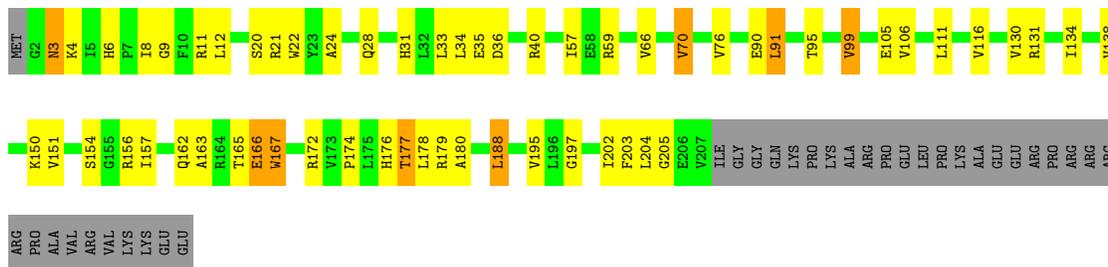
- Molecule 1: 30S Ribosomal RNA rRNA
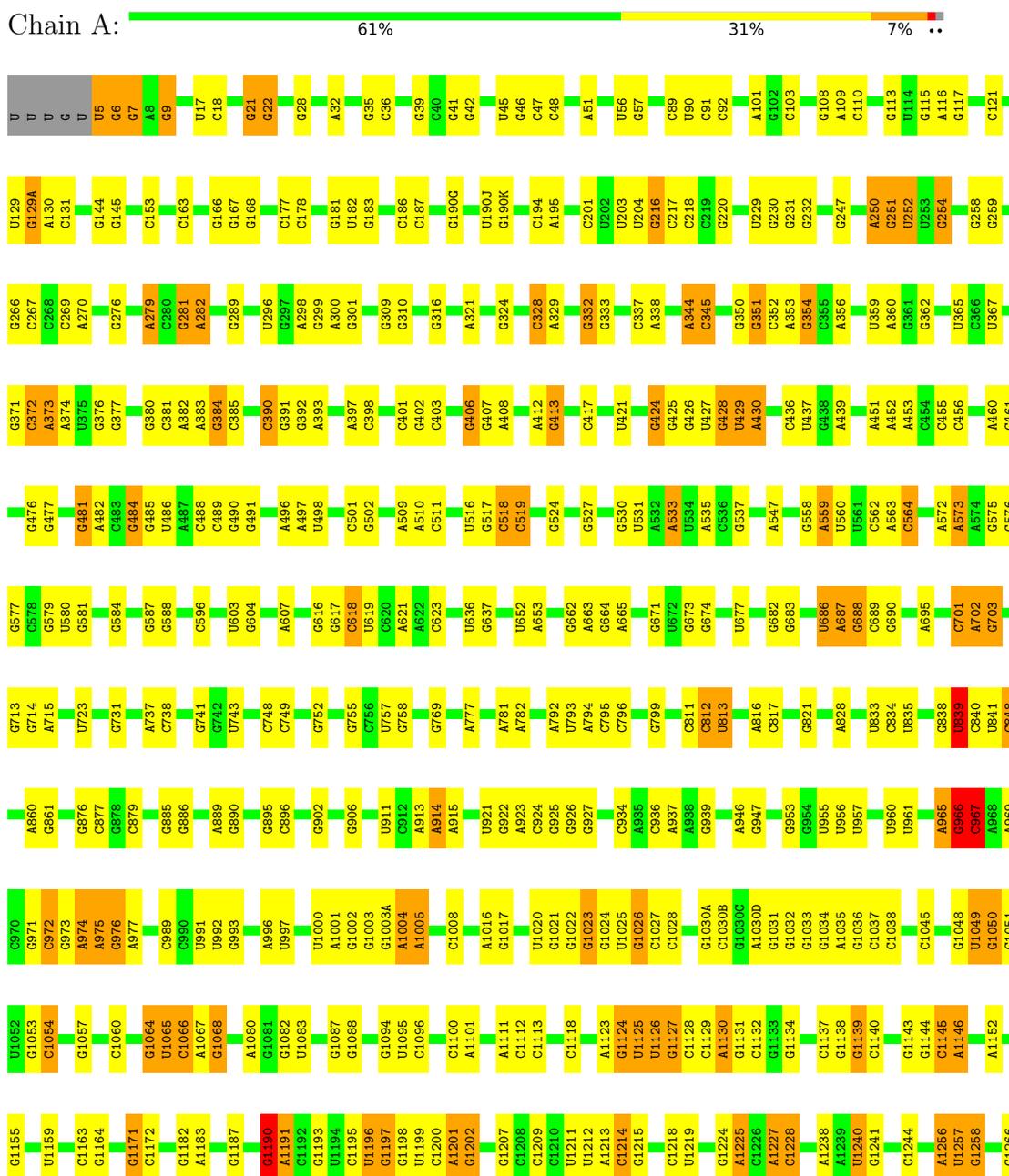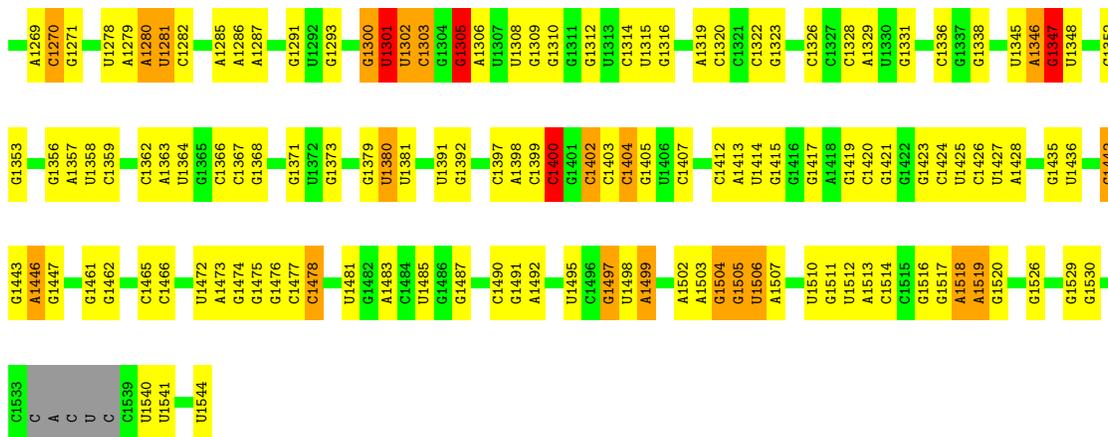
• Molecule 2: 30S ribosomal protein S2

• Molecule 3: 30S ribosomal protein S3

• Molecule 4: 30S ribosomal protein S4

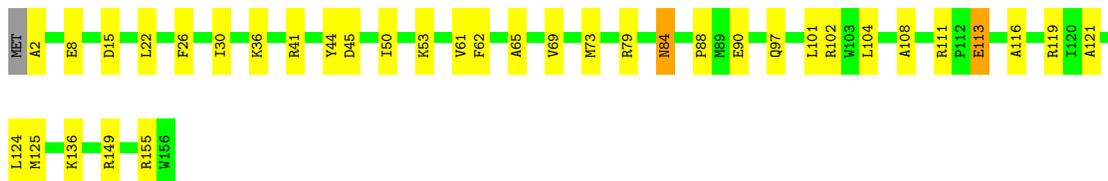• Molecule 5: 30S ribosomal protein S5

Chain E:

67%　　　　　　22%　•　7%

MET PRO GLU THR D5 F6 E7 E8 I11 L12 A17 R18 M19 Q20 R24 R25 F26 R27 F28 L31 V34 V41 G42 L43 K47 Y60 R64 N65 T75 I76 P77 H78 E79 I80 E81 V82 K92 P93 G97 T98 G99 V105 P106 R107 D117 I118

R126 I131 A132 A138 L142 D147 V148 L151 G154 GLU ALA HIS ALA GLN ALA GLN GLY

• Molecule 6: 30S ribosomal protein S6

Chain F:

83%　　　　　　15%　•

M1 L10 N11 E22 K23 E24 Q27 E31 N32 Y33 Y59 W62 Y63 Q64 D70 L75 A76 R77 R82 R86 K92 A101

• Molecule 7: 30S ribosomal protein S7

Chain G:

76%　　　　　　22%　••

MET A2 E8 D15 L22 F26 I30 K36 R41 Y44 D45 I50 K53 V61 F62 A65 V69 M73 R79 N84 P88 M89 E90 Q97 L101 R102 W103 L104 A108 R111 P112 E113 A116 R119 I120 A121

L124 M125 K136 R149 R155 W156

• Molecule 8: 30S ribosomal protein S8

Chain H:

77%　　　　　　17%　7%

M1 L2 T3 D4 P5 I6 T17 R18 V19 Y20 K21 V26 S29 R30 F31 L59 L63 Q70 G71 P72 Q78 R85 I86 S87 K88 R91 R92 K98 L107 L112 S113 D121 R122 E123 L127 L133 I134 C135 E136 V137 W138

• Molecule 9: 30S ribosomal protein S9

Chain I:

69%　　　　　　26%　5%　•

MET E2 T7 G8 R9 R10 K11 E12 A15 R16 G39 A43 V44 A45 A46 R51 L56 T64 V65 R66 G69 Q73 I74 D75 A76 I77 K78 L79 R83 L96 K97 L102 T103 R104 V109 E110 R111 K112 K116 H117 K118 A119 R120 R121 A122 P123

Q124 K127 R128
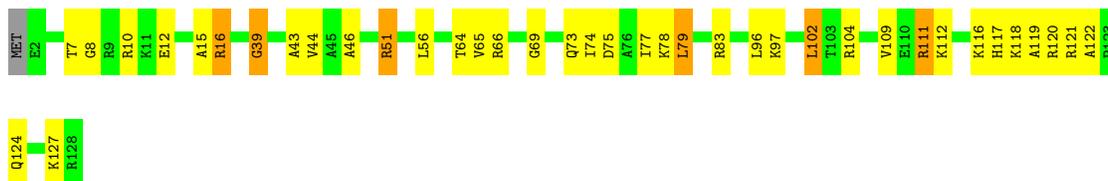
• Molecule 10: 30S ribosomal protein S10

Chain J:

61%　　　　　29%　••　7%

MET PRO K3 I4 R5 I6 K7 L8 K14 S30 S35 G36 P37 I38 P39 L40 R45 V49 I50 R51 G52 P53 F54 K55 H56 K57 R60 E61 H62 T67 N78 R79 K80 T81 I82 E83 Q84 L85 D89 L90 P91 E95 I96 T100 VAL GLY GLY ARG

• Molecule 11: 30S ribosomal protein S11

Chain K:
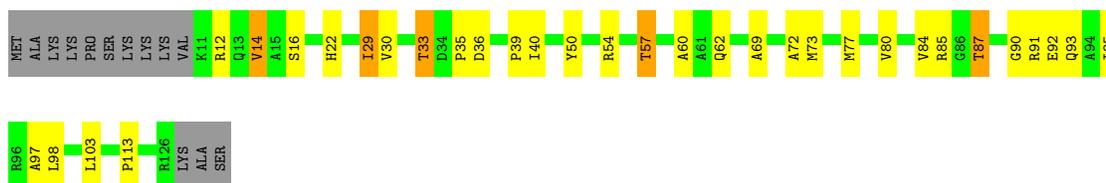
• Molecule 12: 30S ribosomal protein S12

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:
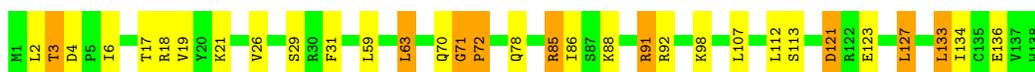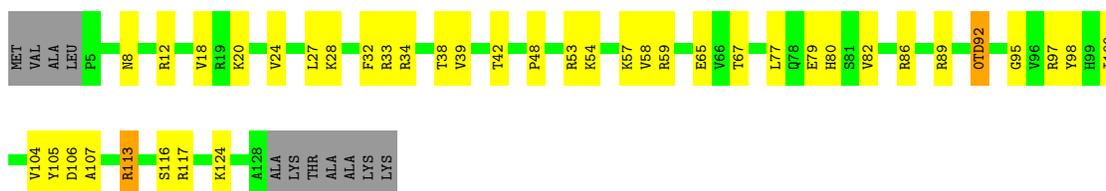
• Molecule 16: 30S ribosomal protein S16

Chain P:

• Molecule 17: 30S ribosomal protein S17

Chain Q:

69% 22% • 6%

MET P2 K3 D13 L22 H29 P30 K34 V35 I36 S39 K40 K41 Y42 H45 D46 P47 L53 I59 I60 E61 S62 R63 P64 I65 S66 K67 R68 K69 R70 V73 L76 D83 K87 L98 S99 K100 ARG GLY GLY LYS ALA

• Molecule 18: 30S ribosomal protein S18

Chain R:

58% 19% • 20%

MET SER THR LYS ASN ALA LYS PRO LYS LYS GLU ALA GLN ARG ARG PRO SER ARG K19 V22 L31 R32 D33 N36 R42 E46 T47 G48 R53 R54 R55 T56 G57 L58 R64 T69 R74 E83 R87 K88

• Molecule 19: 30S ribosomal protein S19

Chain S:

59% 22% 5% 14%

MET P2 R3 S4 L5 K6 V9 F10 V11 D12 D13 H14 L15 K18 L22 K28 R29 L30 I31 S38 T39 V45 T48 I49 A50 H57 I62 T63 E64 N65 M66 T79 Y80 R81 GLY HIS GLY LYS GLU ALA LYS ALA THR LYS LYS LYS

• Molecule 20: 30S ribosomal protein S20

Chain T:

70% 21% • 7%

MET ALA GLN LYS LYS PRO LYS R8 N9 R17 K29 I33 Q45 E50 L53 R57 S61 L62 K68 G69 S70 T71 L72 H73 K74 N75 L84 R89 Q90 L91 L92 E93 G96 I100 G101 G102 G103 L104 S105 A106

• Molecule 21: 30S ribosomal protein Thx

Chain U:

56% 33% 11%

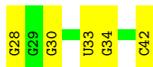MET G2 K3 G4 D5 R6 R10 G11 K12 I13 W14 R22 K25 LYS LYS

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

Chain Y:

50% 33% 17%

U1 U5 U6

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

Chain W:

67% 33%

G28 G29 G30 U33 G34 C42

4 Data and refinement statistics ⓘ

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)	39.88 – 3.60 39.88 – 2.92	Depositor EDS
% Data completeness (in resolution range)	99.0 (39.88-3.60) 71.6 (39.88-2.92)	Depositor EDS
R_{merge}	1.41	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	-0.50 (at 2.90Å)	Xtriage
Refinement program	PHENIX DEV-3318	Depositor
R, R_{free}	0.186 , 0.228 0.186 , 0.228	Depositor DCC
R_{free} test set	2000 reflections (0.68%)	wwPDB-VP
Wilson B-factor (Å^2)	63.2	Xtriage
Anisotropy	0.293	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.21 , 54.7	EDS
L-test for twinning[2]	$< |L| > = 0.44, < L^2 > = 0.27$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	52689	wwPDB-VP
Average B, all atoms (Å^2)	146.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 1.26% 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: 4OC, 5MC, 0TD, 2MG, ZN, UR3, MG, 7MG, MA6, SIS, M2G, PSU

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	0.22	0/36139	0.78	15/56396 (0.0%)
2	B	0.26	0/1935	0.44	0/2609
3	C	0.25	0/1636	0.46	0/2205
4	D	0.25	0/1733	0.41	0/2318
5	E	0.27	0/1162	0.45	0/1564
6	F	0.24	0/856	0.42	0/1154
7	G	0.25	0/1276	0.39	0/1709
8	H	0.26	0/1136	0.45	0/1527
9	I	0.26	0/1029	0.47	1/1379 (0.1%)
10	J	0.25	0/805	0.51	0/1082
11	K	0.26	0/879	0.46	0/1187
12	L	0.25	0/977	0.49	0/1306
13	M	0.24	0/947	0.45	0/1270
14	N	0.26	0/501	0.44	0/664
15	O	0.23	0/740	0.40	0/987
16	P	0.25	0/716	0.44	0/963
17	Q	0.26	0/836	0.47	1/1117 (0.1%)
18	R	0.26	0/579	0.43	0/768
19	S	0.25	0/661	0.48	0/890
20	T	0.24	0/765	0.41	0/1007
21	U	0.23	0/212	0.42	0/277
22	Y	0.22	0/128	0.88	0/196
23	W	0.23	0/357	0.83	0/555
All	All	0.23	0/56005	0.69	17/83130 (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.

W O R L D W I D E
wwPDB
PROTEIN DATA BANK

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 17 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	839	U	C2-N1-C1'	7.87	127.14	117.70
1	A	839	U	N1-C2-O2	6.69	127.48	122.80
1	A	839	U	N3-C2-O2	-6.31	117.78	122.20
1	A	1301	U	P-O3'-C3'	5.63	126.46	119.70
1	A	328	C	C2-N1-C1'	5.50	124.85	118.80

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 ⓘ

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	32644	0	16508	320	0
2	B	1900	0	1951	34	0
3	C	1612	0	1676	35	0
4	D	1703	0	1763	41	0
5	E	1146	0	1207	26	0
6	F	843	0	857	9	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	1177	19	0
9	I	1010	0	1037	27	0
10	J	792	0	835	26	0
11	K	864	0	881	17	0
12	L	972	0	1058	20	0
13	M	937	0	995	21	0
14	N	492	0	529	25	0
15	O	729	0	768	9	0
16	P	700	0	720	7	0
17	Q	823	0	893	15	0
18	R	574	0	644	12	0
19	S	647	0	673	14	0
20	T	763	0	861	12	0
21	U	208	0	221	6	0
22	Y	117	0	62	2	0
23	W	319	0	164	2	0
24	A	234	0	0	0	0
24	B	2	0	0	0	0
24	C	3	0	0	0	0
24	D	3	0	0	0	0
24	E	1	0	0	0	0
24	F	1	0	0	0	0
24	J	1	0	0	0	0
24	P	2	0	0	0	0
24	Q	1	0	0	0	0
24	S	1	0	0	0	0
25	A	31	0	37	3	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	1	0	0	0	0
27	E	5	0	0	0	0
27	K	1	0	0	0	0
27	L	2	0	0	0	0
27	N	2	0	0	0	0
27	Q	2	0	0	0	0
27	T	2	0	0	0	0
All	All	52689	0	36813	638	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 7.

The worst 5 of 638 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	1.02	1.00
8:H:86:ILE:HD11	8:H:136:GLU:HG3	1.57	0.86
1:A:664:G:H22	1:A:741:G:H1	1.26	0.81
1:A:1028:C:H42	1:A:1033:G:H1	1.26	0.79
3:C:33:LEU:HD21	14:N:53:LEU:HD22	1.65	0.78

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	232/256 (91%)	218 (94%)	14 (6%)	0	100	100
3	C	204/239 (85%)	184 (90%)	20 (10%)	0	100	100
4	D	206/209 (99%)	202 (98%)	4 (2%)	0	100	100
5	E	148/162 (91%)	141 (95%)	7 (5%)	0	100	100
6	F	99/101 (98%)	97 (98%)	2 (2%)	0	100	100
7	G	153/156 (98%)	147 (96%)	6 (4%)	0	100	100
8	H	136/138 (99%)	130 (96%)	4 (3%)	2 (2%)	10	47
9	I	125/128 (98%)	114 (91%)	11 (9%)	0	100	100
10	J	96/105 (91%)	81 (84%)	12 (12%)	3 (3%)	4	32
11	K	114/129 (88%)	109 (96%)	5 (4%)	0	100	100
12	L	121/135 (90%)	111 (92%)	10 (8%)	0	100	100
13	M	116/126 (92%)	103 (89%)	13 (11%)	0	100	100
14	N	58/61 (95%)	49 (84%)	9 (16%)	0	100	100
15	O	85/89 (96%)	83 (98%)	2 (2%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
16	P	81/88 (92%)	76 (94%)	5 (6%)	0	100	100
17	Q	97/105 (92%)	94 (97%)	3 (3%)	0	100	100
18	R	68/88 (77%)	64 (94%)	4 (6%)	0	100	100
19	S	78/93 (84%)	71 (91%)	7 (9%)	0	100	100
20	T	97/106 (92%)	88 (91%)	9 (9%)	0	100	100
21	U	22/27 (82%)	22 (100%)	0	0	100	100
All	All	2336/2541 (92%)	2184 (94%)	147 (6%)	5 (0%)	47	79

All (5) Ramachandran outliers are listed below:

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

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	202/220 (92%)	175 (87%)	27 (13%)	4	23
3	C	160/188 (85%)	142 (89%)	18 (11%)	6	30
4	D	180/181 (99%)	162 (90%)	18 (10%)	7	35
5	E	115/123 (94%)	103 (90%)	12 (10%)	7	33
6	F	90/90 (100%)	86 (96%)	4 (4%)	28	63
7	G	126/127 (99%)	117 (93%)	9 (7%)	14	48
8	H	119/119 (100%)	108 (91%)	11 (9%)	9	39
9	I	98/99 (99%)	86 (88%)	12 (12%)	5	26
10	J	87/92 (95%)	77 (88%)	10 (12%)	5	29

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
11	K	88/99 (89%)	79 (90%)	9 (10%)	7	34
12	L	103/110 (94%)	92 (89%)	11 (11%)	6	32
13	M	94/101 (93%)	85 (90%)	9 (10%)	8	37
14	N	49/50 (98%)	47 (96%)	2 (4%)	30	64
15	O	79/80 (99%)	66 (84%)	13 (16%)	2	15
16	P	72/74 (97%)	66 (92%)	6 (8%)	11	42
17	Q	94/97 (97%)	88 (94%)	6 (6%)	17	52
18	R	61/77 (79%)	57 (93%)	4 (7%)	16	51
19	S	71/80 (89%)	61 (86%)	10 (14%)	3	21
20	T	76/82 (93%)	65 (86%)	11 (14%)	3	20
21	U	19/22 (86%)	18 (95%)	1 (5%)	22	58
All	All	1983/2111 (94%)	1780 (90%)	203 (10%)	7	34

5 of 203 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
10	J	60	ARG
13	M	17	VAL
20	T	91	LEU
10	J	89	ASP
12	L	18	VAL

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
2	B	19	HIS
2	B	204	ASN
4	D	129	ASN
5	E	78	HIS

5.3.3 RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
1	A	1504/1522 (98%)	237 (15%)	46 (3%)
22	Y	5/6 (83%)	2 (40%)	0
23	W	14/15 (93%)	3 (21%)	0

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
All	All	1523/1543 (98%)	242 (15%)	46 (3%)

5 of 242 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
1	A	6	G
1	A	7	G
1	A	9	G
1	A	22	G
1	A	32	A

5 of 46 RNA pucker outliers are listed below:

Mol	Chain	Res	Type
1	A	992	U
1	A	1201	A
1	A	1049	U
1	A	1145	C
1	A	1257	U

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

17 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	MA6	A	1518[A]	1	19,26,27	0.97	1 (5%)	18,38,41	0.69	0
1	M2G	A	966	1	20,27,28	2.09	4 (20%)	22,40,43	2.41	4 (18%)
1	MA6	A	1519[B]	1	19,26,27	1.25	3 (15%)	18,38,41	0.59	0
12	0TD	L	92	12	4,9,10	0.83	0	3,11,13	1.78	1 (33%)
1	UR3	A	1498	1	14,22,23	1.02	2 (14%)	15,32,35	1.08	1 (6%)
1	PSU	A	1541	1,24	17,21,22	1.06	2 (11%)	20,30,33	3.17	6 (30%)
1	4OC	A	1402	1	16,23,24	0.92	1 (6%)	17,32,35	0.78	0

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	1.02	1 (5%)	20,30,33	3.16	7 (35%)
1	2MG	A	1207	1	19,26,27	2.36	4 (21%)	21,38,41	1.99	3 (14%)
1	5MC	A	1407	1	15,22,23	1.04	1 (6%)	19,32,35	1.07	1 (5%)
1	7MG	A	527	1	22,26,27	2.15	6 (27%)	28,39,42	1.76	6 (21%)
1	5MC	A	1400	1	15,22,23	0.93	0	19,32,35	1.03	2 (10%)
1	5MC	A	1404	1	15,22,23	0.96	1 (6%)	19,32,35	1.10	2 (10%)
1	MA6	A	1518[B]	1	19,26,27	1.20	2 (10%)	18,38,41	0.75	0
1	5MC	A	967	1	15,22,23	0.91	0	19,32,35	1.12	3 (15%)
1	MA6	A	1519[A]	1	19,26,27	1.03	1 (5%)	18,38,41	0.66	0
1	PSU	A	516	1,24	17,21,22	1.07	2 (11%)	20,30,33	3.15	6 (30%)

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	MA6	A	1518[A]	1	-	0/7/29/30	0/3/3/3
1	M2G	A	966	1	-	3/7/29/30	0/3/3/3
1	MA6	A	1519[B]	1	-	3/7/29/30	0/3/3/3
12	0TD	L	92	12	-	1/3/12/14	-
1	UR3	A	1498	1	-	0/5/25/26	0/2/2/2
1	PSU	A	1541	1,24	-	1/7/25/26	0/2/2/2
1	4OC	A	1402	1	-	2/9/29/30	0/2/2/2
1	PSU	A	1540	1	-	0/7/25/26	0/2/2/2
1	2MG	A	1207	1	-	0/5/27/28	0/3/3/3
1	5MC	A	1407	1	-	0/5/25/26	0/2/2/2
1	7MG	A	527	1	-	2/7/37/38	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	MA6	A	1518[B]	1	-	2/7/29/30	0/3/3/3
1	5MC	A	967	1	-	2/5/25/26	0/2/2/2
1	MA6	A	1519[A]	1	-	4/7/29/30	0/3/3/3
1	PSU	A	516	1,24	-	0/7/25/26	0/2/2/2

The worst 5 of 31 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	1207	2MG	C2-N2	7.52	1.40	1.34
1	A	966	M2G	C6-N1	5.98	1.43	1.33
1	A	1207	2MG	C6-N1	5.87	1.43	1.33
1	A	527	7MG	C4-N3	5.67	1.41	1.34
1	A	527	7MG	C2-N2	4.97	1.43	1.33

The worst 5 of 42 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	1541	PSU	N3-C2-N1	-10.97	119.71	128.43
1	A	1540	PSU	N3-C2-N1	-10.80	119.85	128.43
1	A	516	PSU	N3-C2-N1	-10.75	119.88	128.43
1	A	966	M2G	C5-C6-N1	-8.03	112.44	123.43
1	A	1207	2MG	C5-C6-N1	-7.52	113.14	123.43

There are no chirality outliers.

5 of 22 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
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'
1	A	1402	4OC	O4'-C4'-C5'-O5'

There are no ring outliers.

10 monomers are involved in 14 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	A	1518[A]	MA6	1	0
1	A	966	M2G	2	0
1	A	1519[B]	MA6	3	0
12	L	92	0TD	2	0
1	A	1402	4OC	1	0
1	A	1400	5MC	1	0
1	A	1404	5MC	1	0
1	A	1518[B]	MA6	4	0
1	A	967	5MC	1	0
1	A	1519[A]	MA6	1	0

5.5 Carbohydrates ⓘ

There are no monosaccharides in this entry.

5.6 Ligand geometry ⓘ

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	#$|Z| > 2$
25	SIS	A	1835	-	28,33,33	1.29	2 (7%)	28,49,49	0.90	1 (3%)

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
25	SIS	A	1835	-	-	2/9/65/65	0/3/3/3

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
25	A	1835	SIS	C31-C41	-4.86	1.39	1.50
25	A	1835	SIS	C61-C51	-3.74	1.39	1.48

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
25	A	1835	SIS	C93-N33-C33	-3.17	109.76	114.38

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
25	A	1835	SIS	C23-C33-N33-C93
25	A	1835	SIS	C52-C42-O11-C11

There are no ring outliers.

1 monomer is involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
25	A	1835	SIS	3	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 SIS A 1835

Bond lengths

Bond angles

Torsions

Rings
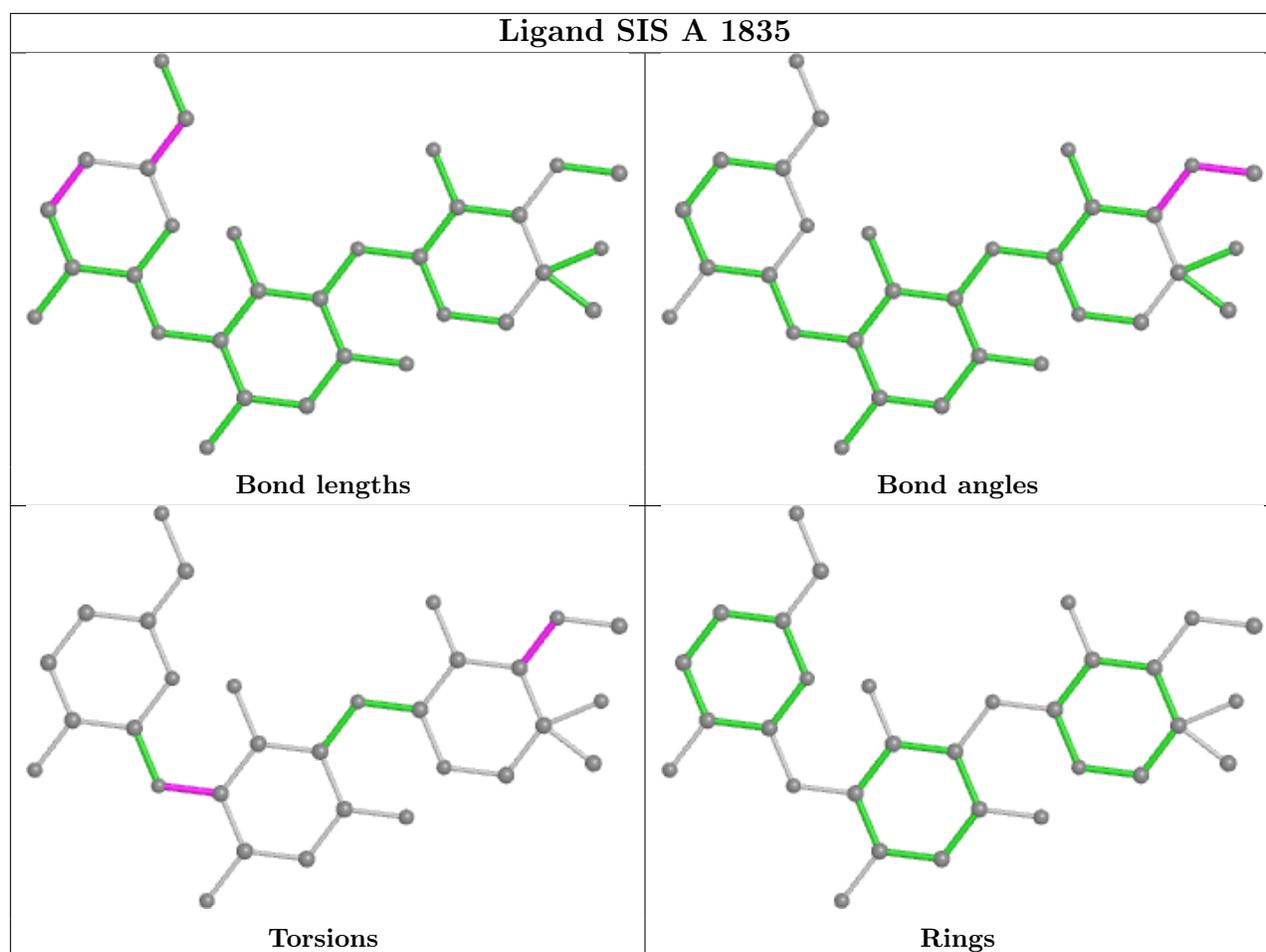
5.7 Other polymers ⓘ

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 ⓘ

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

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

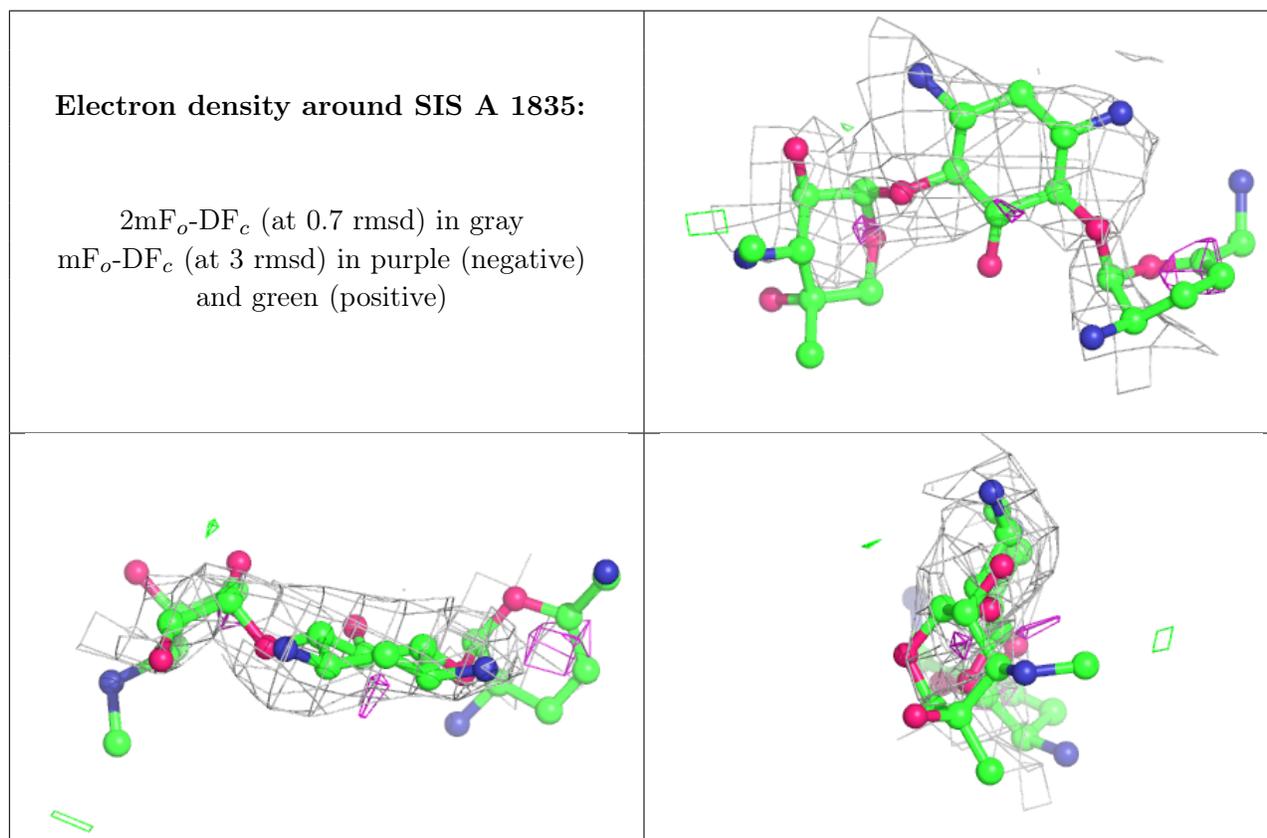Unable to reproduce the depositors R factor - this section is therefore empty.

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 SIS A 1835:

$2mF_o$-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.