Full wwPDB X-ray Structure Validation Report ⓘ

Oct 14, 2023 – 03:17 PM EDT

PDB ID	:	8EG3
Title	:	Structure of human placental steroid (estrone/DHEA) sulfatase at 2.0 angstrom resolution
Authors	:	Ghosh, D.
Deposited on	:	2022-09-10
Resolution	:	2.04 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 types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

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

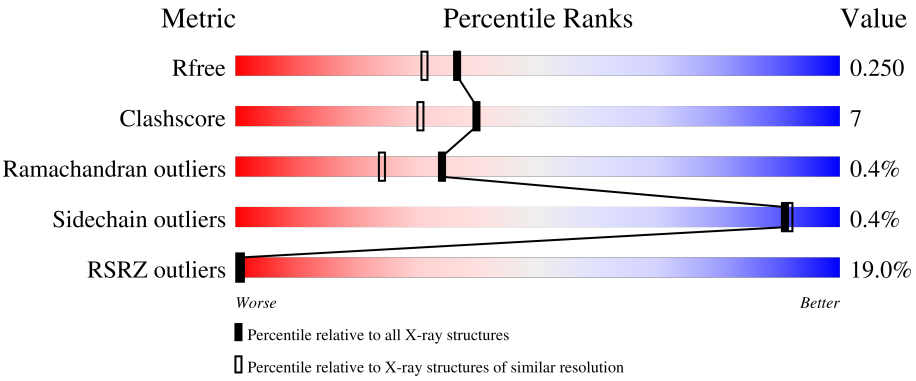
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

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 2.04 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	1692 (2.04-2.04)
Clashscore	141614	1773 (2.04-2.04)
Ramachandran outliers	138981	1752 (2.04-2.04)
Sidechain outliers	138945	1752 (2.04-2.04)
RSRZ outliers	127900	1672 (2.04-2.04)

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	583	18% 80% 15% 5%

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
3	NAG	A	602	-	-	-	X

2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 4575 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 Steryl-sulfatase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	554	4368	2803	750	789	26	0	0	0

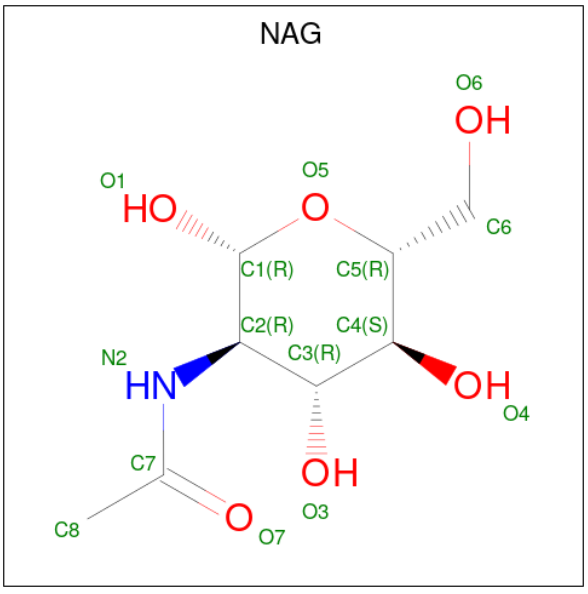
There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	75	ALS	CYS	conflict	UNP P08842

- Molecule 2 is CALCIUM ION (three-letter code: CA) (formula: Ca).

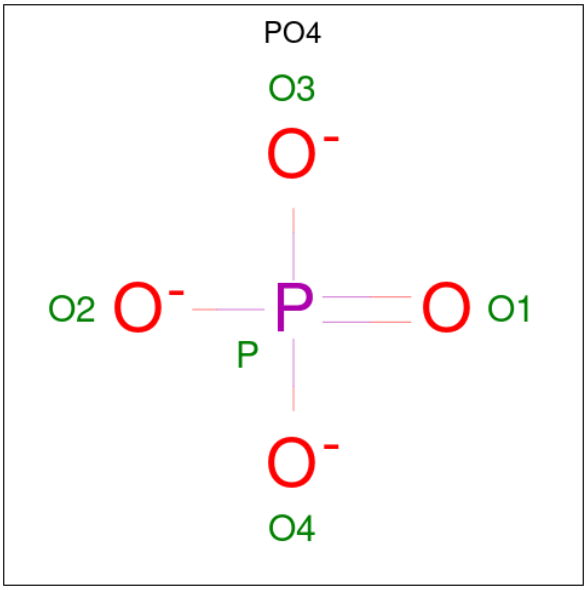
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Ca		
2	A	1	1	1	0	0

- Molecule 3 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula: $C_8H_{15}NO_6$).



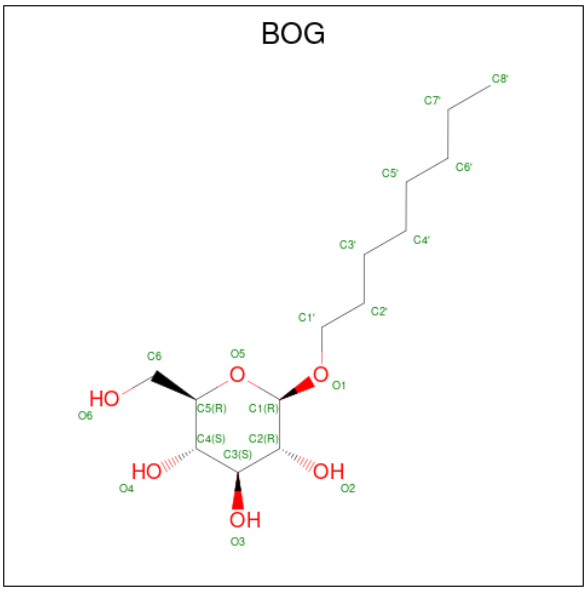
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total C N O 14 8 1 5	0	0
3	A	1	Total C N O 14 8 1 5	0	0

- Molecule 4 is PHOSPHATE ION (three-letter code: PO4) (formula: O_4P) (labeled as "Ligand of Interest" by depositor).



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

- Molecule 5 is octyl beta-D-glucopyranoside (three-letter code: BOG) (formula: $C_{14}H_{28}O_6$) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	1	Total C O 20 14 6	0	0
5	A	1	Total C O 20 14 6	0	0

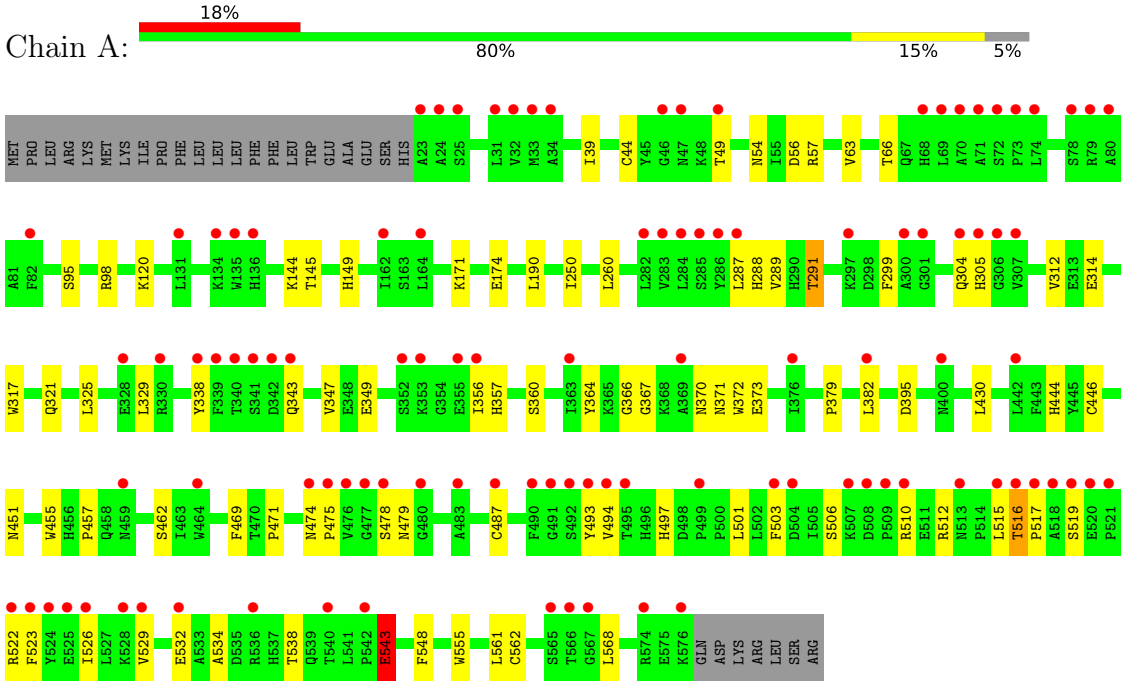
- Molecule 6 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	133	Total O 133 133	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: Steryl-sulfatase



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 3 2 1	Depositor
Cell constants a, b, c, α, β, γ	117.19Å 117.19Å 102.68Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	32.13 – 2.04 32.13 – 2.04	Depositor EDS
% Data completeness (in resolution range)	98.8 (32.13-2.04) 98.8 (32.13-2.04)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.13 (at 2.05Å)	Xtriage
Refinement program	PHENIX 1.18.2_3874	Depositor
R, R_{free}	0.220 , 0.251 0.220 , 0.250	Depositor DCC
R_{free} test set	2535 reflections (4.95%)	wwPDB-VP
Wilson B-factor (Å^2)	36.4	Xtriage
Anisotropy	0.631	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.36 , 71.2	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	0.027 for -h,-k,l	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	4575	wwPDB-VP
Average B, all atoms (Å^2)	67.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.61% 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: NAG, PO4, BOG, ALS, CA

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	0.45	1/4482 (0.0%)	0.67	3/6092 (0.0%)

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

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	2

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	543	GLU	CD-OE2	7.20	1.33	1.25

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	543	GLU	CA-CB-CG	12.63	141.19	113.40
1	A	543	GLU	CB-CA-C	6.75	123.89	110.40
1	A	190	LEU	CA-CB-CG	5.04	126.90	115.30

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	291	THR	Peptide
1	A	543	GLU	Sidechain

5.2 Too-close contacts (i)

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

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	4368	0	4260	63	0
2	A	1	0	0	0	0
3	A	28	0	26	1	0
4	A	5	0	0	0	0
5	A	40	0	56	8	0
6	A	133	0	0	2	0
All	All	4575	0	4342	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 7.

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:A:510:ARG:HD3	1:A:512:ARG:HD2	1.61	0.83
1:A:370:ASN:HD22	1:A:372:TRP:HD1	1.25	0.82
1:A:370:ASN:ND2	1:A:372:TRP:HD1	1.85	0.74
1:A:366:GLY:HA3	1:A:370:ASN:HD21	1.58	0.69
1:A:555:TRP:HH2	5:A:604:BOG:HO3	1.40	0.69
1:A:555:TRP:CD1	5:A:605:BOG:H61	2.34	0.62
1:A:66:THR:HB	1:A:395:ASP:HA	1.83	0.61
1:A:370:ASN:ND2	1:A:372:TRP:CD1	2.63	0.61
1:A:364:TYR:HB3	1:A:373:GLU:HB3	1.83	0.61
1:A:291:THR:HG21	1:A:349:GLU:HG2	1.83	0.61
1:A:349:GLU:HG3	1:A:357:HIS:HB2	1.82	0.60
1:A:98:ARG:NH2	1:A:493:TYR:OH	2.35	0.59
1:A:555:TRP:CG	5:A:605:BOG:H61	2.39	0.57
1:A:260:LEU:HD21	1:A:287:LEU:HD22	1.88	0.56
1:A:382:LEU:HG	1:A:430:LEU:HD21	1.89	0.54
1:A:299:PHE:CD1	1:A:314:GLU:HG3	2.42	0.54
1:A:503:PHE:HA	1:A:515:LEU:HD23	1.88	0.54
1:A:95:SER:HB3	1:A:548:PHE:CB	2.38	0.54
1:A:291:THR:HG21	1:A:349:GLU:CD	2.28	0.54
1:A:519:SER:HB3	1:A:522:ARG:HE	1.72	0.54
1:A:291:THR:HG21	1:A:349:GLU:CG	2.37	0.53

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:304:GLN:NE2	1:A:305:HIS:HB2	2.24	0.53
1:A:304:GLN:CD	1:A:305:HIS:HB2	2.30	0.52
1:A:54:ASN:OD1	1:A:57:ARG:NH2	2.42	0.52
1:A:95:SER:HB3	1:A:548:PHE:HB3	1.91	0.51
1:A:149:HIS:HD2	6:A:751:HOH:O	1.93	0.50
1:A:516:THR:H	1:A:517:PRO:CD	2.25	0.50
1:A:49:THR:HB	1:A:304:GLN:HG3	1.94	0.49
1:A:562:CYS:O	5:A:604:BOG:H8'3	2.12	0.48
1:A:144:LYS:HE2	5:A:604:BOG:H5'1	1.96	0.48
1:A:360:SER:HB2	3:A:601:NAG:H4	1.96	0.48
1:A:474:ASN:HB2	1:A:475:PRO:HD3	1.95	0.48
1:A:250:ILE:HD11	1:A:561:LEU:HB3	1.96	0.47
1:A:367:GLY:O	1:A:370:ASN:OD1	2.32	0.47
1:A:44:CYS:HB3	1:A:56:ASP:CB	2.43	0.47
1:A:497:HIS:CD2	1:A:501:LEU:HD13	2.49	0.47
1:A:510:ARG:HG3	1:A:512:ARG:HB2	1.97	0.47
1:A:568:LEU:O	5:A:604:BOG:O6	2.32	0.47
1:A:120:LYS:HD2	6:A:816:HOH:O	2.15	0.46
1:A:462:SER:HA	1:A:506:SER:HB3	1.97	0.46
1:A:475:PRO:HB2	1:A:478:SER:HB2	1.95	0.46
1:A:347:VAL:HG11	1:A:479:ASN:HB3	1.99	0.45
1:A:325:LEU:HG	1:A:329:LEU:HD23	1.98	0.45
1:A:455:TRP:HZ3	1:A:526:ILE:HB	1.82	0.45
1:A:289:VAL:HG11	1:A:312:VAL:HG22	1.98	0.44
1:A:529:VAL:O	1:A:532:GLU:HG3	2.18	0.44
1:A:501:LEU:HD23	1:A:503:PHE:HE2	1.83	0.44
5:A:604:BOG:H8'2	5:A:604:BOG:H5'2	1.78	0.43
1:A:455:TRP:O	1:A:457:PRO:HD3	2.18	0.43
1:A:305:HIS:NE2	1:A:356:ILE:HG21	2.32	0.43
1:A:371:ASN:OD1	1:A:444:HIS:HB3	2.18	0.42
1:A:534:ALA:O	1:A:538:THR:HG23	2.19	0.42
1:A:39:ILE:HA	1:A:379:PRO:HD3	2.01	0.42
1:A:446:CYS:HB2	1:A:451:ASN:HD22	1.85	0.42
1:A:317:TRP:O	1:A:321:GLN:HG2	2.19	0.42
1:A:562:CYS:H	5:A:604:BOG:H7'1	1.84	0.42
1:A:171:LYS:HD3	1:A:174:GLU:HG3	2.01	0.41
1:A:288:HIS:HA	1:A:289:VAL:HA	1.76	0.41
1:A:144:LYS:HG3	1:A:145:THR:HG23	2.02	0.41
1:A:338:TYR:HB2	1:A:382:LEU:HD13	2.01	0.41
1:A:471:PRO:HB2	1:A:487:CYS:SG	2.60	0.41
1:A:63:VAL:CG1	1:A:382:LEU:HB2	2.51	0.40

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:49:THR:CB	1:A:304:GLN:HG3	2.51	0.40
1:A:469:PHE:HB2	1:A:494:VAL:HG22	2.03	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	551/583 (94%)	512 (93%)	37 (7%)	2 (0%)	34 24

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	516	THR
1	A	523	PHE

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	473/501 (94%)	471 (100%)	2 (0%)	91 91

All (2) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	343	GLN
1	A	543	GLU

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

Mol	Chain	Res	Type
1	A	149	HIS
1	A	273	GLN
1	A	451	ASN

5.3.3 RNA (i)

There are no RNA molecules in this entry.

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

1 non-standard protein/DNA/RNA residue is 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	ALS	A	75	1,2	7,10,11	0.98	0	6,14,16	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
1	ALS	A	75	1,2	-	1/3/11/13	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (1) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
1	A	75	ALS	N-CA-CB-OG

There are no ring outliers.

No monomer is involved in short contacts.

5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

Of 6 ligands modelled in this entry, 1 is monoatomic - leaving 5 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	$\#\|Z\| > 2$
3	NAG	A	602	1	14,14,15	0.27	0	17,19,21	0.45	0
4	PO4	A	603	-	4,4,4	3.54	3 (75%)	6,6,6	0.79	0
5	BOG	A	605	-	20,20,20	1.43	2 (10%)	25,25,25	2.61	10 (40%)
5	BOG	A	604	-	20,20,20	1.39	2 (10%)	25,25,25	0.99	2 (8%)
3	NAG	A	601	1	14,14,15	0.22	0	17,19,21	0.46	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
5	BOG	A	604	-	-	8/11/31/31	0/1/1/1
3	NAG	A	602	1	-	2/6/23/26	0/1/1/1
5	BOG	A	605	-	-	9/11/31/31	0/1/1/1
3	NAG	A	601	1	-	3/6/23/26	0/1/1/1

All (7) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	A	603	PO4	P-O1	5.74	1.64	1.50
5	A	604	BOG	O5-C1	4.64	1.53	1.41
5	A	605	BOG	O5-C1	4.49	1.53	1.41
4	A	603	PO4	P-O3	2.98	1.63	1.54
4	A	603	PO4	P-O4	-2.62	1.46	1.54
5	A	605	BOG	O1-C1	-2.38	1.36	1.40
5	A	604	BOG	O1-C1	-2.15	1.36	1.40

All (12) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
5	A	605	BOG	C3-C4-C5	5.55	120.14	110.24
5	A	605	BOG	C4-C3-C2	4.92	119.40	110.82
5	A	605	BOG	C6-C5-C4	-4.87	101.61	113.00
5	A	605	BOG	O5-C5-C4	4.00	116.96	109.69
5	A	605	BOG	O1-C1-C2	-3.74	102.46	108.30
5	A	605	BOG	O3-C3-C4	-3.66	101.90	110.35
5	A	605	BOG	O3-C3-C2	-3.54	102.16	110.35
5	A	605	BOG	O5-C1-C2	2.96	116.61	110.35
5	A	605	BOG	C1-O5-C5	2.69	118.97	113.69
5	A	605	BOG	C1-C2-C3	2.66	115.53	110.00
5	A	604	BOG	O3-C3-C2	-2.29	105.06	110.35
5	A	604	BOG	C6-C5-C4	-2.12	108.05	113.00

There are no chirality outliers.

All (22) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	A	604	BOG	C2-C1-O1-C1'
5	A	604	BOG	O5-C1-O1-C1'
5	A	605	BOG	C2-C1-O1-C1'
5	A	605	BOG	O5-C1-O1-C1'
5	A	604	BOG	O5-C5-C6-O6
3	A	601	NAG	O5-C5-C6-O6
3	A	601	NAG	C4-C5-C6-O6
5	A	605	BOG	O5-C5-C6-O6
5	A	604	BOG	C4-C5-C6-O6
5	A	605	BOG	C4-C5-C6-O6
5	A	604	BOG	C3'-C4'-C5'-C6'
5	A	605	BOG	C2'-C1'-O1-C1
5	A	605	BOG	C1'-C2'-C3'-C4'

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms
3	A	602	NAG	C4-C5-C6-O6
3	A	602	NAG	O5-C5-C6-O6
5	A	604	BOG	C4'-C5'-C6'-C7'
5	A	605	BOG	C3'-C4'-C5'-C6'
5	A	604	BOG	C5'-C6'-C7'-C8'
5	A	605	BOG	O1-C1'-C2'-C3'
5	A	604	BOG	C1'-C2'-C3'-C4'
3	A	601	NAG	C1-C2-N2-C7
5	A	605	BOG	C2'-C3'-C4'-C5'

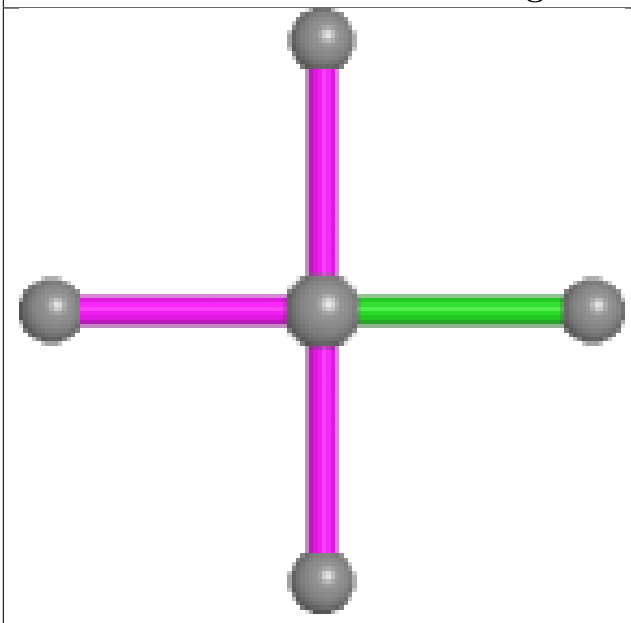
There are no ring outliers.

3 monomers are involved in 9 short contacts:

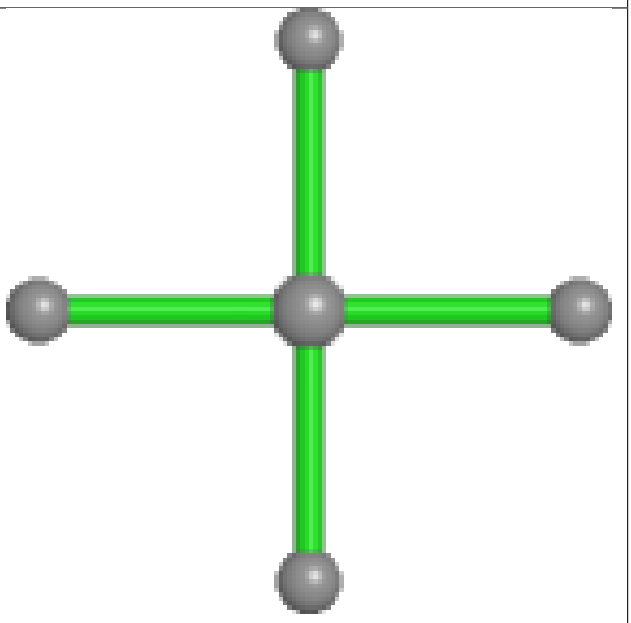
Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	A	605	BOG	2	0
5	A	604	BOG	6	0
3	A	601	NAG	1	0

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.

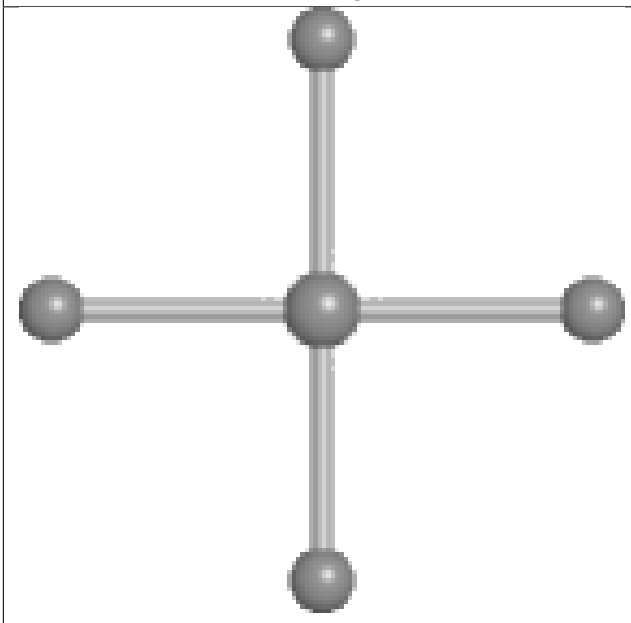
Ligand PO4 A 603



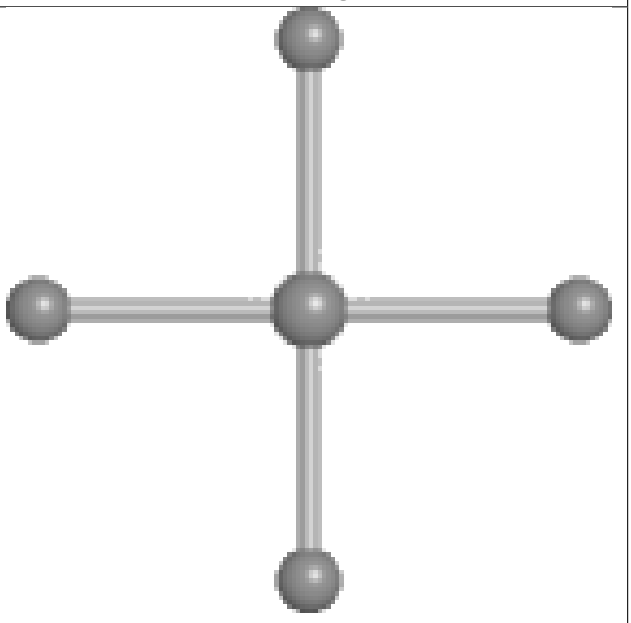
Bond lengths



Bond angles

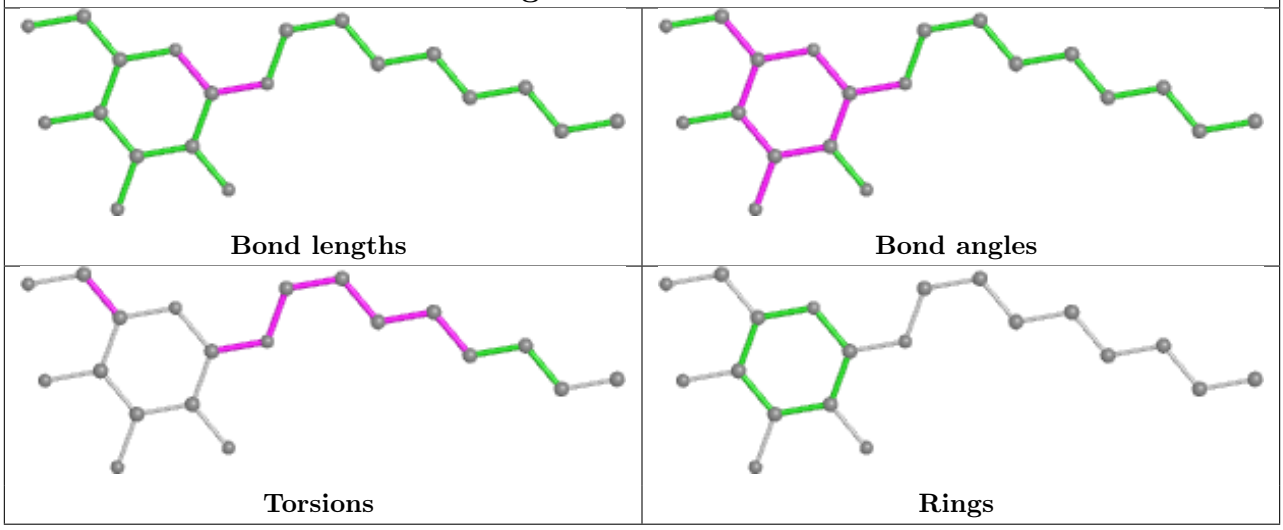


Torsions

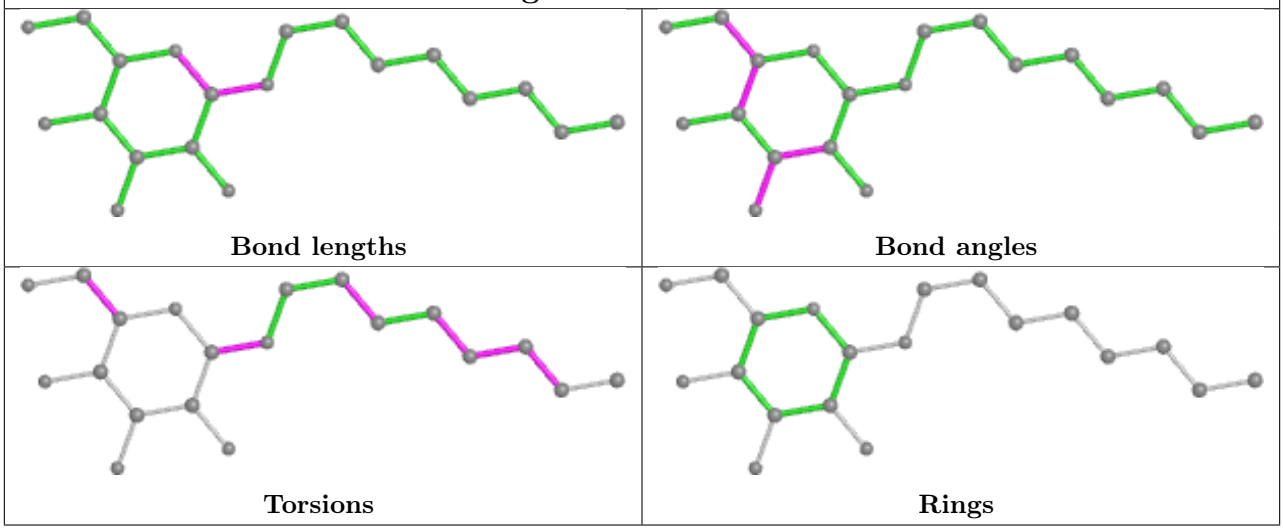


Rings

Ligand BOG A 605



Ligand BOG A 604



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	553/583 (94%)	0.70	105 (18%) 1 0	31, 60, 127, 195	0

All (105) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	476	VAL	8.4
1	A	480	GLY	8.4
1	A	517	PRO	6.3
1	A	524	TYR	6.0
1	A	521	PRO	5.8
1	A	24	ALA	5.6
1	A	23	ALA	5.3
1	A	499	PRO	4.9
1	A	353	LYS	4.8
1	A	518	ALA	4.8
1	A	519	SER	4.8
1	A	529	VAL	4.7
1	A	477	GLY	4.6
1	A	576	LYS	4.6
1	A	49	THR	4.4
1	A	474	ASN	4.4
1	A	69	LEU	4.3
1	A	565	SER	4.3
1	A	510	ARG	4.2
1	A	71	ALA	4.0
1	A	78	SER	4.0
1	A	46	GLY	3.9
1	A	523	PHE	3.9
1	A	507	LYS	3.8
1	A	82	PHE	3.7
1	A	478	SER	3.7
1	A	338	TYR	3.6

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	475	PRO	3.6
1	A	34	ALA	3.6
1	A	516	THR	3.6
1	A	536	ARG	3.6
1	A	532	GLU	3.5
1	A	522	ARG	3.5
1	A	459	ASN	3.5
1	A	70	ALA	3.4
1	A	283	VAL	3.4
1	A	33	MET	3.3
1	A	566	THR	3.3
1	A	304	GLN	3.3
1	A	341	SER	3.3
1	A	306	GLY	3.2
1	A	493	TYR	3.2
1	A	376	ILE	3.2
1	A	164	LEU	3.1
1	A	286	TYR	3.1
1	A	79	ARG	3.1
1	A	491	GLY	3.0
1	A	339	PHE	3.0
1	A	287	LEU	3.0
1	A	495	THR	2.9
1	A	526	ILE	2.9
1	A	330	ARG	2.9
1	A	490	PHE	2.8
1	A	528	LYS	2.8
1	A	340	THR	2.8
1	A	31	LEU	2.8
1	A	32	VAL	2.8
1	A	47	ASN	2.7
1	A	508	ASP	2.7
1	A	483	ALA	2.7
1	A	494	VAL	2.7
1	A	162	ILE	2.7
1	A	574	ARG	2.7
1	A	400	ASN	2.7
1	A	285	SER	2.6
1	A	356	ILE	2.6
1	A	74	LEU	2.6
1	A	352	SER	2.6
1	A	487	CYS	2.6

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	284	LEU	2.6
1	A	305	HIS	2.5
1	A	25	SER	2.5
1	A	355	GLU	2.5
1	A	328	GLU	2.5
1	A	300	ALA	2.5
1	A	131	LEU	2.4
1	A	301	GLY	2.4
1	A	134	LYS	2.4
1	A	343	GLN	2.4
1	A	342	ASP	2.4
1	A	542	PRO	2.4
1	A	72	SER	2.4
1	A	363	ILE	2.4
1	A	540	THR	2.3
1	A	73	PRO	2.3
1	A	297	LYS	2.3
1	A	492	SER	2.3
1	A	504	ASP	2.2
1	A	369	ALA	2.2
1	A	282	LEU	2.2
1	A	515	LEU	2.2
1	A	567	GLY	2.2
1	A	135	TRP	2.2
1	A	525	GLU	2.2
1	A	513	ASN	2.1
1	A	503	PHE	2.1
1	A	382	LEU	2.1
1	A	307	VAL	2.1
1	A	520	GLU	2.1
1	A	136	HIS	2.1
1	A	509	PRO	2.1
1	A	68	HIS	2.1
1	A	464	TRP	2.1
1	A	80	ALA	2.0
1	A	442	LEU	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	ALS	A	75	11/12	0.97	0.16	35,39,47,50	0

6.3 Carbohydrates (i)

There are no monosaccharides in this entry.

6.4 Ligands (i)

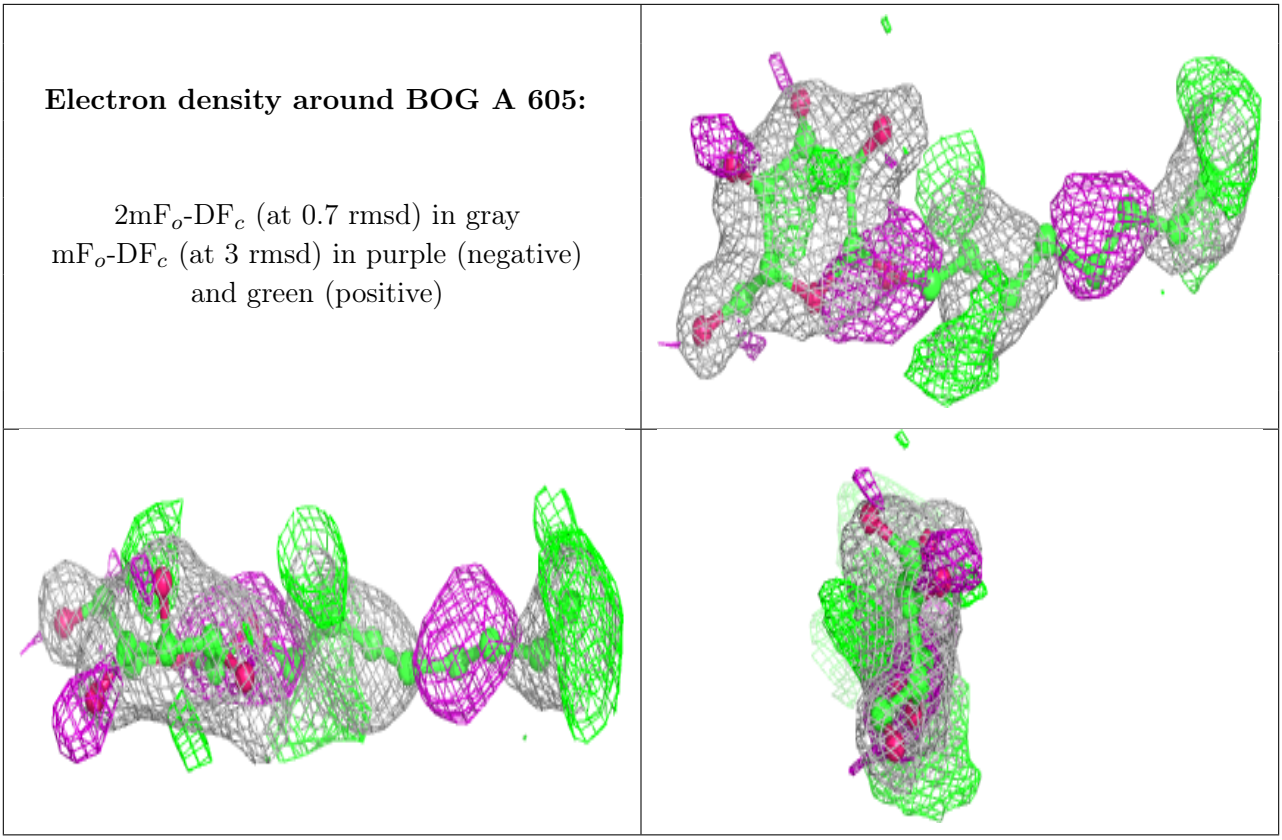
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' 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
5	BOG	A	605	20/20	0.72	0.25	38,50,54,61	0
3	NAG	A	602	14/15	0.77	0.45	65,67,70,71	14
3	NAG	A	601	14/15	0.81	0.28	81,89,99,100	14
5	BOG	A	604	20/20	0.86	0.29	52,60,68,69	0
2	CA	A	600	1/1	0.97	0.17	46,46,46,46	0
4	PO4	A	603	5/5	0.97	0.17	41,44,45,50	5

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.

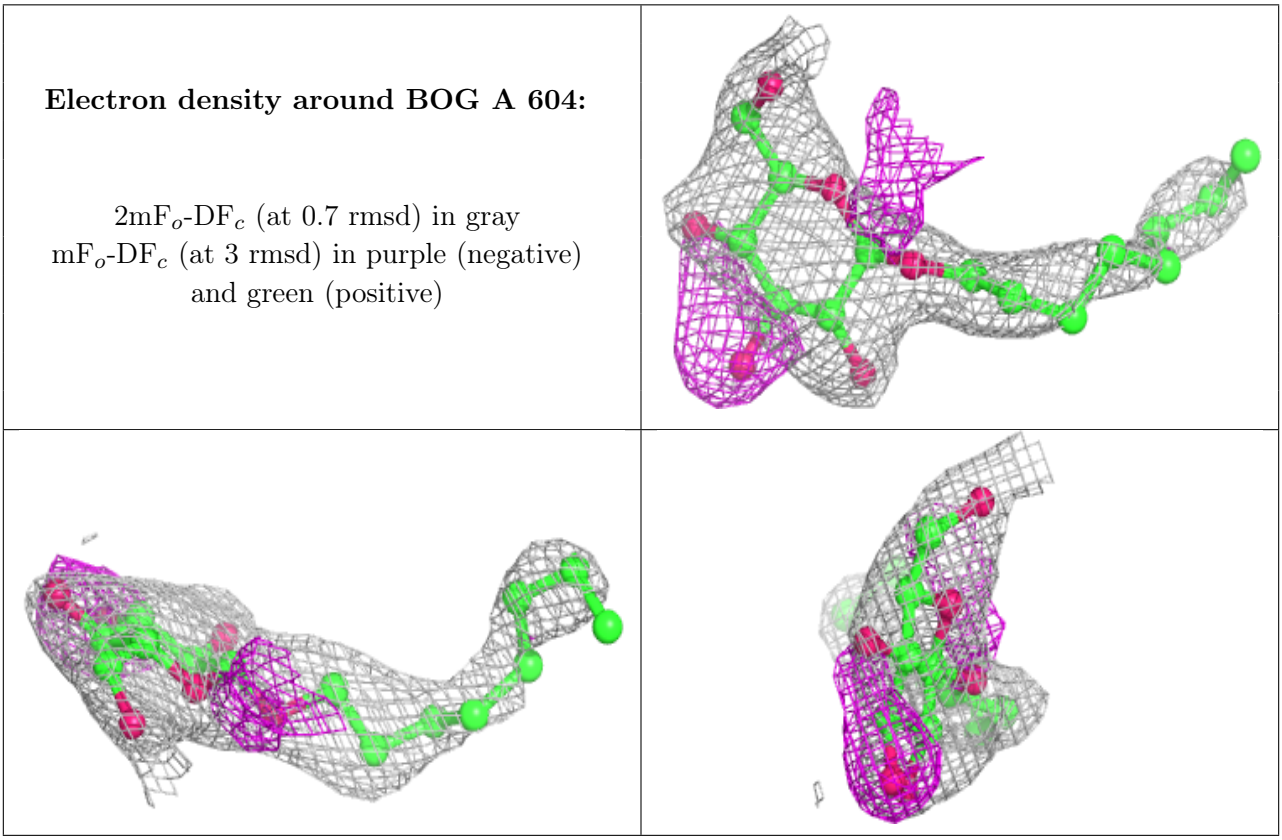
Electron density around BOG A 605:

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



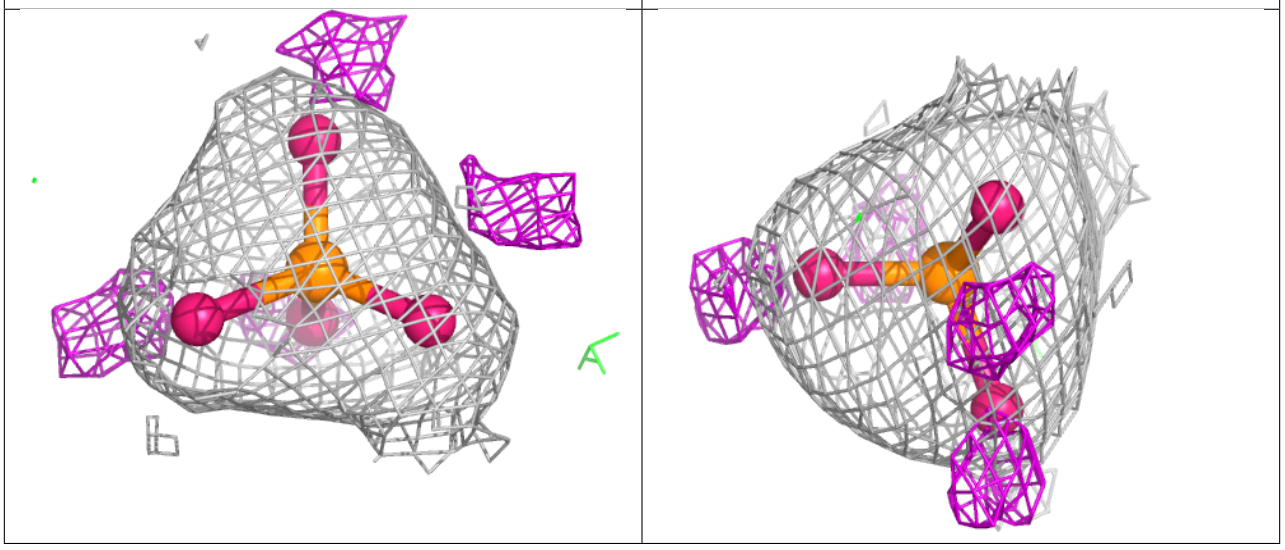
Electron density around BOG A 604:

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



Electron density around PO4 A 603:

$2mF_o$-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.