



Full wwPDB X-ray Structure Validation Report ⓘ

Apr 27, 2022 – 05:04 PM EDT

PDB ID : 7TVF
Title : Crystal structure of the SHOC2-MRAS-PP1CA (SMP) complex to a resolution of 2.17 Angstrom
Authors : Bonsor, D.A.; Simanshu, D.K.
Deposited on : 2022-02-04
Resolution : 2.17 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 : 2.28.1
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) : 2.28.1

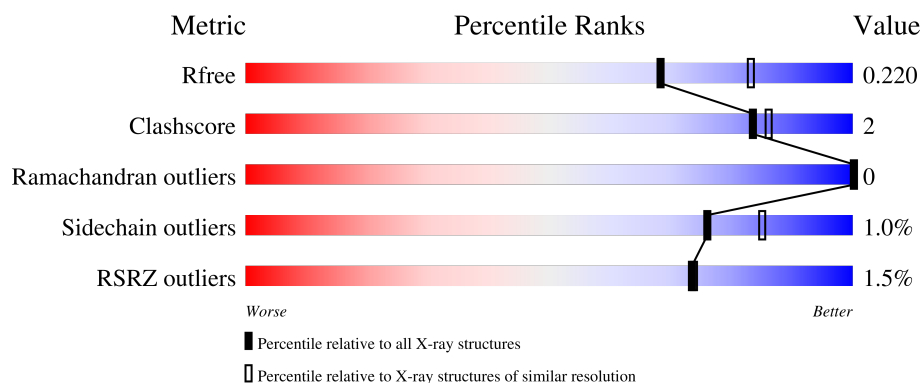
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.17 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	6864 (2.20-2.16)
Clashscore	141614	7689 (2.20-2.16)
Ramachandran outliers	138981	7564 (2.20-2.16)
Sidechain outliers	138945	7564 (2.20-2.16)
RSRZ outliers	127900	6738 (2.20-2.16)

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 $\leq 5\%$. 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	C	329	<div> <div>3%</div> <div>84%</div> <div>7%</div> <div>9%</div> </div>
1	F	329	<div> <div>81%</div> <div>8%</div> <div>11%</div> </div>
2	B	179	<div> <div>%</div> <div>92%</div> <div>6%</div> <div>.</div> </div>
2	E	179	<div> <div>%</div> <div>90%</div> <div>6%</div> <div>..</div> </div>
3	A	582	<div> <div>%</div> <div>84%</div> <div>.</div> <div>12%</div> </div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	D	582	<div><div></div><div>2%</div><div>82%</div><div>7%</div><div>12%</div></div>

2 Entry composition

There are 12 unique types of molecules in this entry. The entry contains 16911 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 Serine/threonine-protein phosphatase PP1-alpha catalytic subunit.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	F	293	Total	C	N	O	S	0	3	0
			2368	1521	397	432	18			
1	C	299	Total	C	N	O	S	0	1	0
			2383	1527	397	441	18			

- Molecule 2 is a protein called Ras-related protein M-Ras.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	E	173	Total	C	N	O	S	0	1	0
			1403	897	239	263	4			
2	B	175	Total	C	N	O	S	0	0	0
			1413	902	240	266	5			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
E	0	GLY	-	expression tag	UNP O14807
E	71	LEU	GLN	engineered mutation	UNP O14807
B	0	GLY	-	expression tag	UNP O14807
B	71	LEU	GLN	engineered mutation	UNP O14807

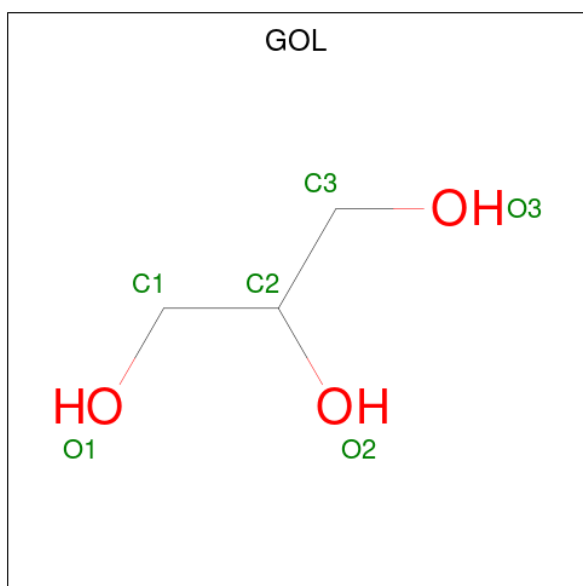
- Molecule 3 is a protein called Leucine-rich repeat protein SHOC-2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	A	513	Total	C	N	O	S	0	6	0
			4063	2569	699	775	20			
3	D	515	Total	C	N	O	S	0	1	0
			4035	2553	694	770	18			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	GLY	-	expression tag	UNP Q9UQ13
D	1	GLY	-	expression tag	UNP Q9UQ13

- Molecule 4 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	F	1	Total	C	O	0	0
			6	3	3		
4	F	1	Total	C	O	0	0
			6	3	3		
4	C	1	Total	C	O	0	0
			6	3	3		
4	E	1	Total	C	O	0	0
			6	3	3		
4	E	1	Total	C	O	0	0
			6	3	3		
4	B	1	Total	C	O	0	0
			6	3	3		
4	A	1	Total	C	O	0	0
			6	3	3		
4	A	1	Total	C	O	0	0
			6	3	3		
4	A	1	Total	C	O	0	0
			6	3	3		
4	A	1	Total	C	O	0	0
			6	3	3		
4	D	1	Total	C	O	0	0
			6	3	3		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	D	1	Total	C	O	0	0
			6	3	3		
4	D	1	Total	C	O	0	0
			6	3	3		

- Molecule 5 is SULFATE ION (three-letter code: SO4) (formula: O₄S).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	F	1	Total	O	S	0	0
			5	4	1		
5	F	1	Total	O	S	0	0
			5	4	1		
5	C	1	Total	O	S	0	0
			5	4	1		
5	E	1	Total	O	S	0	0
			5	4	1		
5	B	1	Total	O	S	0	0
			5	4	1		
5	A	1	Total	O	S	0	0
			5	4	1		
5	D	1	Total	O	S	0	0
			5	4	1		
5	D	1	Total	O	S	0	0
			5	4	1		

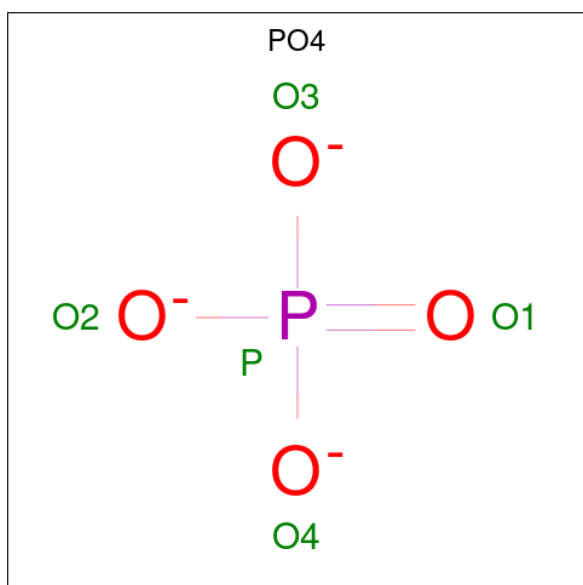
- Molecule 6 is MANGANESE (II) ION (three-letter code: MN) (formula: Mn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	F	2	Total	Mn	0	0
			2	2		
6	C	2	Total	Mn	0	0
			2	2		

- Molecule 7 is SODIUM ION (three-letter code: NA) (formula: Na).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	F	1	Total	Na	0	0
			1	1		
7	E	1	Total	Na	0	0
			1	1		

- Molecule 8 is PHOSPHATE ION (three-letter code: PO4) (formula: O₄P).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
8	F	1	Total	O	P	0	0
			5	4	1		

- Molecule 9 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

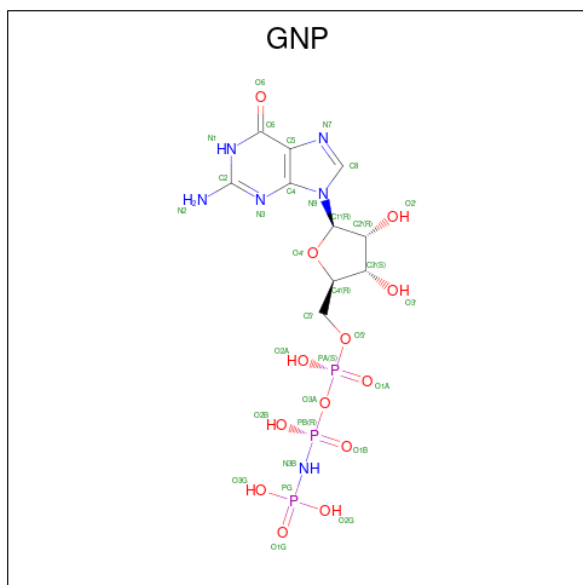
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
9	F	1	Total	Cl	0	0
			1	1		
9	A	3	Total	Cl	0	0
			3	3		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
9	D	1	Total	Cl	0	0
			1	1		

- Molecule 10 is PHOSPHOAMINOPHOSPHONIC ACID-GUANYLATE ESTER (three-letter code: GNP) (formula: $C_{10}H_{17}N_6O_{13}P_3$) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
10	E	1	Total	C	N	O	P	0	0
			32	10	6	13	3		
10	B	1	Total	C	N	O	P	0	0
			32	10	6	13	3		

- Molecule 11 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
11	E	1	Total	Mg	0	0
			1	1		
11	B	1	Total	Mg	0	0
			1	1		

- Molecule 12 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
12	F	170	Total	O	0	0
			170	170		

Continued on next page...


Continued from previous page...

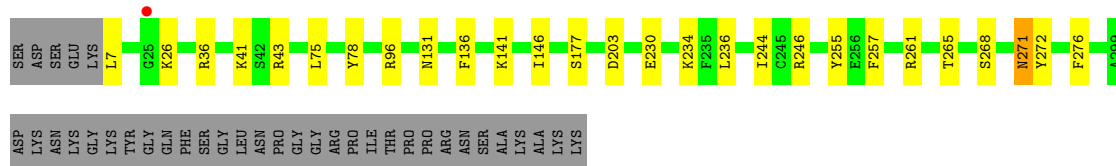
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
12	C	116	Total 116	O 116	0	0
12	E	80	Total 80	O 80	0	0
12	B	176	Total 176	O 176	0	0
12	A	333	Total 333	O 333	0	0
12	D	171	Total 171	O 171	0	0

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.

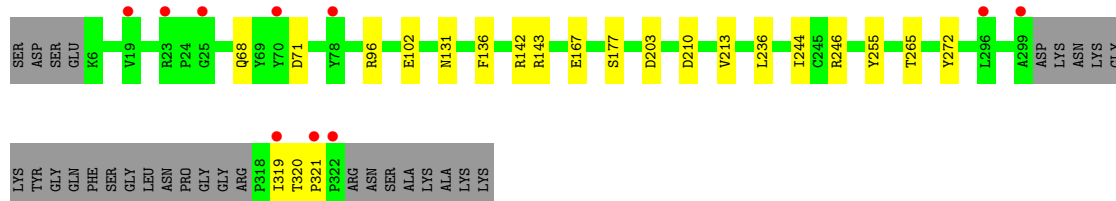
- Molecule 1: Serine/threonine-protein phosphatase PP1-alpha catalytic subunit

Chain F: 

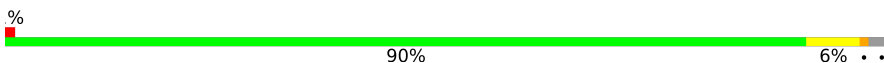


- Molecule 1: Serine/threonine-protein phosphatase PP1-alpha catalytic subunit

Chain C: 

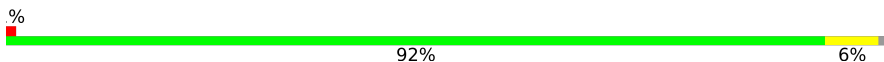


- Molecule 2: Ras-related protein M-Ras

Chain E: 

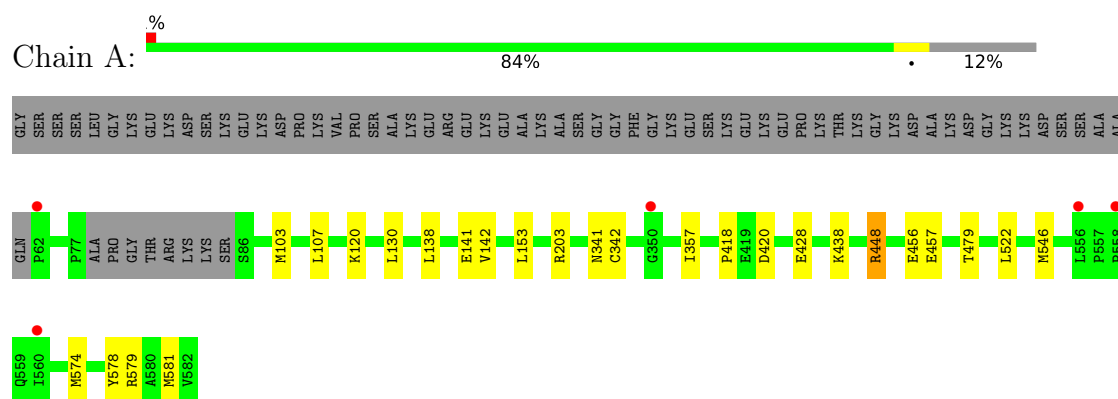


- Molecule 2: Ras-related protein M-Ras

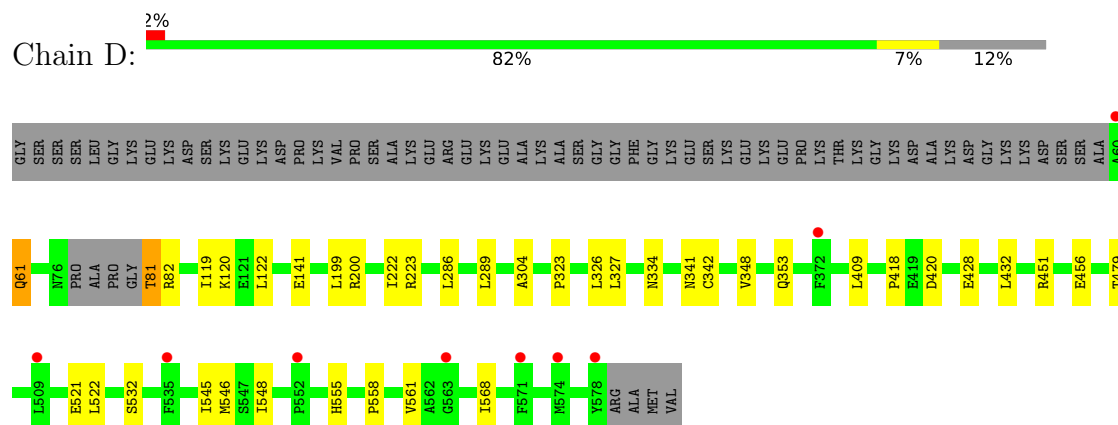
Chain B: 



- Molecule 3: Leucine-rich repeat protein SHOC-2



- Molecule 3: Leucine-rich repeat protein SHOC-2



4 Data and refinement statistics

Property	Value	Source
Space group	P 41 21 2	Depositor
Cell constants a, b, c, α , β , γ	129.94Å 129.94Å 326.80Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	101.70 – 2.17 120.74 – 2.17	Depositor EDS
% Data completeness (in resolution range)	99.9 (101.70-2.17) 99.9 (120.74-2.17)	Depositor EDS
R_{merge}	0.14	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.18 (at 2.16Å)	Xtriage
Refinement program	PHENIX 1.19.2_4158	Depositor
R, R_{free}	0.199 , 0.226 0.193 , 0.220	Depositor DCC
R_{free} test set	7492 reflections (5.05%)	wwPDB-VP
Wilson B-factor (Å ²)	43.5	Xtriage
Anisotropy	0.028	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.32 , 39.2	EDS
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.96	EDS
Total number of atoms	16911	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 2.09% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

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

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: SO4, GOL, PO4, CL, NA, MN, GNP, MG

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	C	0.25	0/2441	0.48	0/3303
1	F	0.25	0/2428	0.48	0/3280
2	B	0.24	0/1443	0.48	0/1952
2	E	0.24	0/1435	0.49	0/1941
3	A	0.23	0/4137	0.48	0/5606
3	D	0.24	0/4097	0.48	0/5556
All	All	0.24	0/15981	0.48	0/21638

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry-related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	C	2383	0	2332	12	0
1	F	2368	0	2339	17	0
2	B	1413	0	1406	5	0
2	E	1403	0	1399	5	0
3	A	4063	0	4229	17	0
3	D	4035	0	4180	21	0
4	A	24	0	32	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	B	6	0	8	0	0
4	C	6	0	8	0	0
4	D	18	0	24	0	0
4	E	12	0	16	0	0
4	F	12	0	16	0	0
5	A	5	0	0	0	0
5	B	5	0	0	0	0
5	C	5	0	0	0	0
5	D	10	0	0	0	0
5	E	5	0	0	0	0
5	F	10	0	0	1	0
6	C	2	0	0	0	0
6	F	2	0	0	0	0
7	E	1	0	0	0	0
7	F	1	0	0	0	0
8	F	5	0	0	0	0
9	A	3	0	0	1	0
9	D	1	0	0	1	0
9	F	1	0	0	0	0
10	B	32	0	13	2	0
10	E	32	0	13	1	0
11	B	1	0	0	0	0
11	E	1	0	0	0	0
12	A	333	0	0	2	0
12	B	176	0	0	0	0
12	C	116	0	0	1	0
12	D	171	0	0	1	0
12	E	80	0	0	1	0
12	F	170	0	0	5	0
All	All	16911	0	16015	76	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 (76) 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)
3:D:327:LEU:HB2	3:D:353:GLN:HB2	1.68	0.76
2:E:8:SER:HB2	2:E:11:LEU:HD21	1.70	0.73
3:A:456:GLU:HG2	3:A:457:GLU:HG3	1.69	0.71
1:F:36[A]:ARG:NH1	12:F:501:HOH:O	2.24	0.69

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
12:F:526:HOH:O	9:D:606:CL:CL	2.48	0.69
2:E:23:GLY:H	10:E:1201:GNP:HNB3	1.39	0.67
1:C:142:ARG:NH1	12:C:501:HOH:O	2.25	0.67
3:D:428:GLU:OE1	3:D:451:ARG:NH1	2.29	0.65
2:B:23:GLY:H	10:B:1201:GNP:HNB3	1.48	0.60
3:D:200:ARG:HG2	3:D:223:ARG:HG2	1.84	0.60
1:F:271:ASN:ND2	1:F:276:PHE:O	2.33	0.60
3:A:579:ARG:NH2	12:A:704:HOH:O	2.35	0.60
1:F:7:LEU:N	12:F:504:HOH:O	2.34	0.59
3:D:456:GLU:HG3	3:D:479:THR:HB	1.85	0.58
2:E:177:GLN:NE2	12:E:1302:HOH:O	2.36	0.57
3:D:546:MET:HE3	3:D:548:ILE:HD11	1.87	0.56
3:D:418:PRO:HB2	3:D:420:ASP:OD1	2.06	0.55
3:A:456:GLU:HG3	3:A:479:THR:HB	1.87	0.54
1:C:236:LEU:HD21	1:C:244:ILE:HG13	1.89	0.54
3:D:61:GLN:NE2	12:D:707:HOH:O	2.37	0.54
1:C:68:GLN:NE2	1:C:71:ASP:OD2	2.39	0.54
1:C:210:ASP:HB3	1:C:213:VAL:HG12	1.90	0.54
3:D:304:ALA:HB2	3:D:326:LEU:HD12	1.90	0.53
1:F:43:ARG:NH1	12:F:505:HOH:O	2.39	0.51
3:A:120:LYS:HB3	3:A:141:GLU:HB3	1.92	0.50
3:D:199:LEU:HB2	3:D:222:ILE:HG22	1.93	0.50
3:D:323:PRO:HD2	3:D:326:LEU:HD22	1.94	0.49
3:D:120:LYS:HB3	3:D:141:GLU:HB3	1.92	0.49
1:F:236:LEU:HD21	1:F:244:ILE:HG13	1.94	0.49
3:D:521:GLU:HG2	3:D:545:ILE:HD12	1.94	0.49
2:B:128:VAL:HG11	2:B:154:GLU:HB3	1.95	0.49
3:D:561:VAL:HA	3:D:568:ILE:HD11	1.94	0.49
3:A:574:MET:O	3:A:579:ARG:NH1	2.46	0.49
1:F:36[A]:ARG:NH2	12:F:507:HOH:O	2.43	0.48
1:C:319:ILE:HG22	1:C:320:THR:HG23	1.95	0.48
1:C:167:GLU:OE2	3:A:203:ARG:NH2	2.40	0.48
3:D:119:ILE:HD12	3:D:122:LEU:HD12	1.96	0.48
3:D:558:PRO:HA	3:D:561:VAL:HG22	1.95	0.48
1:F:230:GLU:HG3	1:F:234:LYS:HE3	1.95	0.48
2:E:97:LYS:NZ	2:E:101:GLU:OE2	2.31	0.48
2:B:114:LYS:HE2	2:B:119:PHE:CZ	2.49	0.47
3:A:438:LYS:NZ	12:A:719:HOH:O	2.48	0.47
1:C:320:THR:HB	1:C:321:PRO:HD2	1.97	0.46
1:F:131:ASN:HB2	1:F:136:PHE:HB3	1.96	0.46
1:F:177:SER:HB2	1:F:203:ASP:HB2	1.98	0.46

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:D:522:LEU:O	3:D:546:MET:HA	2.15	0.46
1:C:131:ASN:HB2	1:C:136:PHE:HB3	1.98	0.46
3:D:409:LEU:HB2	3:D:432:LEU:HD23	1.98	0.46
3:A:522:LEU:O	3:A:546:MET:HA	2.16	0.45
2:B:132:HIS:NE2	3:A:428:GLU:OE2	2.39	0.45
3:D:81:THR:HG23	3:D:82:ARG:H	1.81	0.45
1:F:75:LEU:HD11	1:F:268:SER:HB3	1.98	0.45
1:F:26:LYS:HB2	1:F:26:LYS:HE3	1.79	0.44
3:A:448:ARG:HA	3:A:448:ARG:HD3	1.82	0.44
1:F:141:LYS:HD2	1:F:146:ILE:HD11	1.99	0.44
3:A:138:LEU:HG	3:A:142:VAL:HG21	1.99	0.44
1:F:255:TYR:HA	1:F:265:THR:O	2.17	0.44
2:B:127:LYS:HG2	10:B:1201:GNP:C6	2.48	0.44
3:D:341:ASN:HB3	3:D:342:CYS:H	1.72	0.43
1:F:261:ARG:NH2	5:F:403:SO4:O2	2.40	0.43
3:A:130:LEU:HB2	3:A:153:LEU:HD23	1.99	0.43
3:A:418:PRO:HB2	3:A:420:ASP:OD1	2.18	0.43
2:E:114:LYS:HE2	2:E:119:PHE:CZ	2.54	0.43
3:A:107:LEU:HB2	3:A:130:LEU:HD23	2.01	0.42
3:D:286:LEU:HD21	3:D:289:LEU:HD13	2.00	0.42
1:F:257:PHE:CG	1:F:261:ARG:HG2	2.54	0.42
1:F:41[A]:LYS:HA	1:F:41[A]:LYS:HD2	1.83	0.42
1:C:102:GLU:OE2	1:C:143:ARG:NH1	2.51	0.42
1:C:96:ARG:HG3	1:C:272:TYR:OH	2.19	0.42
1:C:177:SER:HB2	1:C:203:ASP:HB2	2.02	0.41
1:F:96:ARG:HG3	1:F:272:TYR:OH	2.20	0.41
3:A:357:ILE:HG23	9:A:607:CL:CL	2.58	0.41
3:A:341:ASN:HB3	3:A:342[B]:CYS:H	1.79	0.41
1:C:255:TYR:HA	1:C:265:THR:O	2.22	0.40
3:A:578:TYR:HB3	3:A:581:MET:HG3	2.03	0.40
3:D:532:SER:HB2	3:D:555:HIS:HB2	2.03	0.40

There are no symmetry-related clashes.

5.3 Torsion angles ⓘ

5.3.1 Protein backbone ⓘ

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	C	296/329 (90%)	283 (96%)	13 (4%)	0	100	100
1	F	294/329 (89%)	281 (96%)	13 (4%)	0	100	100
2	B	173/179 (97%)	172 (99%)	1 (1%)	0	100	100
2	E	172/179 (96%)	171 (99%)	1 (1%)	0	100	100
3	A	515/582 (88%)	481 (93%)	34 (7%)	0	100	100
3	D	512/582 (88%)	481 (94%)	31 (6%)	0	100	100
All	All	1962/2180 (90%)	1869 (95%)	93 (5%)	0	100	100

There are no Ramachandran outliers to report.

5.3.2 Protein sidechains ⓘ

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	C	260/287 (91%)	259 (100%)	1 (0%)	91	95
1	F	259/287 (90%)	256 (99%)	3 (1%)	71	81
2	B	157/159 (99%)	154 (98%)	3 (2%)	57	68
2	E	155/159 (98%)	150 (97%)	5 (3%)	39	47
3	A	481/530 (91%)	479 (100%)	2 (0%)	91	95
3	D	473/530 (89%)	469 (99%)	4 (1%)	81	89
All	All	1785/1952 (91%)	1767 (99%)	18 (1%)	76	85

All (18) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	F	78	TYR
1	F	246	ARG
1	F	271	ASN
1	C	246	ARG

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	E	6	VAL
2	E	121	MET
2	E	159	ASP
2	E	165	ASP
2	E	177	GLN
2	B	52	LYS
2	B	159	ASP
2	B	165	ASP
3	A	103	MET
3	A	448	ARG
3	D	61	GLN
3	D	81	THR
3	D	334	ASN
3	D	348	VAL

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
2	E	108	GLN

5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

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 37 ligands modelled in this entry, 13 are monoatomic - leaving 24 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	SO4	D	604	-	4,4,4	0.14	0	6,6,6	0.05	0
10	GNP	E	1201	11	28,34,34	2.36	10 (35%)	30,54,54	1.81	4 (13%)
4	GOL	A	603	-	5,5,5	0.78	0	5,5,5	1.17	1 (20%)
5	SO4	E	1204	-	4,4,4	0.14	0	6,6,6	0.05	0
4	GOL	D	602	-	5,5,5	0.92	0	5,5,5	0.99	0
5	SO4	B	1203	-	4,4,4	0.14	0	6,6,6	0.08	0
4	GOL	A	601	-	5,5,5	0.91	0	5,5,5	0.97	0
10	GNP	B	1201	11	28,34,34	2.33	10 (35%)	30,54,54	1.78	3 (10%)
4	GOL	E	1203	-	5,5,5	0.86	0	5,5,5	1.02	0
5	SO4	D	605	-	4,4,4	0.14	0	6,6,6	0.07	0
8	PO4	F	408	-	4,4,4	0.92	0	6,6,6	0.42	0
4	GOL	D	601	-	5,5,5	0.92	0	5,5,5	0.97	0
5	SO4	C	402	-	4,4,4	0.15	0	6,6,6	0.05	0
4	GOL	A	604	-	5,5,5	0.91	0	5,5,5	1.01	0
4	GOL	E	1202	-	5,5,5	0.88	0	5,5,5	1.00	0
4	GOL	B	1202	-	5,5,5	0.88	0	5,5,5	1.01	0
4	GOL	F	402	-	5,5,5	0.98	0	5,5,5	0.88	0
5	SO4	F	403	-	4,4,4	0.14	0	6,6,6	0.06	0
4	GOL	D	603	-	5,5,5	0.85	0	5,5,5	1.05	0
4	GOL	F	401	-	5,5,5	0.90	0	5,5,5	1.00	0
5	SO4	A	605	-	4,4,4	0.14	0	6,6,6	0.05	0
5	SO4	F	404	-	4,4,4	0.13	0	6,6,6	0.07	0
4	GOL	C	401	-	5,5,5	0.77	0	5,5,5	1.05	0
4	GOL	A	602	-	5,5,5	0.88	0	5,5,5	1.01	0

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
4	GOL	B	1202	-	-	0/4/4/4	-
4	GOL	A	601	-	-	1/4/4/4	-
4	GOL	F	402	-	-	2/4/4/4	-
4	GOL	E	1203	-	-	4/4/4/4	-
4	GOL	D	601	-	-	0/4/4/4	-
4	GOL	A	604	-	-	2/4/4/4	-

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
4	GOL	A	603	-	-	2/4/4/4	-
10	GNP	E	1201	11	-	3/17/38/38	0/3/3/3
10	GNP	B	1201	11	-	3/17/38/38	0/3/3/3
4	GOL	D	603	-	-	2/4/4/4	-
4	GOL	E	1202	-	-	2/4/4/4	-
4	GOL	F	401	-	-	0/4/4/4	-
4	GOL	D	602	-	-	1/4/4/4	-
4	GOL	C	401	-	-	2/4/4/4	-
4	GOL	A	602	-	-	0/4/4/4	-

All (20) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
10	E	1201	GNP	C4-N9	-7.39	1.37	1.47
10	B	1201	GNP	C4-N9	-7.35	1.37	1.47
10	B	1201	GNP	C5-C6	-4.58	1.44	1.52
10	E	1201	GNP	C5-C6	-4.55	1.45	1.52
10	E	1201	GNP	PB-O3A	4.51	1.64	1.59
10	B	1201	GNP	PB-O3A	4.17	1.64	1.59
10	E	1201	GNP	PB-O1B	3.22	1.51	1.46
10	E	1201	GNP	C6-N1	3.15	1.38	1.33
10	B	1201	GNP	PB-O1B	3.15	1.51	1.46
10	B	1201	GNP	C6-N1	3.14	1.38	1.33
10	E	1201	GNP	PG-N3B	2.96	1.71	1.63
10	B	1201	GNP	PG-N3B	2.92	1.71	1.63
10	E	1201	GNP	PG-O1G	2.79	1.50	1.46
10	B	1201	GNP	PG-O1G	2.75	1.50	1.46
10	B	1201	GNP	C5-C4	-2.41	1.38	1.53
10	E	1201	GNP	C5-C4	-2.40	1.38	1.53
10	B	1201	GNP	C8-N9	-2.38	1.37	1.45
10	E	1201	GNP	C8-N9	-2.35	1.37	1.45
10	E	1201	GNP	PB-O2B	-2.17	1.50	1.56
10	B	1201	GNP	PB-O2B	-2.15	1.51	1.56

All (8) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
10	E	1201	GNP	C4-C5-N7	6.12	110.57	102.46
10	B	1201	GNP	C4-C5-N7	6.10	110.55	102.46
10	E	1201	GNP	C5-C6-N1	-5.10	111.90	118.19
10	B	1201	GNP	C5-C6-N1	-5.08	111.92	118.19

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
10	E	1201	GNP	O6-C6-C5	3.73	127.46	119.86
10	B	1201	GNP	O6-C6-C5	3.68	127.38	119.86
4	A	603	GOL	C3-C2-C1	-2.15	103.35	111.70
10	E	1201	GNP	PB-O3A-PA	-2.07	125.33	132.62

There are no chirality outliers.

All (24) torsion outliers are listed below:

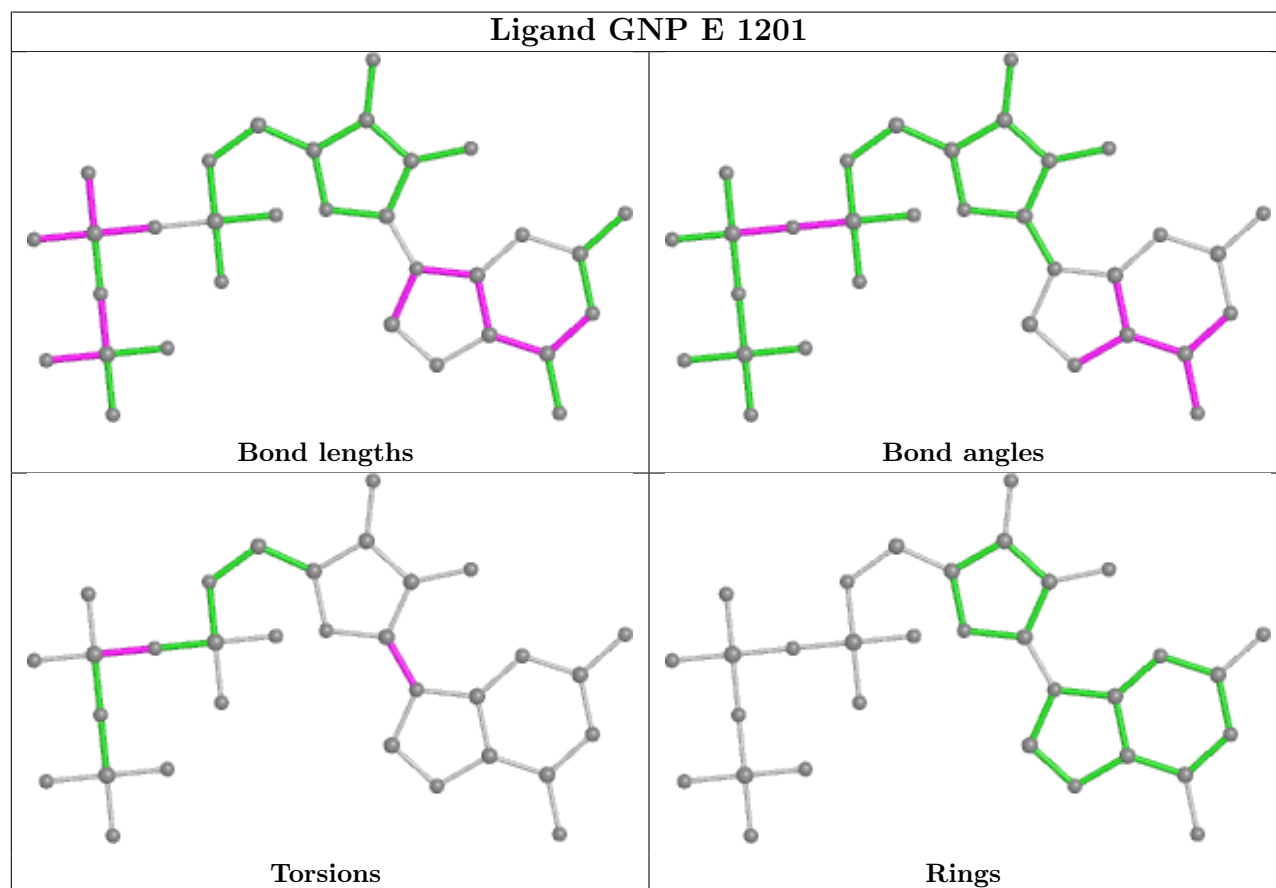
Mol	Chain	Res	Type	Atoms
4	F	402	GOL	O1-C1-C2-C3
4	C	401	GOL	O1-C1-C2-O2
4	C	401	GOL	O1-C1-C2-C3
4	E	1202	GOL	O1-C1-C2-C3
4	E	1203	GOL	C1-C2-C3-O3
4	D	603	GOL	C1-C2-C3-O3
10	E	1201	GNP	PA-O3A-PB-O1B
10	E	1201	GNP	PA-O3A-PB-O2B
10	E	1201	GNP	C2'-C1'-N9-C4
10	B	1201	GNP	PA-O3A-PB-O1B
10	B	1201	GNP	PA-O3A-PB-O2B
10	B	1201	GNP	C2'-C1'-N9-C4
4	F	402	GOL	O1-C1-C2-O2
4	E	1203	GOL	O1-C1-C2-C3
4	A	603	GOL	C1-C2-C3-O3
4	A	604	GOL	O1-C1-C2-C3
4	E	1202	GOL	O1-C1-C2-O2
4	E	1203	GOL	O2-C2-C3-O3
4	A	604	GOL	O1-C1-C2-O2
4	D	603	GOL	O2-C2-C3-O3
4	A	603	GOL	O2-C2-C3-O3
4	E	1203	GOL	O1-C1-C2-O2
4	A	601	GOL	O1-C1-C2-C3
4	D	602	GOL	O1-C1-C2-C3

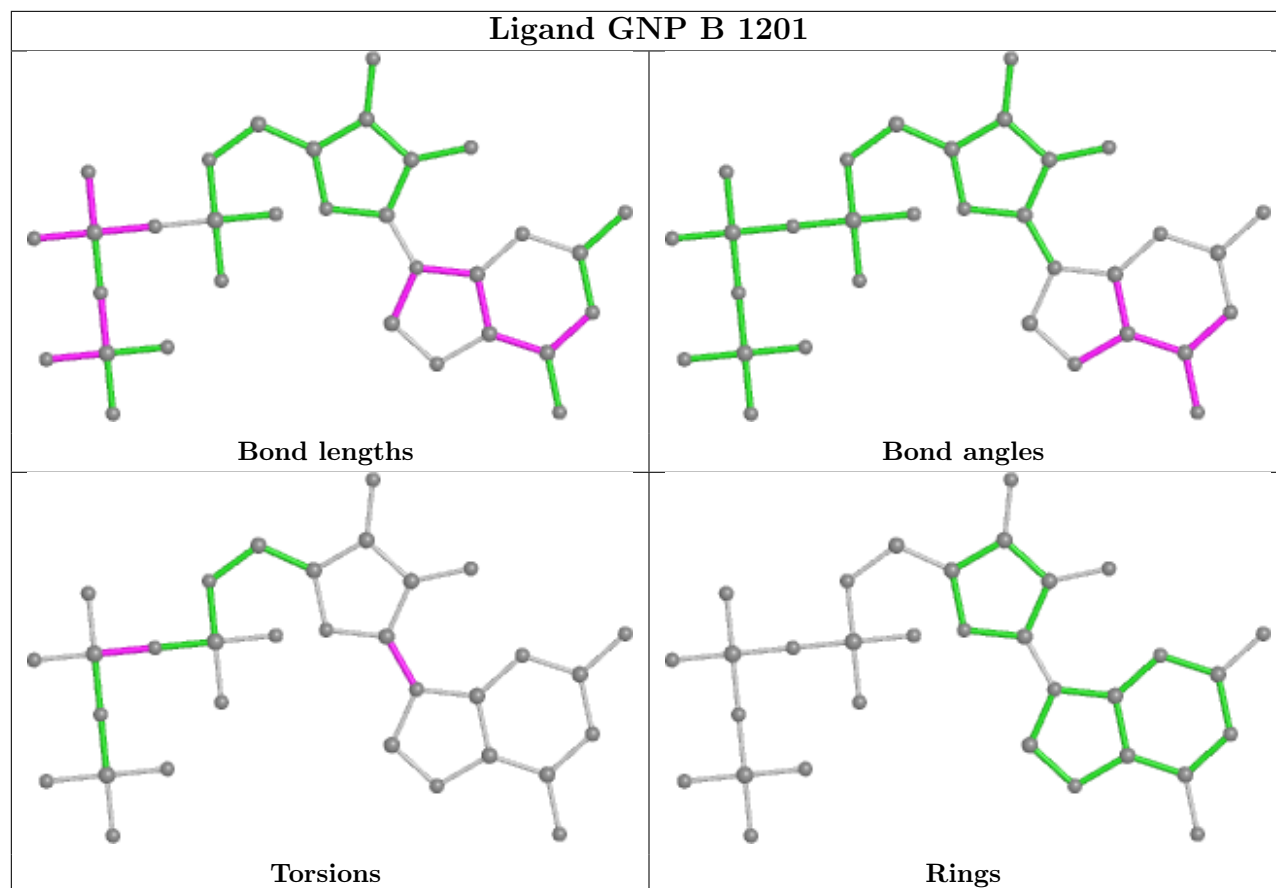
There are no ring outliers.

3 monomers are involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
10	E	1201	GNP	1	0
10	B	1201	GNP	2	0
5	F	403	SO4	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 than 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.





5.7 Other polymers [i](#)

There are no such residues in this entry.

5.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

6 Fit of model and data ⓘ

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, 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	C	299/329 (90%)	0.15	10 (3%)	46 47	29, 50, 90, 114	0
1	F	293/329 (89%)	0.01	1 (0%)	94 94	30, 41, 71, 107	0
2	B	175/179 (97%)	-0.02	2 (1%)	80 80	28, 38, 64, 109	0
2	E	173/179 (96%)	-0.03	2 (1%)	79 79	37, 53, 80, 110	0
3	A	513/582 (88%)	-0.02	5 (0%)	82 82	28, 45, 73, 114	0
3	D	515/582 (88%)	0.07	9 (1%)	70 70	35, 59, 91, 113	0
All	All	1968/2180 (90%)	0.03	29 (1%)	73 74	28, 50, 83, 114	0

All (29) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	4	SER	4.4
1	C	321	PRO	4.1
3	D	552	PRO	3.8
1	F	25	GLY	3.5
1	C	23	ARG	3.3
1	C	19	VAL	3.0
1	C	70	TYR	3.0
3	D	571	PHE	2.9
3	D	509	LEU	2.7
3	D	60	ALA	2.7
2	E	9	ASP	2.6
1	C	25	GLY	2.6
1	C	78	TYR	2.5
1	C	322	PRO	2.5
3	D	578	TYR	2.5
2	E	6	VAL	2.4
3	A	556	LEU	2.4
3	D	574	MET	2.3
1	C	299	ALA	2.3

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
3	A	560	ILE	2.3
1	C	296	LEU	2.2
3	A	350	GLY	2.2
2	B	9	ASP	2.2
3	A	62	PRO	2.2
3	D	372	PHE	2.2
3	D	535	PHE	2.2
3	D	563	GLY	2.1
1	C	319	ILE	2.1
3	A	558	PRO	2.0

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(\AA^2)	Q<0.9
4	GOL	D	603	6/6	0.72	0.17	73,90,94,96	0
5	SO4	D	605	5/5	0.72	0.31	89,110,125,161	0
5	SO4	F	404	5/5	0.73	0.27	81,81,119,155	0
5	SO4	F	403	5/5	0.76	0.20	70,94,109,129	0
4	GOL	E	1203	6/6	0.78	0.20	65,73,94,100	0
4	GOL	D	602	6/6	0.79	0.14	80,84,92,103	0
4	GOL	E	1202	6/6	0.81	0.12	75,88,99,100	0
4	GOL	C	401	6/6	0.84	0.30	60,70,77,84	0
9	CL	A	608	1/1	0.84	0.08	95,95,95,95	0
4	GOL	A	604	6/6	0.85	0.14	89,92,94,97	0
9	CL	F	409	1/1	0.86	0.14	88,88,88,88	0
4	GOL	B	1202	6/6	0.86	0.14	66,73,85,86	0
5	SO4	D	604	5/5	0.88	0.14	82,95,107,123	0

Continued on next page...

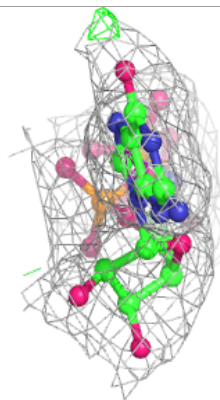
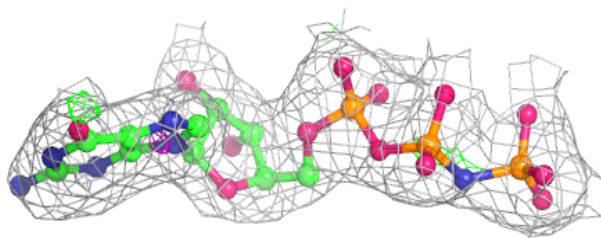
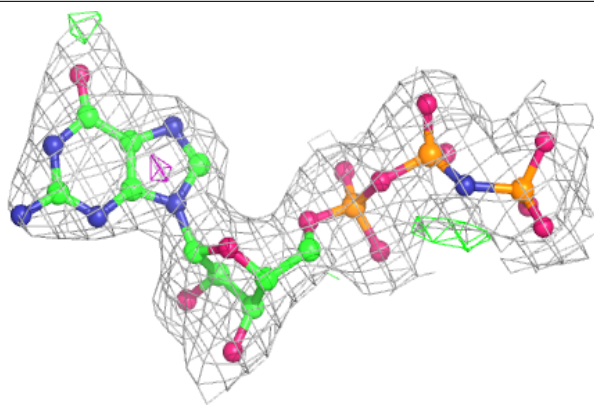
Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(\AA^2)	Q<0.9
8	PO4	F	408	5/5	0.88	0.16	99,105,125,147	0
5	SO4	A	605	5/5	0.89	0.34	102,103,115,150	0
4	GOL	F	401	6/6	0.89	0.16	50,62,66,73	0
5	SO4	C	402	5/5	0.90	0.29	85,94,127,152	0
4	GOL	A	603	6/6	0.91	0.21	54,58,78,82	0
4	GOL	A	602	6/6	0.91	0.16	49,60,63,74	0
7	NA	E	1206	1/1	0.91	0.12	55,55,55,55	0
7	NA	F	407	1/1	0.93	0.09	73,73,73,73	0
9	CL	D	606	1/1	0.93	0.10	82,82,82,82	0
4	GOL	F	402	6/6	0.94	0.19	45,54,62,66	0
9	CL	A	607	1/1	0.95	0.34	61,61,61,61	0
9	CL	A	606	1/1	0.96	0.11	70,70,70,70	0
4	GOL	A	601	6/6	0.96	0.10	32,36,44,47	0
6	MN	F	406	1/1	0.97	0.05	61,61,61,61	0
11	MG	E	1205	1/1	0.97	0.12	39,39,39,39	0
4	GOL	D	601	6/6	0.98	0.10	42,48,56,56	0
5	SO4	E	1204	5/5	0.98	0.09	59,68,69,85	0
10	GNP	E	1201	32/32	0.98	0.13	39,49,63,67	0
6	MN	C	403	1/1	0.98	0.10	52,52,52,52	0
6	MN	F	405	1/1	0.99	0.11	45,45,45,45	0
6	MN	C	404	1/1	0.99	0.06	58,58,58,58	0
10	GNP	B	1201	32/32	0.99	0.13	29,33,36,37	0
5	SO4	B	1203	5/5	0.99	0.11	51,52,63,69	0
11	MG	B	1204	1/1	0.99	0.14	27,27,27,27	0

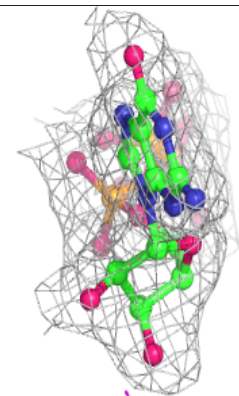
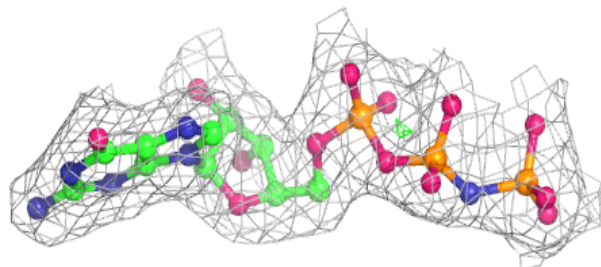
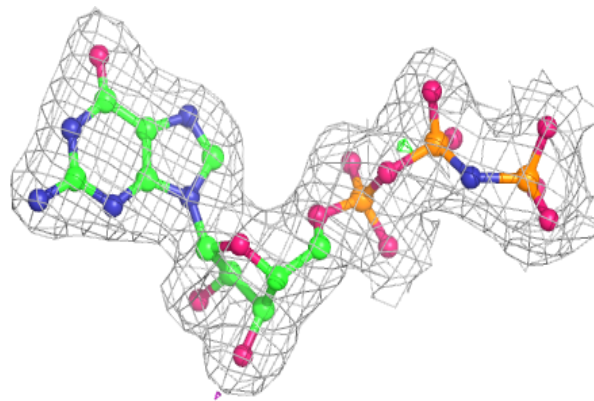
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation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

Electron density around GNP E 1201:

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

**Electron density around GNP B 1201:**

$2mF_o-DF_c$ (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.