wwPDB X-ray Structure Validation Summary Report (i)

Jun 12, 2024 – 05:20 PM EDT

PDB ID	:	1FJG
Title	:	STRUCTURE OF THE THERMUS THERMOPHILUS 30S RIBOSOMAL SUBUNIT IN COMPLEX WITH THE ANTIBIOTICS STREPTOMYCIN, SPECTINOMYCIN, AND PAROMOMYCIN
Authors	:	Carter, A.P.; Clemons Jr., W.M.; Brodersen, D.E.; Wimberly, B.T.; Morgan-Warren, R.J.; Ramakrishnan, V.
Deposited on	:	2000-08-08
Resolution	:	3.00 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 types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

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

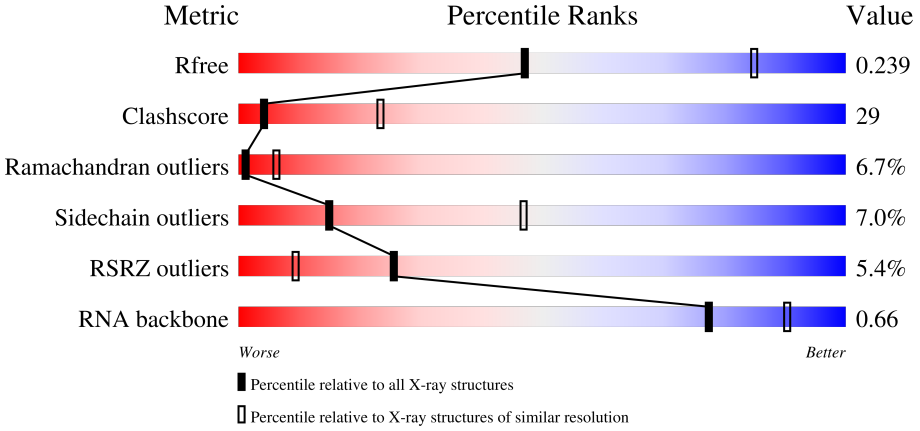
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36.2
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36.2

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 3.00 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	2092 (3.00-3.00)
Clashscore	141614	2416 (3.00-3.00)
Ramachandran outliers	138981	2333 (3.00-3.00)
Sidechain outliers	138945	2336 (3.00-3.00)
RSRZ outliers	127900	1990 (3.00-3.00)
RNA backbone	3102	1173 (3.30-2.70)

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% The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	1522	6%; 36%, 49%, 12%, • •
2	X	6	33%; 83%, 17%
3	B	256	%; 28%, 56%, 8%, 7%
4	C	239	23%, 53%, 9%, •, 14%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
5	D	209	7% 46% 46% 7%
6	E	162	.% 40% 48% 7%
7	F	101	5% 37% 60%
8	G	156	2% 39% 54% 6%
9	H	138	.% 50% 46%
10	I	128	12% 42% 47% 9%
11	J	105	9% 24% 54% 13% 7%
12	K	129	8% 36% 48% 6% 8%
13	L	135	.% 36% 50% 7%
14	M	126	14% 30% 52% 16%
15	N	61	5% 25% 62% 11%
16	O	89	2% 62% 33%
17	P	88	.% 44% 44% 6% 6%
18	Q	105	10% 48% 40% 10%
19	R	88	5% 18% 53% 11% 17%
20	S	93	17% 27% 55% 9% 10%
21	T	106	3% 40% 45% 8% 7%
22	V	26	8% 27% 62% 8%

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	MG	A	1550	-	-	-	X
24	MG	A	1557	-	-	-	X
24	MG	A	1573	-	-	-	X
24	MG	A	1583	-	-	-	X
24	MG	A	1586	-	-	-	X
24	MG	A	1615	-	-	-	X
24	MG	A	1619	-	-	-	X

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
24	MG	A	210	-	-	-	X
24	MG	A	211	-	-	-	X
24	MG	A	212	-	-	-	X

2 Entry composition (i)

There are 27 unique types of molecules in this entry. The entry contains 51995 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 16S RIBOSOMAL RNA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
1	A	1507	32391	14418	6002	10465	1506	22	0	0

- Molecule 2 is a RNA chain called FRAGMENT OF MESSENGER RNA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
2	X	6	117	54	13	45	5	0	0	0

- Molecule 3 is a protein called 30S RIBOSOMAL PROTEIN S2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	B	237	1923	1226	344	348	5	0	0	0

- Molecule 4 is a protein called 30S RIBOSOMAL PROTEIN S3.

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

- Molecule 5 is a protein called 30S RIBOSOMAL PROTEIN S4.

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

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	199	GLN	ASN	CONFLICT	UNP P80373
D	201	ASN	GLN	CONFLICT	UNP P80373

- Molecule 6 is a protein called 30S RIBOSOMAL PROTEIN S5.

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

- Molecule 7 is a protein called 30S RIBOSOMAL PROTEIN S6.

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

- Molecule 8 is a protein called 30S RIBOSOMAL PROTEIN S7.

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

- Molecule 9 is a protein called 30S RIBOSOMAL PROTEIN S8.

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

- Molecule 10 is a protein called 30S RIBOSOMAL PROTEIN S9.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
10	I	127	1011	639	198	174	0	0	0

- Molecule 11 is a protein called 30S RIBOSOMAL PROTEIN S10.

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

- Molecule 12 is a protein called 30S RIBOSOMAL PROTEIN S11.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
12	K	119	885	549	168	165	3	0	0	0

- Molecule 13 is a protein called 30S RIBOSOMAL PROTEIN S12.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
13	L	125	975	614	196	164	1	0	0	0

- Molecule 14 is a protein called 30S RIBOSOMAL PROTEIN S13.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
14	M	125	997	617	207	171	2	0	0	0

- Molecule 15 is a protein called 30S RIBOSOMAL PROTEIN S14.

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

- Molecule 16 is a protein called 30S RIBOSOMAL PROTEIN S15.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
16	O	88	734	459	147	126	2	0	0	0

- Molecule 17 is a protein called 30S RIBOSOMAL PROTEIN S16.

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

- Molecule 18 is a protein called 30S RIBOSOMAL PROTEIN S17.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
18	Q	104	857	547	161	147	2	0	0	0

- Molecule 19 is a protein called 30S RIBOSOMAL PROTEIN S18.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
19	R	73	597	380	118	99	0	0	0

- Molecule 20 is a protein called 30S RIBOSOMAL PROTEIN S19.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
20	S	84	674	430	126	116	2	0	0	0

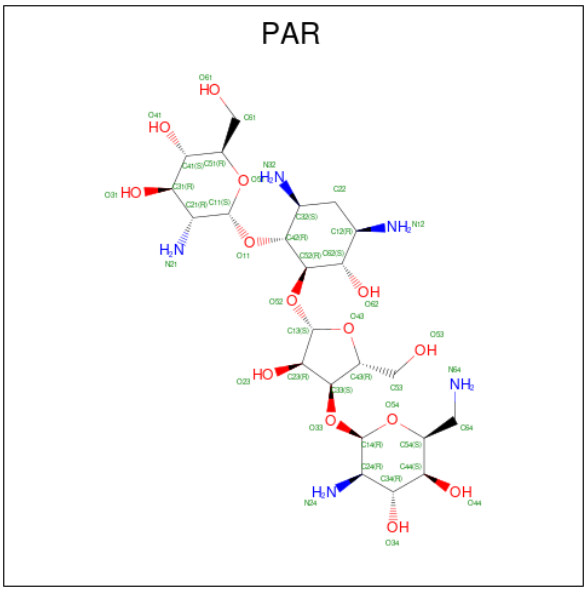
- Molecule 21 is a protein called 30S RIBOSOMAL PROTEIN S20.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
21	T	99	762	469	162	129	2	0	0	0

- Molecule 22 is a protein called 30S RIBOSOMAL PROTEIN THX.

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

- Molecule 23 is PAROMOMYCIN (three-letter code: PAR) (formula: $C_{23}H_{45}N_5O_{14}$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
23	A	1	42	23	5	14	0	0

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

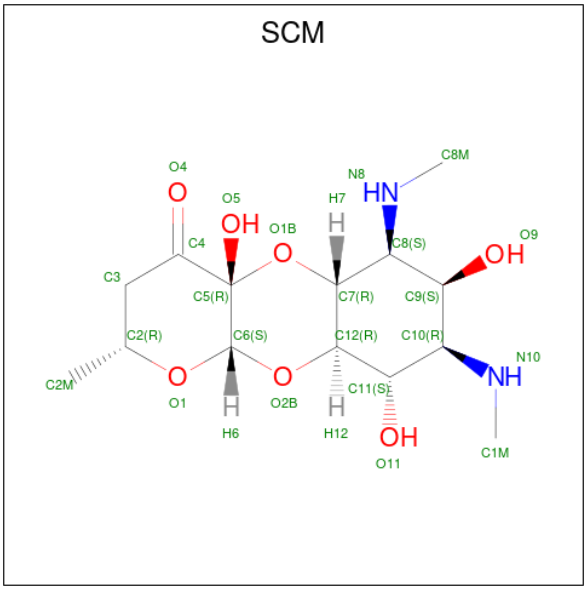
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Mg		
24	A	94	94	94	0	0

Continued on next page...

Continued from previous page...

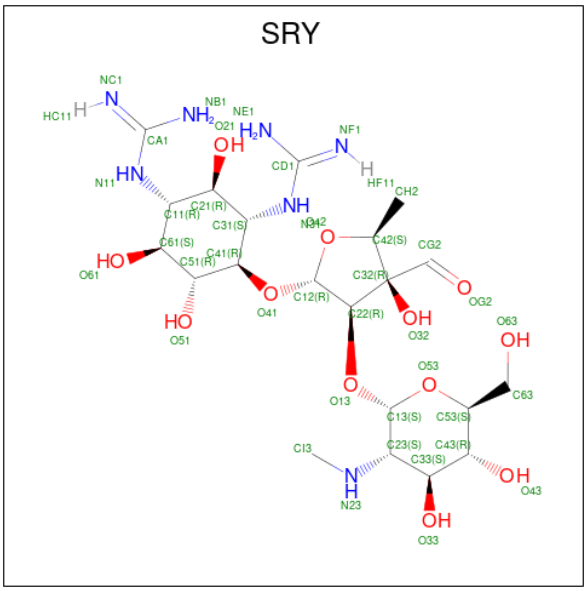
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
24	D	1	Total Mg 1 1	0	0
24	H	1	Total Mg 1 1	0	0

- Molecule 25 is SPECTINOMYCIN (three-letter code: SCM) (formula: $C_{14}H_{24}N_2O_7$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
25	A	1	Total C N O 23 14 2 7	0	0

- Molecule 26 is STREPTOMYCIN (three-letter code: SRY) (formula: $C_{21}H_{39}N_7O_{12}$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
26	A	1	Total C N O 40 21 7 12	0	0

- Molecule 27 is ZINC ION (three-letter code: ZN) (formula: Zn).

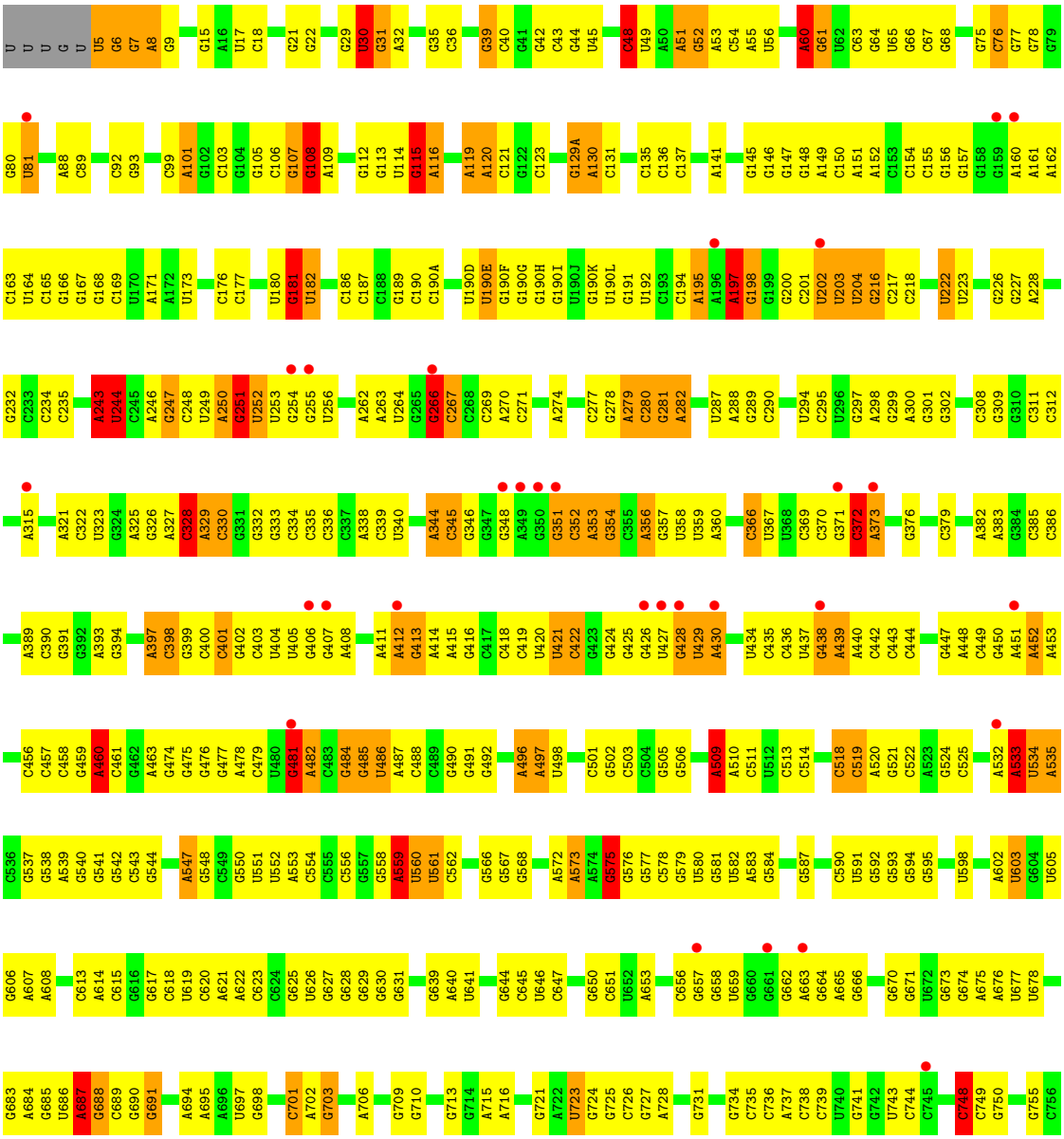
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
27	D	1	Total Zn 1 1	0	0
27	N	1	Total Zn 1 1	0	0

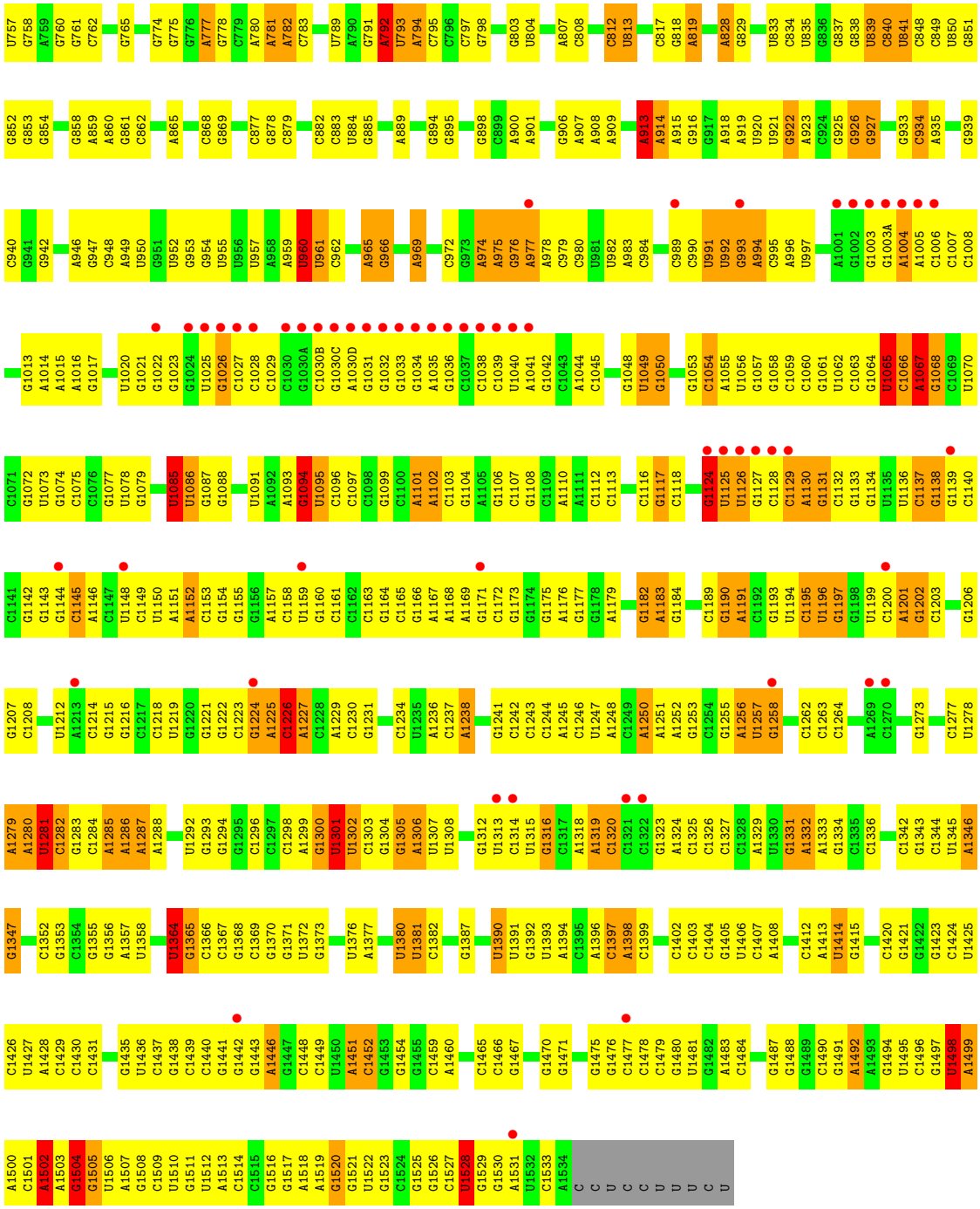
3 Residue-property plots (i)

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 and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density (RSRZ > 2). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

- Molecule 1: 16S RIBOSOMAL RNA

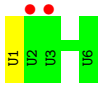
Chain A:





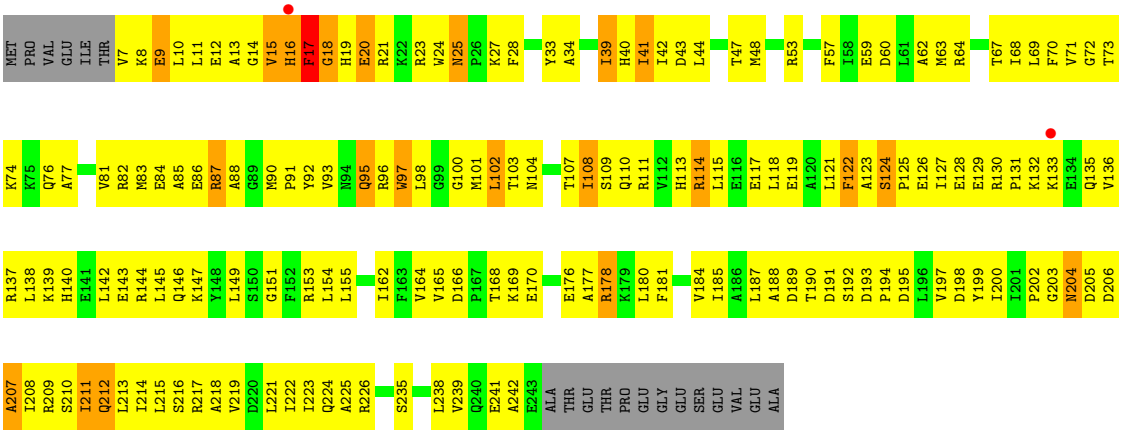
• Molecule 2: FRAGMENT OF MESSENGER RNA

Chain X:



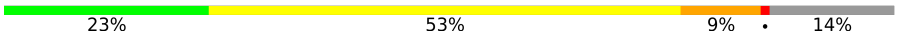
• Molecule 3: 30S RIBOSOMAL PROTEIN S2

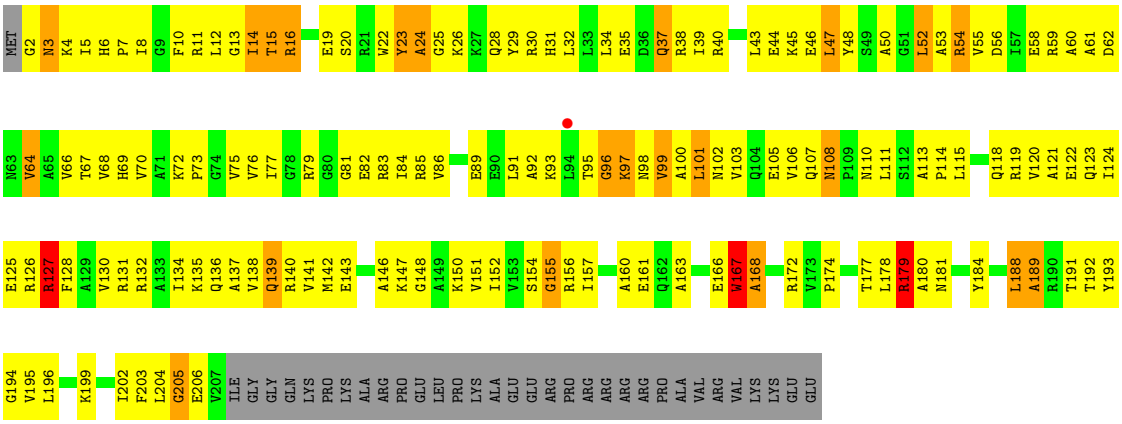
Chain B:



• Molecule 4: 30S RIBOSOMAL PROTEIN S3

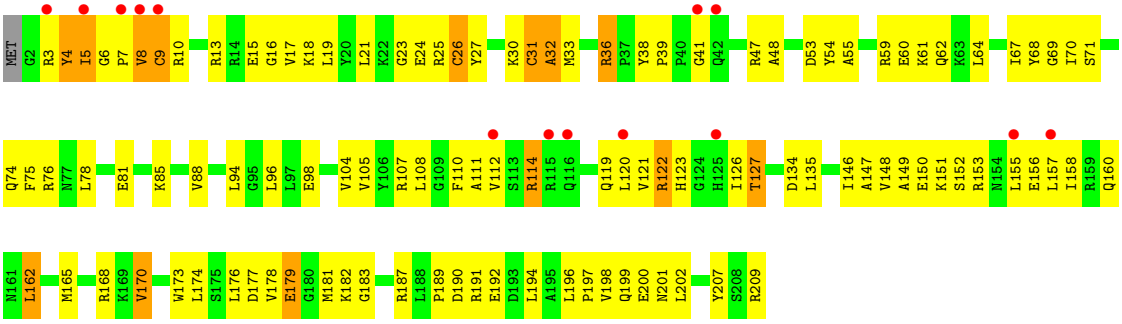
Chain C:





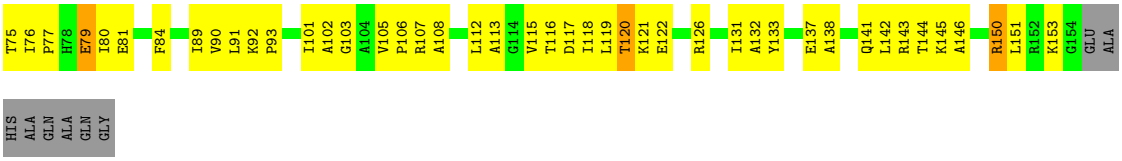
• Molecule 5: 30S RIBOSOMAL PROTEIN S4

Chain D:



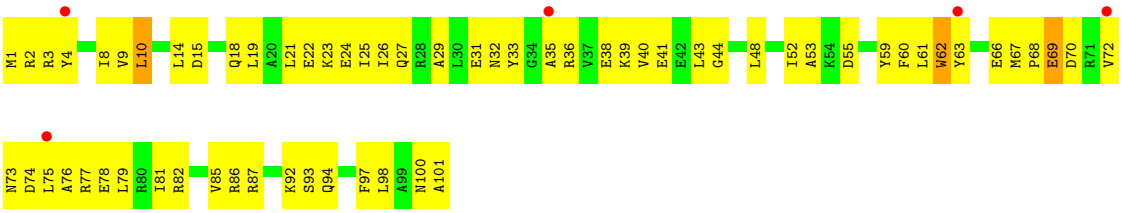
• Molecule 6: 30S RIBOSOMAL PROTEIN S5

Chain E:



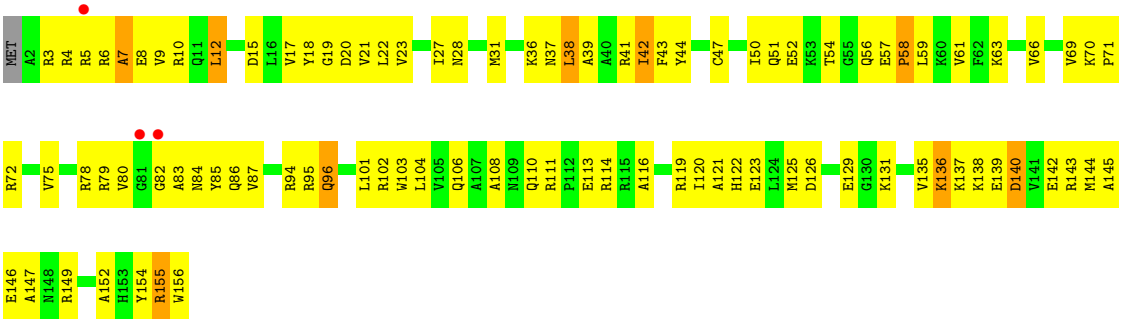
• Molecule 7: 30S RIBOSOMAL PROTEIN S6

Chain F:



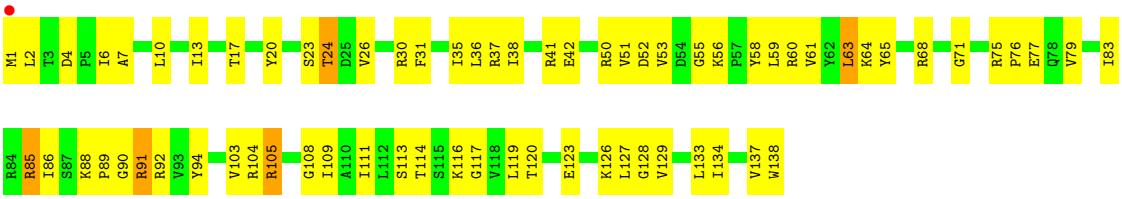
• Molecule 8: 30S RIBOSOMAL PROTEIN S7

Chain G:



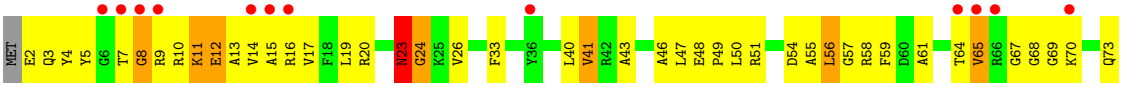
• Molecule 9: 30S RIBOSOMAL PROTEIN S8

Chain H:



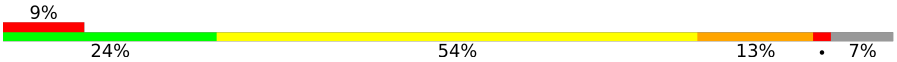
• Molecule 10: 30S RIBOSOMAL PROTEIN S9

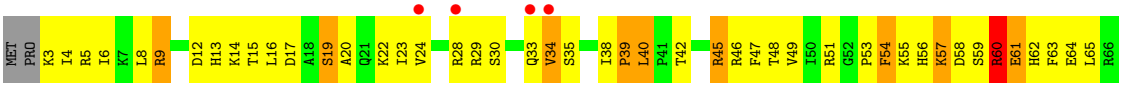
Chain I:

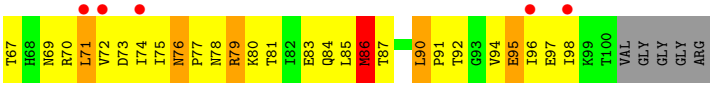


• Molecule 11: 30S RIBOSOMAL PROTEIN S10

Chain J:

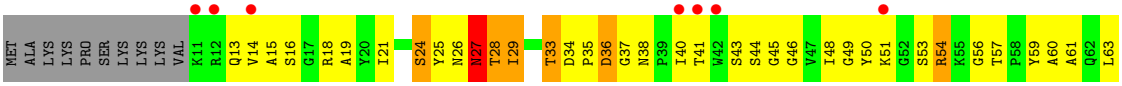


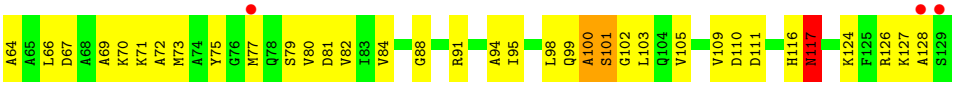




• Molecule 12: 30S RIBOSOMAL PROTEIN S11

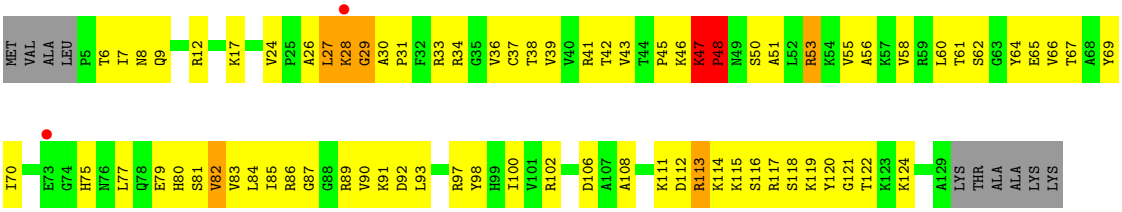
Chain K:





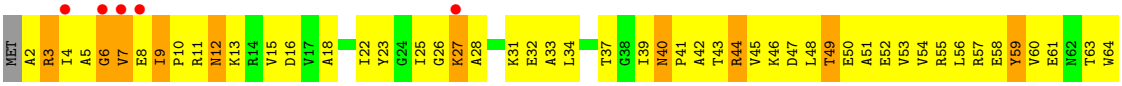
• Molecule 13: 30S RIBOSOMAL PROTEIN S12

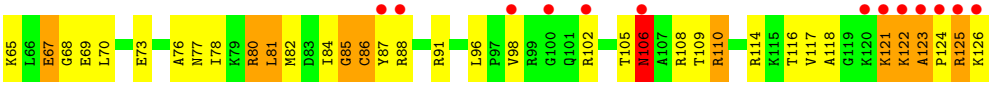
Chain L:



• Molecule 14: 30S RIBOSOMAL PROTEIN S13

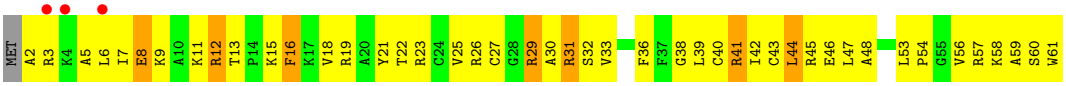
Chain M:





• Molecule 15: 30S RIBOSOMAL PROTEIN S14

Chain N:

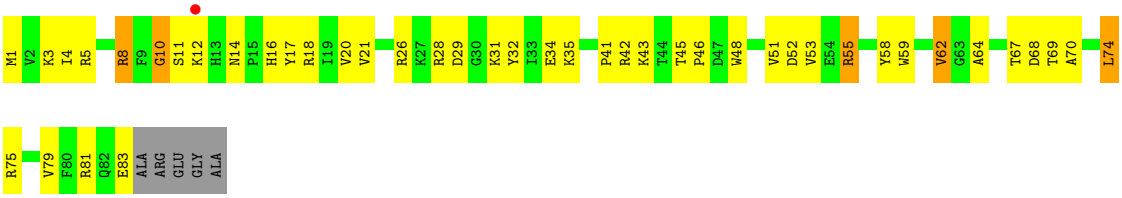


• Molecule 16: 30S RIBOSOMAL PROTEIN S15

Chain O:

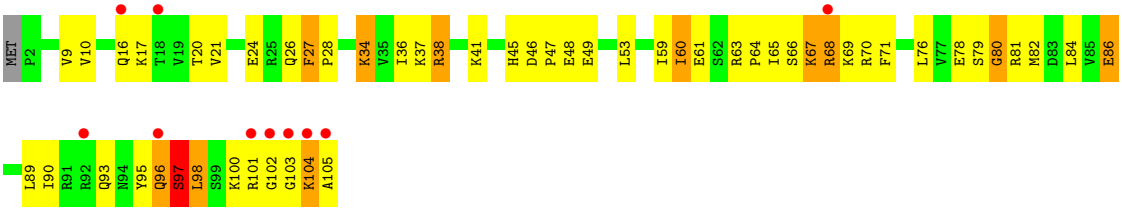
• Molecule 17: 30S RIBOSOMAL PROTEIN S16

Chain P:



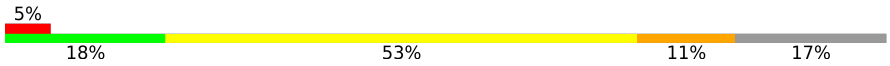
• Molecule 18: 30S RIBOSOMAL PROTEIN S17

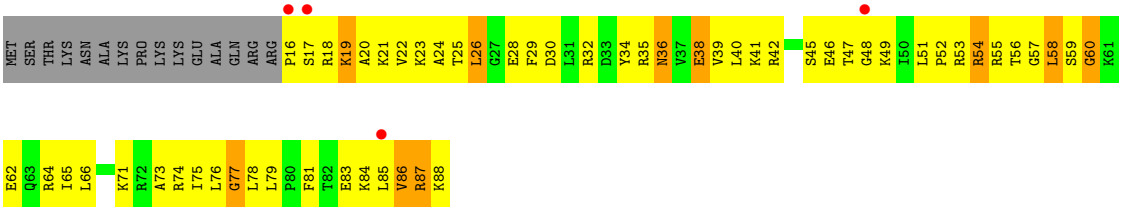
Chain Q:



• Molecule 19: 30S RIBOSOMAL PROTEIN S18

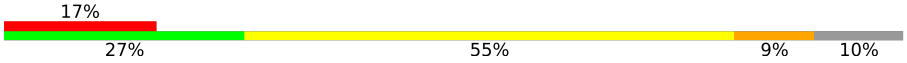
Chain R:

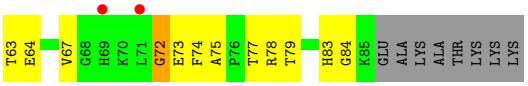




• Molecule 20: 30S RIBOSOMAL PROTEIN S19

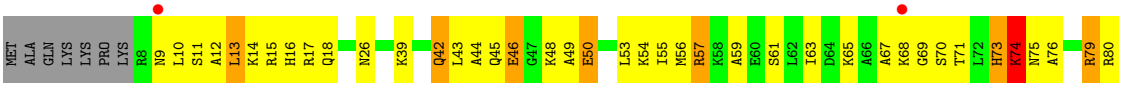
Chain S:

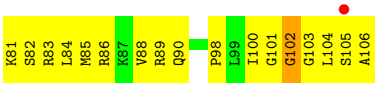




• Molecule 21: 30S RIBOSOMAL PROTEIN S20

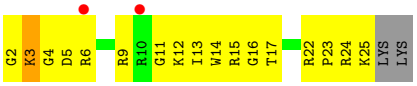
Chain T:





• Molecule 22: 30S RIBOSOMAL PROTEIN THX

Chain V:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 41 21 2	Depositor
Cell constants a, b, c, α, β, γ	402.00Å 402.00Å 176.49Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	95.35 – 3.00 97.50 – 3.00	Depositor EDS
% Data completeness (in resolution range)	91.1 (95.35-3.00) 91.0 (97.50-3.00)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.95 (at 3.01Å)	Xtriage
Refinement program	CNS	Depositor
R, R_{free}	0.221 , 0.255 0.207 , 0.239	Depositor DCC
R_{free} test set	13237 reflections (5.09%)	wwPDB-VP
Wilson B-factor (Å^2)	72.5	Xtriage
Anisotropy	0.235	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.30 , 78.5	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.87	EDS
Total number of atoms	51995	wwPDB-VP
Average B, all atoms (Å^2)	73.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.14% 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 (i)

5.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: PAR, MG, ZN, SCM, SRY

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	0.48	0/36259	0.74	39/56593 (0.1%)
2	X	1.41	1/128 (0.8%)	1.72	5/196 (2.6%)
3	B	0.34	0/1958	0.65	0/2640
4	C	0.37	0/1636	0.66	0/2205
5	D	0.38	0/1733	0.63	0/2318
6	E	0.46	0/1162	0.71	0/1564
7	F	0.32	0/856	0.60	0/1154
8	G	0.34	0/1276	0.60	0/1709
9	H	0.44	0/1136	0.71	0/1527
10	I	0.36	0/1029	0.64	0/1378
11	J	0.35	0/805	0.66	1/1082 (0.1%)
12	K	0.36	0/900	0.65	0/1213
13	L	0.41	0/991	0.79	1/1327 (0.1%)
14	M	0.32	0/1008	0.65	0/1347
15	N	0.39	0/501	0.69	0/664
16	O	0.36	0/745	0.62	1/992 (0.1%)
17	P	0.43	0/716	0.73	0/963
18	Q	0.45	0/870	0.77	1/1159 (0.1%)
19	R	0.37	0/603	0.64	0/799
20	S	0.31	0/689	0.72	2/926 (0.2%)
21	T	0.38	0/764	0.71	0/1006
22	V	0.42	0/212	0.65	0/277
All	All	0.45	1/55977 (0.0%)	0.73	50/83039 (0.1%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand.A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	4	39

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	X	1	U	C4-O4	13.50	1.34	1.23

The worst 5 of 50 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	X	1	U	N3-C4-C5	12.17	121.90	114.60
2	X	1	U	C2-N3-C4	-11.66	120.00	127.00
1	A	1498	U	C2'-C3'-O3'	9.77	130.99	109.50
2	X	1	U	C5-C4-O4	-9.60	120.14	125.90
1	A	559	A	C2'-C3'-O3'	9.52	130.44	109.50

All (4) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
1	A	115	G	C3'
1	A	559	A	C3'
1	A	1498	U	C3'
1	A	1528	U	C3'

5 of 39 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	108	G	Sidechain
1	A	197	A	Sidechain
1	A	222	U	Sidechain
1	A	251	G	Sidechain
1	A	76	C	Sidechain

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	32391	0	16348	1037	0
2	X	117	0	63	0	0
3	B	1923	0	1968	208	0
4	C	1612	0	1677	188	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
5	D	1703	0	1763	128	0
6	E	1146	0	1207	104	0
7	F	843	0	857	74	0
8	G	1257	0	1296	108	0
9	H	1116	0	1177	75	0
10	I	1011	0	1043	104	0
11	J	792	0	835	110	0
12	K	885	0	904	68	0
13	L	975	0	1062	90	0
14	M	997	0	1072	113	0
15	N	492	0	529	67	0
16	O	734	0	771	34	0
17	P	700	0	720	45	0
18	Q	857	0	930	88	0
19	R	597	0	668	83	0
20	S	674	0	699	72	0
21	T	762	0	859	73	0
22	V	208	0	221	22	0
23	A	42	0	45	4	0
24	A	94	0	0	0	0
24	D	1	0	0	0	0
24	H	1	0	0	0	0
25	A	23	0	24	2	0
26	A	40	0	37	5	0
27	D	1	0	0	0	0
27	N	1	0	0	0	0
All	All	51995	0	36775	2610	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 29.

The worst 5 of 2610 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:243:A:H4'	1:A:244:U:H5'	1.30	1.13
5:D:23:GLY:HA3	5:D:112:VAL:HG12	1.31	1.12
11:J:51:ARG:HB2	11:J:59:SER:HB3	1.27	1.11
12:K:54:ARG:HB3	12:K:54:ARG:HH11	1.13	1.09
1:A:761:G:C2	18:Q:105:ALA:HA	1.88	1.08

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
3	B	235/256 (92%)	172 (73%)	43 (18%)	20 (8%)	1 4
4	C	204/239 (85%)	135 (66%)	43 (21%)	26 (13%)	0 1
5	D	206/209 (99%)	173 (84%)	25 (12%)	8 (4%)	3 17
6	E	148/162 (91%)	130 (88%)	16 (11%)	2 (1%)	11 43
7	F	99/101 (98%)	81 (82%)	14 (14%)	4 (4%)	3 17
8	G	153/156 (98%)	123 (80%)	25 (16%)	5 (3%)	4 21
9	H	136/138 (99%)	125 (92%)	10 (7%)	1 (1%)	22 60
10	I	125/128 (98%)	96 (77%)	18 (14%)	11 (9%)	1 3
11	J	96/105 (91%)	64 (67%)	18 (19%)	14 (15%)	0 1
12	K	117/129 (91%)	91 (78%)	19 (16%)	7 (6%)	1 9
13	L	123/135 (91%)	95 (77%)	19 (15%)	9 (7%)	1 5
14	M	123/126 (98%)	92 (75%)	14 (11%)	17 (14%)	0 1
15	N	58/61 (95%)	44 (76%)	11 (19%)	3 (5%)	2 12
16	O	86/89 (97%)	69 (80%)	12 (14%)	5 (6%)	1 10
17	P	81/88 (92%)	72 (89%)	8 (10%)	1 (1%)	13 48
18	Q	102/105 (97%)	86 (84%)	11 (11%)	5 (5%)	2 13
19	R	71/88 (81%)	49 (69%)	16 (22%)	6 (8%)	1 4
20	S	82/93 (88%)	60 (73%)	16 (20%)	6 (7%)	1 5
21	T	97/106 (92%)	77 (79%)	12 (12%)	8 (8%)	1 4
22	V	22/26 (85%)	17 (77%)	4 (18%)	1 (4%)	2 14
All	All	2364/2540 (93%)	1851 (78%)	354 (15%)	159 (7%)	1 6

5 of 159 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	B	15	VAL

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
3	B	16	HIS
3	B	17	PHE
3	B	21	ARG
3	B	124	SER

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	
3	B	204/220 (93%)	193 (95%)	11 (5%)	22	57
4	C	160/188 (85%)	147 (92%)	13 (8%)	11	40
5	D	180/181 (99%)	169 (94%)	11 (6%)	18	53
6	E	115/123 (94%)	104 (90%)	11 (10%)	8	32
7	F	90/90 (100%)	88 (98%)	2 (2%)	52	81
8	G	126/127 (99%)	115 (91%)	11 (9%)	10	37
9	H	119/119 (100%)	113 (95%)	6 (5%)	24	60
10	I	98/99 (99%)	91 (93%)	7 (7%)	14	46
11	J	87/92 (95%)	80 (92%)	7 (8%)	12	40
12	K	90/99 (91%)	82 (91%)	8 (9%)	9	35
13	L	104/111 (94%)	98 (94%)	6 (6%)	20	55
14	M	100/101 (99%)	91 (91%)	9 (9%)	9	35
15	N	49/50 (98%)	43 (88%)	6 (12%)	5	21
16	O	79/80 (99%)	76 (96%)	3 (4%)	33	69
17	P	72/74 (97%)	67 (93%)	5 (7%)	15	48
18	Q	96/97 (99%)	88 (92%)	8 (8%)	11	39
19	R	64/77 (83%)	59 (92%)	5 (8%)	12	42
20	S	73/80 (91%)	69 (94%)	4 (6%)	21	57
21	T	76/82 (93%)	69 (91%)	7 (9%)	9	34
22	V	19/21 (90%)	19 (100%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
All	All	2001/2111 (95%)	1861 (93%)	140 (7%)	15 47

5 of 140 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
17	P	74	LEU
18	Q	68	ARG
20	S	53	ASN
8	G	8	GLU
7	F	69	GLU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 60 such sidechains are listed below:

Mol	Chain	Res	Type
7	F	94	GLN
19	R	36	ASN
10	I	23	ASN
18	Q	16	GLN
21	T	18	GLN

5.3.3 RNA (i)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
1	A	1507/1522 (99%)	213 (14%)	88 (5%)
2	X	5/6 (83%)	0	0
All	All	1512/1528 (98%)	213 (14%)	88 (5%)

5 of 213 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
1	A	6	G
1	A	8	A
1	A	9	G
1	A	30	U
1	A	31	G

5 of 88 RNA pucker outliers are listed below:

Mol	Chain	Res	Type
1	A	1085	U

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	1257	U
1	A	1101	A
1	A	1196	U
1	A	1301	U

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

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

Of 101 ligands modelled in this entry, 98 are monoatomic - leaving 3 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	SCM	A	1633	-	23,25,25	0.97	1 (4%)	26,39,39	1.37	3 (11%)
26	SRY	A	1634	-	40,42,42	1.23	3 (7%)	49,63,63	1.70	9 (18%)
23	PAR	A	1545	-	45,45,45	1.54	9 (20%)	64,67,67	1.18	5 (7%)

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	SCM	A	1633	-	-	2/4/57/57	0/3/3/3
26	SRY	A	1634	-	-	6/20/87/87	0/3/3/3
23	PAR	A	1545	-	-	3/18/94/94	0/4/4/4

The worst 5 of 13 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
26	A	1634	SRY	C31-N31	4.37	1.52	1.45
23	A	1545	PAR	C64-C54	3.93	1.57	1.52
23	A	1545	PAR	O54-C14	3.93	1.51	1.41
26	A	1634	SRY	C11-N11	3.69	1.51	1.45
23	A	1545	PAR	C31-C21	3.12	1.57	1.53

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

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
26	A	1634	SRY	C13-O13-C22	-5.96	105.90	116.25
25	A	1633	SCM	C1M-N10-C10	-4.97	107.15	114.38
26	A	1634	SRY	CI3-N23-C23	-4.89	107.26	114.38
26	A	1634	SRY	O41-C12-O42	-4.24	106.84	111.43
23	A	1545	PAR	O33-C14-C24	3.36	114.00	108.22

There are no chirality outliers.

5 of 11 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
25	A	1633	SCM	C11-C10-N10-C1M
26	A	1634	SRY	NB1-CA1-N11-C11
26	A	1634	SRY	NC1-CA1-N11-C11
26	A	1634	SRY	C13-C23-N23-CI3
26	A	1634	SRY	C43-C53-C63-O63

There are no ring outliers.

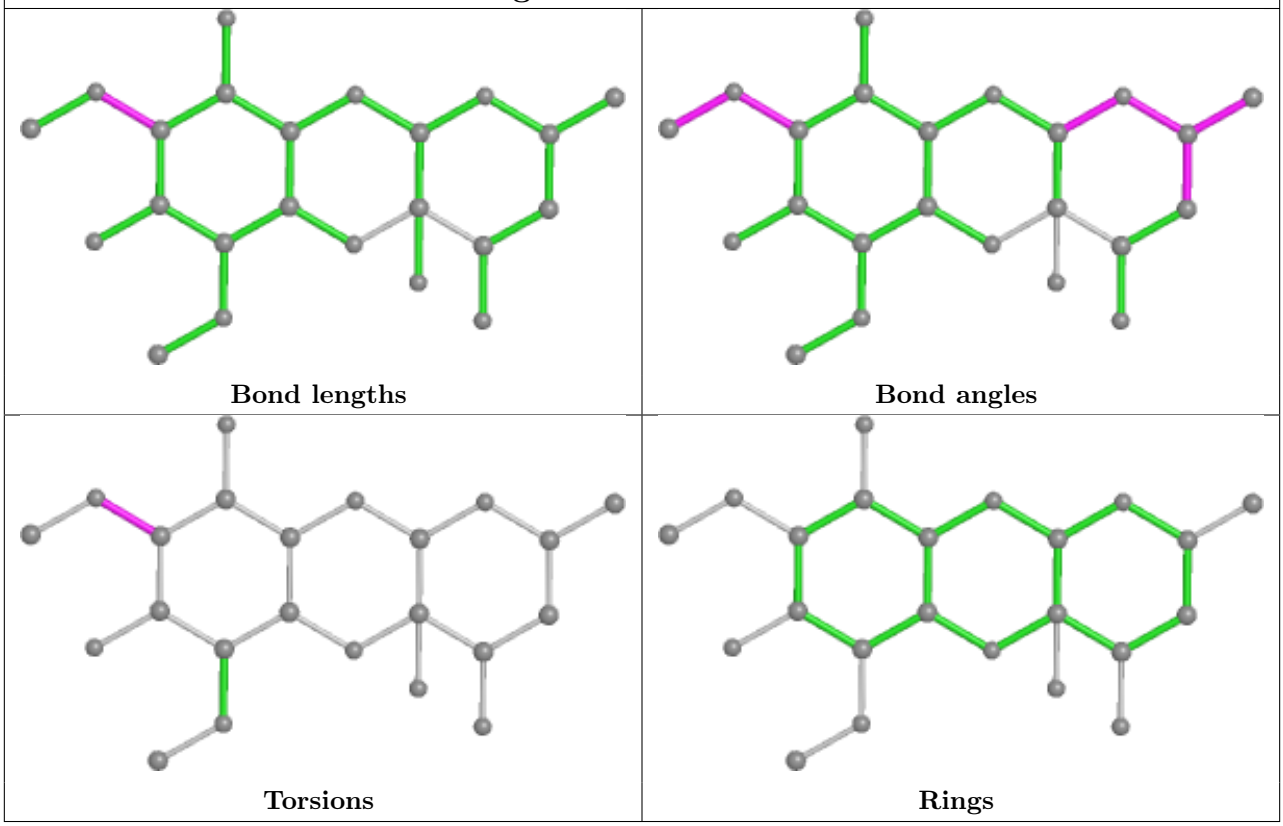
3 monomers are involved in 11 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
25	A	1633	SCM	2	0
26	A	1634	SRY	5	0
23	A	1545	PAR	4	0

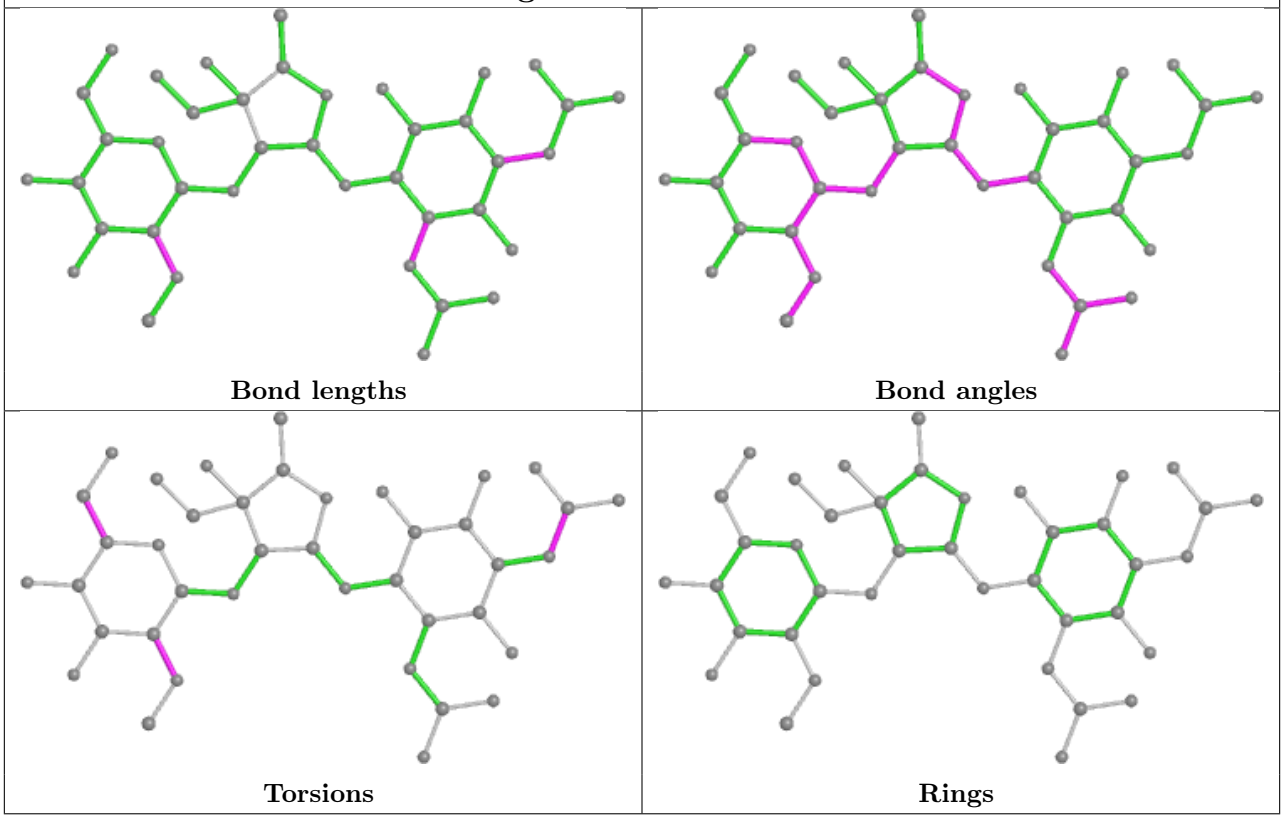
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. 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.

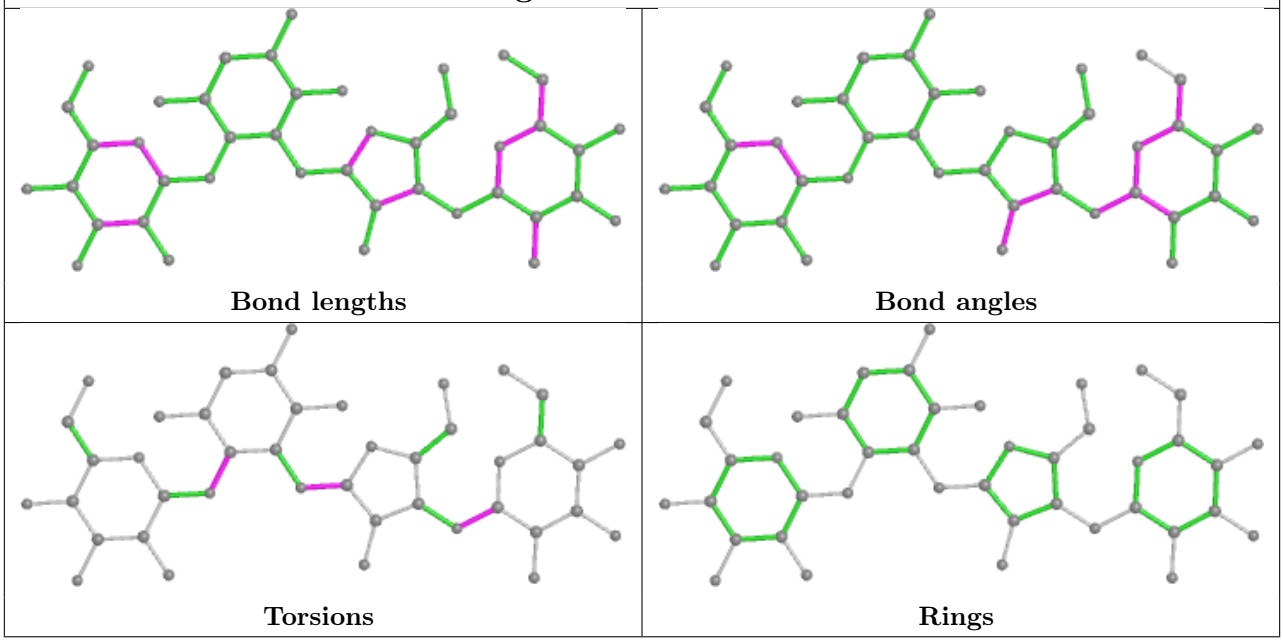
Ligand SCM A 1633



Ligand SRY A 1634



Ligand PAR A 1545



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)

In the following table, the column labelled '#RSRZ> 2' contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95th percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled 'Q< 0.9' lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2			OWAB(Å^2)	Q<0.9
1	A	1506/1522 (98%)	0.73	86 (5%)	23	8	34, 60, 137, 201	0
2	X	6/6 (100%)	1.38	2 (33%)	0	0	57, 71, 116, 118	0
3	B	237/256 (92%)	0.05	2 (0%)	86	65	35, 83, 162, 185	0
4	C	206/239 (86%)	-0.00	1 (0%)	91	75	32, 79, 136, 190	0
5	D	208/209 (99%)	0.27	14 (6%)	17	5	40, 63, 114, 143	0
6	E	150/162 (92%)	0.14	1 (0%)	87	69	29, 52, 95, 146	0
7	F	101/101 (100%)	0.33	5 (4%)	28	10	49, 89, 138, 165	0
8	G	155/156 (99%)	-0.11	3 (1%)	66	37	40, 72, 141, 177	0
9	H	138/138 (100%)	0.05	1 (0%)	87	69	30, 51, 90, 148	0
10	I	127/128 (99%)	0.58	15 (11%)	4	1	33, 82, 137, 177	0
11	J	98/105 (93%)	0.49	9 (9%)	9	3	42, 109, 179, 201	0
12	K	119/129 (92%)	0.48	10 (8%)	11	3	40, 69, 125, 194	0
13	L	125/135 (92%)	0.21	2 (1%)	72	44	21, 59, 134, 182	0
14	M	125/126 (99%)	0.89	18 (14%)	2	1	47, 83, 146, 185	0
15	N	60/61 (98%)	0.26	3 (5%)	28	10	38, 73, 119, 200	0
16	O	88/89 (98%)	0.13	2 (2%)	60	31	39, 66, 127, 191	0
17	P	83/88 (94%)	0.10	1 (1%)	79	54	37, 53, 90, 159	0
18	Q	104/105 (99%)	0.70	10 (9%)	8	2	26, 58, 143, 201	0
19	R	73/88 (82%)	0.24	4 (5%)	25	9	43, 69, 129, 184	0
20	S	84/93 (90%)	1.22	16 (19%)	1	0	71, 96, 145, 201	0
21	T	99/106 (93%)	0.08	3 (3%)	50	22	39, 64, 119, 166	0
22	V	24/26 (92%)	0.98	2 (8%)	11	3	43, 68, 102, 119	0
All	All	3916/4068 (96%)	0.46	210 (5%)	25	9	21, 66, 142, 201	0

The worst 5 of 210 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
20	S	3	ARG	14.3
14	M	123	ALA	13.8
14	M	124	PRO	12.5
18	Q	102	GLY	11.3
1	A	1002	G	11.2

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

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates (i)

There are no monosaccharides in this entry.

6.4 Ligands (i)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
24	MG	A	1615	1/1	0.34	0.69	38,38,38,38	1
24	MG	A	1583	1/1	0.41	1.11	38,38,38,38	1
24	MG	A	210	1/1	0.43	0.47	38,38,38,38	1
24	MG	A	1586	1/1	0.50	0.41	38,38,38,38	1
24	MG	A	211	1/1	0.63	0.41	38,38,38,38	1
24	MG	A	1571	1/1	0.71	0.35	38,38,38,38	1
24	MG	A	1550	1/1	0.73	0.40	38,38,38,38	1
24	MG	A	1557	1/1	0.73	0.48	38,38,38,38	1
24	MG	A	1619	1/1	0.74	0.61	38,38,38,38	1
24	MG	A	1573	1/1	0.78	0.78	38,38,38,38	1
24	MG	A	212	1/1	0.78	0.74	38,38,38,38	1
24	MG	A	1620	1/1	0.80	0.15	38,38,38,38	1
24	MG	A	1616	1/1	0.81	1.00	38,38,38,38	1
24	MG	A	1613	1/1	0.81	0.19	38,38,38,38	1
24	MG	A	1582	1/1	0.81	0.20	38,38,38,38	1
24	MG	A	1567	1/1	0.82	0.35	38,38,38,38	0
24	MG	A	1598	1/1	0.83	0.37	38,38,38,38	0
24	MG	A	1600	1/1	0.83	0.23	38,38,38,38	0
24	MG	A	1610	1/1	0.83	0.31	38,38,38,38	1

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
24	MG	A	1591	1/1	0.83	0.51	38,38,38,38	1
24	MG	A	1621	1/1	0.83	0.26	38,38,38,38	1
24	MG	A	1561	1/1	0.84	0.44	38,38,38,38	1
24	MG	A	1551	1/1	0.85	0.94	38,38,38,38	1
24	MG	A	1580	1/1	0.85	0.31	38,38,38,38	0
24	MG	A	1565	1/1	0.86	0.43	38,38,38,38	1
24	MG	A	1602	1/1	0.86	0.10	38,38,38,38	0
24	MG	A	1601	1/1	0.87	0.26	38,38,38,38	0
24	MG	A	1627	1/1	0.87	0.81	38,38,38,38	1
24	MG	D	215	1/1	0.88	0.35	38,38,38,38	0
24	MG	A	1593	1/1	0.89	0.20	38,38,38,38	0
24	MG	A	1611	1/1	0.89	0.19	38,38,38,38	1
24	MG	A	1546	1/1	0.89	0.17	38,38,38,38	1
24	MG	A	1560	1/1	0.89	0.42	38,38,38,38	0
24	MG	A	1552	1/1	0.89	0.35	38,38,38,38	0
24	MG	A	1630	1/1	0.90	0.45	38,38,38,38	1
24	MG	A	1612	1/1	0.90	0.16	38,38,38,38	0
24	MG	A	1595	1/1	0.91	0.29	38,38,38,38	0
24	MG	A	1604	1/1	0.91	0.34	38,38,38,38	0
24	MG	A	1605	1/1	0.91	0.20	38,38,38,38	0
24	MG	A	1563	1/1	0.91	0.48	38,38,38,38	0
24	MG	A	1577	1/1	0.91	0.40	38,38,38,38	0
24	MG	A	1594	1/1	0.91	0.31	38,38,38,38	0
24	MG	A	1564	1/1	0.92	0.09	38,38,38,38	0
23	PAR	A	1545	42/42	0.92	0.24	23,23,23,23	0
24	MG	A	71	1/1	0.92	0.51	38,38,38,38	0
24	MG	A	1596	1/1	0.92	0.34	38,38,38,38	0
24	MG	A	1553	1/1	0.92	0.21	38,38,38,38	1
24	MG	H	213	1/1	0.92	0.61	38,38,38,38	1
24	MG	A	1614	1/1	0.93	0.44	38,38,38,38	1
24	MG	A	1607	1/1	0.93	0.56	38,38,38,38	0
24	MG	A	1629	1/1	0.93	0.44	38,38,38,38	1
24	MG	A	1547	1/1	0.93	0.29	38,38,38,38	0
24	MG	A	214	1/1	0.93	0.20	38,38,38,38	0
24	MG	A	1609	1/1	0.93	0.26	38,38,38,38	0
24	MG	A	1581	1/1	0.94	0.27	38,38,38,38	0
24	MG	A	1587	1/1	0.94	0.40	38,38,38,38	0
24	MG	A	1589	1/1	0.94	0.23	38,38,38,38	0
24	MG	A	1631	1/1	0.94	0.34	38,38,38,38	1
24	MG	A	1549	1/1	0.94	0.34	38,38,38,38	0
24	MG	A	1568	1/1	0.94	0.46	38,38,38,38	0
24	MG	A	1628	1/1	0.95	0.45	38,38,38,38	1

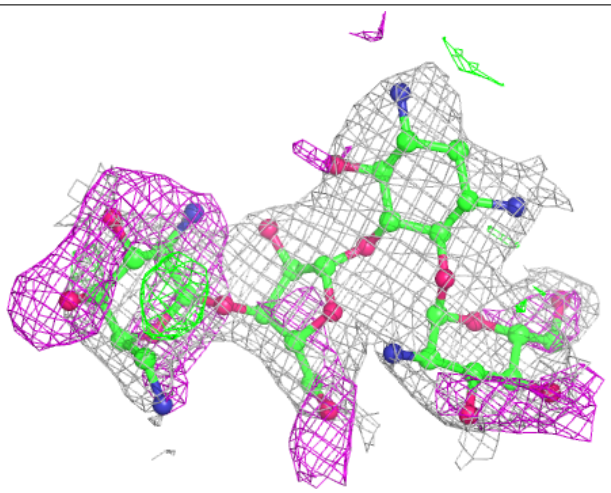
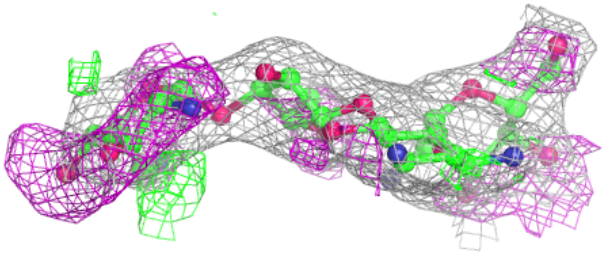
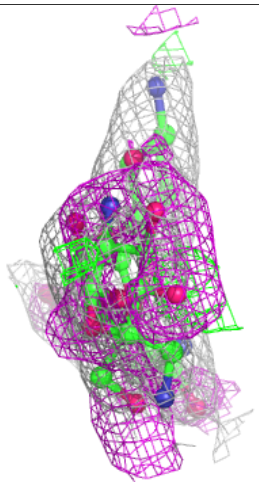
Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
24	MG	A	1618	1/1	0.95	0.14	38,38,38,38	1
24	MG	A	1597	1/1	0.95	0.59	38,38,38,38	1
24	MG	A	1579	1/1	0.95	0.31	38,38,38,38	0
24	MG	A	1584	1/1	0.95	0.11	38,38,38,38	1
24	MG	A	1588	1/1	0.95	0.26	38,38,38,38	0
26	SRY	A	1634	40/40	0.95	0.21	23,23,23,23	0
24	MG	A	1555	1/1	0.96	0.32	38,38,38,38	0
24	MG	A	1592	1/1	0.96	0.60	38,38,38,38	1
24	MG	A	1575	1/1	0.96	0.38	38,38,38,38	0
24	MG	A	1576	1/1	0.96	0.07	38,38,38,38	1
24	MG	A	1606	1/1	0.96	0.15	38,38,38,38	0
24	MG	A	1554	1/1	0.96	0.29	38,38,38,38	0
24	MG	A	1569	1/1	0.96	0.39	38,38,38,38	0
24	MG	A	1632	1/1	0.96	0.40	38,38,38,38	1
24	MG	A	1559	1/1	0.96	0.34	38,38,38,38	0
24	MG	A	1572	1/1	0.96	0.43	38,38,38,38	0
25	SCM	A	1633	23/23	0.96	0.21	23,23,23,23	0
24	MG	A	1590	1/1	0.96	0.28	38,38,38,38	1
24	MG	A	1585	1/1	0.97	0.46	38,38,38,38	0
24	MG	A	1617	1/1	0.97	0.15	38,38,38,38	1
24	MG	A	86	1/1	0.97	0.30	38,38,38,38	0
24	MG	A	1599	1/1	0.97	0.33	38,38,38,38	1
24	MG	A	1548	1/1	0.97	0.35	38,38,38,38	0
24	MG	A	1608	1/1	0.97	0.43	38,38,38,38	1
24	MG	A	1624	1/1	0.97	0.53	38,38,38,38	1
24	MG	A	1626	1/1	0.97	0.30	38,38,38,38	1
24	MG	A	1578	1/1	0.97	0.26	38,38,38,38	0
27	ZN	D	300	1/1	0.97	0.31	47,47,47,47	0
24	MG	A	1574	1/1	0.98	0.50	38,38,38,38	0
24	MG	A	1558	1/1	0.98	0.52	38,38,38,38	0
24	MG	A	1625	1/1	0.98	0.34	38,38,38,38	1
24	MG	A	87	1/1	0.98	0.45	38,38,38,38	1
24	MG	A	1556	1/1	0.98	0.23	38,38,38,38	0
24	MG	A	1562	1/1	0.99	0.69	38,38,38,38	0
24	MG	A	1622	1/1	0.99	0.30	38,38,38,38	1
24	MG	A	1566	1/1	0.99	0.64	38,38,38,38	0
24	MG	A	1603	1/1	0.99	0.41	38,38,38,38	0
27	ZN	N	190	1/1	0.99	0.21	47,47,47,47	1
24	MG	A	1623	1/1	1.00	0.37	38,38,38,38	1
24	MG	A	1570	1/1	1.00	0.22	38,38,38,38	1

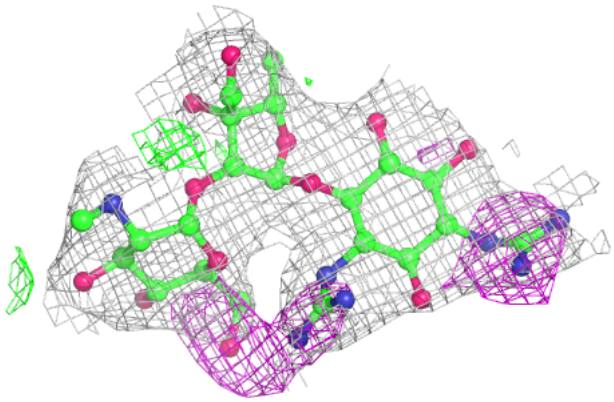
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers

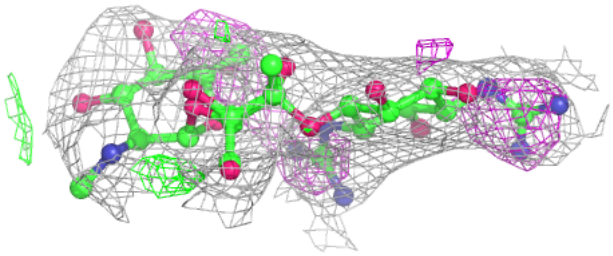
as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation to approximate a three-dimensional view.

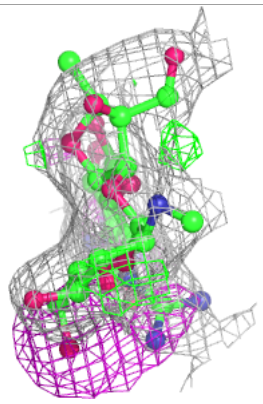
Electron density around PAR A 1545: 2mF_o-DF_c (at 0.7 rmsd) in gray mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)	
	

Electron density around SRY A 1634:

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

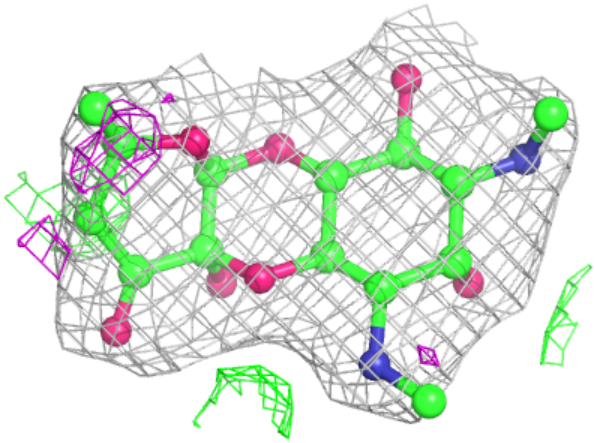


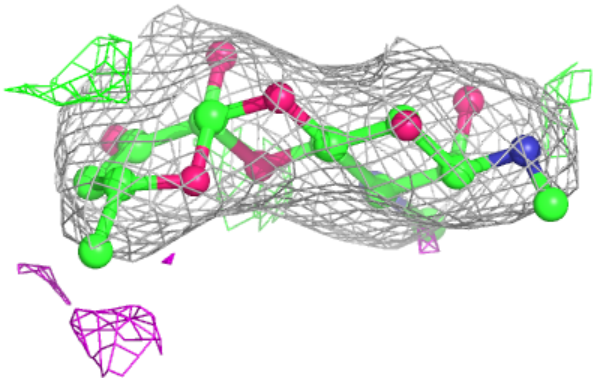


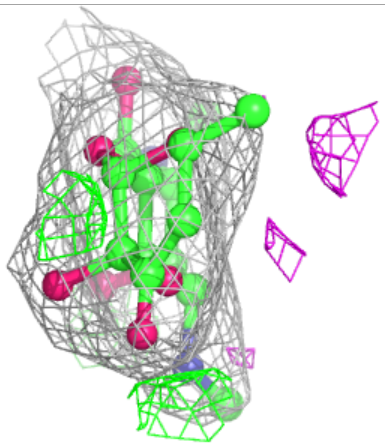


Electron density around SCM A 1633:

$2mF_o$-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 (i)

There are no such residues in this entry.