wwPDB X-ray Structure Validation Summary Report (i)

Oct 3, 2023 – 01:58 PM JST

PDB ID	:	8H2V
Title	:	Cellodextrin phosphorylase from Clostridium thermocellum mutant - all cysteine residues were substituted with serines
Authors	:	Kuga, T.; Sunagawa, N.; Igarashi, K.
Deposited on	:	2022-10-07
Resolution	:	1.68 Å(reported)

This is a wwPDB X-ray Structure Validation Summary Report for a publicly released PDB entry.

We welcome your comments at *validation@mail.wwpdb.org*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

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

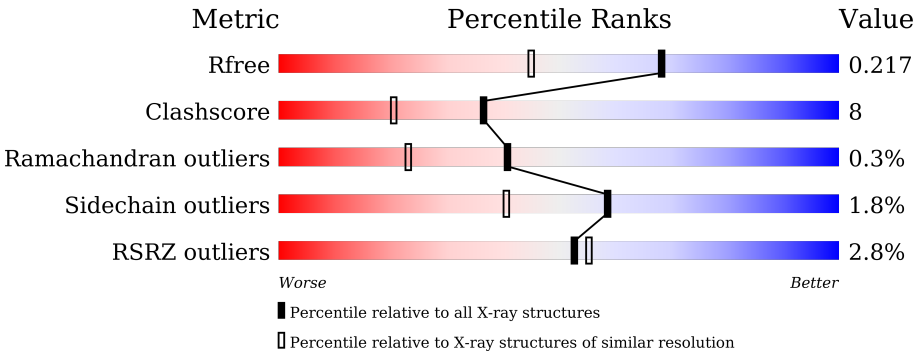
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.35.1
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35.1

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 1.68 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	6780 (1.70-1.66)
Clashscore	141614	7310 (1.70-1.66)
Ramachandran outliers	138981	7173 (1.70-1.66)
Sidechain outliers	138945	7172 (1.70-1.66)
RSRZ outliers	127900	6661 (1.70-1.66)

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	993	3% 85% 13%
1	B	993	2% 86% 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
4	GOL	B	1004	-	-	X	-
5	ACT	A	1005	-	-	X	-
5	ACT	B	1005	-	-	X	-
5	ACT	B	1007	-	-	X	-
5	ACT	B	1015	-	-	X	-
6	PEG	B	1013	-	-	X	-

2 Entry composition (i)

There are 8 unique types of molecules in this entry. The entry contains 32463 atoms, of which 14635 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 Cellodextrin phosphorylase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	984	Total	C	H	N	O	S	190	19	0
			15225	5141	7193	1336	1533	22			
1	B	984	Total	C	H	N	O	S	188	22	0
			15305	5154	7259	1333	1537	22			

There are 44 discrepancies between the modelled and reference sequences:

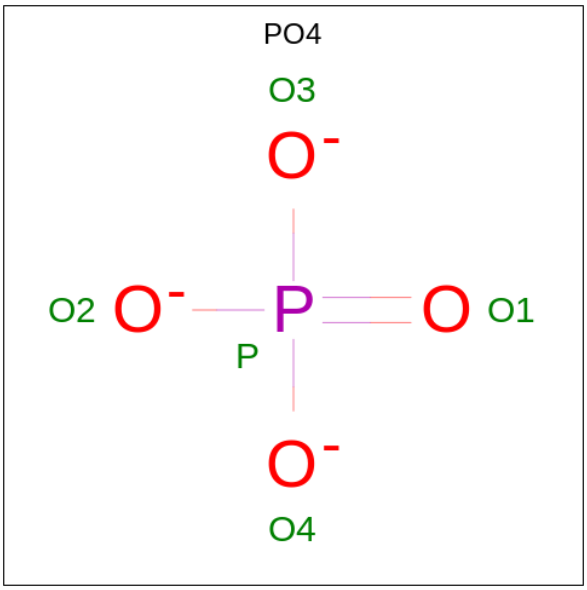
Chain	Residue	Modelled	Actual	Comment	Reference
A	1	MET	-	initiating methionine	UNP Q93HT8
A	2	GLY	-	expression tag	UNP Q93HT8
A	65	SER	CYS	engineered mutation	UNP Q93HT8
A	81	SER	CYS	engineered mutation	UNP Q93HT8
A	226	SER	CYS	engineered mutation	UNP Q93HT8
A	230	SER	CYS	engineered mutation	UNP Q93HT8
A	241	SER	CYS	engineered mutation	UNP Q93HT8
A	354	SER	CYS	engineered mutation	UNP Q93HT8
A	373	SER	CYS	engineered mutation	UNP Q93HT8
A	607	SER	CYS	engineered mutation	UNP Q93HT8
A	626	SER	CYS	engineered mutation	UNP Q93HT8
A	630	ASP	ALA	engineered mutation	UNP Q93HT8
A	873	SER	CYS	engineered mutation	UNP Q93HT8
A	935	SER	CYS	engineered mutation	UNP Q93HT8
A	986	LEU	-	expression tag	UNP Q93HT8
A	987	GLU	-	expression tag	UNP Q93HT8
A	988	HIS	-	expression tag	UNP Q93HT8
A	989	HIS	-	expression tag	UNP Q93HT8
A	990	HIS	-	expression tag	UNP Q93HT8
A	991	HIS	-	expression tag	UNP Q93HT8
A	992	HIS	-	expression tag	UNP Q93HT8
A	993	HIS	-	expression tag	UNP Q93HT8
B	1	MET	-	initiating methionine	UNP Q93HT8
B	2	GLY	-	expression tag	UNP Q93HT8
B	65	SER	CYS	engineered mutation	UNP Q93HT8

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	81	SER	CYS	engineered mutation	UNP Q93HT8
B	226	SER	CYS	engineered mutation	UNP Q93HT8
B	230	SER	CYS	engineered mutation	UNP Q93HT8
B	241	SER	CYS	engineered mutation	UNP Q93HT8
B	354	SER	CYS	engineered mutation	UNP Q93HT8
B	373	SER	CYS	engineered mutation	UNP Q93HT8
B	607	SER	CYS	engineered mutation	UNP Q93HT8
B	626	SER	CYS	engineered mutation	UNP Q93HT8
B	630	ASP	ALA	engineered mutation	UNP Q93HT8
B	873	SER	CYS	engineered mutation	UNP Q93HT8
B	935	SER	CYS	engineered mutation	UNP Q93HT8
B	986	LEU	-	expression tag	UNP Q93HT8
B	987	GLU	-	expression tag	UNP Q93HT8
B	988	HIS	-	expression tag	UNP Q93HT8
B	989	HIS	-	expression tag	UNP Q93HT8
B	990	HIS	-	expression tag	UNP Q93HT8
B	991	HIS	-	expression tag	UNP Q93HT8
B	992	HIS	-	expression tag	UNP Q93HT8
B	993	HIS	-	expression tag	UNP Q93HT8

- Molecule 2 is PHOSPHATE ION (three-letter code: PO4) (formula: O_4P) (labeled as "Ligand of Interest" by depositor).



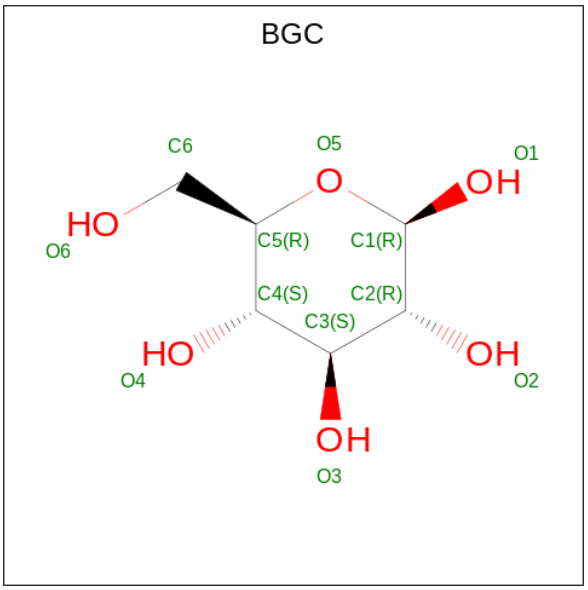
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	1	Total O P 5 4 1	0	0

Continued on next page...

Continued from previous page...

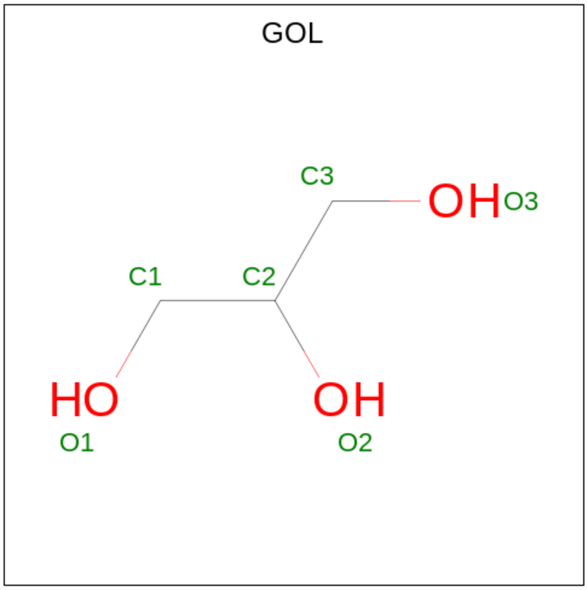
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	B	1	Total O P 5 4 1	0	0

- Molecule 3 is beta-D-glucopyranose (three-letter code: BGC) (formula: $C_6H_{12}O_6$) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total C O 12 6 6	0	0
3	B	1	Total C O 12 6 6	0	0

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



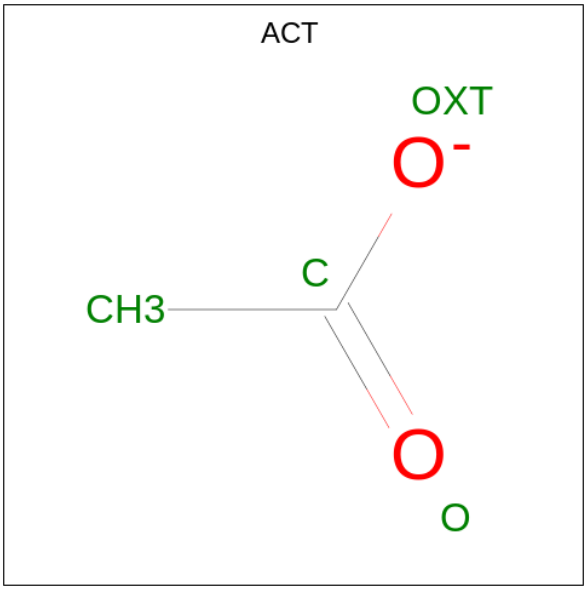
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total C H O 14 3 8 3	0	0
4	A	1	Total C H O 14 3 8 3	0	0
4	A	1	Total C H O 14 3 8 3	0	0
4	A	1	Total C H O 14 3 8 3	0	0
4	A	1	Total C H O 14 3 8 3	0	0
4	A	1	Total C H O 14 3 8 3	0	0
4	B	1	Total C H O 14 3 8 3	0	0
4	B	1	Total C H O 14 3 8 3	0	0
4	B	1	Total C H O 14 3 8 3	0	0
4	B	1	Total C H O 14 3 8 3	0	0
4	B	1	Total C H O 14 3 8 3	0	0
4	B	1	Total C H O 14 3 8 3	0	0
4	B	1	Total C H O 14 3 8 3	0	0
4	B	1	Total C H O 14 3 8 3	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	B	1	Total C H O 14 3 8 3	0	0

- Molecule 5 is ACETATE ION (three-letter code: ACT) (formula: $C_2H_3O_2$).



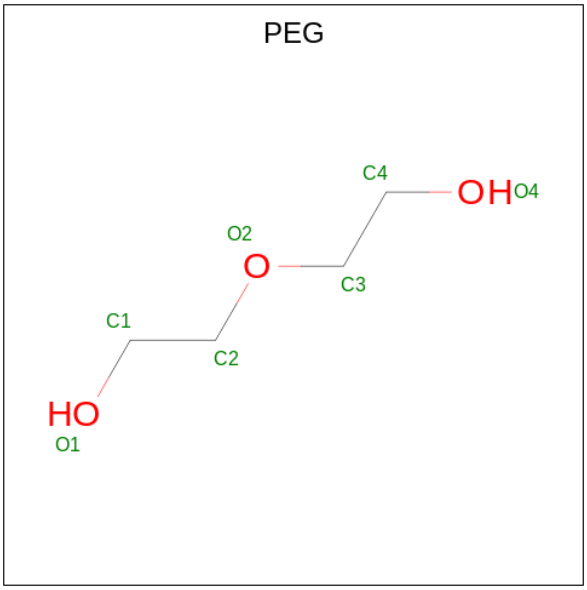
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	1	Total C H O 7 2 3 2	0	0
5	A	1	Total C H O 7 2 3 2	0	0
5	A	1	Total C H O 7 2 3 2	0	0
5	A	1	Total C H O 7 2 3 2	0	0
5	A	1	Total C H O 7 2 3 2	0	0
5	A	1	Total C H O 7 2 3 2	0	0
5	B	1	Total C H O 7 2 3 2	0	0
5	B	1	Total C H O 7 2 3 2	0	0
5	B	1	Total C H O 7 2 3 2	0	0
5	B	1	Total C H O 7 2 3 2	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	B	1	Total C H O 7 2 3 2	0	0

- Molecule 6 is DI(HYDROXYETHYL)ETHER (three-letter code: PEG) (formula: $C_4H_{10}O_3$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	1	Total C H O 17 4 10 3	0	0
6	A	1	Total C H O 17 4 10 3	0	0
6	B	1	Total C H O 17 4 10 3	0	0

- Molecule 7 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	A	2	Total Cl 2 2	0	0
7	B	2	Total Cl 2 2	0	0

- Molecule 8 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
8	A	794	Total O 794 794	0	0

Continued on next page...

Continued from previous page...

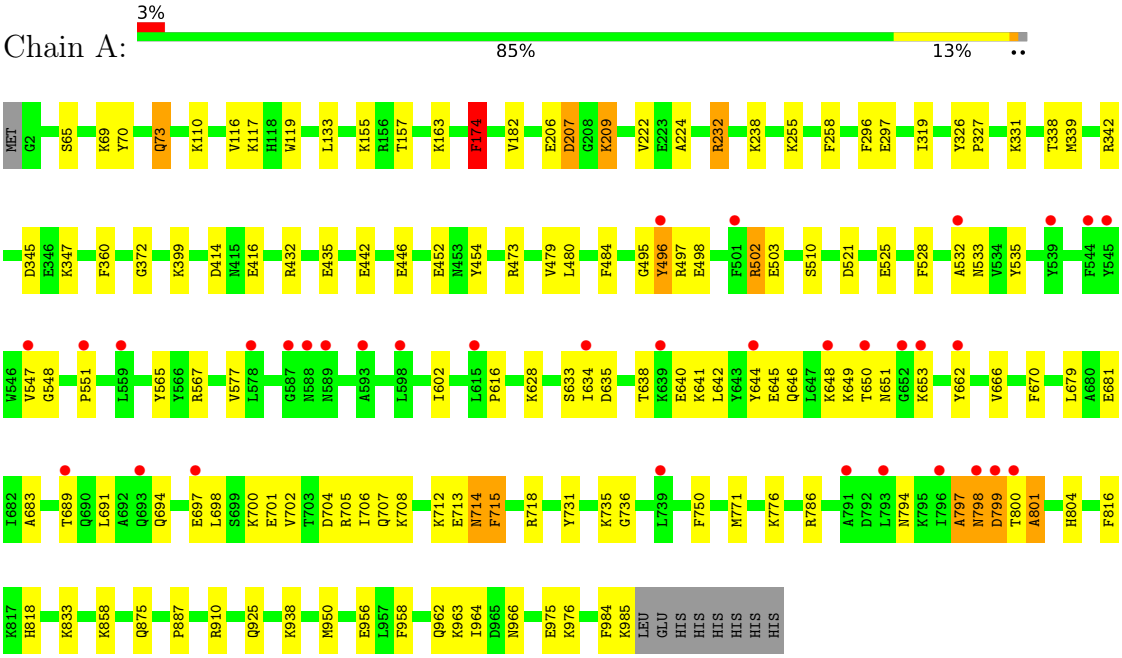
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	B	763	Total 763	O 763	0	0

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.

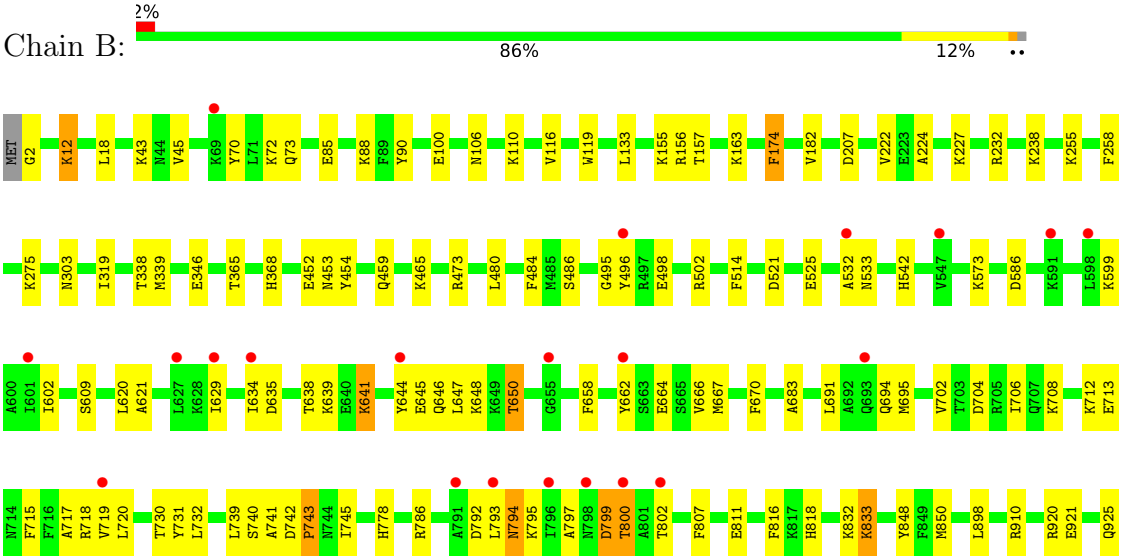
• Molecule 1: Cellodextrin phosphorylase

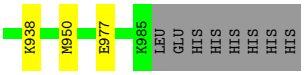
Chain A:



• Molecule 1: Cellodextrin phosphorylase

Chain B:




K938
M950
E977
K985
LEU
GLU
HIS
HIS
HIS
HIS
HIS
HIS

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α, β, γ	82.28Å 88.52Å 88.54Å 98.41° 110.69° 110.65°	Depositor
Resolution (Å)	43.44 – 1.68 43.44 – 1.68	Depositor EDS
% Data completeness (in resolution range)	97.0 (43.44-1.68) 96.9 (43.44-1.68)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.24 (at 1.68Å)	Xtriage
Refinement program	PHENIX 1.20_4459	Depositor
R, R_{free}	0.184 , 0.217 0.183 , 0.217	Depositor DCC
R_{free} test set	11539 reflections (5.02%)	wwPDB-VP
Wilson B-factor (Å^2)	23.4	Xtriage
Anisotropy	0.224	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.37 , 43.3	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	0.479 for -h,-l,-k	Xtriage
F_o,F_c correlation	0.97	EDS
Total number of atoms	32463	wwPDB-VP
Average B, all atoms (Å^2)	35.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.53% 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: PEG, PO4, CL, ACT, BGC, GOL

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	0.55	0/8210	0.72	2/11106 (0.0%)
1	B	0.54	1/8224 (0.0%)	0.71	0/11130
All	All	0.54	1/16434 (0.0%)	0.72	2/22236 (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
1	B	0	2
All	All	0	4

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	743	PRO	N-CD	9.80	1.61	1.47

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	174	PHE	CB-CG-CD1	5.28	124.49	120.80
1	A	174	PHE	CB-CG-CD2	-5.14	117.20	120.80

There are no chirality outliers.

All (4) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	495	GLY	Peptide
1	A	496[B]	TYR	Peptide
1	B	495	GLY	Peptide
1	B	496[B]	TYR	Peptide

5.2 Too-close contacts (i)

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

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	8032	7193	7882	129	0
1	B	8046	7259	7894	119	0
2	A	5	0	0	0	0
2	B	5	0	0	0	0
3	A	12	0	12	0	0
3	B	12	0	12	0	0
4	A	36	48	48	4	0
4	B	54	72	72	9	0
5	A	24	18	18	4	0
5	B	20	15	15	10	0
6	A	14	20	20	2	0
6	B	7	10	10	6	0
7	A	2	0	0	1	0
7	B	2	0	0	1	0
8	A	794	0	0	32	0
8	B	763	0	0	45	0
All	All	17828	14635	15983	255	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 8.

The worst 5 of 255 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:718[B]:ARG:NH1	7:A:1015:CL:CL	2.30	1.01
1:A:65:SER:O	1:A:69:LYS:HD3	1.63	0.97
1:B:704:ASP:HB3	1:B:708:LYS:HE2	1.48	0.94
1:B:634:ILE:HD11	8:B:1793:HOH:O	1.73	0.89

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:521:ASP:O	1:A:525:GLU:HG3	1.73	0.89

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	1001/993 (101%)	956 (96%)	42 (4%)	3 (0%)	41 23
1	B	1004/993 (101%)	966 (96%)	35 (4%)	3 (0%)	41 23
All	All	2005/1986 (101%)	1922 (96%)	77 (4%)	6 (0%)	41 23

5 of 6 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	799	ASP
1	A	801	ALA
1	B	498	GLU
1	B	800	THR
1	B	799	ASP

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	869/859 (101%)	854 (98%)	15 (2%)	60 43

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
1	B	872/859 (102%)	856 (98%)	16 (2%)	59 40
All	All	1741/1718 (101%)	1710 (98%)	31 (2%)	59 40

5 of 31 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	798	ASN
1	B	794	ASN
1	B	163	LYS
1	B	833	LYS
1	B	641	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 8 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	925	GLN
1	B	778	HIS
1	B	31	ASN
1	A	798	ASN
1	B	694	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, 4 are monoatomic - leaving 33 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	$\#\|Z\| > 2$
2	PO4	B	1001	-	4,4,4	0.68	0	6,6,6	1.15	0
4	GOL	A	1003	-	5,5,5	1.22	0	5,5,5	0.87	0
2	PO4	A	1001	-	4,4,4	1.18	0	6,6,6	1.19	1 (16%)
5	ACT	A	1005	-	3,3,3	0.80	0	3,3,3	1.81	2 (66%)
4	GOL	B	1012	-	5,5,5	1.22	0	5,5,5	0.92	0
3	BGC	B	1002	-	12,12,12	0.48	0	17,17,17	0.52	0
5	ACT	A	1006	-	3,3,3	1.18	0	3,3,3	1.34	0
5	ACT	A	1012	-	3,3,3	1.05	0	3,3,3	1.55	0
4	GOL	B	1011	-	5,5,5	0.77	0	5,5,5	0.91	0
5	ACT	A	1004	-	3,3,3	1.60	1 (33%)	3,3,3	1.39	0
4	GOL	A	1016	-	5,5,5	0.93	0	5,5,5	0.78	0
4	GOL	A	1008	-	5,5,5	0.87	0	5,5,5	1.22	0
4	GOL	B	1010	-	5,5,5	1.21	0	5,5,5	1.08	0
4	GOL	B	1006	-	5,5,5	0.71	0	5,5,5	0.96	0
5	ACT	B	1003	-	3,3,3	1.09	0	3,3,3	1.82	1 (33%)
4	GOL	B	1019	-	5,5,5	0.87	0	5,5,5	0.80	0
4	GOL	B	1014	-	5,5,5	1.25	1 (20%)	5,5,5	1.18	1 (20%)
5	ACT	B	1007	-	3,3,3	0.97	0	3,3,3	1.64	1 (33%)
4	GOL	B	1009	-	5,5,5	0.97	0	5,5,5	0.95	0
4	GOL	B	1018	-	5,5,5	1.01	0	5,5,5	0.39	0
5	ACT	B	1008	-	3,3,3	1.43	1 (33%)	3,3,3	1.37	0
3	BGC	A	1002	-	12,12,12	0.58	0	17,17,17	0.63	0
6	PEG	B	1013	-	6,6,6	0.23	0	5,5,5	0.15	0
5	ACT	A	1007	-	3,3,3	1.54	1 (33%)	3,3,3	1.43	0
5	ACT	A	1010	-	3,3,3	1.61	1 (33%)	3,3,3	1.18	0
6	PEG	A	1011	-	6,6,6	0.33	0	5,5,5	0.12	0
5	ACT	B	1015	-	3,3,3	1.18	0	3,3,3	1.48	0
4	GOL	B	1004	-	5,5,5	0.70	0	5,5,5	0.83	0
6	PEG	A	1013	-	6,6,6	0.17	0	5,5,5	0.17	0
4	GOL	A	1009	-	5,5,5	0.90	0	5,5,5	0.94	0
4	GOL	A	1017	-	5,5,5	0.74	0	5,5,5	0.73	0
5	ACT	B	1005	-	3,3,3	1.26	0	3,3,3	1.53	0
4	GOL	A	1018	-	5,5,5	0.73	0	5,5,5	0.97	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	A	1003	-	-	4/4/4/4	-
4	GOL	B	1012	-	-	4/4/4/4	-
3	BGC	B	1002	-	-	2/2/22/22	0/1/1/1
4	GOL	B	1011	-	-	2/4/4/4	-
4	GOL	A	1016	-	-	4/4/4/4	-
4	GOL	A	1008	-	-	2/4/4/4	-
4	GOL	B	1010	-	-	2/4/4/4	-
4	GOL	B	1006	-	-	0/4/4/4	-
4	GOL	B	1019	-	-	2/4/4/4	-
4	GOL	B	1014	-	-	2/4/4/4	-
4	GOL	B	1009	-	-	0/4/4/4	-
4	GOL	B	1018	-	-	0/4/4/4	-
3	BGC	A	1002	-	-	2/2/22/22	0/1/1/1
6	PEG	B	1013	-	-	3/4/4/4	-
6	PEG	A	1011	-	-	2/4/4/4	-
4	GOL	B	1004	-	-	4/4/4/4	-
6	PEG	A	1013	-	-	2/4/4/4	-
4	GOL	A	1009	-	-	2/4/4/4	-
4	GOL	A	1017	-	-	4/4/4/4	-
4	GOL	A	1018	-	-	2/4/4/4	-

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	A	1004	ACT	CH3-C	2.46	1.59	1.49
5	A	1007	ACT	CH3-C	2.21	1.58	1.49
5	B	1008	ACT	CH3-C	2.20	1.58	1.49
5	A	1010	ACT	CH3-C	2.18	1.58	1.49
4	B	1014	GOL	C3-C2	2.03	1.60	1.51

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

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
5	B	1003	ACT	O-C-CH3	-2.39	113.03	122.33
4	B	1014	GOL	C3-C2-C1	-2.27	102.86	111.70

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
5	A	1005	ACT	O-C-CH3	-2.23	113.64	122.33
5	A	1005	ACT	OXT-C-O	2.19	130.14	122.05
2	A	1001	PO4	O4-P-O3	2.13	114.81	107.97

There are no chirality outliers.

5 of 45 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	A	1003	GOL	O1-C1-C2-C3
4	A	1008	GOL	C1-C2-C3-O3
4	A	1008	GOL	O2-C2-C3-O3
4	A	1009	GOL	C1-C2-C3-O3
4	A	1017	GOL	C1-C2-C3-O3

There are no ring outliers.

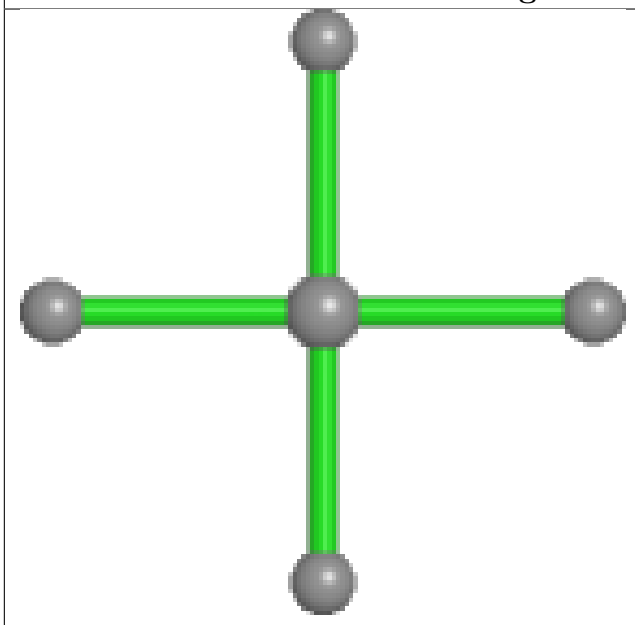
16 monomers are involved in 35 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	1003	GOL	1	0
5	A	1005	ACT	2	0
4	B	1011	GOL	1	0
5	A	1004	ACT	1	0
4	A	1008	GOL	1	0
4	B	1010	GOL	2	0
5	B	1007	ACT	2	0
4	B	1018	GOL	2	0
6	B	1013	PEG	6	0
5	A	1010	ACT	1	0
6	A	1011	PEG	2	0
5	B	1015	ACT	2	0
4	B	1004	GOL	4	0
4	A	1009	GOL	1	0
4	A	1017	GOL	1	0
5	B	1005	ACT	6	0

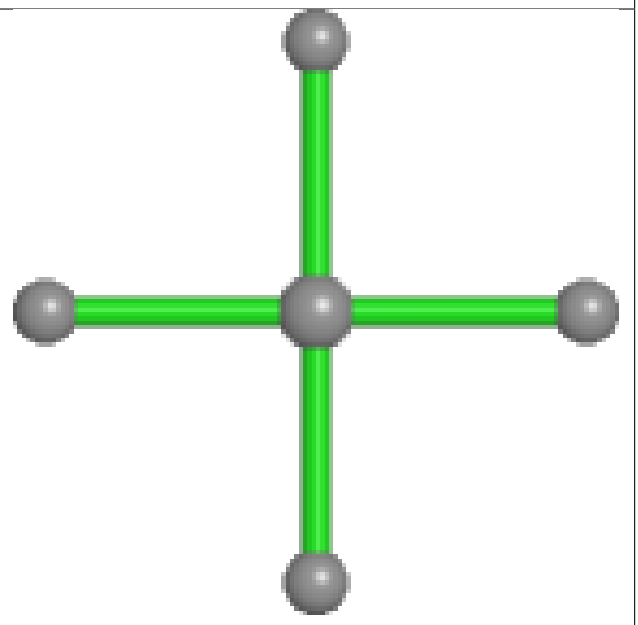
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included.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

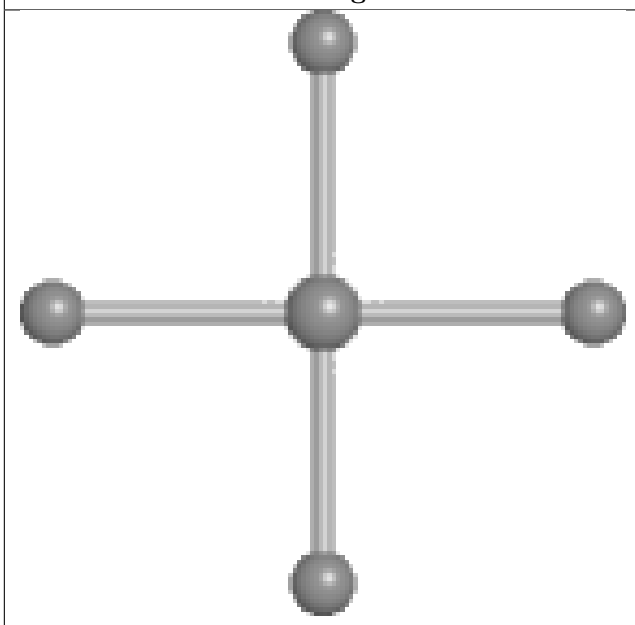
Ligand PO4 B 1001



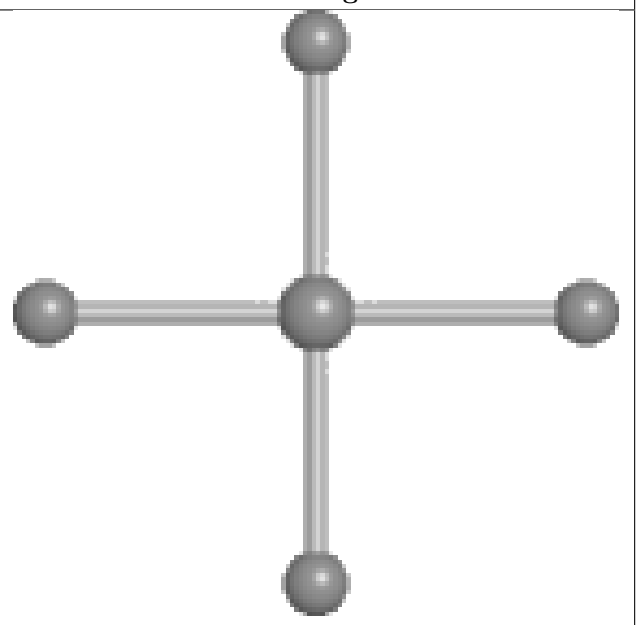
Bond lengths



Bond angles

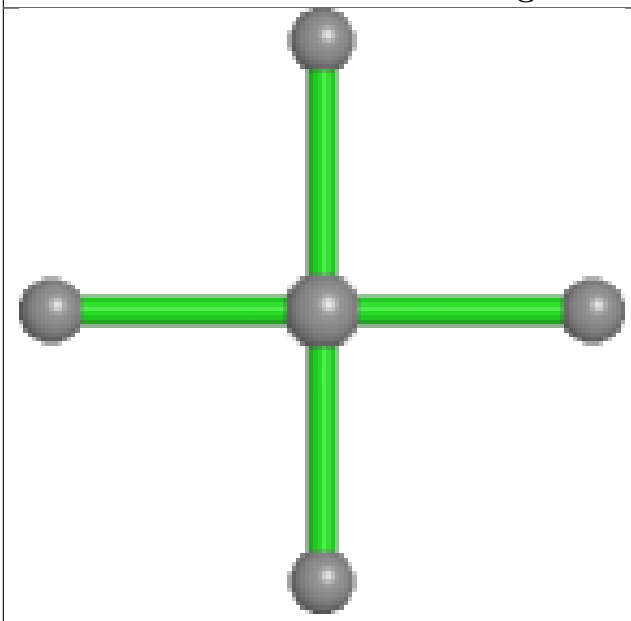


Torsions

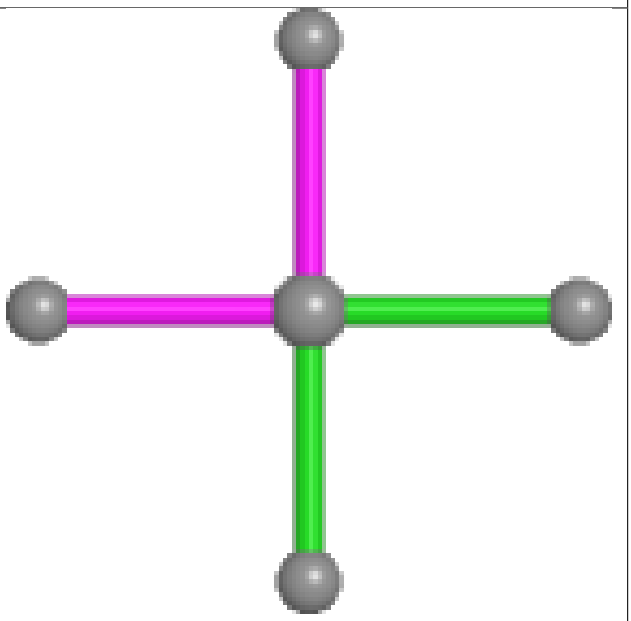


Rings

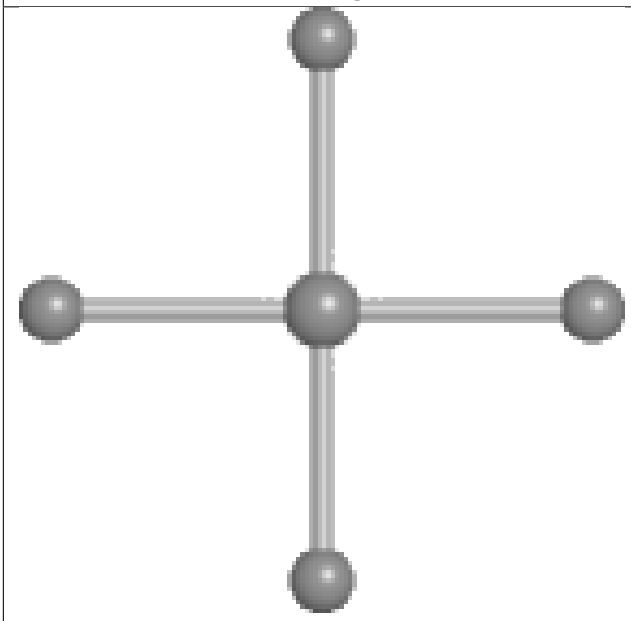
Ligand PO4 A 1001



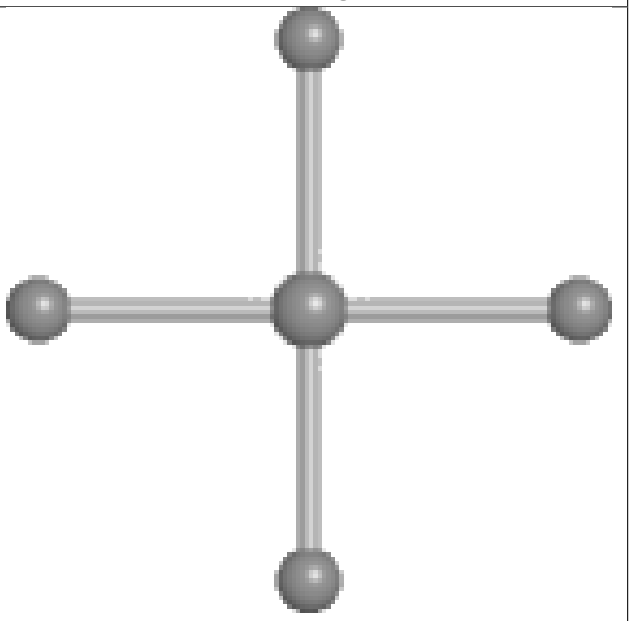
Bond lengths



Bond angles

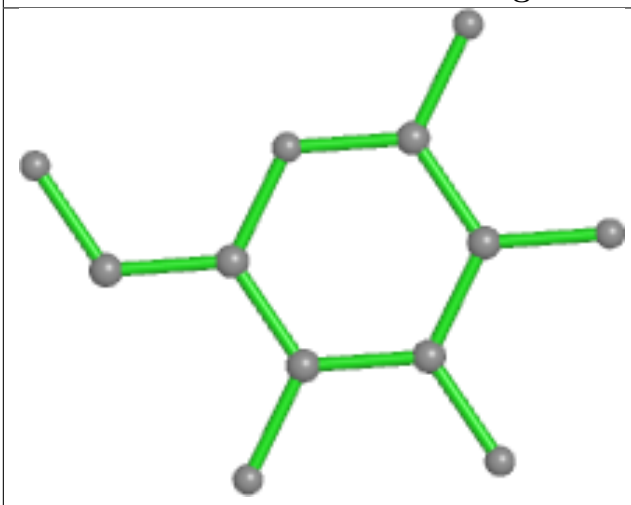


Torsions

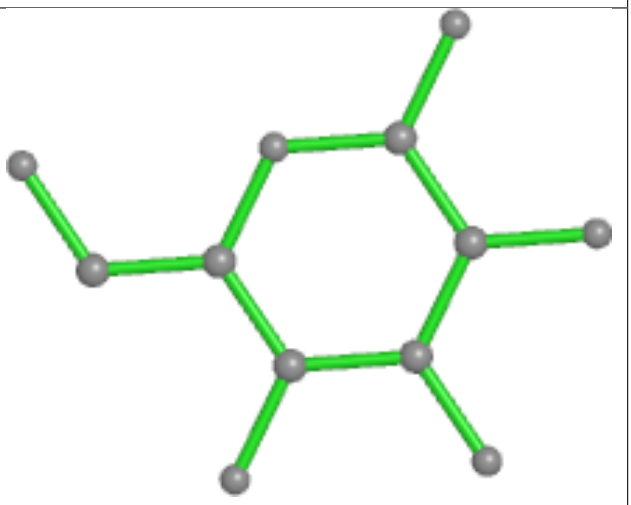


Rings

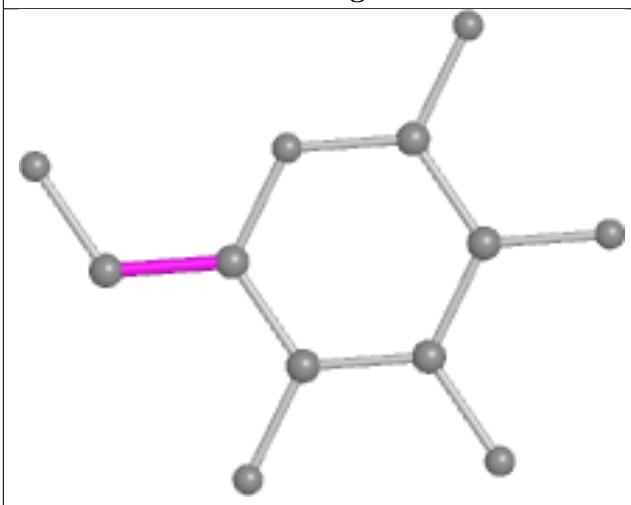
Ligand BGC B 1002



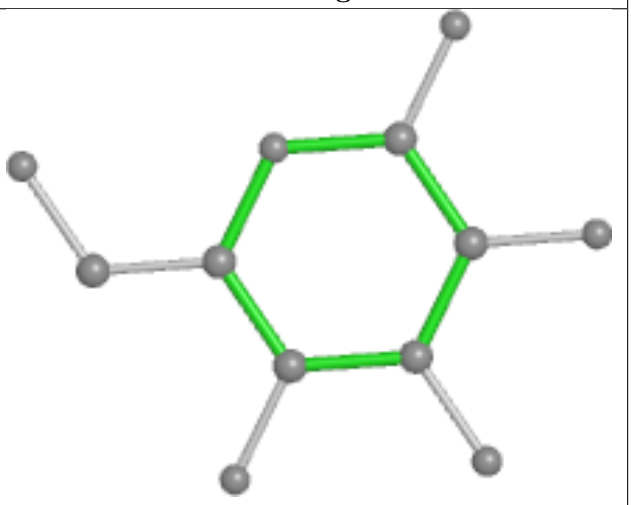
Bond lengths



Bond angles

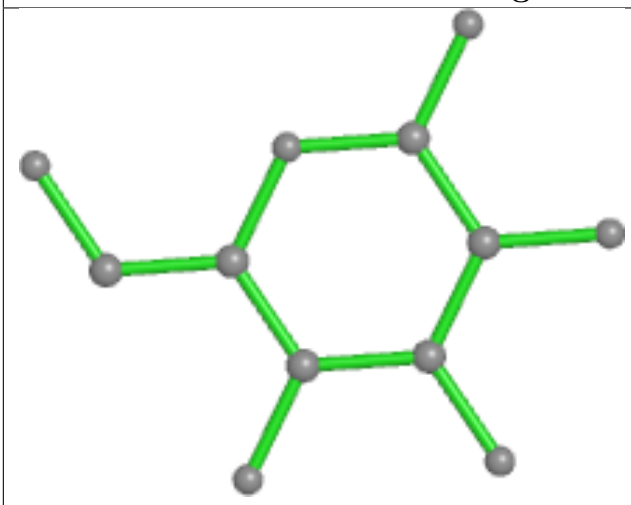


Torsions

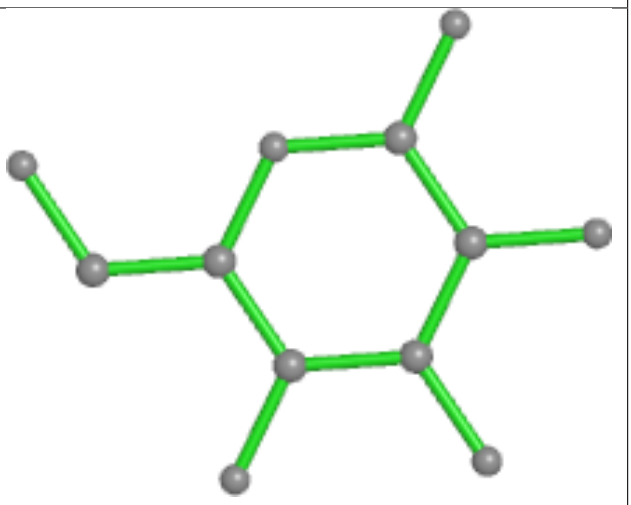


Rings

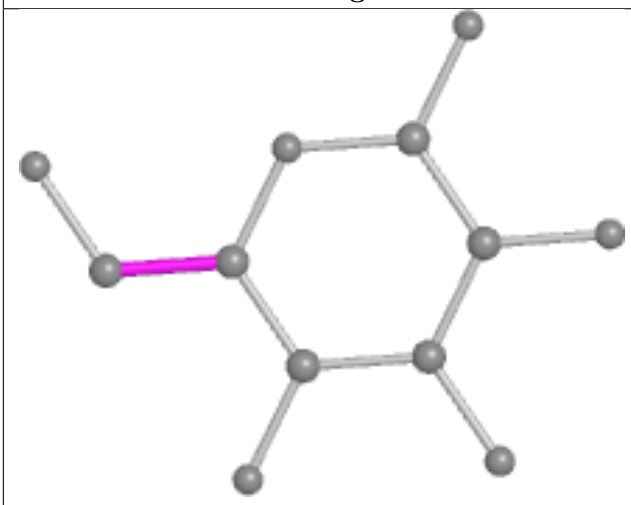
Ligand BGC A 1002



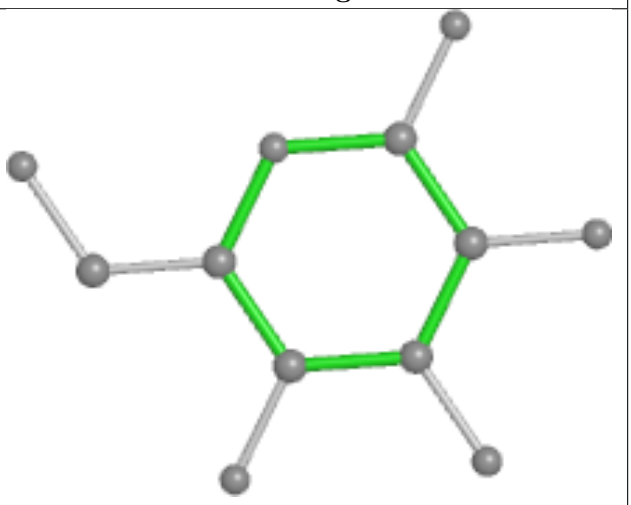
Bond lengths



Bond angles



Torsions



Rings

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	984/993 (99%)	0.02	34 (3%)	44	47	15, 31, 52, 96	1 (0%)
1	B	984/993 (99%)	0.00	21 (2%)	63	67	16, 31, 52, 82	0
All	All	1968/1986 (99%)	0.01	55 (2%)	53	55	15, 31, 52, 96	1 (0%)

The worst 5 of 55 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	802	THR	8.0
1	A	496[A]	TYR	7.3
1	B	496[A]	TYR	5.7
1	A	796	ILE	5.7
1	B	796	ILE	5.1

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

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

6.3 Carbohydrates

There are no monosaccharides in this entry.

6.4 Ligands

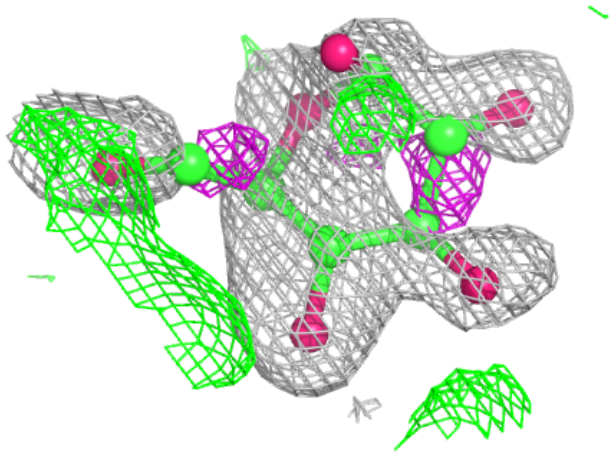
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9'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

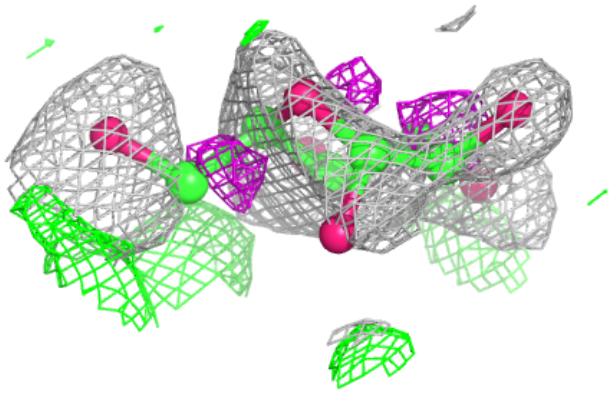
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
3	BGC	B	1002	12/12	0.64	0.29	39,52,60,65	0
3	BGC	A	1002	12/12	0.69	0.27	41,51,58,59	0
4	GOL	A	1018	6/6	0.75	0.15	54,71,79,85	0
6	PEG	A	1013	7/7	0.78	0.21	56,70,84,84	0
4	GOL	A	1003	6/6	0.82	0.10	47,56,67,81	0
4	GOL	B	1018	6/6	0.83	0.15	61,73,85,86	0
5	ACT	B	1015	4/4	0.84	0.11	41,48,49,52	0
5	ACT	B	1005	4/4	0.84	0.18	43,51,52,52	0
4	GOL	A	1017	6/6	0.85	0.18	46,61,75,75	0
4	GOL	B	1006	6/6	0.85	0.10	52,62,72,75	0
4	GOL	B	1019	6/6	0.86	0.11	62,75,79,79	0
6	PEG	A	1011	7/7	0.87	0.14	39,56,70,70	0
5	ACT	A	1012	4/4	0.88	0.11	58,64,74,74	0
4	GOL	B	1012	6/6	0.89	0.09	49,62,74,74	0
4	GOL	B	1011	6/6	0.89	0.11	51,63,76,76	0
5	ACT	B	1007	4/4	0.90	0.13	33,39,40,40	0
5	ACT	A	1006	4/4	0.90	0.14	48,56,57,57	0
6	PEG	B	1013	7/7	0.90	0.15	30,44,59,59	0
5	ACT	A	1004	4/4	0.91	0.08	46,56,57,59	0
5	ACT	B	1008	4/4	0.91	0.10	47,52,56,56	0
4	GOL	B	1014	6/6	0.91	0.10	35,49,64,70	0
5	ACT	A	1010	4/4	0.91	0.11	42,51,52,53	0
4	GOL	A	1016	6/6	0.91	0.14	31,46,58,64	0
4	GOL	B	1010	6/6	0.91	0.15	34,50,64,64	0
5	ACT	A	1007	4/4	0.92	0.13	46,55,56,57	0
5	ACT	A	1005	4/4	0.94	0.12	38,50,61,61	0
4	GOL	A	1009	6/6	0.94	0.09	33,40,48,53	0
5	ACT	B	1003	4/4	0.95	0.11	25,30,46,51	0
4	GOL	B	1004	6/6	0.95	0.09	26,39,52,55	0
7	CL	B	1017	1/1	0.97	0.05	28,28,28,28	0
2	PO4	A	1001	5/5	0.98	0.08	29,29,33,36	0
4	GOL	A	1008	6/6	0.98	0.09	24,37,44,50	0
4	GOL	B	1009	6/6	0.98	0.08	23,35,44,44	0
7	CL	A	1015	1/1	0.98	0.04	29,29,29,29	0
2	PO4	B	1001	5/5	0.98	0.07	26,30,32,35	0
7	CL	B	1016	1/1	1.00	0.07	18,18,18,18	0
7	CL	A	1014	1/1	1.00	0.08	18,18,18,18	0

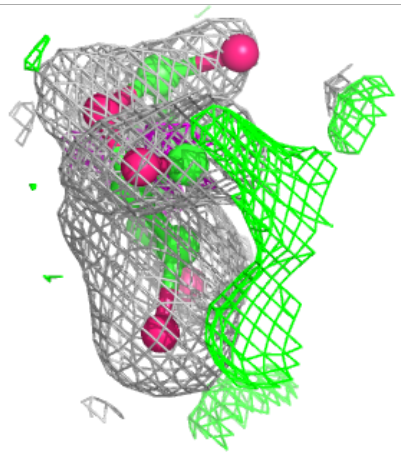
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation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

Electron density around BGC B 1002:

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

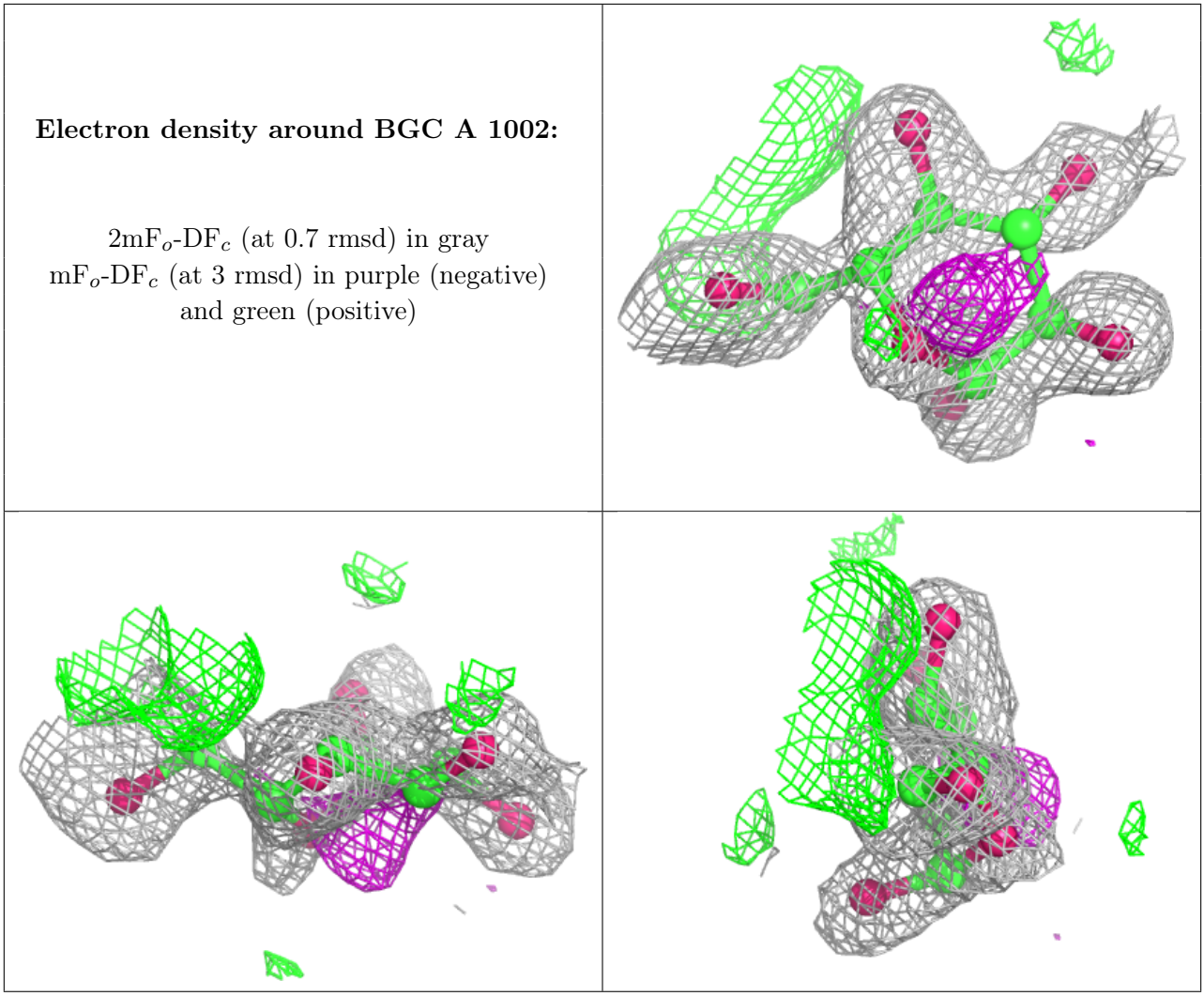






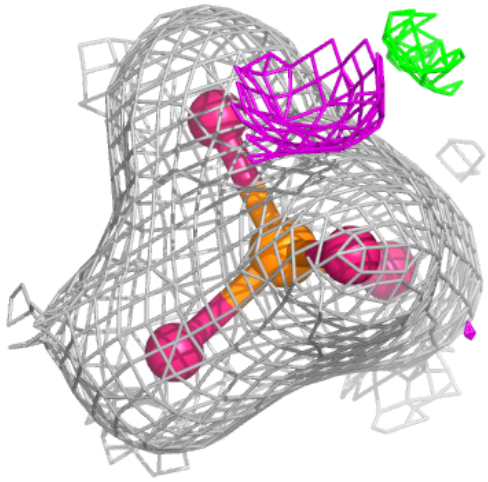
Electron density around BGC A 1002:

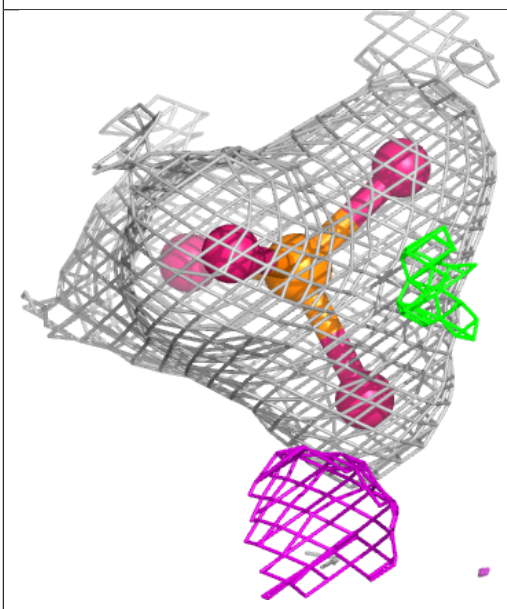
$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)

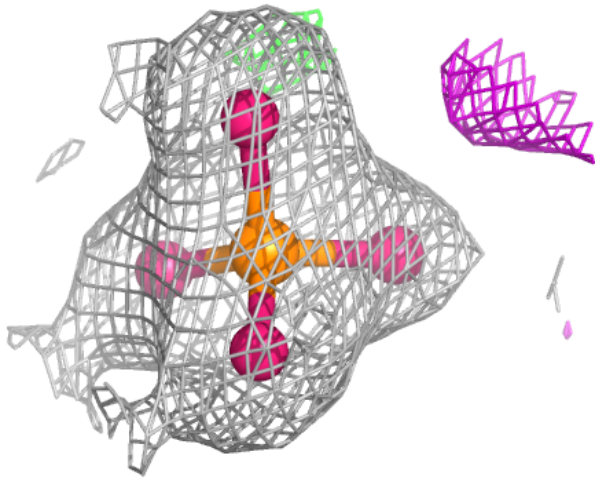


Electron density around PO4 A 1001:

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

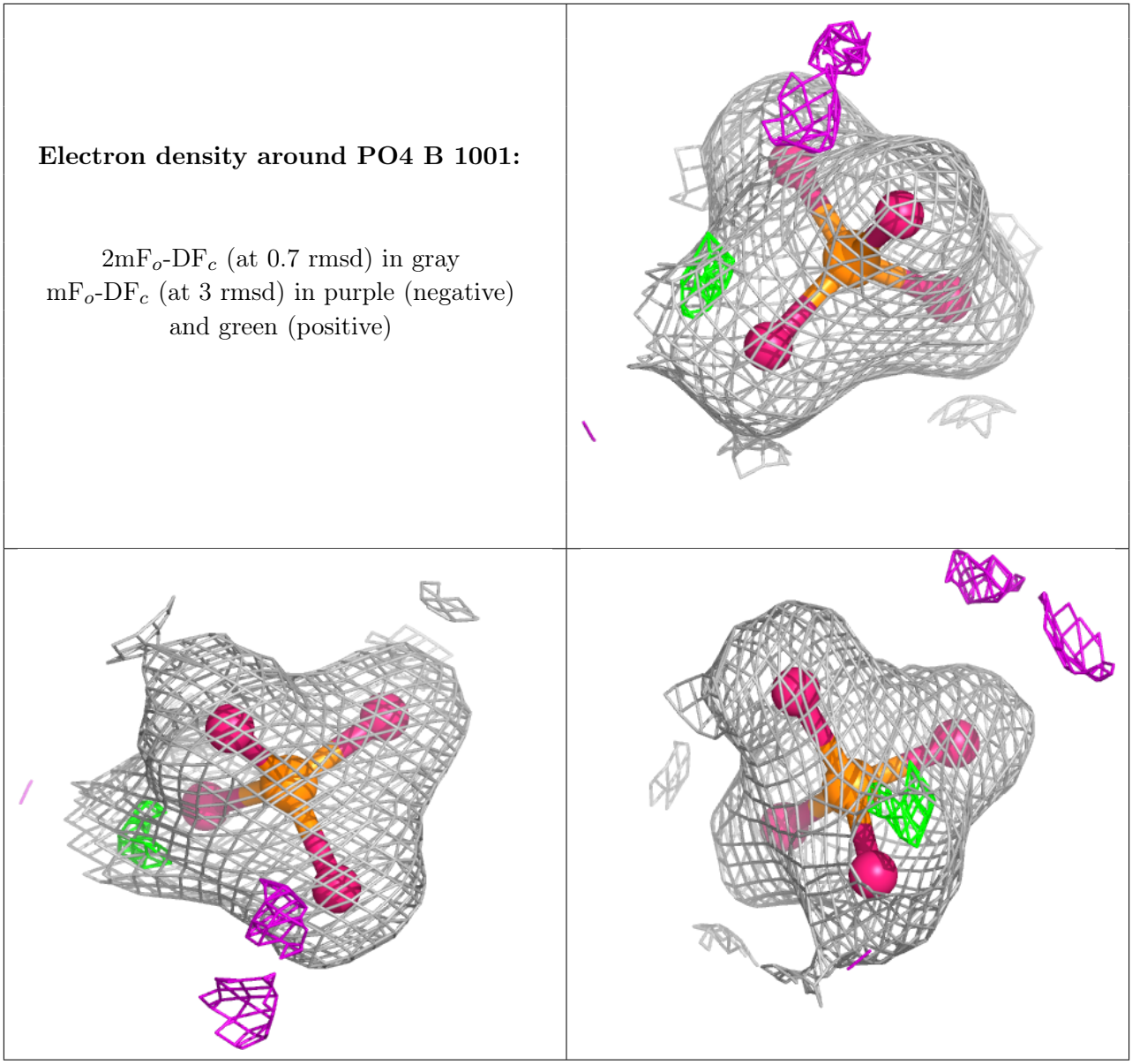






Electron density around PO4 B 1001:

$2mF_o$-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.