



# wwPDB X-ray Structure Validation Summary Report ⓘ

Jan 3, 2022 – 02:17 PM JST

PDB ID : 7DUJ  
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''Bz bound  
Authors : DeMirici, 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](mailto: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 ⓘ symbol.

---

The following versions of software and data (see [references ⓘ](#)) 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

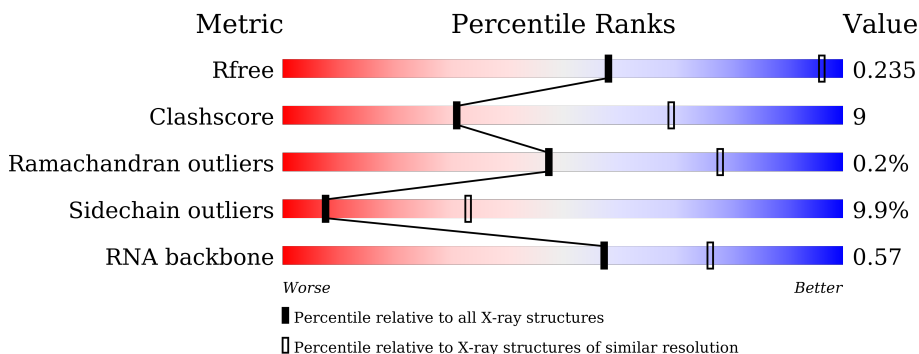
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.75 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  $\geq 3$ , 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  $\leq 5\%$

Mol	Chain	Length	Quality of chain
1	A	1522	53% (green), 38% (yellow), 7% (orange), 2% (red), 0% (grey)
2	B	256	62% (green), 26% (yellow), 9% (orange), 3% (red), 0% (grey)
3	C	239	58% (green), 24% (yellow), 14% (orange), 4% (red), 0% (grey)
4	D	209	71% (green), 26% (yellow), 3% (orange), 0% (red), 0% (grey)
5	E	162	70% (green), 17% (yellow), 6% (orange), 7% (red), 0% (grey)
6	F	101	75% (green), 23% (yellow), 2% (orange), 0% (red), 0% (grey)

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
7	G	156	74% 24% ..
8	H	138	78% 17% 5%
9	I	128	62% 36% ..
10	J	105	53% 35% . . 7%
11	K	129	62% 23% 5% 10%
12	L	135	60% 27% 5% 8%
13	M	126	60% 31% . 6%
14	N	61	49% 48% . .
15	O	89	69% 24% 6% .
16	P	88	76% 18% 6%
17	Q	105	66% 25% . . 6%
18	R	88	57% 20% . 20%
19	S	93	57% 23% 6% 14%
20	T	106	65% 25% . 7%
21	U	27	63% 26% 11%
22	Y	6	50% 33% 17%
23	W	15	73% 27%

## 2 Entry composition [i](#)

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
			Total	C	N	O	P			
1	A	1512	32644	14540	6040	10546	1518	0	6	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
			Total	C	N	O	S			
7	G	155	1257	781	252	218	6	0	0	0

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

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

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

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

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

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

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

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

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

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

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
13	M	118	937	579	193	163	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
			Total	C	N	O			
21	U	24	208	128	50	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
			Total	C	N	O	P			
22	Y	6	117	54	12	46	5	0	0	0

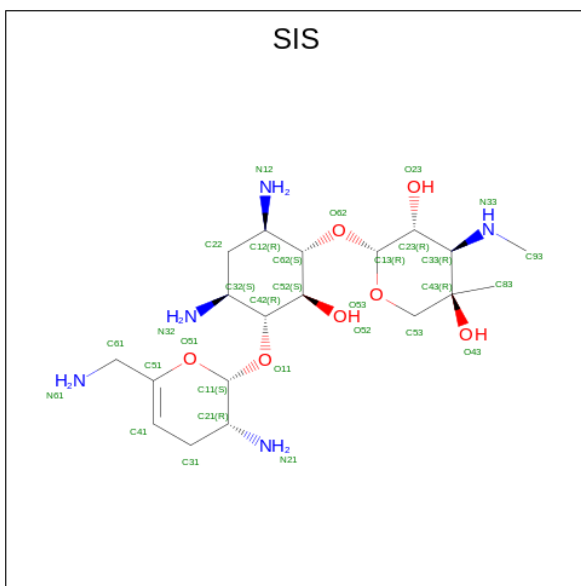
- Molecule 23 is a RNA chain called RNA (5'-R(\*GP\*GP\*GP\*AP\*UP\*UP\*GP\*AP\*AP\*AP\*AP\*UP\*CP\*CP\*C)-3').

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

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

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

- 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<sub>19</sub>H<sub>37</sub>N<sub>5</sub>O<sub>7</sub>) (labeled as "Ligand of Interest" by depositor).



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

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

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

- Molecule 27 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
27	A	221	221	221	0	0
27	C	1	1	1	0	0
27	D	1	1	1	0	0
27	E	6	6	6	0	0

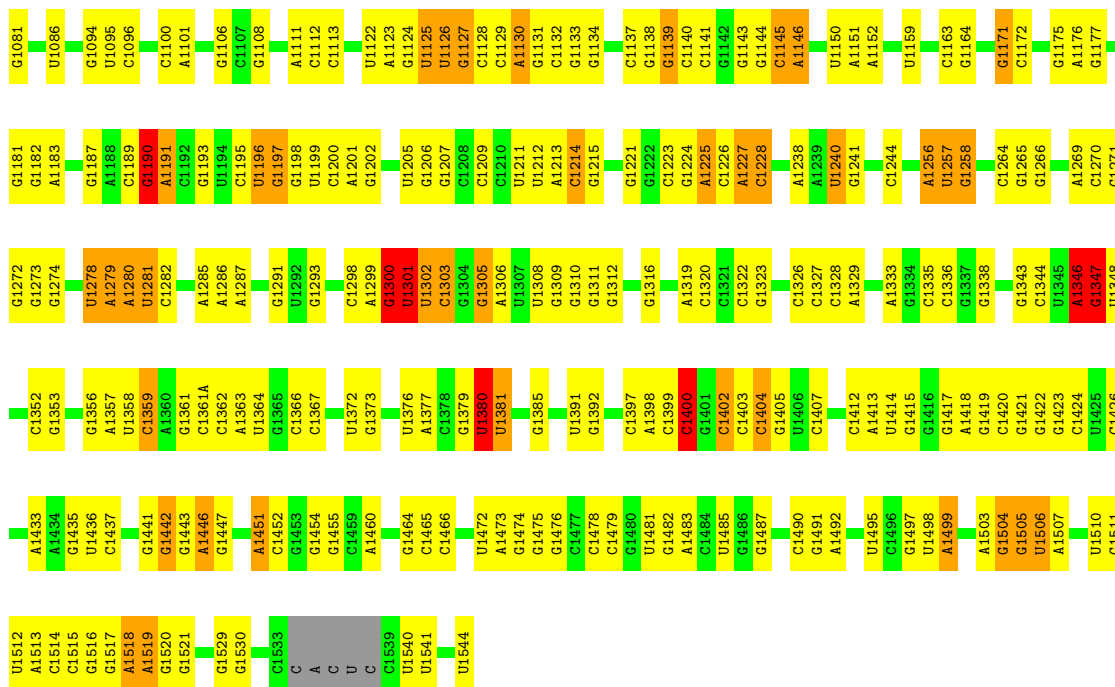
*Continued on next page...*



*Continued from previous page...*

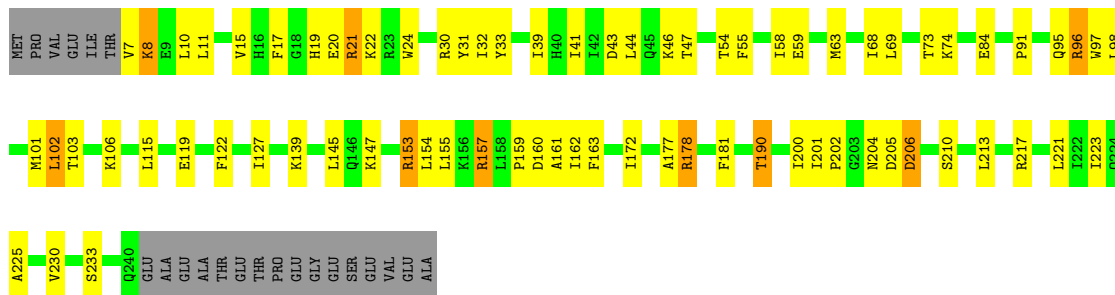
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
27	K	1	Total O 1 1	0	0
27	L	3	Total O 3 3	0	0
27	N	2	Total O 2 2	0	0
27	O	1	Total O 1 1	0	0
27	Q	1	Total O 1 1	0	0
27	T	2	Total O 2 2	0	0





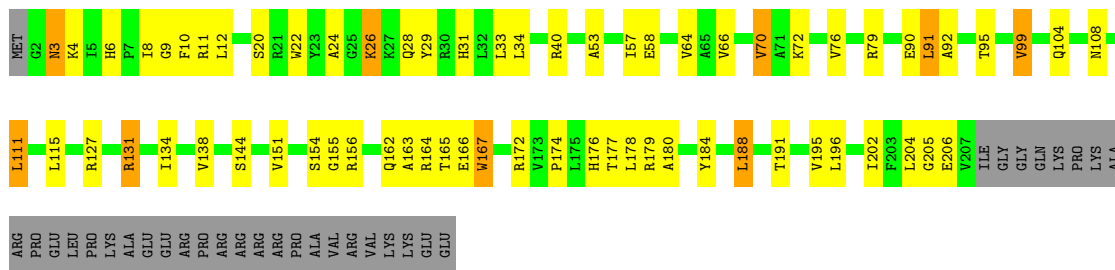
- Molecule 2: 30S ribosomal protein S2

Chain B: 62% 26% 9%



- Molecule 3: 30S ribosomal protein S3

Chain C: 58% 24% 14%



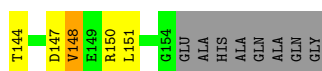
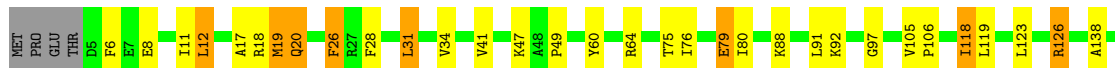
- Molecule 4: 30S ribosomal protein S4

Chain D: 71% 26%



- Molecule 5: 30S ribosomal protein S5

Chain E: 70% 17% 6% 7%



- Molecule 6: 30S ribosomal protein S6

Chain F: 75% 23%



- Molecule 7: 30S ribosomal protein S7

Chain G: 74% 24%



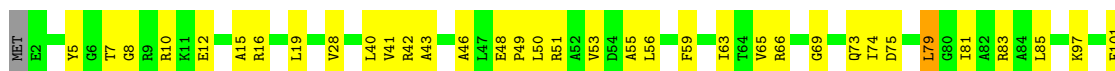
- Molecule 8: 30S ribosomal protein S8

Chain H: 78% 17% 5%



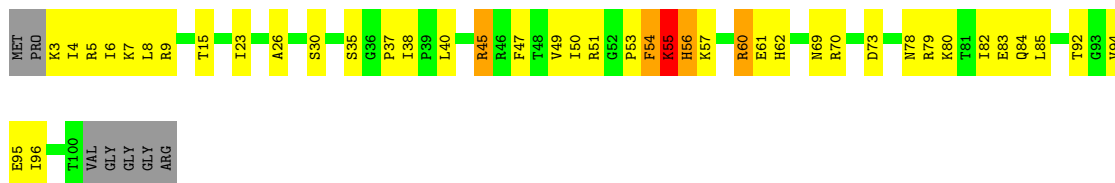
- Molecule 9: 30S ribosomal protein S9

Chain I: 62% 36%



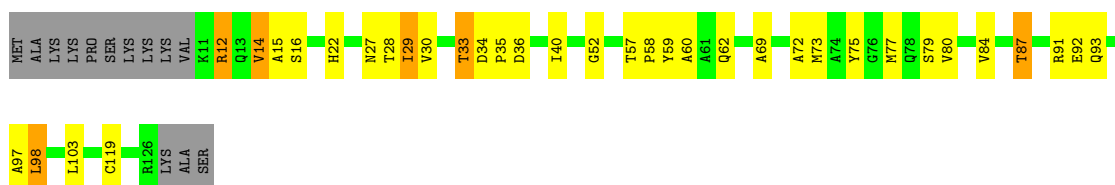
- Molecule 10: 30S ribosomal protein S10

Chain J:  53% 35% 7%



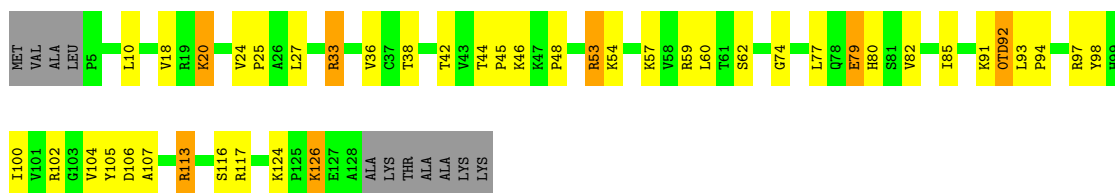
- Molecule 11: 30S ribosomal protein S11

Chain K:  62% 23% 5% 10%



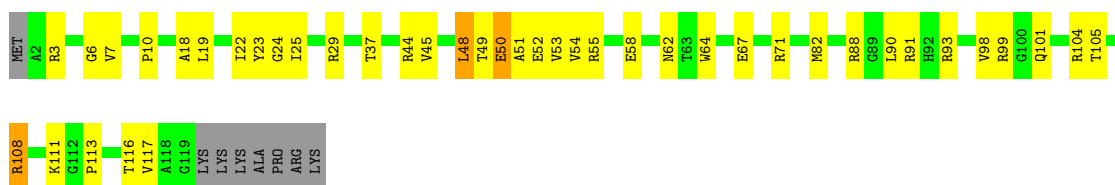
- Molecule 12: 30S ribosomal protein S12

Chain L:  60% 27% 5% 8%



- Molecule 13: 30S ribosomal protein S13

Chain M:  60% 31% 6%



- Molecule 14: 30S ribosomal protein S14 type Z

Chain N:  49% 48% 2%




- Molecule 15: 30S ribosomal protein S15

Chain O:  69% 24% 6%



- Molecule 16: 30S ribosomal protein S16

Chain P:  76% 18% 6%



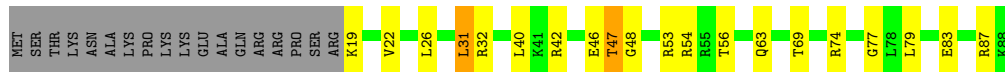
- Molecule 17: 30S ribosomal protein S17

Chain Q:  66% 25% 6%



- Molecule 18: 30S ribosomal protein S18

Chain R:  57% 20% 20%



- Molecule 19: 30S ribosomal protein S19

Chain S:  57% 23% 6% 14%



- Molecule 20: 30S ribosomal protein S20

Chain T:  65% 25% 7%



- Molecule 21: 30S ribosomal protein Thx

Chain U:  63% 26% 11%



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

Chain Y:  50% 33% 17%



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

Chain W:  73% 27%



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 41 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	400.00Å 400.00Å 173.00Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	39.80 – 3.75 39.80 – 2.91	Depositor EDS
% Data completeness (in resolution range)	99.3 (39.80-3.75) 62.6 (39.80-2.91)	Depositor EDS
$R_{merge}$	1.21	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	-0.83 (at 2.90Å)	Xtrriage
Refinement program	PHENIX DEV-3318	Depositor
R, $R_{free}$	0.189 , 0.235 0.189 , 0.235	Depositor DCC
$R_{free}$ test set	2000 reflections (0.68%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	49.2	Xtrriage
Anisotropy	0.151	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.12 , -10.0	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.45$ , $\langle L^2 \rangle = 0.28$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.92	EDS
Total number of atoms	52689	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	181.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 1.15% of the height of the origin peak. No significant pseudotranslation is detected.*

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

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.



## 5 Model quality i

### 5.1 Standard geometry i

Bond lengths and bond angles in the following residue types are not validated in this section: 2MG, SIS, MA6, ZN, 4OC, PSU, 0TD, 5MC, M2G, UR3, MG, 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	0.24	0/36139	0.80	20/56396 (0.0%)
2	B	0.26	0/1935	0.43	0/2609
3	C	0.26	0/1636	0.46	0/2205
4	D	0.26	0/1733	0.41	0/2318
5	E	0.27	0/1162	0.45	0/1564
6	F	0.25	0/856	0.42	0/1154
7	G	0.25	0/1276	0.39	0/1709
8	H	0.28	0/1136	0.46	0/1527
9	I	0.26	0/1029	0.47	0/1379
10	J	0.26	0/805	0.52	0/1082
11	K	0.26	0/879	0.46	0/1187
12	L	0.26	0/977	0.49	0/1306
13	M	0.25	0/947	0.46	0/1270
14	N	0.26	0/501	0.44	0/664
15	O	0.24	0/740	0.40	0/987
16	P	0.25	0/716	0.44	0/963
17	Q	0.27	0/836	0.47	1/1117 (0.1%)
18	R	0.26	0/579	0.43	0/768
19	S	0.26	0/661	0.49	0/890
20	T	0.25	0/765	0.40	0/1007
21	U	0.24	0/212	0.40	0/277
22	Y	0.24	0/128	0.91	0/196
23	W	0.26	0/357	0.83	0/555
All	All	0.25	0/56005	0.71	21/83130 (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.

Mol	Chain	#Chirality outliers	#Planarity outliers
3	C	0	1
8	H	0	1
10	J	0	2
13	M	0	3
All	All	0	7

There are no bond length outliers.

The worst 5 of 21 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	839	U	C2-N1-C1'	7.52	126.73	117.70
1	A	839	U	N1-C2-O2	7.37	127.95	122.80
1	A	839	U	N3-C2-O2	-6.64	117.55	122.20
1	A	1190	G	P-O3'-C3'	5.66	126.49	119.70
1	A	254	G	O5'-P-OP1	-5.53	100.73	105.70

There are no chirality outliers.

5 of 7 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	32644	0	16508	429	0
2	B	1900	0	1951	37	0
3	C	1612	0	1677	46	0
4	D	1703	0	1763	37	0
5	E	1146	0	1207	25	0
6	F	843	0	857	14	0
7	G	1257	0	1296	27	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	19	0
9	I	1010	0	1037	29	0
10	J	792	0	835	32	0
11	K	864	0	881	20	0
12	L	972	0	1058	31	0
13	M	937	0	995	28	0
14	N	492	0	529	30	0
15	O	729	0	768	15	0
16	P	700	0	720	8	0
17	Q	823	0	893	18	0
18	R	574	0	644	14	0
19	S	647	0	673	16	0
20	T	763	0	861	13	0
21	U	208	0	221	4	0
22	Y	117	0	62	3	0
23	W	319	0	164	2	0
24	A	232	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	H	1	0	0	0	0
24	J	1	0	0	0	0
24	K	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	2	0
26	D	1	0	0	0	0
26	N	1	0	0	0	0
27	A	221	0	0	2	0
27	C	1	0	0	0	0
27	D	1	0	0	0	0
27	E	6	0	0	0	0
27	K	1	0	0	0	0
27	L	3	0	0	0	0
27	N	2	0	0	0	0
27	O	1	0	0	0	0
27	Q	1	0	0	0	0
27	T	2	0	0	1	0
All	All	52689	0	36813	791	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 791 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	0.96	0.95
1:A:1028:C:H42	1:A:1033:G:H1	1.21	0.87
20:T:100:ILE:HG22	20:T:102:GLY:H	1.43	0.84
1:A:1516[B]:G:N2	1:A:1519[B]:MA6:OP2	2.12	0.82
1:A:145:G:H1	1:A:177:C:H42	1.26	0.80

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%)	217 (94%)	15 (6%)	0	100	100
3	C	204/239 (85%)	180 (88%)	24 (12%)	0	100	100
4	D	206/209 (99%)	197 (96%)	9 (4%)	0	100	100
5	E	148/162 (91%)	143 (97%)	5 (3%)	0	100	100
6	F	99/101 (98%)	98 (99%)	1 (1%)	0	100	100
7	G	153/156 (98%)	144 (94%)	9 (6%)	0	100	100
8	H	136/138 (99%)	129 (95%)	5 (4%)	2 (2%)	10	45
9	I	125/128 (98%)	114 (91%)	11 (9%)	0	100	100
10	J	96/105 (91%)	81 (84%)	12 (12%)	3 (3%)	4	33
11	K	114/129 (88%)	107 (94%)	7 (6%)	0	100	100
12	L	121/135 (90%)	110 (91%)	11 (9%)	0	100	100
13	M	116/126 (92%)	101 (87%)	15 (13%)	0	100	100

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
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%)	79 (98%)	2 (2%)	0	100	100
17	Q	97/105 (92%)	92 (95%)	5 (5%)	0	100	100
18	R	68/88 (77%)	65 (96%)	3 (4%)	0	100	100
19	S	78/93 (84%)	71 (91%)	7 (9%)	0	100	100
20	T	97/106 (92%)	87 (90%)	10 (10%)	0	100	100
21	U	22/27 (82%)	21 (96%)	1 (4%)	0	100	100
All	All	2336/2541 (92%)	2169 (93%)	162 (7%)	5 (0%)	47	78

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	72	PRO
8	H	71	GLY

### 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%)	177 (88%)	25 (12%)	4	24
3	C	160/188 (85%)	140 (88%)	20 (12%)	4	24
4	D	180/181 (99%)	165 (92%)	15 (8%)	11	41
5	E	115/123 (94%)	103 (90%)	12 (10%)	7	31
6	F	90/90 (100%)	85 (94%)	5 (6%)	21	53
7	G	126/127 (99%)	116 (92%)	10 (8%)	12	43
8	H	119/119 (100%)	110 (92%)	9 (8%)	13	44

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
9	I	98/99 (99%)	90 (92%)	8 (8%)	11	41
10	J	87/92 (95%)	79 (91%)	8 (9%)	9	36
11	K	88/99 (89%)	79 (90%)	9 (10%)	7	32
12	L	103/110 (94%)	92 (89%)	11 (11%)	6	30
13	M	94/101 (93%)	85 (90%)	9 (10%)	8	34
14	N	49/50 (98%)	46 (94%)	3 (6%)	18	51
15	O	79/80 (99%)	69 (87%)	10 (13%)	4	24
16	P	72/74 (97%)	66 (92%)	6 (8%)	11	41
17	Q	94/97 (97%)	87 (93%)	7 (7%)	13	44
18	R	61/77 (79%)	57 (93%)	4 (7%)	16	48
19	S	71/80 (89%)	59 (83%)	12 (17%)	2	14
20	T	76/82 (93%)	64 (84%)	12 (16%)	2	17
21	U	19/22 (86%)	17 (90%)	2 (10%)	7	30
All	All	1983/2111 (94%)	1786 (90%)	197 (10%)	8	33

5 of 197 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
11	K	33	THR
14	N	25	VAL
11	K	98	LEU
12	L	113	ARG
15	O	70	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
2	B	204	ASN
3	C	37	GLN
9	I	3	GLN

### 5.3.3 RNA [i](#)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
1	A	1504/1522 (98%)	242 (16%)	46 (3%)

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
22	Y	5/6 (83%)	2 (40%)	0
23	W	14/15 (93%)	2 (14%)	0
All	All	1523/1543 (98%)	246 (16%)	46 (3%)

5 of 246 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
1	A	6	G
1	A	9	G
1	A	13	C
1	A	21	G
1	A	22	G

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 [i](#)

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	5MC	A	967	1	15,22,23	0.94	0	19,32,35	1.13	3 (15%)
1	M2G	A	966	1	20,27,28	1.94	4 (20%)	22,40,43	2.38	4 (18%)
1	4OC	A	1402	1	16,23,24	0.94	2 (12%)	17,32,35	0.73	0
1	MA6	A	1519[A]	1	19,26,27	1.04	1 (5%)	18,38,41	0.58	0
1	5MC	A	1404	1	15,22,23	0.92	0	19,32,35	1.09	2 (10%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
1	PSU	A	516	24,1	17,21,22	1.05	2 (11%)	20,30,33	3.11	6 (30%)
1	7MG	A	527	1	22,26,27	2.18	6 (27%)	28,39,42	1.77	6 (21%)
1	PSU	A	1541	1	17,21,22	1.05	2 (11%)	20,30,33	3.14	6 (30%)
1	MA6	A	1518[A]	1	19,26,27	0.95	1 (5%)	18,38,41	0.74	0
1	MA6	A	1519[B]	1	19,26,27	1.31	3 (15%)	18,38,41	0.60	0
1	UR3	A	1498	1	14,22,23	1.10	2 (14%)	15,32,35	1.02	1 (6%)
1	MA6	A	1518[B]	1	19,26,27	1.22	2 (10%)	18,38,41	0.77	0
1	5MC	A	1400	1	15,22,23	0.88	0	19,32,35	1.05	2 (10%)
1	5MC	A	1407	1	15,22,23	0.99	1 (6%)	19,32,35	1.13	2 (10%)
12	0TD	L	92	12	4,9,10	0.81	0	3,11,13	1.76	1 (33%)
1	2MG	A	1207	1	19,26,27	2.30	4 (21%)	21,38,41	2.01	3 (14%)
1	PSU	A	1540	1	17,21,22	1.02	1 (5%)	20,30,33	3.22	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	5MC	A	967	1	-	2/5/25/26	0/2/2/2
1	M2G	A	966	1	-	3/7/29/30	0/3/3/3
1	4OC	A	1402	1	-	3/9/29/30	0/2/2/2
1	MA6	A	1519[A]	1	-	5/7/29/30	0/3/3/3
1	5MC	A	1404	1	-	0/5/25/26	0/2/2/2
1	PSU	A	516	24,1	-	0/7/25/26	0/2/2/2
1	7MG	A	527	1	-	1/7/37/38	0/3/3/3
1	PSU	A	1541	1	-	1/7/25/26	0/2/2/2
1	MA6	A	1518[A]	1	-	1/7/29/30	0/3/3/3
1	MA6	A	1519[B]	1	-	3/7/29/30	0/3/3/3
1	UR3	A	1498	1	-	0/5/25/26	0/2/2/2
1	MA6	A	1518[B]	1	-	0/7/29/30	0/3/3/3
1	5MC	A	1400	1	-	2/5/25/26	0/2/2/2
1	5MC	A	1407	1	-	0/5/25/26	0/2/2/2
12	0TD	L	92	12	-	1/3/12/14	-
1	2MG	A	1207	1	-	0/5/27/28	0/3/3/3
1	PSU	A	1540	1	-	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	7.19	1.40	1.34
1	A	527	7MG	C4-N3	5.85	1.41	1.34
1	A	1207	2MG	C6-N1	5.85	1.43	1.33
1	A	966	M2G	C6-N1	5.81	1.43	1.33
1	A	527	7MG	C2-N2	4.85	1.43	1.33

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	1540	PSU	N3-C2-N1	-10.87	119.79	128.43
1	A	1541	PSU	N3-C2-N1	-10.79	119.85	128.43
1	A	516	PSU	N3-C2-N1	-10.74	119.89	128.43
1	A	966	M2G	C5-C6-N1	-8.06	112.41	123.43
1	A	1207	2MG	C5-C6-N1	-7.55	113.11	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	1400	5MC	O4'-C4'-C5'-O5'
1	A	1402	4OC	O4'-C4'-C5'-O5'
1	A	1519[A]	MA6	C3'-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	967	5MC	2	0
1	A	966	M2G	1	0
1	A	1402	4OC	1	0
1	A	1519[A]	MA6	1	0
1	A	1404	5MC	2	0
1	A	1518[A]	MA6	1	0
1	A	1519[B]	MA6	2	0
1	A	1518[B]	MA6	2	0
1	A	1400	5MC	1	0
12	L	92	0TD	2	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
25	SIS	A	1833	-	28,33,33	1.32	2 (7%)	28,49,49	0.93	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	1833	-	-	1/9/65/65	0/3/3/3

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
25	A	1833	SIS	C31-C41	-4.85	1.39	1.50
25	A	1833	SIS	C61-C51	-3.75	1.39	1.48

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
25	A	1833	SIS	C93-N33-C33	-3.11	109.86	114.38

There are no chirality outliers.

All (1) torsion outliers are listed below:

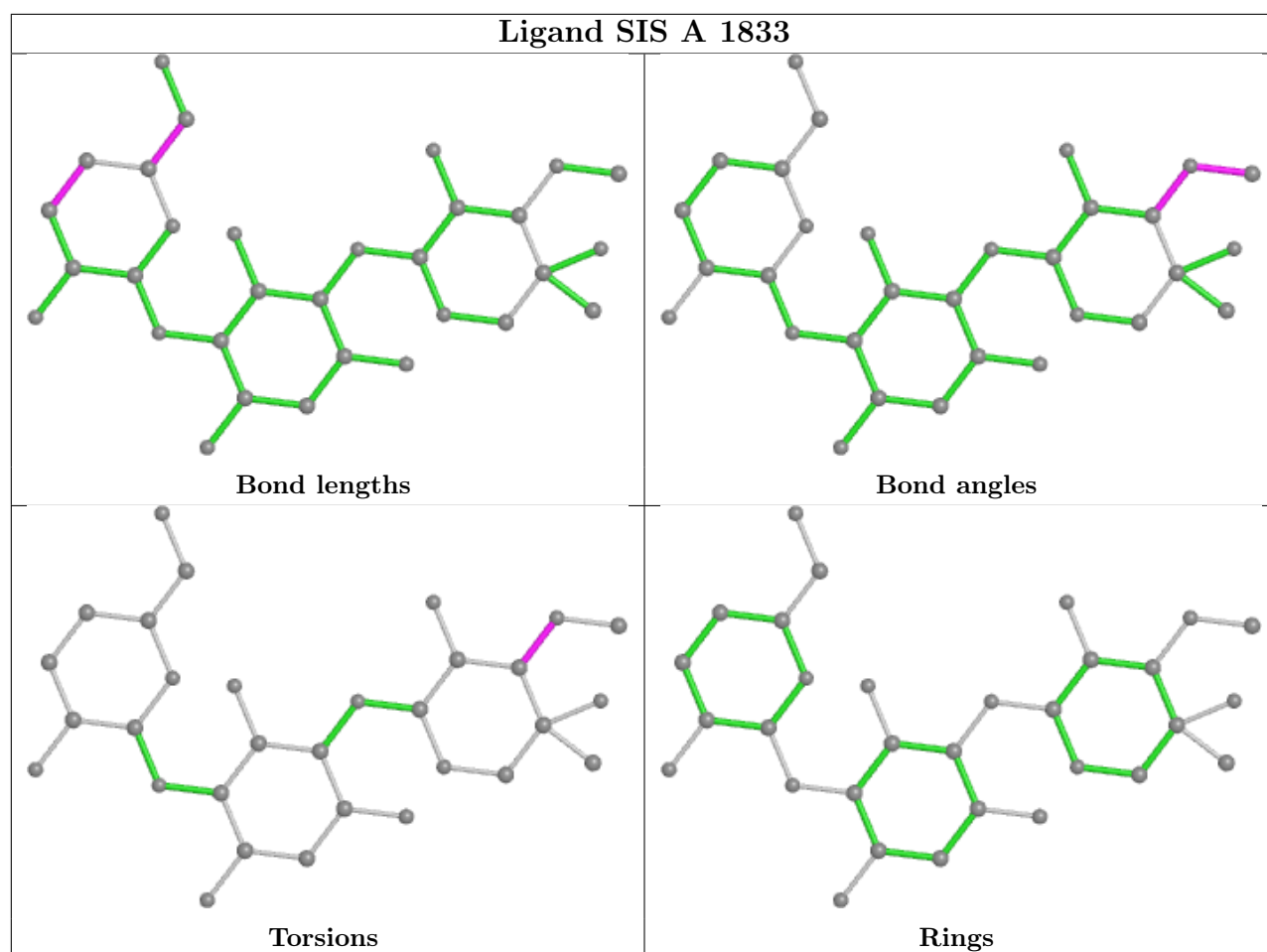
Mol	Chain	Res	Type	Atoms
25	A	1833	SIS	C23-C33-N33-C93

There are no ring outliers.

1 monomer is involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
25	A	1833	SIS	2	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 than 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.



## 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 [i](#)

### 6.1 Protein, DNA and RNA chains [i](#)

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

### 6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

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

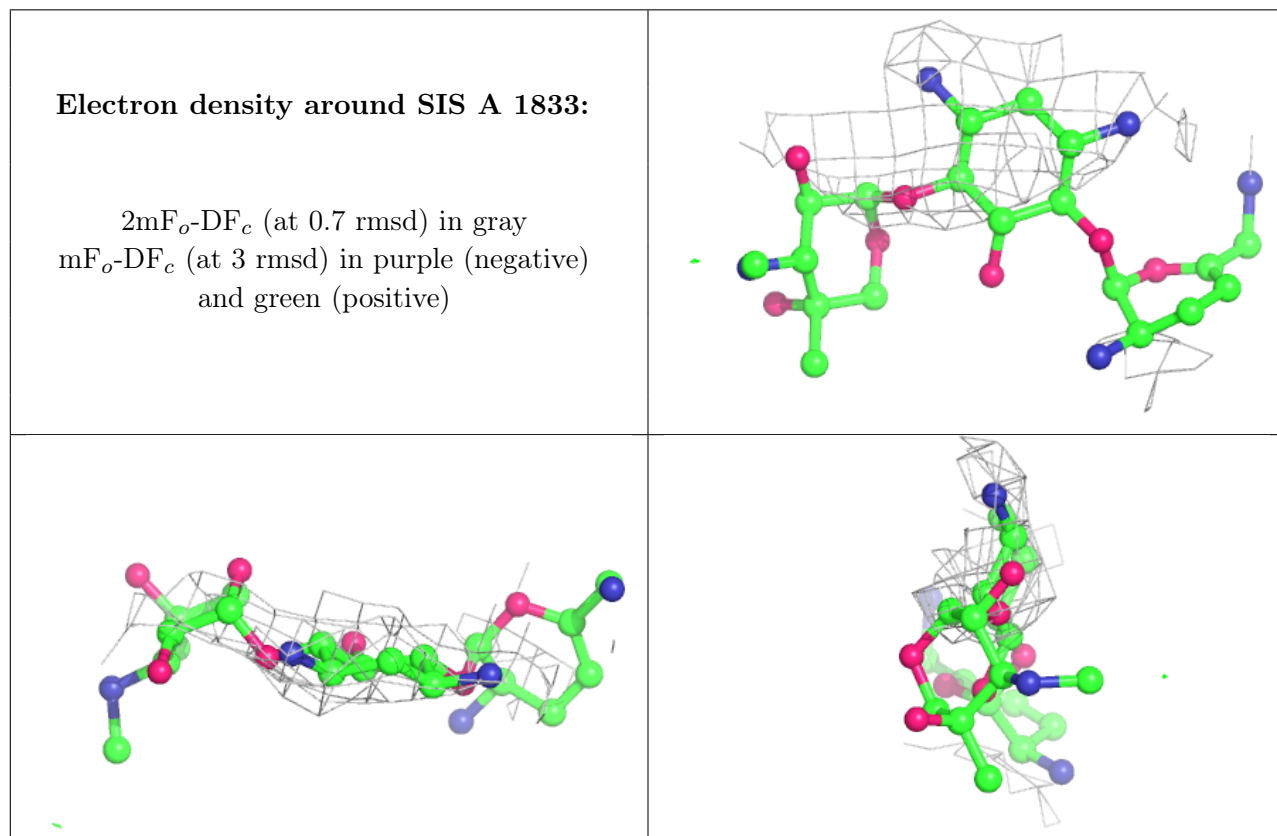
### 6.3 Carbohydrates [i](#)

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

### 6.4 Ligands [i](#)

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.



## 6.5 Other polymers

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