wwPDB X-ray Structure Validation Summary Report (i)

Nov 17, 2024 – 04:42 AM EST

PDB ID	:	4JYA
Title	:	Crystal structures of pseudouridinilated stop codons with ASLs
Authors	:	Fernandez, I.S.; Ng, C.L.; Kelley, A.C.; Guowei, W.; Yu, Y.T.; Ramakrishnan, V.
Deposited on	:	2013-03-29
Resolution	:	3.10 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:

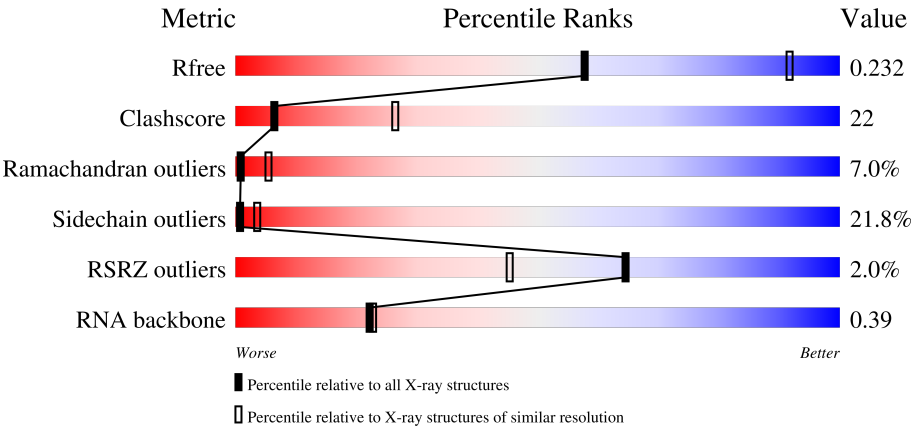
MolProbity	:	4.02b-467
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	1.20.1
EDS	:	3.0
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.003 (Gargrove)
Density-Fitness	:	1.0.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.39

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 3.10 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.



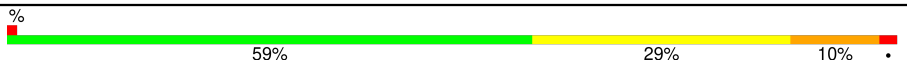
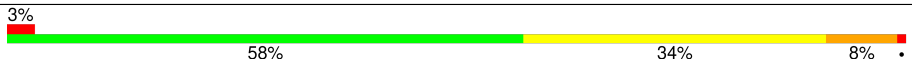
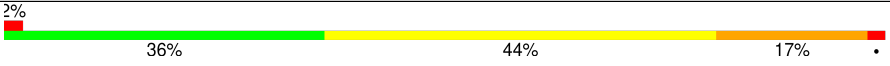
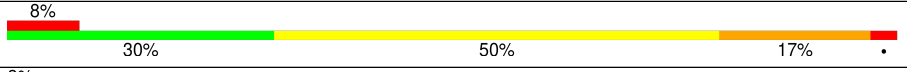
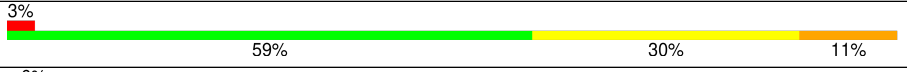
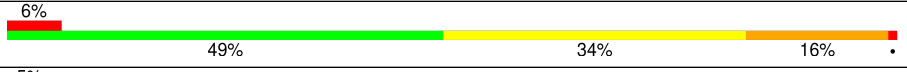
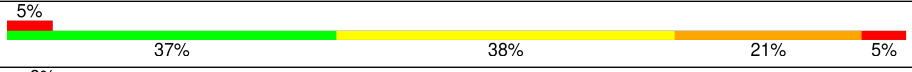
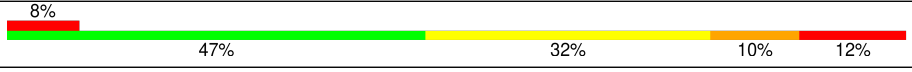
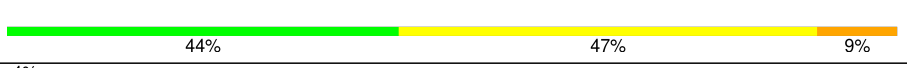
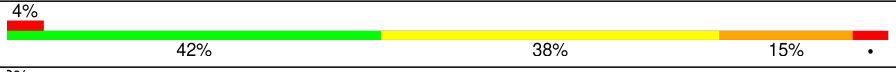
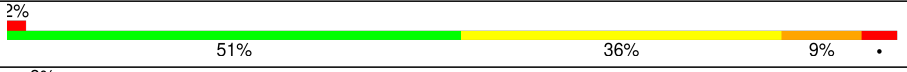
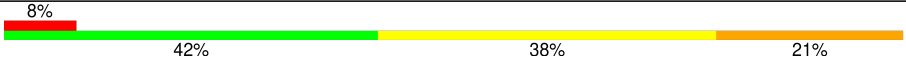
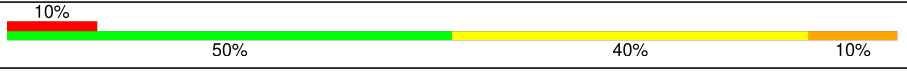
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	164625	1351 (3.10-3.10)
Clashscore	180529	1454 (3.10-3.10)
Ramachandran outliers	177936	1391 (3.10-3.10)
Sidechain outliers	177891	1391 (3.10-3.10)
RSRZ outliers	164620	1351 (3.10-3.10)
RNA backbone	3690	1021 (3.36-2.84)

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	1516	% 40% 40% 17% ·
2	B	234	3% 38% 43% 18% ·
3	C	206	% 40% 40% 16% ·
4	D	208	3% 50% 35% 13% ·

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
5	E	150	
6	F	101	
7	G	155	
8	H	138	
9	I	127	
10	J	98	
11	K	119	
12	L	125	
13	M	120	
14	N	60	
15	O	88	
16	P	83	
17	Q	99	
18	R	70	
19	S	78	
20	T	99	
21	U	24	
22	X	6	
23	Y	10	

2 Entry composition (i)

There are 25 unique types of molecules in this entry. The entry contains 52087 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	1516	32570	14499	6025	10531	1515	0	0	0

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	79	A	G	conflict	GB 55771382

- 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
6	F	101	Total	C	N	O	S	0	0	0
			843	531	155	154	3			

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

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

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

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

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

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

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
I	58	ARG	HIS	conflict	UNP P80374

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
10	J	98	Total	C	N	O	S	0	0	0
			794	499	156	138	1			

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

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

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

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

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
13	M	120	Total	C	N	O	S	0	0	0
			955	591	197	165	2			

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

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

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

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

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

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

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
17	Q	99	Total	C	N	O	S	0	0	0
			823	528	151	142	2			

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

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

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
19	S	78	Total	C	N	O	S	0	0	0
			629	403	114	110	2			

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

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

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

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

- Molecule 22 is a RNA chain called mRNA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
22	X	6	Total	C	N	O	P	0	0	0
			127	58	24	40	5			

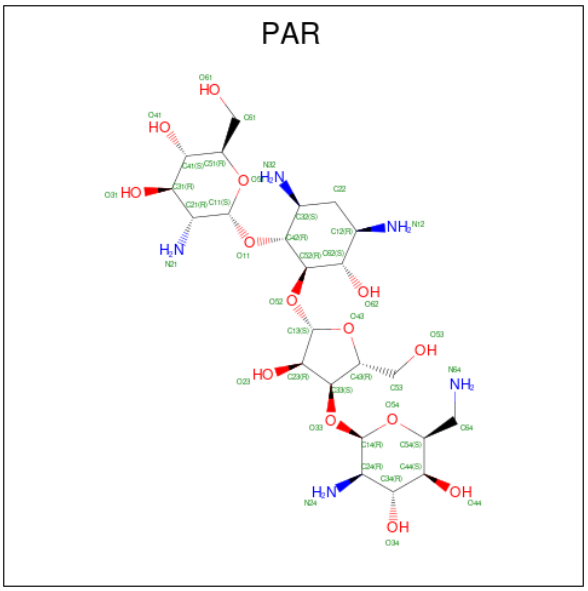
- Molecule 23 is a RNA chain called ASL-tRNA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
23	Y	10	Total	C	N	O	P	0	0	0
			214	96	38	70	10			

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

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
24	A	13	Total	Mg	0	0
			13	13		
24	X	1	Total	Mg	0	0
			1	1		

- Molecule 25 is PAROMOMYCIN (three-letter code: PAR) (formula: $C_{23}H_{45}N_5O_{14}$).



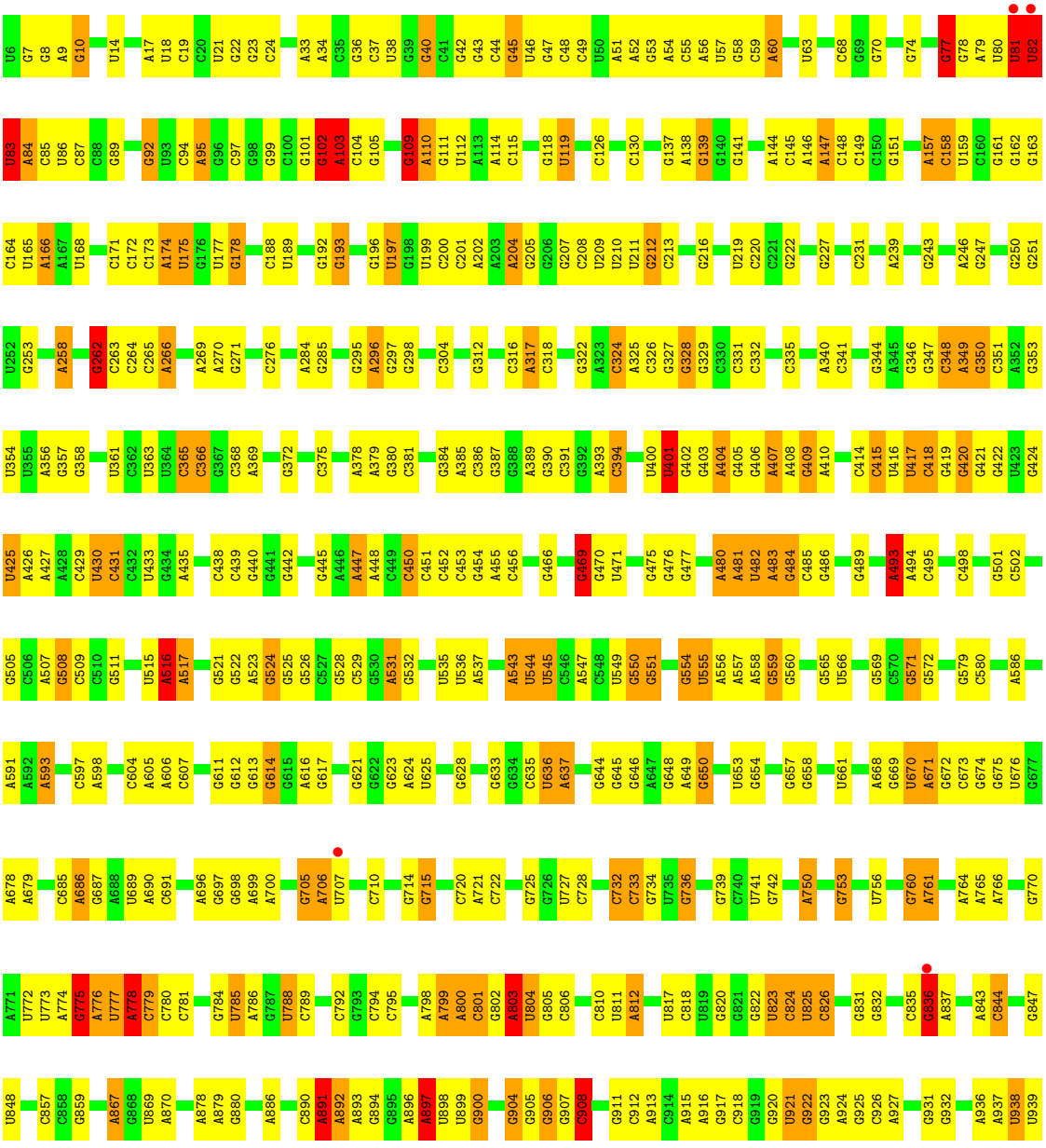
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
25	A	1	Total	C	N	O	0	0
			42	23	5	14		

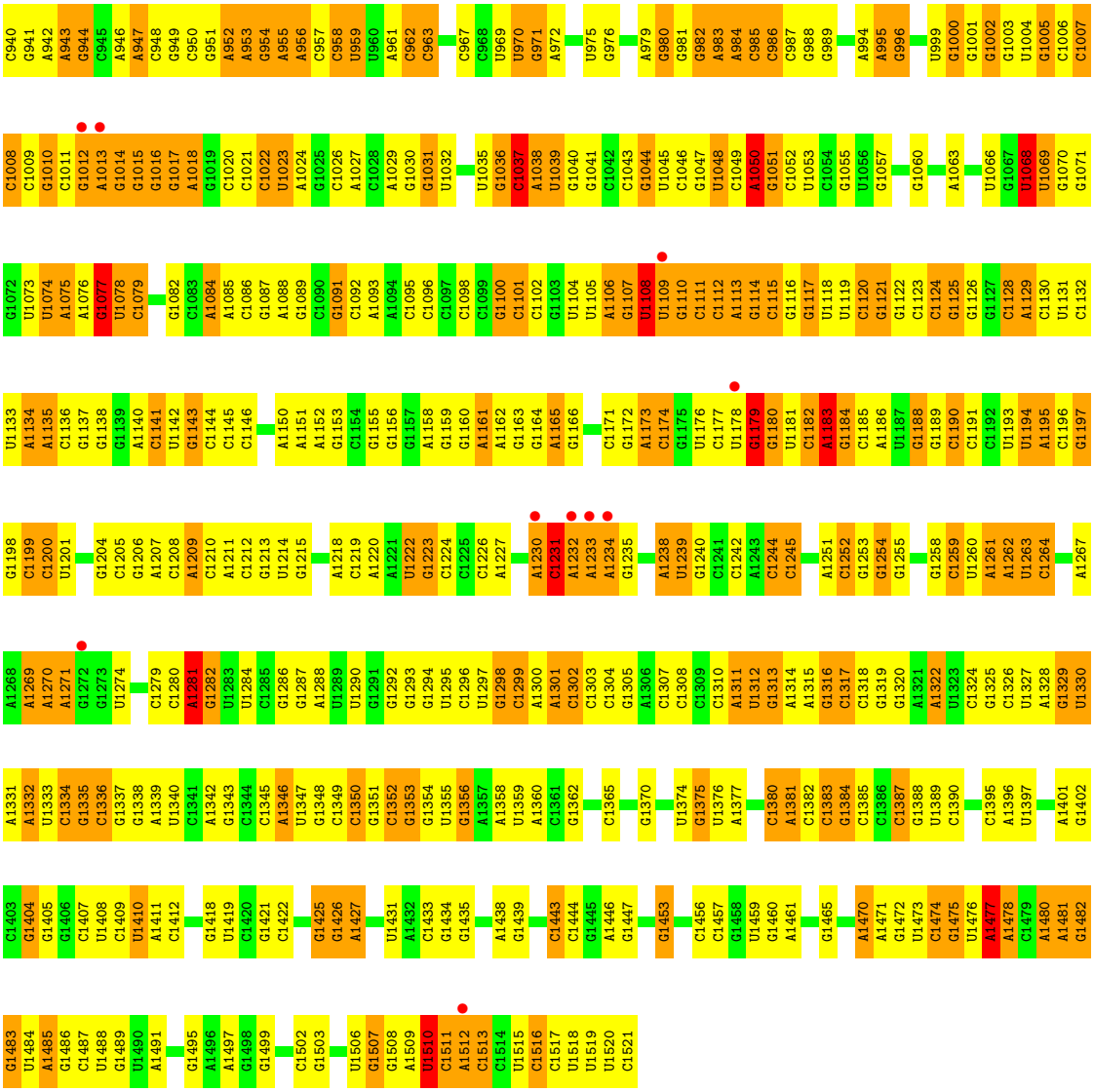
3 Residue-property plots (i)

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

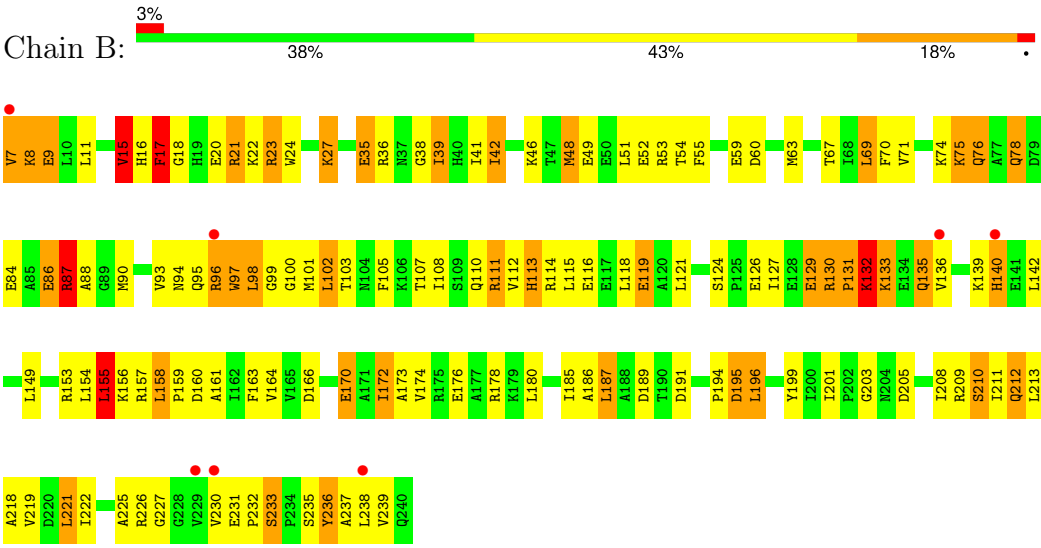
Chain A:





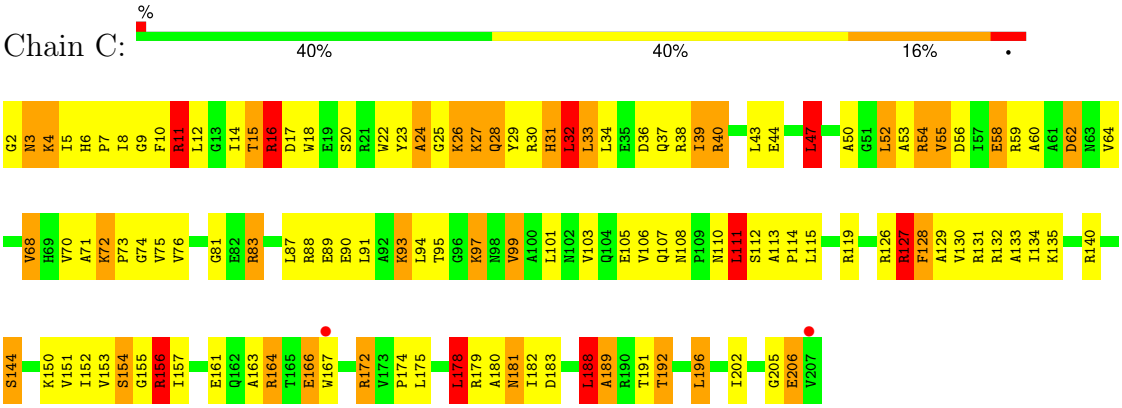
• Molecule 2: 30S ribosomal protein S2

Chain B:



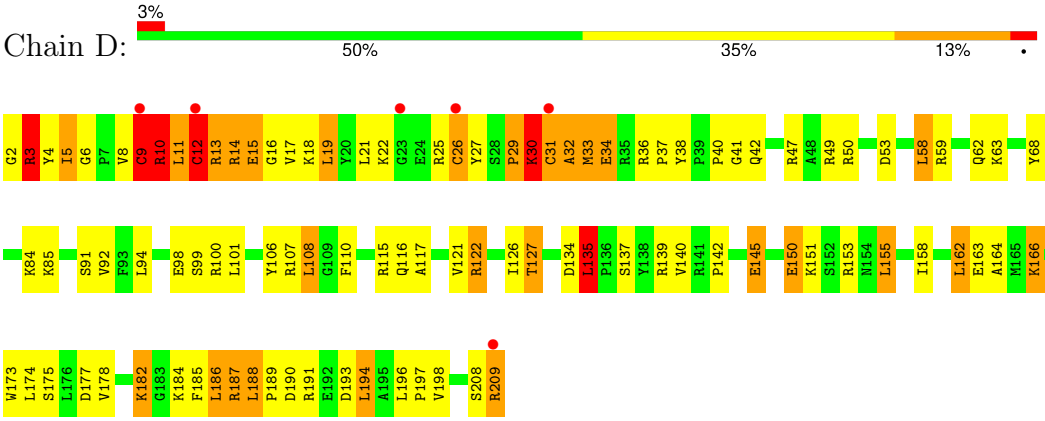
• Molecule 3: 30S ribosomal protein S3

Chain C:



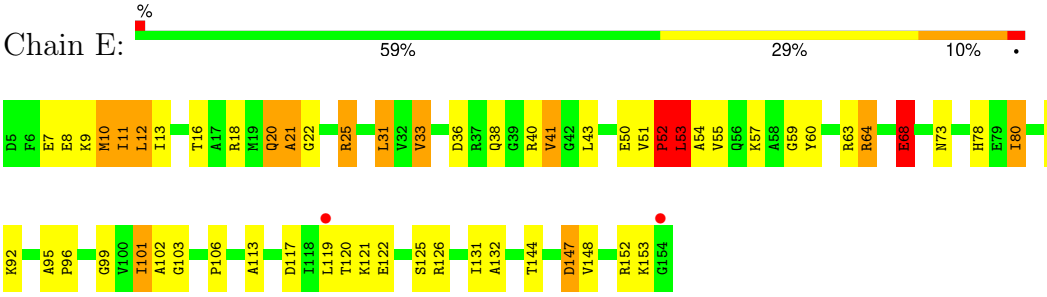
• Molecule 4: 30S ribosomal protein S4

Chain D:



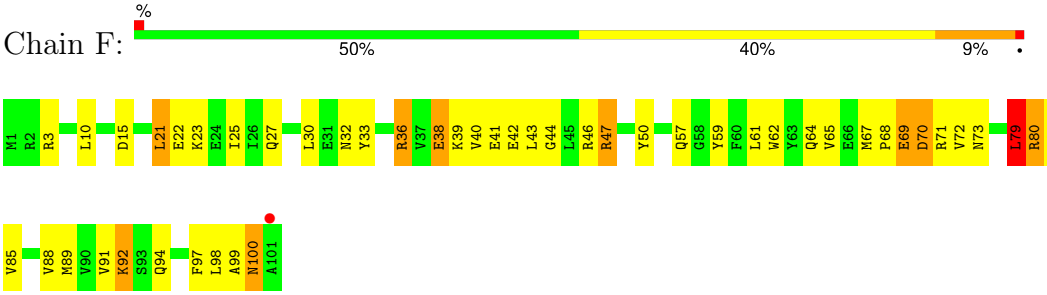
• Molecule 5: 30S ribosomal protein S5

Chain E:



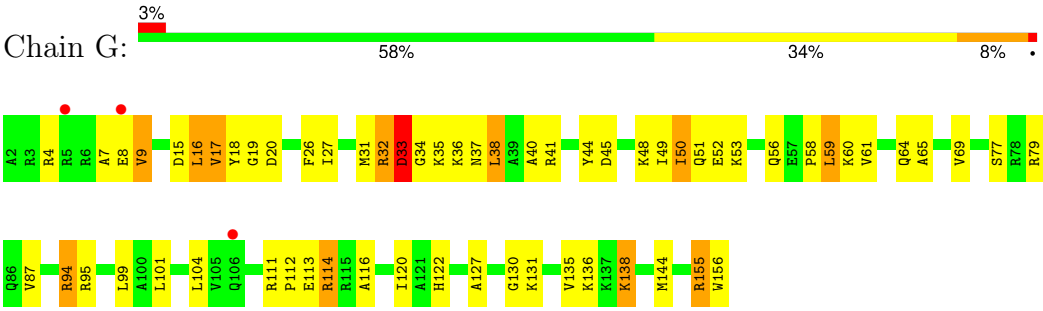
• Molecule 6: 30S ribosomal protein S6

Chain F:



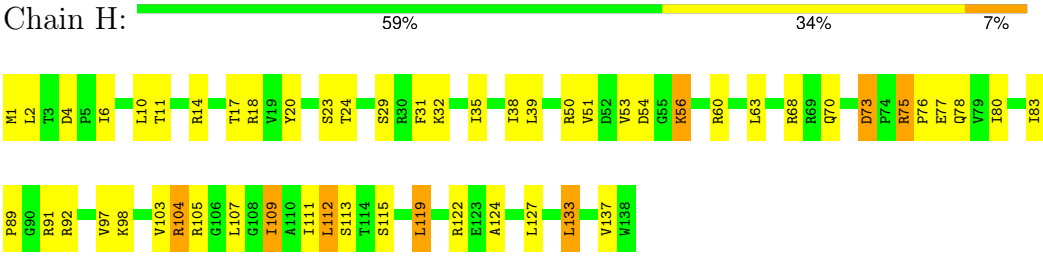
• Molecule 7: 30S ribosomal protein S7

Chain G:



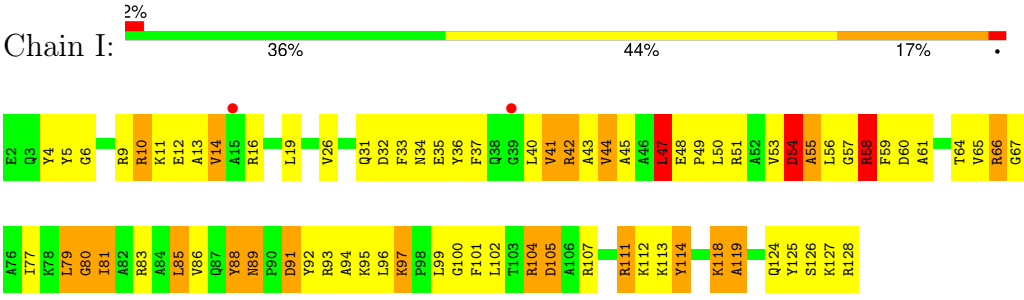
• Molecule 8: 30S ribosomal protein S8

Chain H:



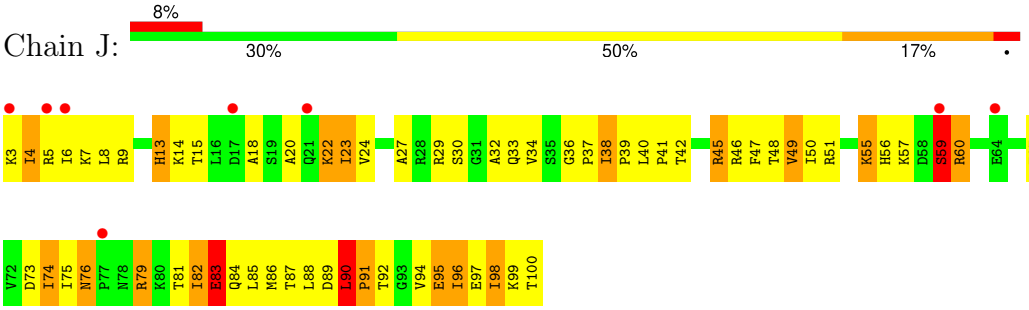
• Molecule 9: 30S ribosomal protein S9

Chain I:



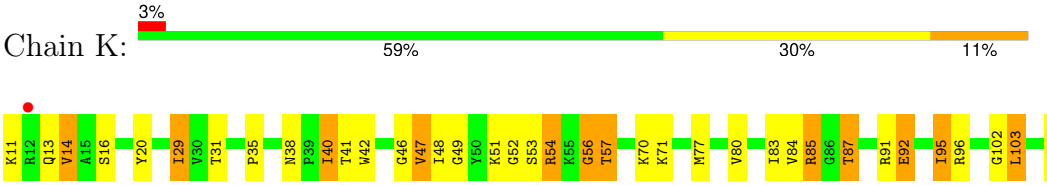
• Molecule 10: 30S ribosomal protein S10

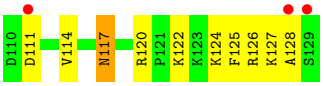
Chain J:



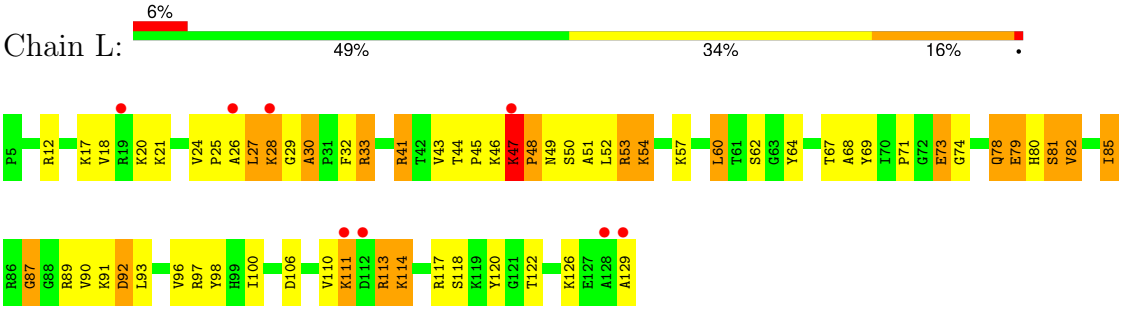
• Molecule 11: 30S ribosomal protein S11

Chain K:

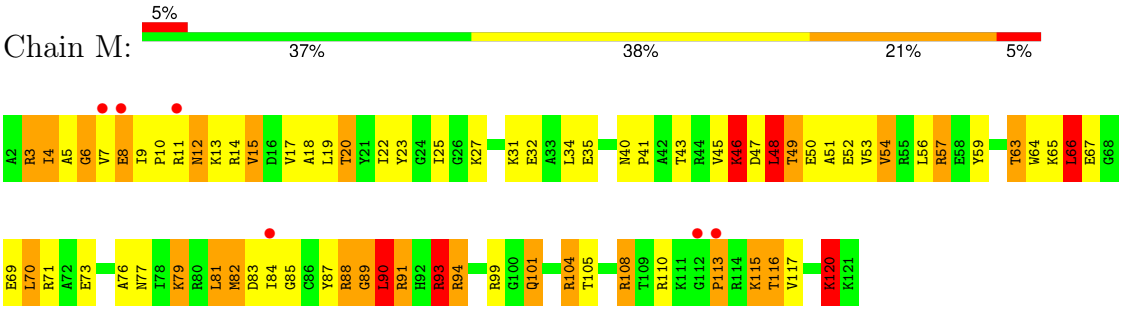




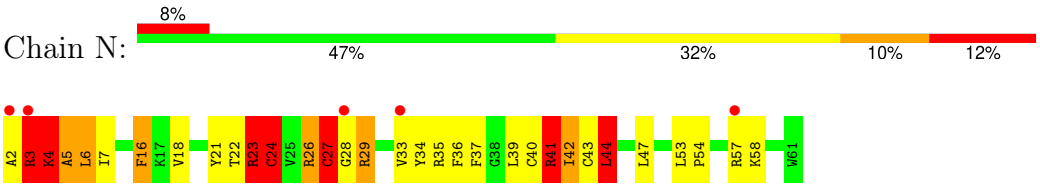
• Molecule 12: 30S ribosomal protein S12



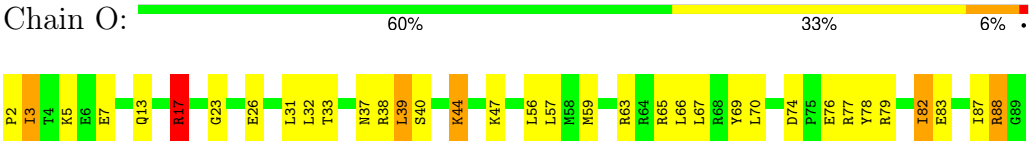
• Molecule 13: 30S ribosomal protein S13



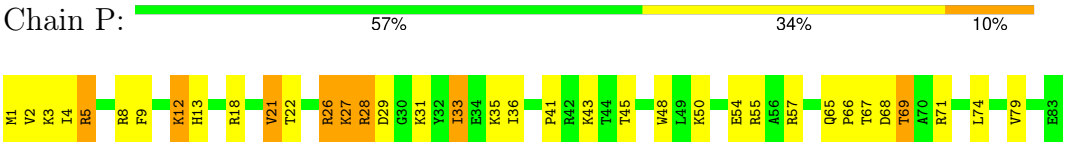
• Molecule 14: 30S ribosomal protein S14



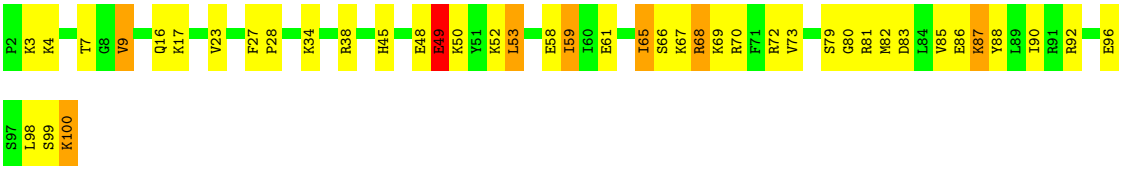
• Molecule 15: 30S ribosomal protein S15



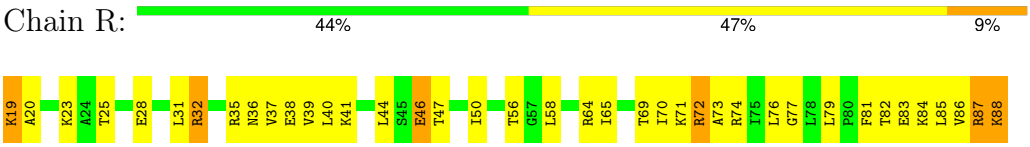
• Molecule 16: 30S ribosomal protein S16



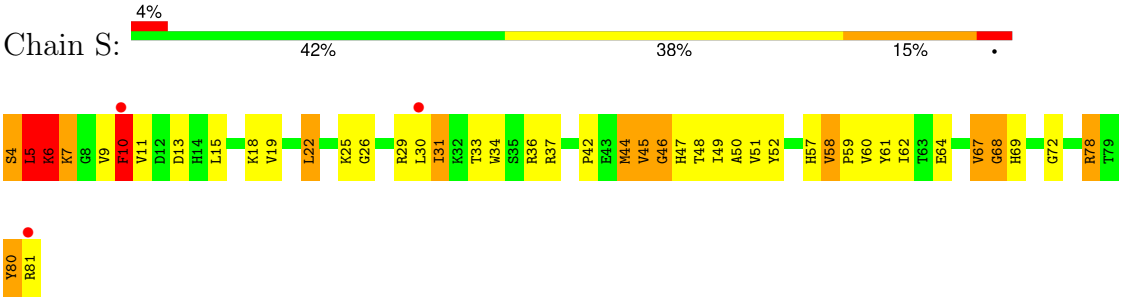
• Molecule 17: 30S ribosomal protein S17



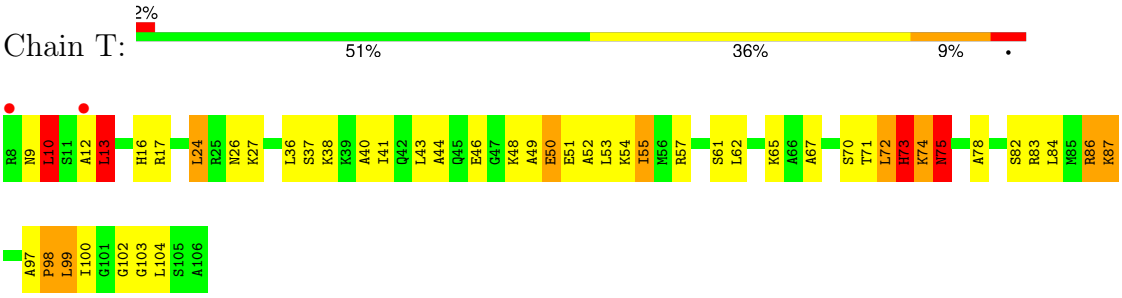
• Molecule 18: 30S ribosomal protein S18



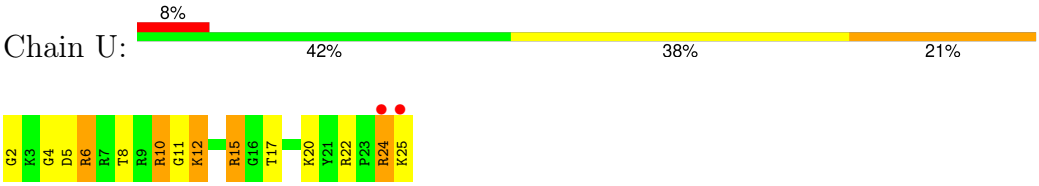
• Molecule 19: 30S ribosomal protein S19



• Molecule 20: 30S ribosomal protein S20



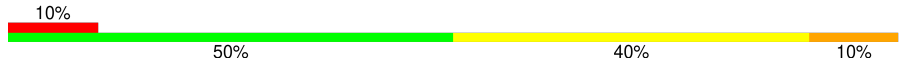
• Molecule 21: 30S ribosomal protein THX



• Molecule 22: mRNA

• Molecule 23: ASL-tRNA

Chain Y:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 41 21 2	Depositor
Cell constants a, b, c, α, β, γ	402.32Å 402.32Å 174.93Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	48.93 – 3.10 48.93 – 3.10	Depositor EDS
% Data completeness (in resolution range)	99.0 (48.93-3.10) 98.9 (48.93-3.10)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.73 (at 3.12Å)	Xtriage
Refinement program	REFMAC 5.8.0021	Depositor
R, R_{free}	0.183 , 0.232 0.183 , 0.232	Depositor DCC
R_{free} test set	12741 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å^2)	78.5	Xtriage
Anisotropy	0.000	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.28 , 47.7	EDS
L-test for twinning[2]	$< \|L\| > = 0.45$, $< L^2 > = 0.28$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	52087	wwPDB-VP
Average B, all atoms (Å^2)	83.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 1.40% 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: PSU, MG, PAR

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	0.51	10/36456 (0.0%)	0.90	100/56897 (0.2%)
2	B	0.55	0/1935	0.86	1/2609 (0.0%)
3	C	0.63	0/1636	0.90	5/2205 (0.2%)
4	D	0.63	1/1733 (0.1%)	0.99	8/2318 (0.3%)
5	E	0.67	0/1162	0.95	2/1564 (0.1%)
6	F	0.49	0/856	0.82	1/1154 (0.1%)
7	G	0.51	0/1276	0.77	0/1709
8	H	0.63	0/1136	0.93	0/1527
9	I	0.55	0/1029	0.89	1/1378 (0.1%)
10	J	0.60	0/807	0.86	0/1085
11	K	0.53	0/900	0.84	1/1213 (0.1%)
12	L	0.67	0/991	1.02	2/1327 (0.2%)
13	M	0.57	0/965	0.93	3/1292 (0.2%)
14	N	0.73	0/501	1.11	3/664 (0.5%)
15	O	0.52	0/745	0.86	1/992 (0.1%)
16	P	0.60	0/716	0.91	1/963 (0.1%)
17	Q	0.63	0/836	0.98	2/1117 (0.2%)
18	R	0.52	0/579	0.86	1/768 (0.1%)
19	S	0.62	0/642	0.93	1/865 (0.1%)
20	T	0.58	0/765	1.00	1/1007 (0.1%)
21	U	0.69	0/212	0.96	0/277
22	X	0.83	0/102	1.65	2/160 (1.2%)
23	Y	0.42	0/239	0.84	1/370 (0.3%)
All	All	0.54	11/56219 (0.0%)	0.90	137/83461 (0.2%)

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
2	B	0	1
4	D	0	2
12	L	0	1
13	M	0	1
20	T	0	5
All	All	0	10

The worst 5 of 11 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	82	U	O3'-P	11.41	1.74	1.61
1	A	81	U	O3'-P	9.90	1.73	1.61
1	A	83	U	O3'-P	9.05	1.72	1.61
1	A	750	A	P-OP2	7.96	1.62	1.49
1	A	836	G	O3'-P	6.39	1.68	1.61

The worst 5 of 137 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
22	X	3	A	O5'-P-OP1	-13.46	93.58	105.70
1	A	1477	A	O5'-P-OP1	-12.08	94.83	105.70
1	A	103	A	O5'-P-OP1	-9.71	96.96	105.70
1	A	891	A	C2'-C3'-O3'	9.37	130.10	109.50
1	A	1486	G	O5'-P-OP2	-9.13	97.48	105.70

There are no chirality outliers.

5 of 10 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	B	191	ASP	Peptide
4	D	30	LYS	Peptide
4	D	31	CYS	Peptide
12	L	87	GLY	Peptide
13	M	66	LEU	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	32570	0	16445	880	0
2	B	1900	0	1951	146	0
3	C	1612	0	1677	135	0
4	D	1703	0	1763	96	0
5	E	1146	0	1207	53	0
6	F	843	0	857	36	0
7	G	1257	0	1296	53	0
8	H	1116	0	1177	40	0
9	I	1011	0	1043	86	0
10	J	794	0	840	92	0
11	K	885	0	904	35	0
12	L	975	0	1062	68	0
13	M	955	0	1021	57	0
14	N	492	0	529	66	0
15	O	734	0	771	24	0
16	P	700	0	720	27	0
17	Q	823	0	891	32	0
18	R	574	0	644	39	0
19	S	629	0	652	56	0
20	T	763	0	861	38	0
21	U	208	0	221	15	0
22	X	127	0	65	20	0
23	Y	214	0	107	5	0
24	A	13	0	0	0	0
24	X	1	0	0	0	0
25	A	42	0	45	3	0
All	All	52087	0	36749	1905	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 22.

The worst 5 of 1905 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:1230:A:O2'	1:A:1231:C:H5"	1.32	1.29
1:A:1021:C:H2'	1:A:1022:C:O2	1.40	1.19
10:J:45:ARG:HB3	10:J:45:ARG:NH1	1.57	1.17
12:L:47:LYS:HB3	12:L:48:PRO:CD	1.76	1.16
10:J:8:LEU:HD13	10:J:20:ALA:CB	1.75	1.15

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	232/234 (99%)	168 (72%)	44 (19%)	20 (9%)	0 4
3	C	204/206 (99%)	137 (67%)	44 (22%)	23 (11%)	0 2
4	D	206/208 (99%)	163 (79%)	25 (12%)	18 (9%)	0 4
5	E	148/150 (99%)	126 (85%)	17 (12%)	5 (3%)	3 17
6	F	99/101 (98%)	79 (80%)	15 (15%)	5 (5%)	1 10
7	G	153/155 (99%)	122 (80%)	22 (14%)	9 (6%)	1 8
8	H	136/138 (99%)	120 (88%)	13 (10%)	3 (2%)	5 24
9	I	125/127 (98%)	86 (69%)	25 (20%)	14 (11%)	0 2
10	J	96/98 (98%)	69 (72%)	20 (21%)	7 (7%)	1 5
11	K	117/119 (98%)	101 (86%)	10 (8%)	6 (5%)	1 10
12	L	123/125 (98%)	93 (76%)	21 (17%)	9 (7%)	1 5
13	M	118/120 (98%)	83 (70%)	21 (18%)	14 (12%)	0 1
14	N	58/60 (97%)	37 (64%)	12 (21%)	9 (16%)	0 0
15	O	86/88 (98%)	75 (87%)	9 (10%)	2 (2%)	5 23
16	P	81/83 (98%)	73 (90%)	8 (10%)	0	100 100
17	Q	97/99 (98%)	81 (84%)	14 (14%)	2 (2%)	5 25
18	R	68/70 (97%)	62 (91%)	5 (7%)	1 (2%)	8 33
19	S	76/78 (97%)	54 (71%)	12 (16%)	10 (13%)	0 1
20	T	97/99 (98%)	74 (76%)	15 (16%)	8 (8%)	1 4
21	U	22/24 (92%)	17 (77%)	5 (23%)	0	100 100
All	All	2342/2382 (98%)	1820 (78%)	357 (15%)	165 (7%)	1 5

5 of 165 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	15	VAL

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	B	17	PHE
2	B	78	GLN
2	B	87	ARG
2	B	130	ARG

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	202/202 (100%)	150 (74%)	52 (26%)	0	1
3	C	160/160 (100%)	119 (74%)	41 (26%)	0	1
4	D	180/180 (100%)	145 (81%)	35 (19%)	1	5
5	E	115/115 (100%)	91 (79%)	24 (21%)	1	4
6	F	90/90 (100%)	73 (81%)	17 (19%)	1	5
7	G	126/126 (100%)	107 (85%)	19 (15%)	2	10
8	H	119/119 (100%)	98 (82%)	21 (18%)	1	7
9	I	98/98 (100%)	73 (74%)	25 (26%)	0	1
10	J	88/88 (100%)	65 (74%)	23 (26%)	0	1
11	K	90/90 (100%)	74 (82%)	16 (18%)	1	6
12	L	104/104 (100%)	87 (84%)	17 (16%)	2	8
13	M	96/96 (100%)	61 (64%)	35 (36%)	0	0
14	N	49/49 (100%)	34 (69%)	15 (31%)	0	0
15	O	79/79 (100%)	66 (84%)	13 (16%)	2	8
16	P	72/72 (100%)	57 (79%)	15 (21%)	1	4
17	Q	94/94 (100%)	77 (82%)	17 (18%)	1	6
18	R	61/61 (100%)	49 (80%)	12 (20%)	1	5
19	S	69/69 (100%)	54 (78%)	15 (22%)	1	4
20	T	76/76 (100%)	62 (82%)	14 (18%)	1	6
21	U	19/19 (100%)	12 (63%)	7 (37%)	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
All	All	1987/1987 (100%)	1554 (78%)	433 (22%)	1 3

5 of 433 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
9	I	125	TYR
12	L	85	ILE
19	S	6	LYS
10	J	38	ILE
11	K	47	VAL

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 21 such sidechains are listed below:

Mol	Chain	Res	Type
9	I	124	GLN
15	O	37	ASN
20	T	90	GLN
20	T	16	HIS
13	M	62	ASN

5.3.3 RNA (i)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
1	A	1515/1516 (99%)	423 (27%)	87 (5%)
22	X	5/6 (83%)	5 (100%)	0
23	Y	9/10 (90%)	2 (22%)	0
All	All	1529/1532 (99%)	430 (28%)	87 (5%)

5 of 430 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
1	A	7	G
1	A	10	G
1	A	21	U
1	A	33	A
1	A	40	G

5 of 87 RNA pucker outliers are listed below:

Mol	Chain	Res	Type
1	A	1068	U
1	A	1233	A
1	A	1107	G
1	A	1165	A
1	A	1279	C

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

2 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	$\#\|Z\| > 2$
22	PSU	X	1	22	18,18,22	2.04	8 (44%)	22,26,33	3.31	13 (59%)
22	PSU	X	4	22,23	18,21,22	1.51	3 (16%)	21,30,33	2.14	4 (19%)

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
22	PSU	X	1	22	-	2/6/22/26	0/2/2/2
22	PSU	X	4	22,23	-	2/7/25/26	0/2/2/2

The worst 5 of 11 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
22	X	1	PSU	C6-C5	4.54	1.40	1.35
22	X	4	PSU	C6-C5	4.18	1.39	1.35
22	X	1	PSU	C4-C5	2.97	1.52	1.44
22	X	1	PSU	C2-N1	2.94	1.40	1.36
22	X	1	PSU	C2'-C1'	-2.42	1.50	1.53

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

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
22	X	4	PSU	N1-C2-N3	6.76	122.30	115.17
22	X	1	PSU	N1-C2-N3	6.11	121.61	115.17
22	X	1	PSU	C4-N3-C2	-5.91	118.22	126.37
22	X	1	PSU	O3'-C3'-C4'	5.77	127.66	111.08
22	X	1	PSU	O5'-C5'-C4'	4.84	127.81	111.33

There are no chirality outliers.

All (4) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
22	X	1	PSU	C3'-C4'-C5'-O5'
22	X	1	PSU	O4'-C4'-C5'-O5'
22	X	4	PSU	C3'-C4'-C5'-O5'
22	X	4	PSU	O4'-C4'-C5'-O5'

There are no ring outliers.

2 monomers are involved in 5 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
22	X	1	PSU	1	0
22	X	4	PSU	4	0

5.5 Carbohydrates (i)

There are no oligosaccharides in this entry.

5.6 Ligand geometry (i)

Of 15 ligands modelled in this entry, 14 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	PAR	A	1614	-	44,45,45	1.30	6 (13%)	63,67,67	2.00	17 (26%)

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	PAR	A	1614	-	-	3/18/94/94	0/4/4/4

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
25	A	1614	PAR	C34-C24	-4.00	1.48	1.53
25	A	1614	PAR	C13-C23	-3.28	1.48	1.52
25	A	1614	PAR	C33-C43	-2.57	1.46	1.52
25	A	1614	PAR	C31-C21	-2.52	1.50	1.53
25	A	1614	PAR	C62-C12	-2.39	1.48	1.53

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

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
25	A	1614	PAR	O54-C54-C64	-9.02	88.74	106.07
25	A	1614	PAR	C44-C34-C24	-3.66	104.92	110.99
25	A	1614	PAR	O53-C53-C43	-3.30	100.08	111.33
25	A	1614	PAR	C11-O51-C51	-3.27	107.34	113.72
25	A	1614	PAR	C14-O54-C54	3.14	119.86	113.72

There are no chirality outliers.

All (3) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
25	A	1614	PAR	C41-C51-C61-O61
25	A	1614	PAR	O51-C51-C61-O61
25	A	1614	PAR	C52-C42-O11-C11

There are no ring outliers.

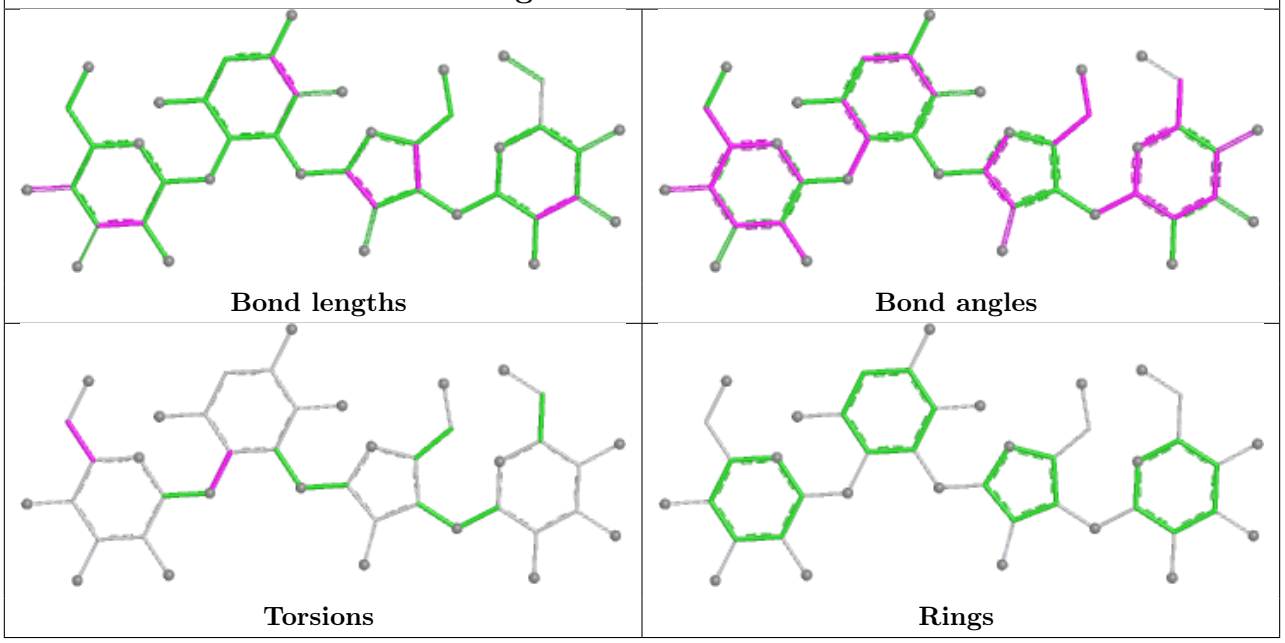
1 monomer is involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
25	A	1614	PAR	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 PAR A 1614



5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

There are no chain breaks in this entry.

6 Fit of model and data

6.1 Protein, DNA and RNA chains

In the following table, the column labelled '#RSRZ> 2' contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95^{th} 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	1516/1516 (100%)	-0.71	14 (0%) 81 66	42, 72, 141, 239	0
2	B	234/234 (100%)	0.06	7 (2%) 52 33	58, 105, 160, 191	0
3	C	206/206 (100%)	-0.17	2 (0%) 79 64	62, 92, 140, 162	0
4	D	208/208 (100%)	-0.28	6 (2%) 54 34	50, 80, 120, 145	0
5	E	150/150 (100%)	-0.66	2 (1%) 74 58	46, 66, 94, 137	0
6	F	101/101 (100%)	-0.44	1 (0%) 79 64	66, 98, 125, 150	0
7	G	155/155 (100%)	-0.31	4 (2%) 57 38	63, 89, 145, 165	0
8	H	138/138 (100%)	-0.72	0 100 100	41, 64, 91, 123	0
9	I	127/127 (100%)	0.10	2 (1%) 70 52	52, 104, 135, 154	0
10	J	98/98 (100%)	0.63	8 (8%) 19 11	57, 126, 180, 221	0
11	K	119/119 (100%)	-0.29	4 (3%) 48 28	44, 76, 113, 168	0
12	L	125/125 (100%)	-0.11	8 (6%) 27 16	39, 73, 116, 168	0
13	M	120/120 (100%)	-0.03	6 (5%) 35 21	62, 90, 130, 159	0
14	N	60/60 (100%)	0.38	5 (8%) 19 10	69, 91, 114, 127	0
15	O	88/88 (100%)	-0.47	0 100 100	51, 79, 116, 161	0
16	P	83/83 (100%)	-0.50	0 100 100	50, 65, 94, 126	0
17	Q	99/99 (100%)	-0.61	0 100 100	42, 67, 102, 126	0
18	R	70/70 (100%)	-0.52	0 100 100	62, 85, 116, 132	0
19	S	78/78 (100%)	0.07	3 (3%) 44 26	67, 107, 161, 176	0
20	T	99/99 (100%)	-0.16	2 (2%) 64 45	45, 70, 113, 128	0
21	U	24/24 (100%)	0.56	2 (8%) 19 10	59, 79, 106, 136	0
22	X	4/6 (66%)	0.09	0 100 100	66, 67, 83, 93	0
23	Y	10/10 (100%)	0.15	1 (10%) 14 8	76, 96, 145, 188	0
All	All	3912/3914 (99%)	-0.40	77 (1%) 64 45	39, 79, 141, 239	0

The worst 5 of 77 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
21	U	25	LYS	10.7
11	K	128	ALA	8.0
11	K	129	SER	7.8
12	L	129	ALA	5.4
1	A	1232	A	5.3

6.2 Non-standard residues in protein, DNA, RNA chains (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, 95^{th} 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
22	PSU	X	1	17/21	0.87	0.14	29,58,70,82	0
22	PSU	X	4	20/21	0.95	0.07	58,67,81,83	0

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, 95^{th} 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	X	101	1/1	0.87	0.30	55,55,55,55	0
24	MG	A	1603	1/1	0.91	0.25	60,60,60,60	0
24	MG	A	1608	1/1	0.92	0.47	52,52,52,52	0
24	MG	A	1602	1/1	0.92	0.20	48,48,48,48	0
24	MG	A	1610	1/1	0.94	0.27	43,43,43,43	0
24	MG	A	1609	1/1	0.95	0.23	47,47,47,47	0
24	MG	A	1607	1/1	0.95	0.36	46,46,46,46	0
24	MG	A	1604	1/1	0.95	0.27	57,57,57,57	0
24	MG	A	1612	1/1	0.96	0.28	44,44,44,44	0
24	MG	A	1611	1/1	0.96	0.39	51,51,51,51	0
25	PAR	A	1614	42/42	0.96	0.08	49,60,88,106	0

Continued on next page...

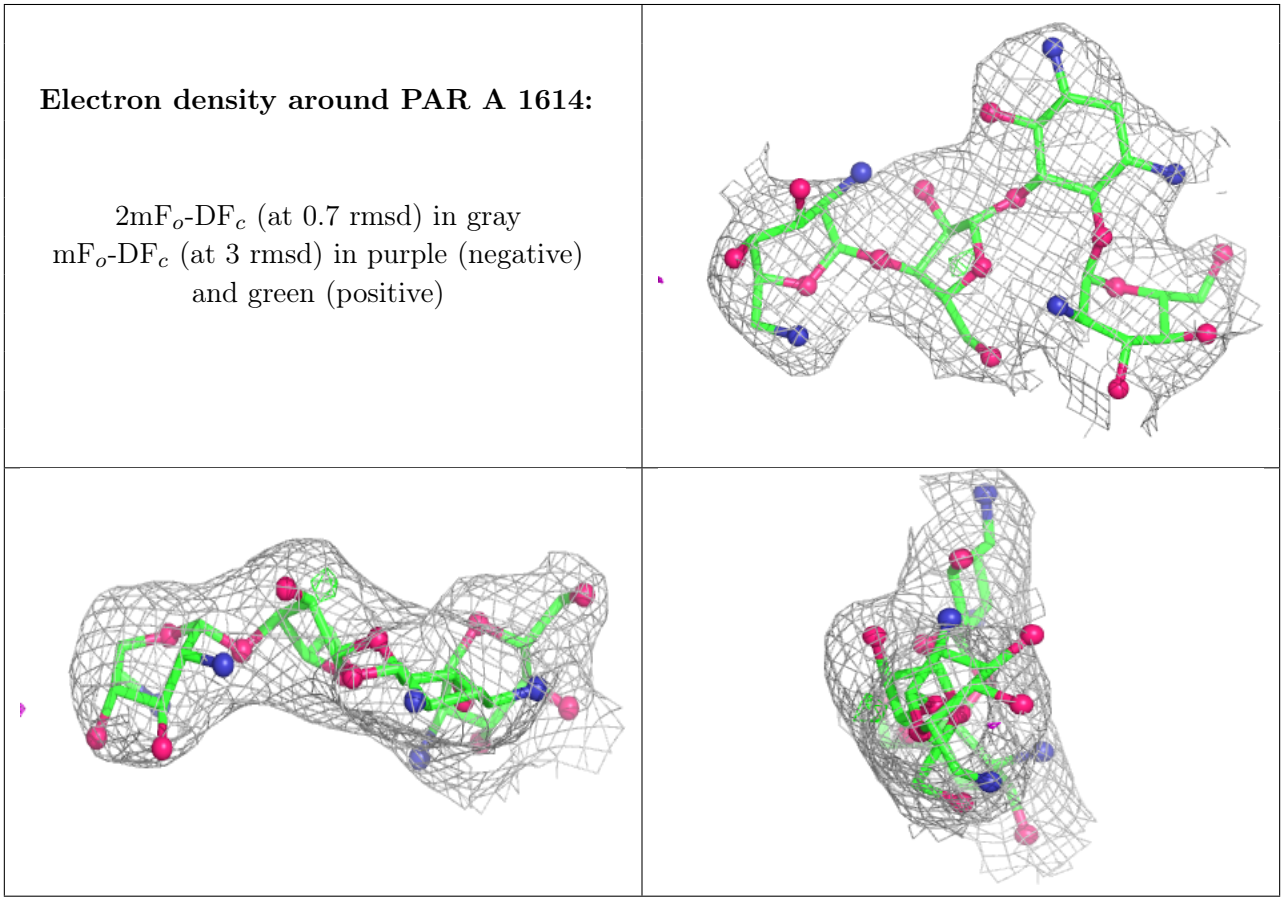
Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
24	MG	A	1613	1/1	0.97	0.13	41,41,41,41	0
24	MG	A	1605	1/1	0.98	0.14	42,42,42,42	0
24	MG	A	1601	1/1	0.98	0.05	49,49,49,49	0
24	MG	A	1606	1/1	0.99	0.05	52,52,52,52	0

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 1614:

$2mF_o$-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.