Full wwPDB X-ray Structure Validation Report ⓘ

May 23, 2020 – 07:40 am BST

PDB ID	:	5L6I
Title	:	Uba1 in complex with Ub-MLN4924 covalent adduct
Authors	:	Misra, M.; Schindelin, H.
Deposited on	:	2016-05-30
Resolution	:	2.76 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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 ⓘ symbol.

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

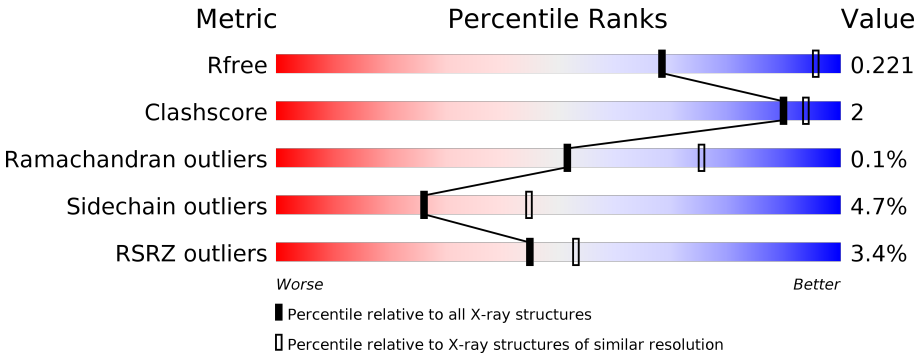
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.11
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

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 2.76 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	1235 (2.78-2.74)
Clashscore	141614	1277 (2.78-2.74)
Ramachandran outliers	138981	1257 (2.78-2.74)
Sidechain outliers	138945	1257 (2.78-2.74)
RSRZ outliers	127900	1207 (2.78-2.74)

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 on 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	1024	2% 90% 7%
1	C	1024	2% 89% 8%
2	B	76	4% 93% 7%
2	D	76	91% 9%
2	E	76	37% 82% 12%

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
5	GOL	A	1113	-	-	-	X

2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 18321 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.

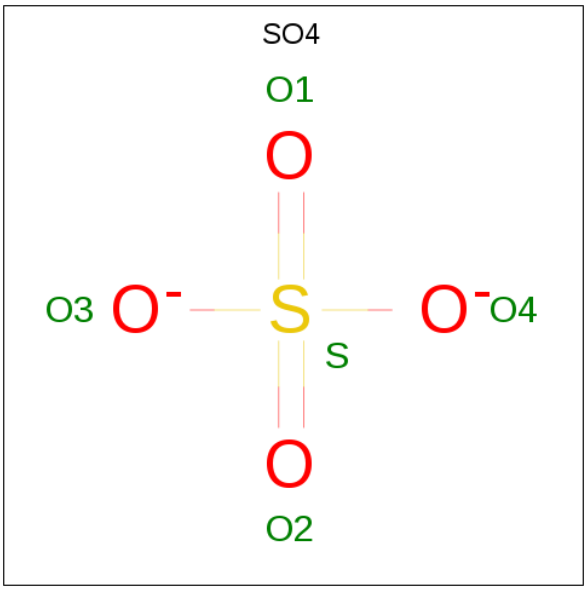
- Molecule 1 is a protein called Ubiquitin-activating enzyme E1 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	1007	Total	C	N	O	S	0	1	0
			7953	5067	1313	1550	23			
1	C	1004	Total	C	N	O	S	0	4	0
			7954	5070	1315	1546	23			

- Molecule 2 is a protein called Ubiquitin-40S ribosomal protein S31.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	76	Total	C	N	O	S	0	0	0
			600	375	105	119	1			
2	D	76	Total	C	N	O	S	0	0	0
			600	375	105	119	1			
2	E	73	Total	C	N	O	S	0	0	0
			577	361	99	116	1			

- Molecule 3 is SULFATE ION (three-letter code: SO4) (formula: O_4S).

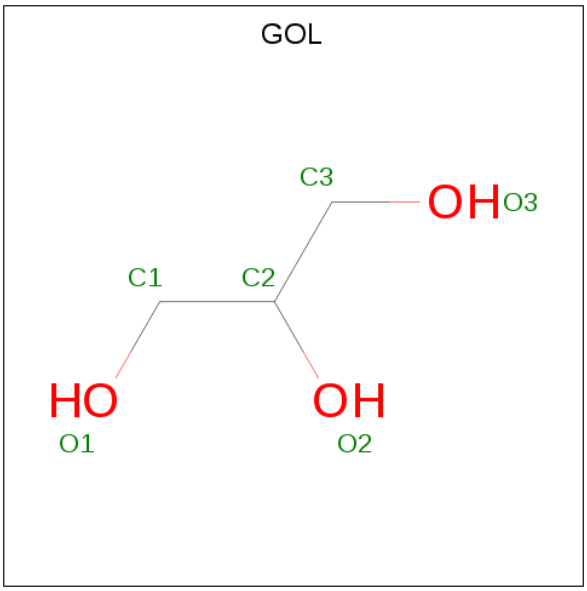


Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total O S 5 4 1	0	0
3	C	1	Total O S 5 4 1	0	0

- Molecule 4 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	B	1	Total Cl 1 1	0	0
4	A	7	Total Cl 7 7	0	0
4	C	8	Total Cl 8 8	0	0

- Molecule 5 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).



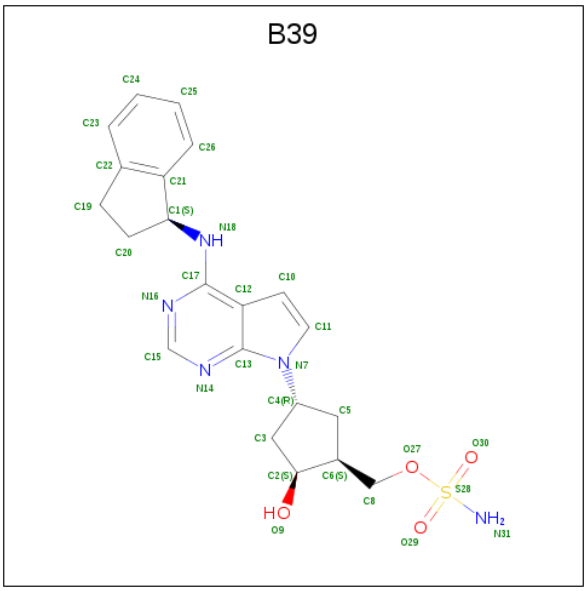
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	1	Total C O 6 3 3	0	0
5	A	1	Total C O 6 3 3	0	0
5	A	1	Total C O 6 3 3	0	0
5	A	1	Total C O 6 3 3	0	0
5	A	1	Total C O 6 3 3	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	1	Total C O 6 3 3	0	0
5	A	1	Total C O 6 3 3	0	0
5	A	1	Total C O 6 3 3	0	0
5	A	1	Total C O 6 3 3	0	0
5	C	1	Total C O 6 3 3	0	0
5	C	1	Total C O 6 3 3	0	0
5	C	1	Total C O 6 3 3	0	0
5	C	1	Total C O 6 3 3	0	0
5	C	1	Total C O 6 3 3	0	0
5	C	1	Total C O 6 3 3	0	0
5	C	1	Total C O 6 3 3	0	0

- Molecule 6 is [(1S,2S,4R)-4-{4-[(1S)-2,3-dihydro-1H-inden-1-ylamino]-7H-pyrrolo[2,3-d]pyrimidin-7-yl}-2-hydroxycyclopentyl]methyl sulfamate (three-letter code: B39) (formula: $C_{21}H_{25}N_5O_4S$).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
6	B	1	Total	C	N	O	S	0	0
			31	21	5	4	1		
6	D	1	Total	C	N	O	S	0	0
			31	21	5	4	1		

- Molecule 7 is water.

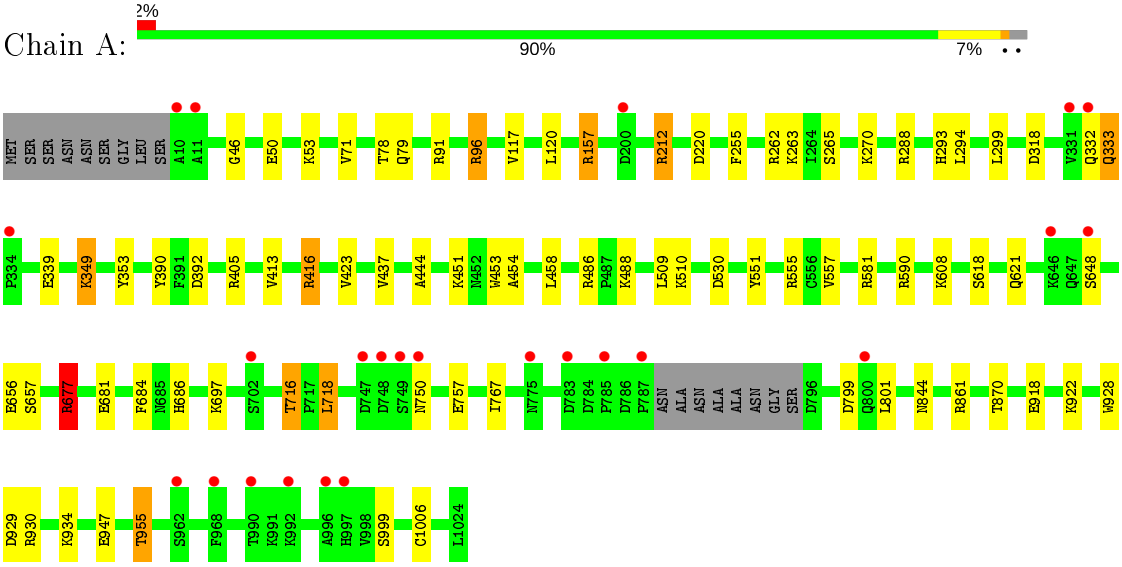
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	223	Total	O	0	0
			223	223		
7	B	12	Total	O	0	0
			12	12		
7	C	205	Total	O	0	0
			205	205		
7	D	12	Total	O	0	0
			12	12		
7	E	1	Total	O	0	0
			1	1		

3 Residue-property plots (i)

These plots are drawn for all protein, RNA and DNA 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.

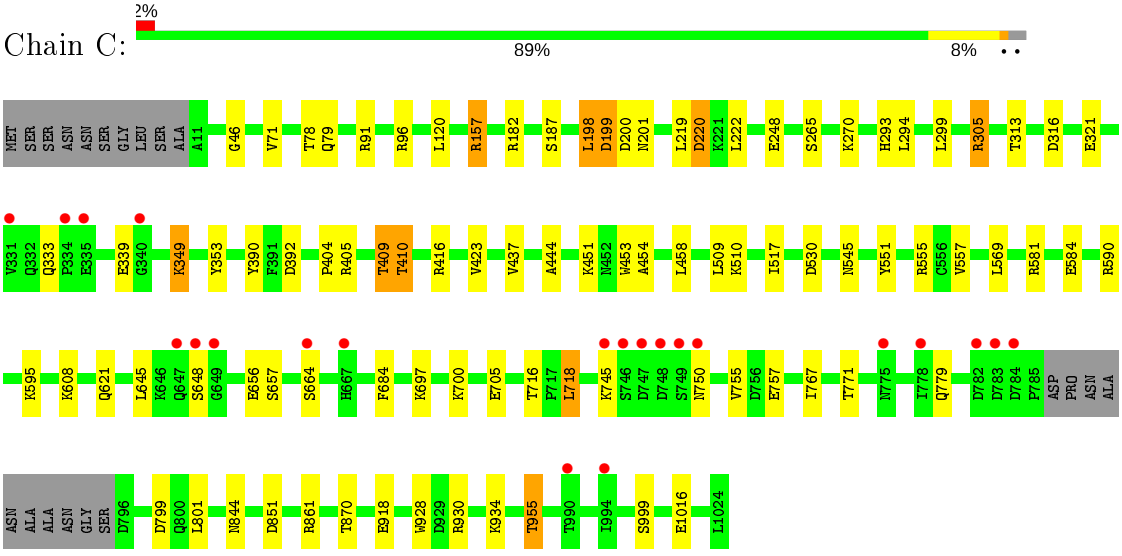
- Molecule 1: Ubiquitin-activating enzyme E1 1

Chain A:



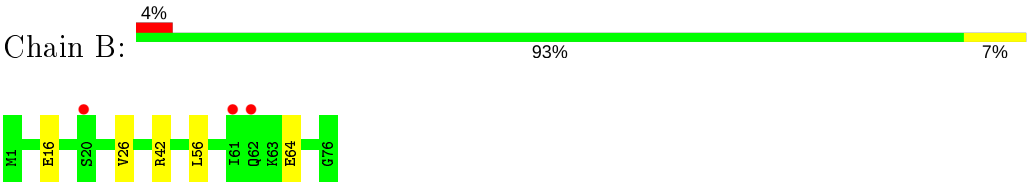
- Molecule 1: Ubiquitin-activating enzyme E1 1

Chain C:



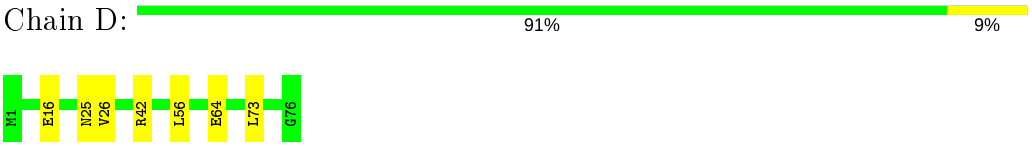
- Molecule 2: Ubiquitin-40S ribosomal protein S31

Chain B:



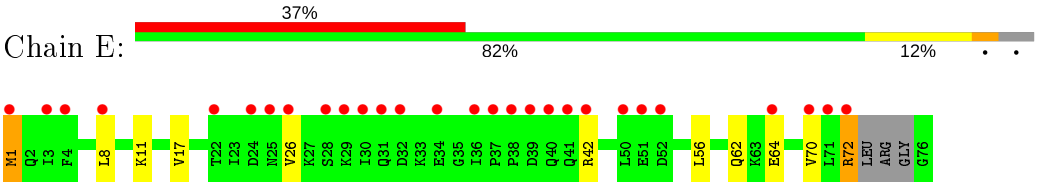
• Molecule 2: Ubiquitin-40S ribosomal protein S31

Chain D:



• Molecule 2: Ubiquitin-40S ribosomal protein S31

Chain E:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 2 21 21	Depositor
Cell constants a, b, c, α, β, γ	72.50Å 191.89Å 230.20Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	20.00 – 2.76 49.35 – 2.76	Depositor EDS
% Data completeness (in resolution range)	99.1 (20.00-2.76) 99.5 (49.35-2.76)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.13	Depositor
$< I/\sigma(I) >$ [1]	2.02 (at 2.77Å)	Xtriage
Refinement program	REFMAC 5.8.0135	Depositor
R, R_{free}	0.174 , 0.220 0.179 , 0.221	Depositor DCC
R_{free} test set	4103 reflections (4.95%)	wwPDB-VP
Wilson B-factor (Å^2)	45.0	Xtriage
Anisotropy	0.544	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.32 , 32.1	EDS
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	18321	wwPDB-VP
Average B, all atoms (Å^2)	52.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.53% 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: SO4, GOL, CSO, B39, CL

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	0.66	0/8113	0.83	11/10978 (0.1%)
1	C	0.66	1/8117 (0.0%)	0.82	3/10981 (0.0%)
2	B	0.61	0/605	0.83	0/812
2	D	0.60	0/605	0.88	0/812
2	E	0.64	0/581	0.94	2/779 (0.3%)
All	All	0.66	1/18021 (0.0%)	0.83	16/24362 (0.1%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	584	GLU	CD-OE1	5.21	1.31	1.25

All (16) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	C	861	ARG	NE-CZ-NH2	-5.96	117.32	120.30
1	A	416	ARG	NE-CZ-NH1	5.88	123.24	120.30
1	A	212	ARG	NE-CZ-NH2	-5.84	117.38	120.30
1	C	182	ARG	NE-CZ-NH2	-5.82	117.39	120.30
1	C	581	ARG	NE-CZ-NH1	5.80	123.20	120.30
2	E	1	MET	CG-SD-CE	5.76	109.42	100.20
1	A	677	ARG	NE-CZ-NH2	-5.74	117.43	120.30
1	A	581	ARG	NE-CZ-NH1	5.70	123.15	120.30
2	E	72	ARG	NE-CZ-NH2	5.64	123.12	120.30
1	A	262	ARG	NE-CZ-NH1	5.50	123.05	120.30
1	A	262	ARG	NE-CZ-NH2	-5.42	117.59	120.30
1	A	861	ARG	NE-CZ-NH2	5.31	122.95	120.30
1	A	929	ASP	CB-CG-OD1	5.26	123.04	118.30
1	A	318	ASP	CB-CG-OD1	5.16	122.94	118.30
1	A	716	THR	N-CA-CB	5.16	120.10	110.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	288	ARG	NE-CZ-NH2	-5.13	117.74	120.30

There are no chirality outliers.

There are no planarity outliers.

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	7953	0	7855	27	0
1	C	7954	0	7860	32	0
2	B	600	0	625	1	0
2	D	600	0	625	2	0
2	E	577	0	597	7	0
3	A	5	0	0	0	0
3	C	5	0	0	0	0
4	A	7	0	0	1	0
4	B	1	0	0	0	0
4	C	8	0	0	1	0
5	A	54	0	72	0	0
5	C	42	0	56	2	0
6	B	31	0	23	0	0
6	D	31	0	23	0	0
7	A	223	0	0	3	0
7	B	12	0	0	0	0
7	C	205	0	0	1	0
7	D	12	0	0	0	0
7	E	1	0	0	0	0
All	All	18321	0	17736	64	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 2.

All (64) 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:C:569:LEU:HB3	2:D:73:LEU:HD22	1.64	0.80
1:A:488:LYS:NZ	7:A:1201:HOH:O	2.20	0.74
1:C:199:ASP:O	2:E:70:VAL:HG11	1.89	0.72
1:A:677:ARG:NH2	1:A:681:GLU:OE2	2.27	0.68
1:C:305:ARG:NH2	1:C:321:GLU:OE2	2.31	0.63
1:C:404:PRO:O	1:C:409:THR:HG21	1.99	0.62
1:C:200:ASP:OD2	2:E:72:ARG:HG3	2.03	0.58
1:C:530:ASP:OD2	1:C:999:SER:OG	2.23	0.56
1:A:530:ASP:OD2	1:A:999:SER:OG	2.24	0.55
1:C:656:GLU:HG3	1:C:801:LEU:HD21	1.90	0.54
1:A:656:GLU:HG3	1:A:801:LEU:HD21	1.89	0.53
4:C:1106:CL:CL	5:C:1114:GOL:O3	2.64	0.52
1:C:198:LEU:HG	2:E:8:LEU:HD11	1.92	0.52
1:A:684:PHE:HB3	1:A:718:LEU:HD22	1.91	0.52
1:C:705:GLU:HG3	2:E:11:LYS:HB2	1.92	0.52
1:C:684:PHE:HB3	1:C:718:LEU:HD22	1.92	0.51
1:C:444:ALA:HB1	1:C:870:THR:HG21	1.93	0.51
1:C:405:ARG:HD2	1:C:423:VAL:O	2.12	0.49
1:C:551:TYR:CZ	1:C:555:ARG:HD2	2.47	0.49
1:C:955:THR:HG21	7:C:1206:HOH:O	2.11	0.49
1:A:551:TYR:CZ	1:A:555:ARG:HD2	2.48	0.49
1:A:405:ARG:HD2	1:A:423:VAL:O	2.13	0.48
1:A:444:ALA:HB1	1:A:870:THR:HG21	1.95	0.48
1:C:437:VAL:HG21	1:C:458:LEU:HD21	1.94	0.48
1:A:157:ARG:HD3	1:A:299:LEU:HD13	1.97	0.47
1:A:437:VAL:HG21	1:A:458:LEU:HD21	1.95	0.47
1:A:557:VAL:HA	1:A:928:TRP:CZ3	2.49	0.47
1:C:316:ASP:OD1	1:C:353:TYR:OH	2.20	0.47
1:C:293:HIS:HA	1:C:390:TYR:CZ	2.50	0.47
1:C:313:THR:CG2	1:C:410:THR:HG23	2.45	0.46
1:A:349:LYS:HD2	1:A:353:TYR:CE2	2.51	0.46
1:C:437:VAL:HG11	1:C:453:TRP:CH2	2.50	0.46
1:C:157:ARG:HD3	1:C:299:LEU:HD13	1.98	0.46
1:C:187:SER:HB3	1:C:198:LEU:HA	1.97	0.45
1:C:557:VAL:HA	1:C:928:TRP:CZ3	2.52	0.45
2:D:26:VAL:HG21	2:D:56:LEU:HD21	1.98	0.45
1:A:413:VAL:O	1:A:413:VAL:HG12	2.16	0.45
1:A:263:LYS:HE2	7:A:1227:HOH:O	2.17	0.45
1:A:96:ARG:NH1	4:A:1102:CL:CL	2.87	0.44
1:C:46:GLY:HA3	1:C:78:THR:OG1	2.17	0.44
1:A:454:ALA:HA	1:A:509:LEU:HD11	2.00	0.44
1:C:199:ASP:N	1:C:199:ASP:OD1	2.50	0.44

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:46:GLY:HA3	1:A:78:THR:OG1	2.17	0.44
2:B:26:VAL:HG21	2:B:56:LEU:HD21	1.99	0.44
1:C:454:ALA:HA	1:C:509:LEU:HD11	2.00	0.44
1:A:437:VAL:HG11	1:A:453:TRP:CH2	2.52	0.44
1:C:851:ASP:HB2	5:C:1111:GOL:H12	1.99	0.43
2:E:26:VAL:HG21	2:E:56:LEU:HD21	1.98	0.43
1:A:293:HIS:HA	1:A:390:TYR:CZ	2.53	0.43
1:C:219:LEU:O	1:C:220:ASP:C	2.57	0.43
1:C:718:LEU:HD13	1:C:844:ASN:ND2	2.34	0.43
2:E:1:MET:HB3	2:E:17:VAL:O	2.19	0.43
1:A:117:VAL:CG2	1:C:755:VAL:HG21	2.49	0.42
1:C:71:VAL:HG22	1:C:91:ARG:HG2	2.00	0.42
1:A:955:THR:HG22	1:A:1006:CSO:HB3	2.02	0.41
1:A:50:GLU:OE2	1:A:53:LYS:NZ	2.52	0.41
1:C:349:LYS:HD2	1:C:353:TYR:CE2	2.56	0.41
1:A:486:ARG:NH2	7:A:1203:HOH:O	2.42	0.41
1:A:718:LEU:HD13	1:A:844:ASN:ND2	2.36	0.41
1:A:255:PHE:CD1	1:A:255:PHE:C	2.93	0.41
1:A:333:GLN:HE21	1:A:333:GLN:HB2	1.63	0.40
2:E:42:ARG:HB3	2:E:70:VAL:CG2	2.51	0.40
1:A:71:VAL:HG22	1:A:91:ARG:HG2	2.02	0.40
1:C:199:ASP:OD2	1:C:201:ASN:ND2	2.54	0.40

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
1	A	1003/1024 (98%)	969 (97%)	33 (3%)	1 (0%)	51 75
1	C	1002/1024 (98%)	968 (97%)	32 (3%)	2 (0%)	47 69
2	B	74/76 (97%)	73 (99%)	1 (1%)	0	100 100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
2	D	74/76 (97%)	73 (99%)	1 (1%)	0	100 100
2	E	70/76 (92%)	66 (94%)	4 (6%)	0	100 100
All	All	2223/2276 (98%)	2149 (97%)	71 (3%)	3 (0%)	51 75

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	220	ASP
1	C	220	ASP
1	C	198	LEU

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
1	A	888/899 (99%)	851 (96%)	37 (4%)	30 49
1	C	888/899 (99%)	842 (95%)	46 (5%)	23 39
2	B	69/69 (100%)	66 (96%)	3 (4%)	29 48
2	D	69/69 (100%)	65 (94%)	4 (6%)	20 35
2	E	67/69 (97%)	65 (97%)	2 (3%)	41 61
All	All	1981/2005 (99%)	1889 (95%)	92 (5%)	26 46

All (92) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	79	GLN
1	A	96	ARG
1	A	120	LEU
1	A	157	ARG
1	A	212	ARG
1	A	265	SER
1	A	270	LYS
1	A	294	LEU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	332	GLN
1	A	333	GLN
1	A	339	GLU
1	A	349	LYS
1	A	392	ASP
1	A	416	ARG
1	A	451	LYS
1	A	510	LYS
1	A	590	ARG
1	A	608	LYS
1	A	618	SER
1	A	621	GLN
1	A	648	SER
1	A	657	SER
1	A	677	ARG
1	A	686	HIS
1	A	697	LYS
1	A	716	THR
1	A	718	LEU
1	A	750	ASN
1	A	757	GLU
1	A	767	ILE
1	A	799	ASP
1	A	918	GLU
1	A	922	LYS
1	A	930	ARG
1	A	934	LYS
1	A	947	GLU
1	A	955	THR
2	B	16	GLU
2	B	42	ARG
2	B	64	GLU
1	C	79	GLN
1	C	96	ARG
1	C	120	LEU
1	C	157	ARG
1	C	199	ASP
1	C	222	LEU
1	C	248	GLU
1	C	265	SER
1	C	270	LYS
1	C	294	LEU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	C	305	ARG
1	C	333	GLN
1	C	339	GLU
1	C	349	LYS
1	C	392	ASP
1	C	409	THR
1	C	410	THR
1	C	416	ARG
1	C	451	LYS
1	C	510	LYS
1	C	517	ILE
1	C	545	ASN
1	C	590	ARG
1	C	595	LYS
1	C	608	LYS
1	C	621	GLN
1	C	645	LEU
1	C	648	SER
1	C	657	SER
1	C	664	SER
1	C	697	LYS
1	C	700	LYS
1	C	716	THR
1	C	718	LEU
1	C	745	LYS
1	C	750	ASN
1	C	757	GLU
1	C	767	ILE
1	C	771	THR
1	C	779	GLN
1	C	799	ASP
1	C	918	GLU
1	C	930	ARG
1	C	934	LYS
1	C	955	THR
1	C	1016	GLU
2	D	16	GLU
2	D	25	ASN
2	D	42	ARG
2	D	64	GLU
2	E	62	GLN
2	E	64	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (8) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	332	GLN
1	A	333	GLN
1	A	452	ASN
1	A	647	GLN
1	A	750	ASN
1	C	647	GLN
1	C	764	HIS
2	D	25	ASN

5.3.3 RNA

There are no RNA molecules in this entry.

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

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	$\#\|Z\| > 2$
1	CSO	A	1006	1	3,6,7	1.28	0	0,6,8	0.00	-
1	CSO	C	1006	1	3,6,7	1.11	0	0,6,8	0.00	-

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	CSO	A	1006	1	-	0/1/5/7	-
1	CSO	C	1006	1	-	0/1/5/7	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	A	1006	CSO	1	0

5.5 Carbohydrates (i)

There are no carbohydrates in this entry.

5.6 Ligand geometry (i)

Of 36 ligands modelled in this entry, 16 are monoatomic - leaving 20 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	$\#\|Z\| > 2$
5	GOL	C	1113	-	5,5,5	0.60	0	5,5,5	0.39	0
5	GOL	A	1113	-	5,5,5	0.58	0	5,5,5	0.37	0
3	SO4	C	1101	-	4,4,4	0.28	0	6,6,6	0.26	0
5	GOL	A	1110	-	5,5,5	0.48	0	5,5,5	0.59	0
5	GOL	A	1117	-	5,5,5	0.51	0	5,5,5	0.47	0
5	GOL	C	1110	-	5,5,5	0.52	0	5,5,5	0.72	0
5	GOL	A	1114	-	5,5,5	0.50	0	5,5,5	0.64	0
5	GOL	C	1114	-	5,5,5	0.25	0	5,5,5	0.41	0
5	GOL	A	1111	-	5,5,5	0.47	0	5,5,5	0.41	0
3	SO4	A	1101	-	4,4,4	0.31	0	6,6,6	0.25	0
5	GOL	C	1111	-	5,5,5	0.61	0	5,5,5	0.88	0
5	GOL	A	1109	-	5,5,5	0.67	0	5,5,5	0.97	0
5	GOL	A	1115	-	5,5,5	0.28	0	5,5,5	0.58	0
5	GOL	C	1112	-	5,5,5	0.43	0	5,5,5	0.39	0
5	GOL	C	1115	-	5,5,5	0.31	0	5,5,5	0.36	0
5	GOL	A	1112	-	5,5,5	0.48	0	5,5,5	0.36	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\|>2$	Counts	RMSZ	$\#\|Z\|>2$
5	GOL	C	1116	-	5,5,5	0.52	0	5,5,5	0.61	0
6	B39	D	101	2	30,35,35	1.75	6 (20%)	36,52,52	3.12	13 (36%)
5	GOL	A	1116	-	5,5,5	0.40	0	5,5,5	0.32	0
6	B39	B	101	2	30,35,35	1.75	6 (20%)	36,52,52	3.36	11 (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
5	GOL	C	1113	-	-	2/4/4/4	-
5	GOL	A	1113	-	-	2/4/4/4	-
5	GOL	A	1110	-	-	4/4/4/4	-
5	GOL	A	1117	-	-	2/4/4/4	-
5	GOL	C	1110	-	-	2/4/4/4	-
5	GOL	A	1114	-	-	2/4/4/4	-
5	GOL	C	1114	-	-	2/4/4/4	-
5	GOL	A	1111	-	-	1/4/4/4	-
5	GOL	C	1111	-	-	2/4/4/4	-
5	GOL	A	1109	-	-	2/4/4/4	-
5	GOL	A	1115	-	-	2/4/4/4	-
5	GOL	C	1112	-	-	2/4/4/4	-
5	GOL	C	1115	-	-	0/4/4/4	-
5	GOL	A	1112	-	-	1/4/4/4	-
5	GOL	C	1116	-	-	0/4/4/4	-
6	B39	D	101	2	-	2/10/35/35	0/5/5/5
5	GOL	A	1116	-	-	2/4/4/4	-
6	B39	B	101	2	-	3/10/35/35	0/5/5/5

All (12) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	D	101	B39	C17-C12	5.10	1.50	1.44
6	B	101	B39	C17-C12	4.93	1.50	1.44
6	D	101	B39	C22-C21	4.68	1.47	1.39
6	B	101	B39	C22-C21	4.57	1.47	1.39
6	B	101	B39	O27-S28	-3.68	1.52	1.57
6	D	101	B39	O27-S28	-3.20	1.53	1.57

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	D	101	B39	C21-C1	-2.64	1.48	1.51
6	B	101	B39	O30-S28	2.44	1.44	1.42
6	B	101	B39	C12-C13	2.26	1.48	1.43
6	B	101	B39	O29-S28	2.22	1.44	1.42
6	D	101	B39	C17-N16	2.12	1.37	1.34
6	D	101	B39	O29-S28	2.07	1.44	1.42

All (24) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
6	B	101	B39	C15-N16-C17	9.51	124.75	116.59
6	B	101	B39	O30-S28-O29	-9.15	111.32	119.97
6	D	101	B39	O30-S28-O29	-8.92	111.53	119.97
6	D	101	B39	C15-N16-C17	8.74	124.09	116.59
6	B	101	B39	C8-O27-S28	7.55	126.67	117.21
6	B	101	B39	C12-C17-N16	-7.40	115.25	121.35
6	B	101	B39	C12-C17-N18	7.26	127.47	120.63
6	D	101	B39	C12-C17-N16	-6.49	116.01	121.35
6	D	101	B39	C12-C17-N18	5.95	126.23	120.63
6	D	101	B39	C8-O27-S28	5.38	123.95	117.21
6	D	101	B39	N14-C15-N16	-4.05	122.35	128.68
6	B	101	B39	N14-C15-N16	-3.79	122.75	128.68
6	D	101	B39	O30-S28-N31	3.78	114.84	109.14
6	D	101	B39	C3-C4-C5	3.25	108.79	103.11
6	D	101	B39	O29-S28-N31	-3.19	104.32	109.14
6	B	101	B39	O27-C8-C6	-3.07	100.62	108.30
6	D	101	B39	C20-C19-C22	-2.75	100.63	103.31
6	D	101	B39	C25-C26-C21	-2.67	117.61	121.01
6	B	101	B39	C2-C3-C4	2.57	107.27	104.94
6	B	101	B39	C19-C20-C1	2.24	109.15	105.54
6	D	101	B39	C24-C23-C22	-2.12	117.73	120.89
6	B	101	B39	C3-C4-C5	2.08	106.75	103.11
6	D	101	B39	C19-C20-C1	2.05	108.83	105.54
6	B	101	B39	C25-C26-C21	-2.02	118.44	121.01

There are no chirality outliers.

All (33) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	C	1112	GOL	O1-C1-C2-C3
5	A	1110	GOL	O1-C1-C2-C3
5	C	1114	GOL	O1-C1-C2-C3

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms
5	A	1114	GOL	O1-C1-C2-C3
5	A	1117	GOL	O1-C1-C2-C3
5	C	1110	GOL	O1-C1-C2-C3
5	A	1116	GOL	C1-C2-C3-O3
5	A	1109	GOL	O1-C1-C2-C3
5	A	1115	GOL	O2-C2-C3-O3
5	C	1113	GOL	O1-C1-C2-O2
5	C	1113	GOL	O1-C1-C2-C3
5	A	1113	GOL	C1-C2-C3-O3
6	B	101	B39	N16-C17-N18-C1
6	B	101	B39	C12-C17-N18-C1
6	B	101	B39	C8-O27-S28-N31
6	D	101	B39	N16-C17-N18-C1
6	D	101	B39	C12-C17-N18-C1
5	A	1110	GOL	O1-C1-C2-O2
5	A	1116	GOL	O2-C2-C3-O3
5	A	1109	GOL	O1-C1-C2-O2
5	A	1110	GOL	C1-C2-C3-O3
5	C	1111	GOL	C1-C2-C3-O3
5	A	1115	GOL	C1-C2-C3-O3
5	C	1112	GOL	O1-C1-C2-O2
5	C	1114	GOL	O1-C1-C2-O2
5	A	1117	GOL	O1-C1-C2-O2
5	C	1110	GOL	O1-C1-C2-O2
5	A	1113	GOL	O2-C2-C3-O3
5	A	1110	GOL	O2-C2-C3-O3
5	A	1114	GOL	O1-C1-C2-O2
5	C	1111	GOL	O2-C2-C3-O3
5	A	1111	GOL	C1-C2-C3-O3
5	A	1112	GOL	C1-C2-C3-O3

There are no ring outliers.

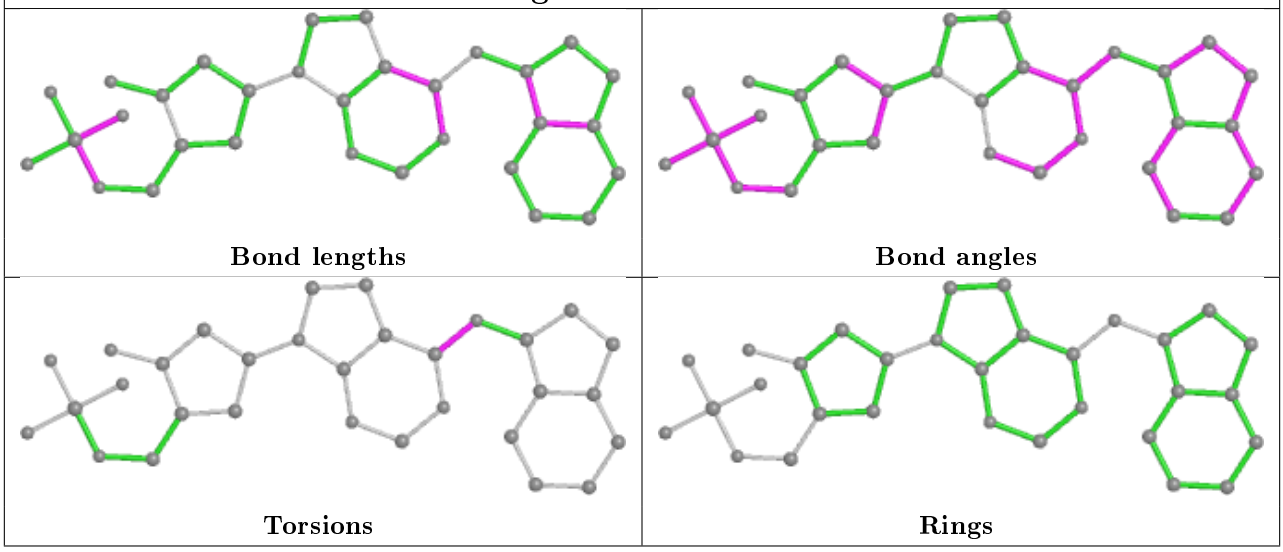
2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	C	1114	GOL	1	0
5	C	1111	GOL	1	0

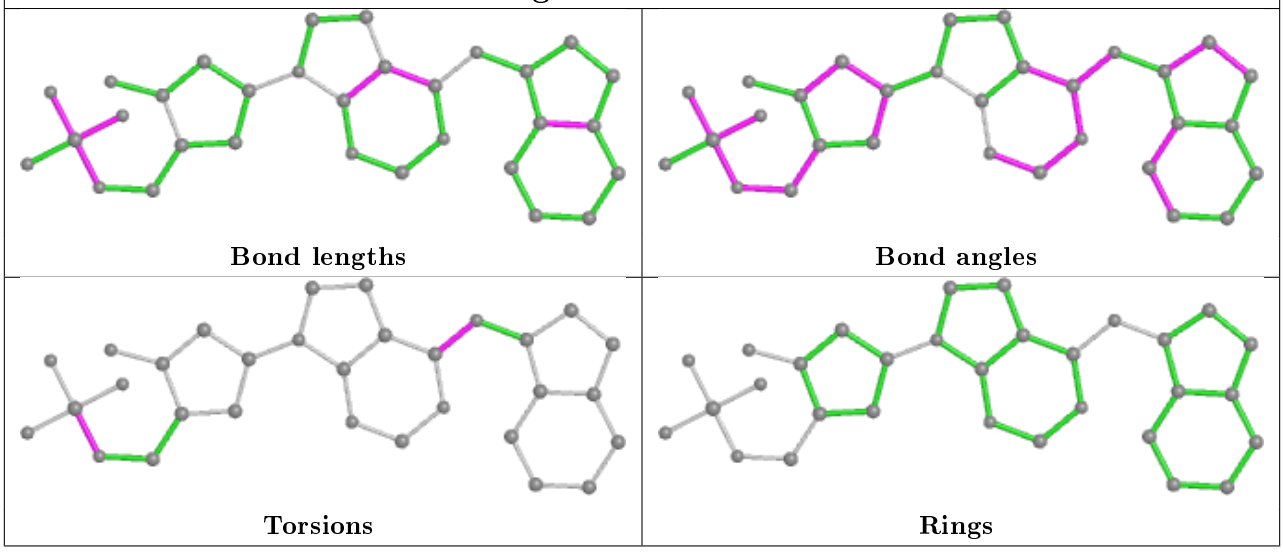
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less then 5% of the Mogul distribution of torsion angles is

within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.

Ligand B39 D 101



Ligand B39 B 101



5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues

There are no chain breaks in this entry.

6 Fit of model and data

6.1 Protein, DNA and RNA chains

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	1006/1024 (98%)	-0.05	24 (2%) 59 68	27, 45, 89, 128	0
1	C	1003/1024 (97%)	-0.07	22 (2%) 62 70	28, 46, 90, 133	0
2	B	76/76 (100%)	0.29	3 (3%) 39 46	33, 50, 79, 84	0
2	D	76/76 (100%)	-0.02	0 100 100	33, 52, 80, 88	0
2	E	73/76 (96%)	2.11	28 (38%) 0 0	66, 93, 124, 135	0
All	All	2234/2276 (98%)	0.02	77 (3%) 45 53	27, 47, 95, 135	0

All (77) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	E	38	PRO	7.0
2	E	40	GLN	5.8
1	C	331	VAL	5.4
1	A	787	PRO	5.0
2	E	30	ILE	4.9
1	C	783	ASP	4.9
1	A	10	ALA	4.7
2	E	31	GLN	4.6
2	E	22	THR	4.6
1	C	750	ASN	4.6
2	E	39	ASP	4.5
2	E	72	ARG	4.2
2	E	25	ASN	4.2
1	A	331	VAL	4.2
2	E	34	GLU	4.1
2	E	8	LEU	4.1
2	E	71	LEU	4.1
1	C	747	ASP	4.1
1	A	750	ASN	3.9
2	E	36	ILE	3.9

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
2	E	32	ASP	3.8
2	E	26	VAL	3.8
1	A	749	SER	3.6
1	C	775	ASN	3.5
1	C	749	SER	3.5
2	E	28	SER	3.4
2	E	24	ASP	3.4
2	E	37	PRO	3.4
1	A	748	ASP	3.3
1	A	785	PRO	3.3
1	A	200	ASP	3.3
2	E	52	ASP	3.2
1	C	784	ASP	3.1
1	A	996	ALA	3.1
1	A	702	SER	3.0
2	E	50	LEU	3.0
2	E	51	GLU	3.0
1	C	745	LYS	3.0
2	E	70	VAL	2.9
1	A	775	ASN	2.9
1	A	11	ALA	2.8
2	E	3	ILE	2.8
2	E	41	GLN	2.8
1	C	748	ASP	2.8
2	E	64	GLU	2.7
1	A	646	LYS	2.7
1	A	783	ASP	2.7
1	C	334	PRO	2.6
2	B	62	GLN	2.6
2	E	42	ARG	2.6
1	A	990	THR	2.5
2	E	4	PHE	2.5
1	A	992	LYS	2.5
2	B	20	SER	2.5
1	C	649	GLY	2.5
1	C	782	ASP	2.4
2	E	29	LYS	2.4
1	A	747	ASP	2.4
1	C	778	ILE	2.3
1	A	800	GLN	2.3
1	C	335	GLU	2.3
1	C	664	SER	2.3

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	332	GLN	2.2
1	C	667	HIS	2.2
1	A	334	PRO	2.2
1	C	994	ILE	2.2
1	A	997	HIS	2.2
2	E	1	MET	2.1
1	A	968	PHE	2.1
1	C	647	GLN	2.1
1	A	648	SER	2.1
1	C	746	SER	2.1
1	C	990	THR	2.1
1	A	962	SER	2.1
1	C	340	GLY	2.1
2	B	61	ILE	2.1
1	C	648	SER	2.0

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
1	CSO	A	1006	7/8	0.94	0.16	47,48,54,56	0
1	CSO	C	1006	7/8	0.97	0.15	38,43,52,59	0

6.3 Carbohydrates (i)

There are no carbohydrates 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
5	GOL	A	1113	6/6	0.77	0.51	75,80,83,84	0

Continued on next page...

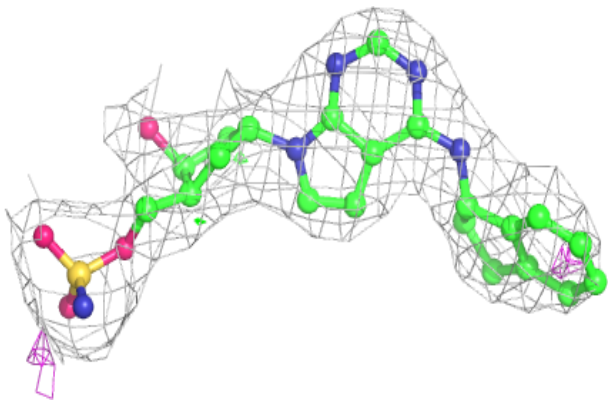
Continued from previous page...

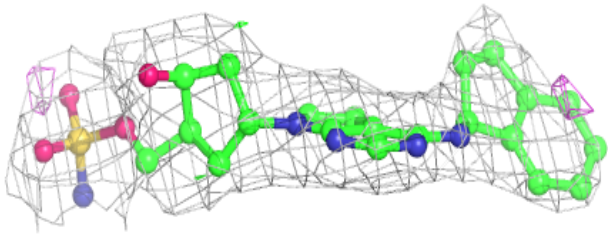
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
4	CL	A	1106	1/1	0.78	0.19	69,69,69,69	0
4	CL	C	1108	1/1	0.81	0.20	69,69,69,69	0
4	CL	A	1107	1/1	0.81	0.18	68,68,68,68	0
4	CL	C	1109	1/1	0.81	0.17	66,66,66,66	0
5	GOL	A	1116	6/6	0.82	0.29	68,69,71,71	0
5	GOL	C	1111	6/6	0.83	0.25	60,67,72,74	0
4	CL	C	1104	1/1	0.83	0.30	73,73,73,73	0
5	GOL	C	1116	6/6	0.86	0.29	64,68,71,74	0
4	CL	C	1107	1/1	0.86	0.13	64,64,64,64	0
5	GOL	A	1117	6/6	0.89	0.41	73,78,81,81	0
5	GOL	C	1114	6/6	0.89	0.21	76,77,79,81	0
5	GOL	A	1110	6/6	0.90	0.26	43,48,53,57	0
4	CL	C	1103	1/1	0.90	0.08	59,59,59,59	0
5	GOL	C	1115	6/6	0.90	0.27	72,73,78,81	0
4	CL	C	1105	1/1	0.91	0.08	52,52,52,52	0
5	GOL	A	1112	6/6	0.91	0.22	58,65,66,66	0
4	CL	A	1108	1/1	0.91	0.17	73,73,73,73	0
5	GOL	A	1115	6/6	0.91	0.27	70,72,72,76	0
5	GOL	C	1113	6/6	0.92	0.19	63,67,69,71	0
5	GOL	C	1112	6/6	0.92	0.31	64,64,67,68	0
4	CL	A	1104	1/1	0.92	0.25	77,77,77,77	0
5	GOL	A	1114	6/6	0.93	0.40	64,65,68,69	0
5	GOL	A	1109	6/6	0.94	0.24	47,54,54,57	0
4	CL	B	102	1/1	0.94	0.09	58,58,58,58	0
4	CL	A	1103	1/1	0.94	0.10	64,64,64,64	0
4	CL	A	1102	1/1	0.95	0.06	46,46,46,46	0
4	CL	C	1102	1/1	0.95	0.27	52,52,52,52	0
5	GOL	C	1110	6/6	0.96	0.19	31,37,38,38	0
5	GOL	A	1111	6/6	0.96	0.18	51,54,55,58	0
6	B39	D	101	31/31	0.96	0.18	31,39,60,62	0
6	B39	B	101	31/31	0.97	0.15	30,37,49,50	0
4	CL	C	1106	1/1	0.97	0.12	55,55,55,55	0
4	CL	A	1105	1/1	0.98	0.26	45,45,45,45	0
3	SO4	C	1101	5/5	0.99	0.11	33,35,36,38	0
3	SO4	A	1101	5/5	0.99	0.12	31,32,34,34	0

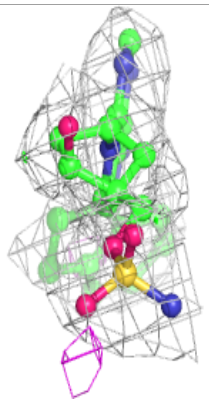
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation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

Electron density around B39 D 101:

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

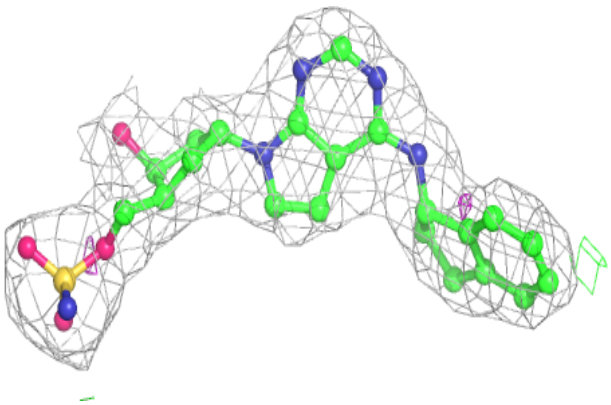


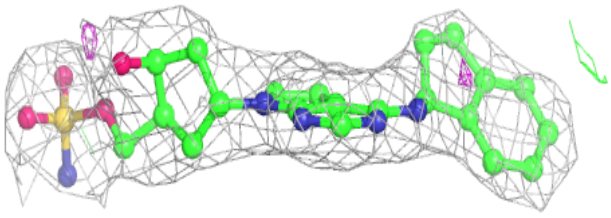


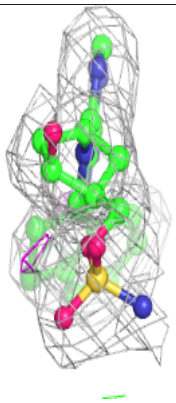


Electron density around B39 B 101:

$2mF_o$-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.