wwPDB X-ray Structure Validation Summary Report (i)

Nov 14, 2023 – 05:59 PM JST

PDB ID	:	5ZZO
Title	:	Crystal structure of CcpE regulatory domain in complex with citrate from Staphyloccocus aureus
Authors	:	Chen, J.; Wang, L.; Shang, F.; Xu, Y.
Deposited on	:	2018-06-04
Resolution	:	2.50 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:

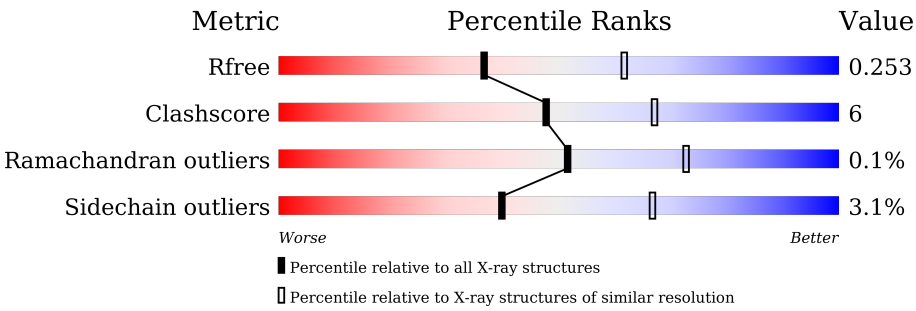
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber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

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.50 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	4661 (2.50-2.50)
Clashscore	141614	5346 (2.50-2.50)
Ramachandran outliers	138981	5231 (2.50-2.50)
Sidechain outliers	138945	5233 (2.50-2.50)

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%

Mol	Chain	Length	Quality of chain
1	A	201	83% 13% •
1	B	201	86% 9% • •
1	C	201	86% 12% •
1	D	201	83% 14% • •
1	E	201	85% 11% •
1	F	201	86% 12% • •

2 Entry composition (i)

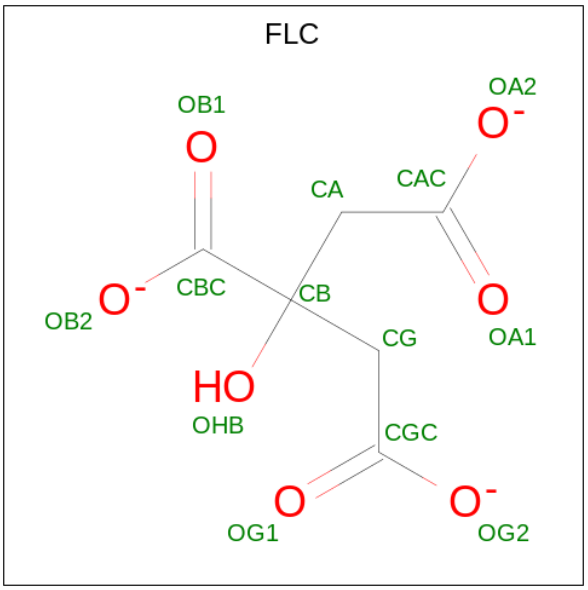
There are 3 unique types of molecules in this entry. The entry contains 19346 atoms, of which 9230 are hydrogens and 0 are deuteriums.

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 LysR family transcriptional regulator.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
1	A	194	3067	995	1516	247	299	10	0	0	0
1	B	194	3064	994	1516	248	296	10	0	0	0
1	C	196	3119	1007	1548	254	300	10	0	0	0
1	D	196	3119	1007	1548	254	300	10	0	0	0
1	E	195	3069	996	1517	248	298	10	0	0	0
1	F	199	3149	1018	1555	258	308	10	0	0	0

- Molecule 2 is CITRATE ANION (three-letter code: FLC) (formula: $C_6H_5O_7$) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	H	O		
2	A	1	18	6	5	7	0	0
2	B	1	18	6	5	7	0	0
2	C	1	18	6	5	7	0	0
2	D	1	18	6	5	7	0	0
2	E	1	18	6	5	7	0	0
2	F	1	18	6	5	7	0	0

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
3	A	117	117	117	0	0
3	B	108	108	108	0	0
3	C	114	114	114	0	0
3	D	101	101	101	0	0
3	E	99	99	99	0	0
3	F	112	112	112	0	0

3 Residue-property plots (i)

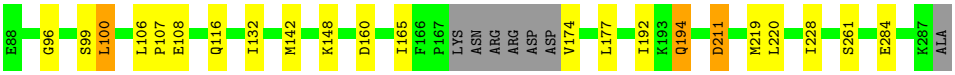
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry. Residues are color-coded 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. 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: LysR family transcriptional regulator

Chain A:

• Molecule 1: LysR family transcriptional regulator

Chain B:



• Molecule 1: LysR family transcriptional regulator

Chain C:

• Molecule 1: LysR family transcriptional regulator

Chain D:

• Molecule 1: LysR family transcriptional regulator

Chain E:

• Molecule 1: LysR family transcriptional regulator

Chain F:

E88
G96
L100
L106
N114
A115
Q116
Q125
T129
E130
K133
M142
P167
LYS
ASN
R170
R171
D172
D173
V174
T175
K176
Q183
P186
I189
I192
Y196
L200
D211
I228
M266
E284
A288

4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	104.79Å 111.66Å 152.97Å 90.00° 95.14° 90.00°	Depositor
Resolution (Å)	31.89 – 2.50 33.05 – 1.95	Depositor EDS
% Data completeness (in resolution range)	93.2 (31.89-2.50) 92.1 (33.05-1.95)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.15	Depositor
$< I/\sigma(I) >$ [1]	2.42 (at 1.95Å)	Xtriage
Refinement program	PHENIX 1.8.4_1496	Depositor
R, R_{free}	0.198 , 0.249 0.206 , 0.253	Depositor DCC
R_{free} test set	2000 reflections (1.70%)	wwPDB-VP
Wilson B-factor (Å^2)	28.5	Xtriage
Anisotropy	0.042	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.45 , 62.5	EDS
L-test for twinning[2]	$< \|L\| > = 0.47$, $< L^2 > = 0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	19346	wwPDB-VP
Average B, all atoms (Å^2)	45.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 96.45 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 5.3624e-10. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

[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 (i)

5.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: FLC

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	0.69	1/1584 (0.1%)	0.69	0/2152
1	B	0.67	0/1581	0.74	0/2148
1	C	0.67	0/1604	0.71	0/2177
1	D	0.68	0/1604	0.71	0/2177
1	E	0.68	1/1585 (0.1%)	0.71	0/2154
1	F	0.65	0/1627	0.70	0/2209
All	All	0.67	2/9585 (0.0%)	0.71	0/13017

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	121	GLU	CG-CD	6.00	1.60	1.51
1	E	211	ASP	CB-CG	-5.12	1.41	1.51

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts (i)

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

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	1551	1516	1516	19	0
1	B	1548	1516	1516	16	0
1	C	1571	1548	1547	18	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	D	1571	1548	1547	16	1
1	E	1552	1517	1517	23	0
1	F	1594	1555	1555	22	0
2	A	13	5	5	0	0
2	B	13	5	5	0	0
2	C	13	5	5	1	0
2	D	13	5	5	1	0
2	E	13	5	5	1	0
2	F	13	5	5	0	0
3	A	117	0	0	7	1
3	B	108	0	0	8	1
3	C	114	0	0	11	0
3	D	101	0	0	5	0
3	E	99	0	0	11	1
3	F	112	0	0	14	1
All	All	10116	9230	9228	114	3

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 6.

The worst 5 of 114 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:F:116:GLN:NE2	3:F:402:HOH:O	1.94	0.98
1:C:116:GLN:NE2	3:C:401:HOH:O	1.99	0.96
1:B:284:GLU:O	3:B:401:HOH:O	1.85	0.95
1:F:284:GLU:O	3:F:401:HOH:O	1.89	0.90
1:C:142:MET:SD	3:C:500:HOH:O	2.31	0.88

All (3) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:B:401:HOH:O	3:F:401:HOH:O[3_445]	1.80	0.40
3:A:401:HOH:O	3:E:401:HOH:O[3_545]	1.81	0.39
1:D:281:SER:OG	1:D:281:SER:OG[2_755]	2.14	0.06

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	190/201 (94%)	184 (97%)	6 (3%)	0	100	100
1	B	190/201 (94%)	182 (96%)	8 (4%)	0	100	100
1	C	192/201 (96%)	184 (96%)	8 (4%)	0	100	100
1	D	192/201 (96%)	186 (97%)	6 (3%)	0	100	100
1	E	191/201 (95%)	183 (96%)	8 (4%)	0	100	100
1	F	195/201 (97%)	188 (96%)	6 (3%)	1 (0%)	29	48
All	All	1150/1206 (95%)	1107 (96%)	42 (4%)	1 (0%)	51	73

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	F	171	ARG

5.3.2 Protein sidechains (i)

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

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

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	178/186 (96%)	172 (97%)	6 (3%)	37	63
1	B	177/186 (95%)	171 (97%)	6 (3%)	37	63
1	C	180/186 (97%)	174 (97%)	6 (3%)	38	64
1	D	180/186 (97%)	174 (97%)	6 (3%)	38	64
1	E	177/186 (95%)	172 (97%)	5 (3%)	43	70
1	F	182/186 (98%)	178 (98%)	4 (2%)	52	77

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
All	All	1074/1116 (96%)	1041 (97%)	33 (3%)	40 67

5 of 33 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	E	211	ASP
1	F	100	LEU
1	F	176	LYS
1	C	100	LEU
1	B	211	ASP

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

Mol	Chain	Res	Type
1	C	194	GLN
1	E	116	GLN
1	F	116	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)

6 ligands are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the

expected value. A bond length (or angle) with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
2	FLC	C	301	-	12,12,12	1.17	0	17,17,17	1.63	3 (17%)
2	FLC	F	301	-	12,12,12	1.17	0	17,17,17	1.97	4 (23%)
2	FLC	D	301	-	12,12,12	1.26	1 (8%)	17,17,17	1.83	4 (23%)
2	FLC	A	301	-	12,12,12	1.11	0	17,17,17	1.71	2 (11%)
2	FLC	E	301	-	12,12,12	1.30	0	17,17,17	1.58	2 (11%)
2	FLC	B	301	-	12,12,12	1.08	0	17,17,17	1.70	3 (17%)

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
2	FLC	C	301	-	-	0/16/16/16	-
2	FLC	F	301	-	-	3/16/16/16	-
2	FLC	D	301	-	-	2/16/16/16	-
2	FLC	A	301	-	-	2/16/16/16	-
2	FLC	E	301	-	-	2/16/16/16	-
2	FLC	B	301	-	-	2/16/16/16	-

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	D	301	FLC	CB-CBC	2.40	1.55	1.53

The worst 5 of 18 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	F	301	FLC	OB2-CBC-CB	5.96	123.40	113.05
2	D	301	FLC	OB2-CBC-CB	5.13	121.97	113.05
2	A	301	FLC	OB2-CBC-CB	4.70	121.21	113.05
2	B	301	FLC	OB2-CBC-CB	4.11	120.18	113.05
2	E	301	FLC	OB2-CBC-CB	3.95	119.90	113.05

There are no chirality outliers.

5 of 11 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	E	301	FLC	CB-CA-CAC-OA1
2	E	301	FLC	CB-CA-CAC-OA2
2	B	301	FLC	CG-CB-CBC-OB1
2	B	301	FLC	CG-CB-CBC-OB2
2	F	301	FLC	CG-CB-CBC-OB2

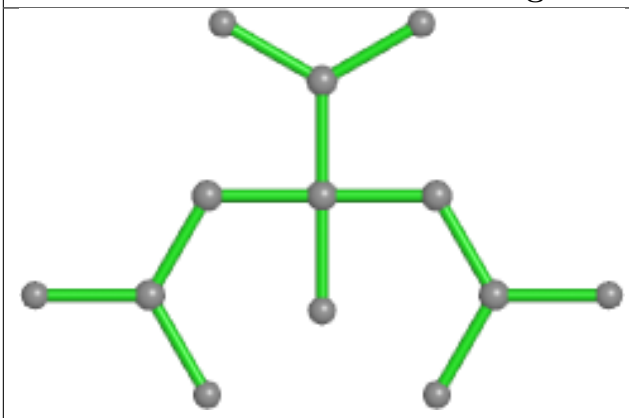
There are no ring outliers.

3 monomers are involved in 3 short contacts:

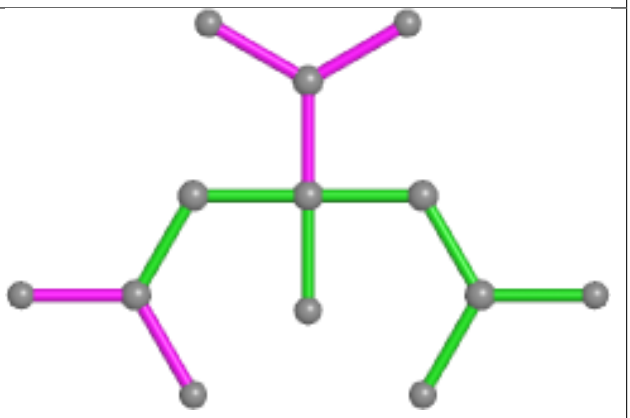
Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	C	301	FLC	1	0
2	D	301	FLC	1	0
2	E	301	FLC	1	0

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.

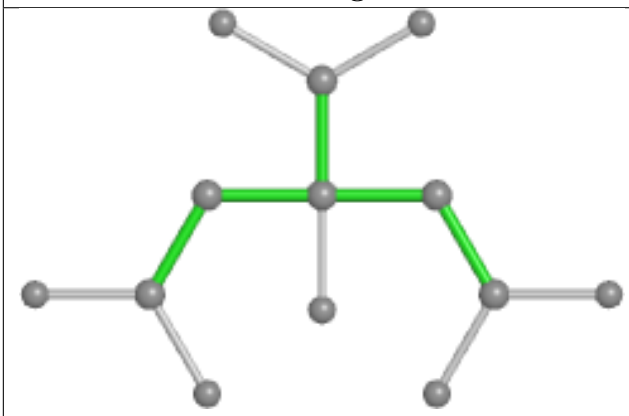
Ligand FLC C 301



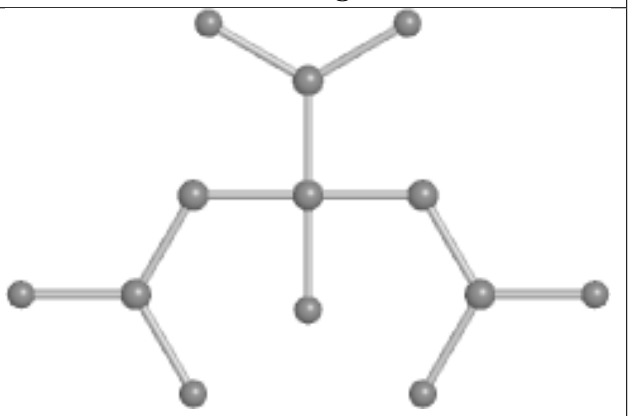
Bond lengths



Bond angles

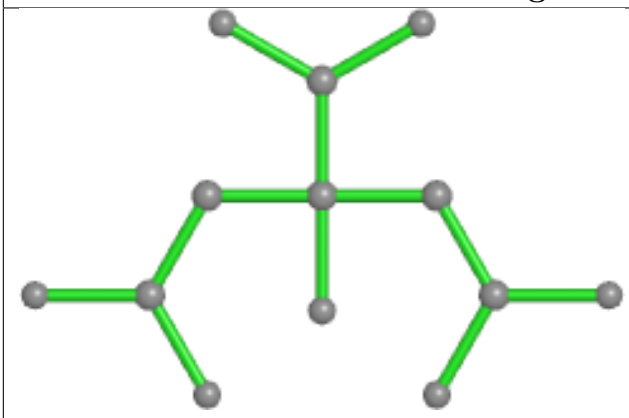


Torsions

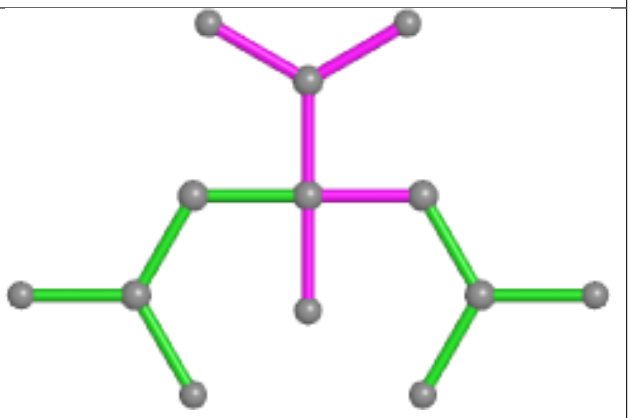


Rings

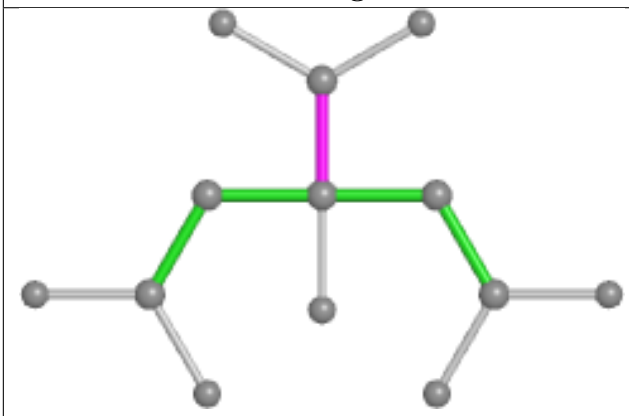
Ligand FLC F 301



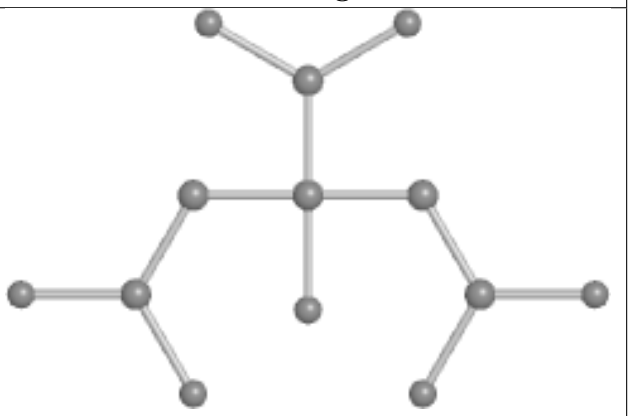
Bond lengths



Bond angles

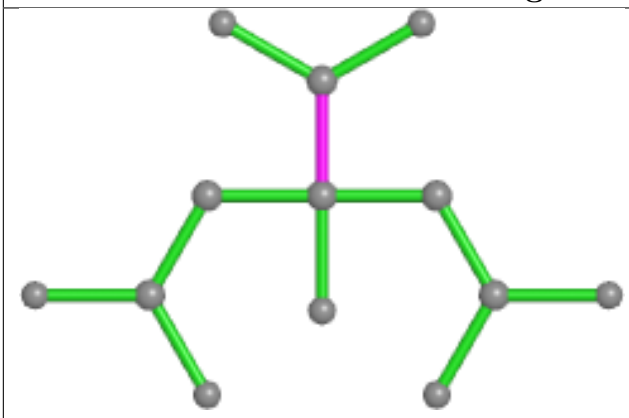


Torsions

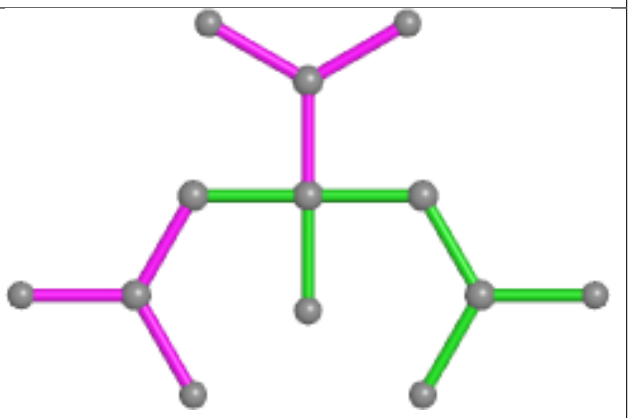


Rings

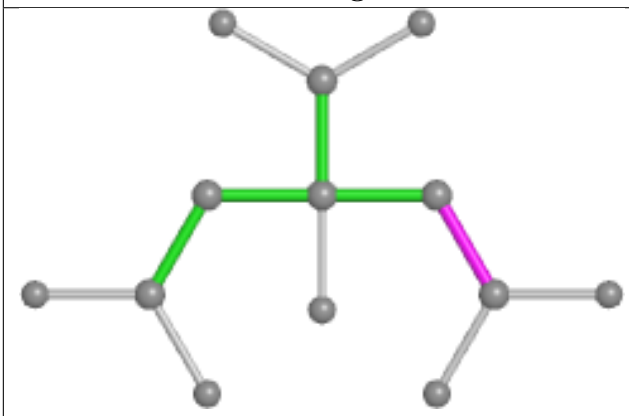
Ligand FLC D 301



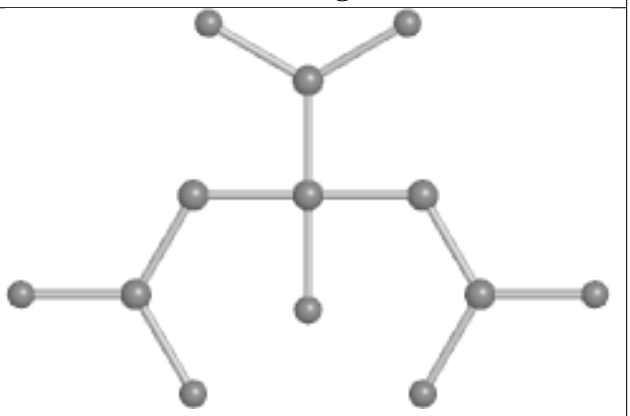
Bond lengths



Bond angles

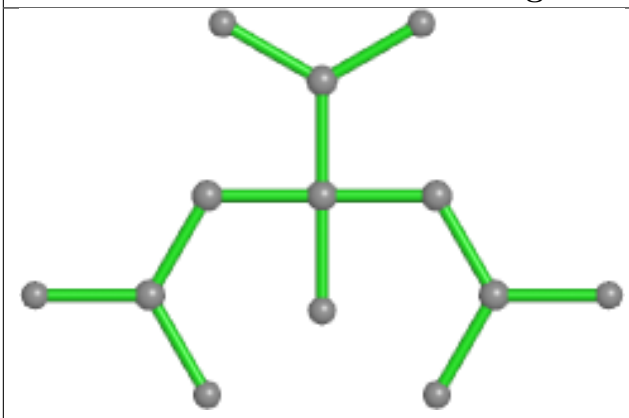


Torsions

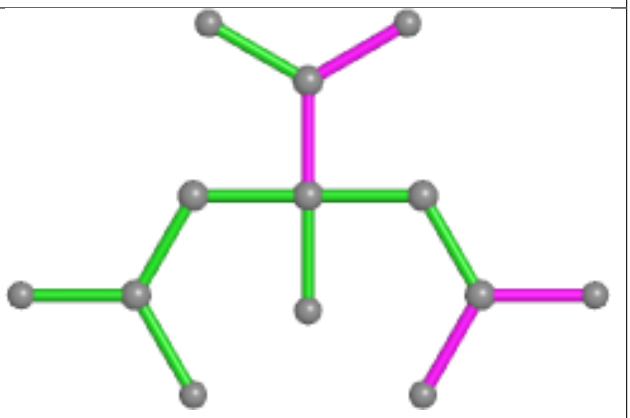


Rings

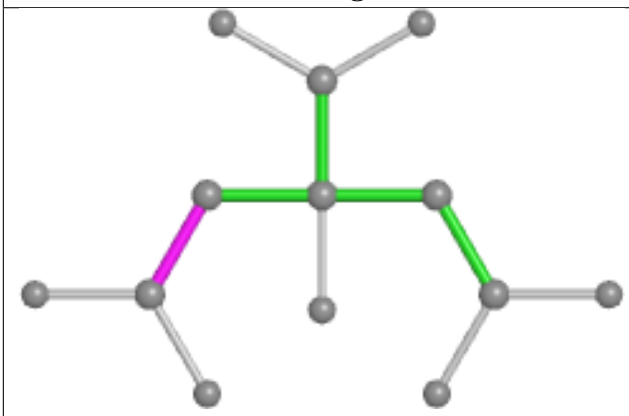
Ligand FLC A 301



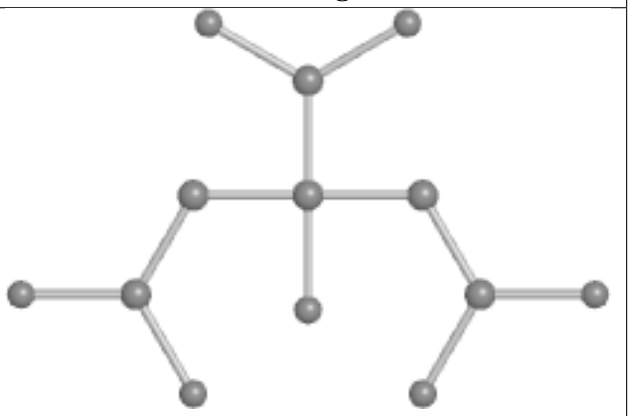
Bond lengths



Bond angles

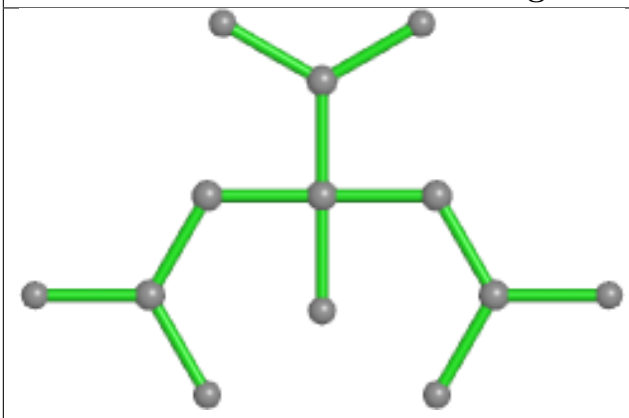


Torsions

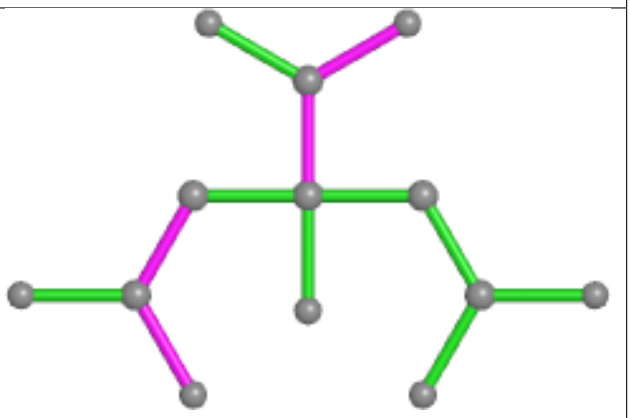


Rings

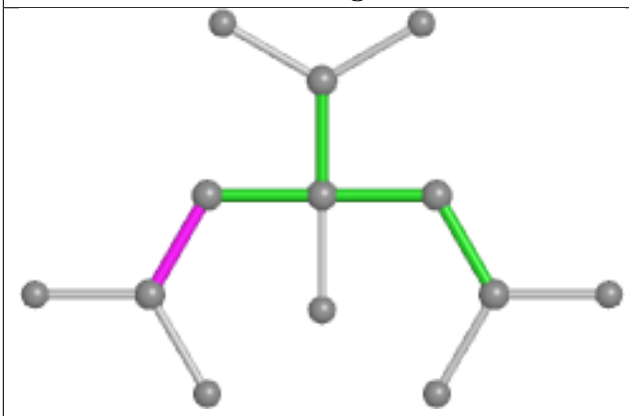
Ligand FLC E 301



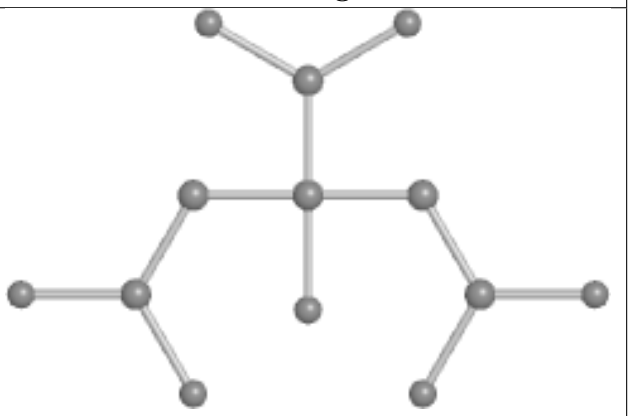
Bond lengths



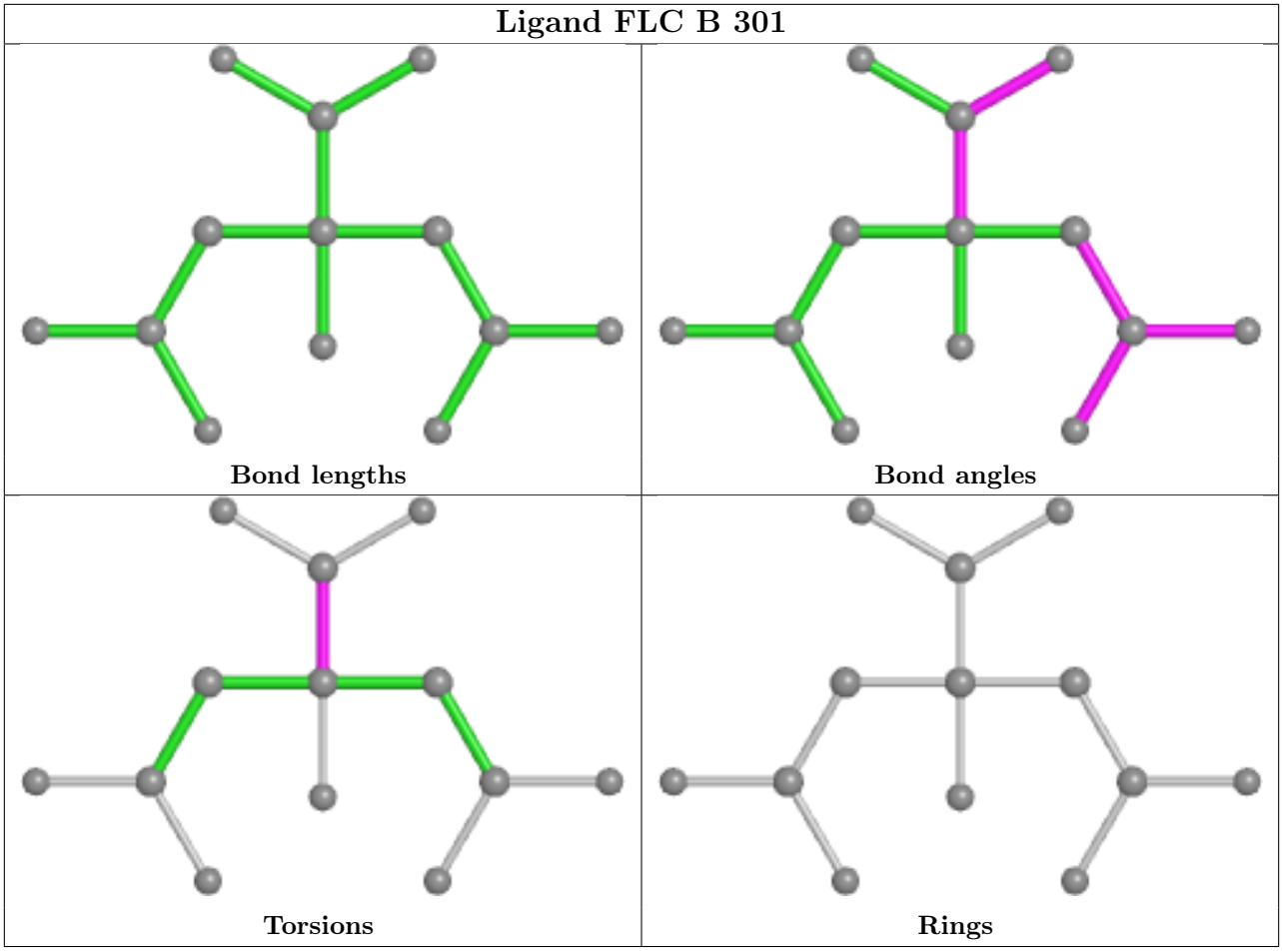
Bond angles



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 (i)

6.1 Protein, DNA and RNA chains (i)

Unable to reproduce the depositors R factor - this section is therefore empty.

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

Unable to reproduce the depositors R factor - this section is therefore empty.

6.3 Carbohydrates (i)

Unable to reproduce the depositors R factor - this section is therefore empty.

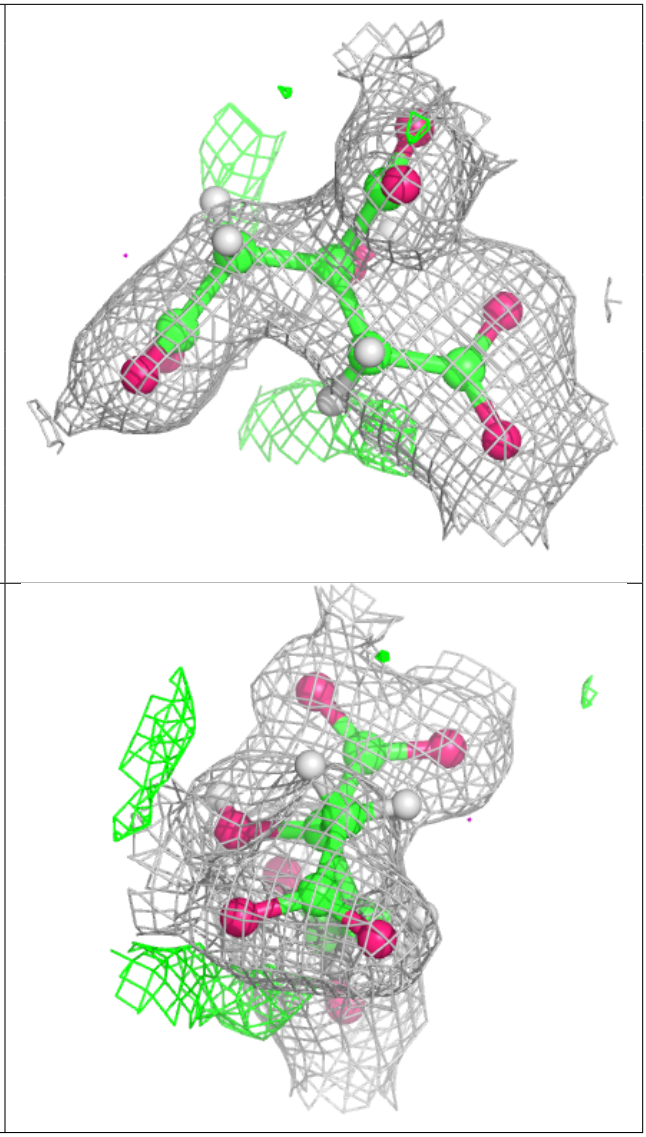
6.4 Ligands (i)

Unable to reproduce the depositors R factor - this section is therefore empty.

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.

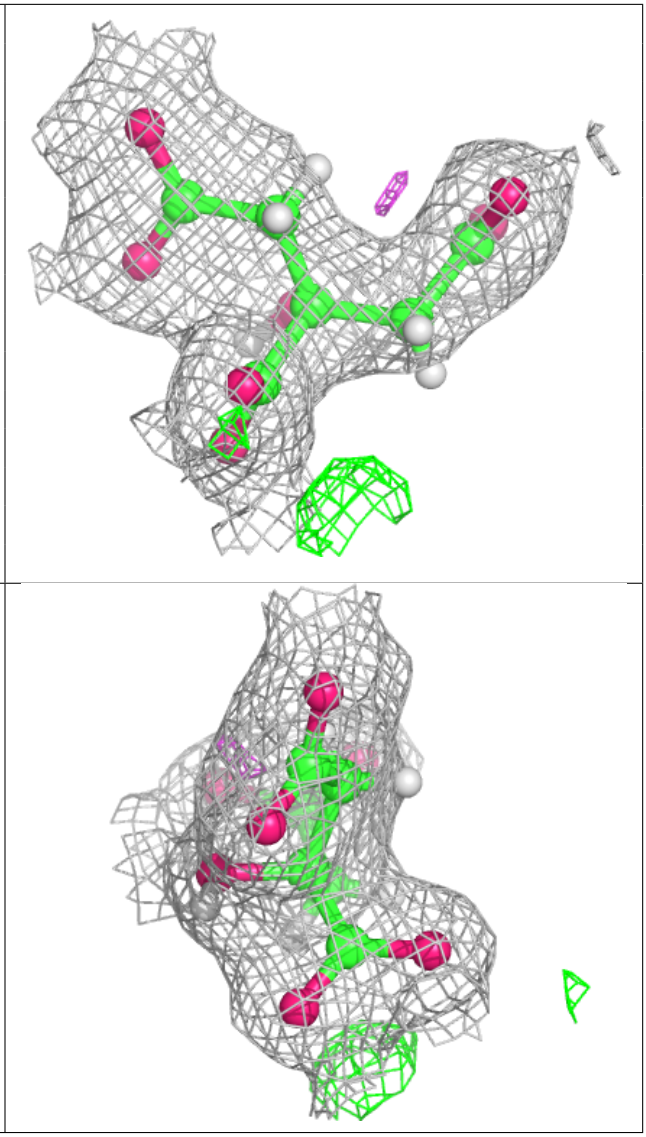
Electron density around FLC A 301:

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



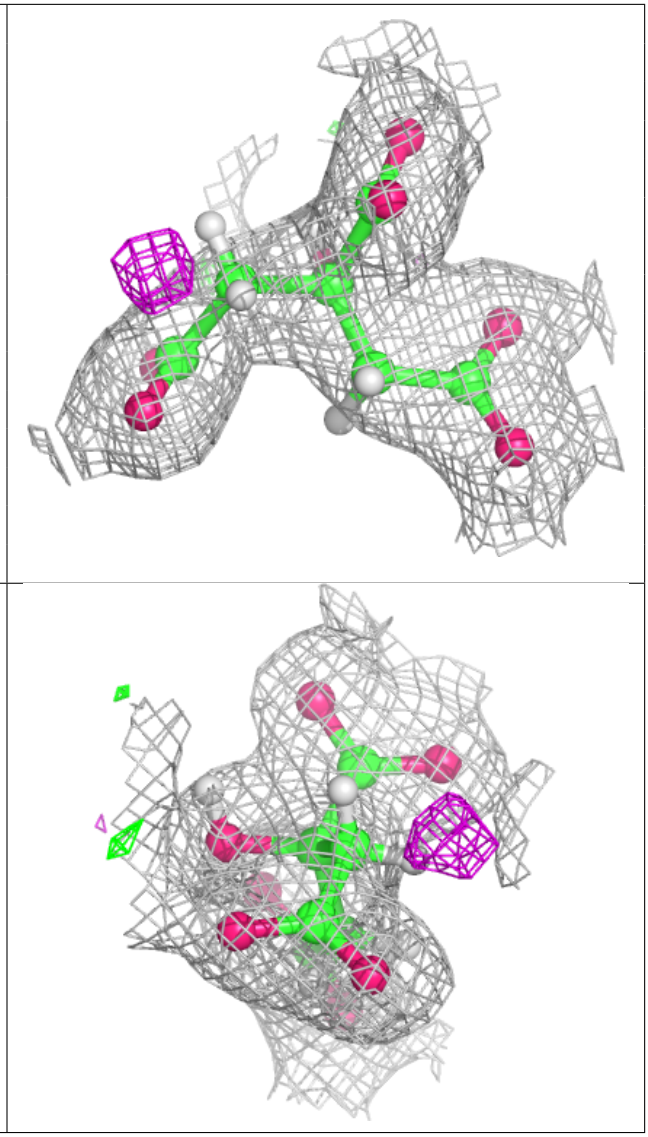
Electron density around FLC B 301:

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



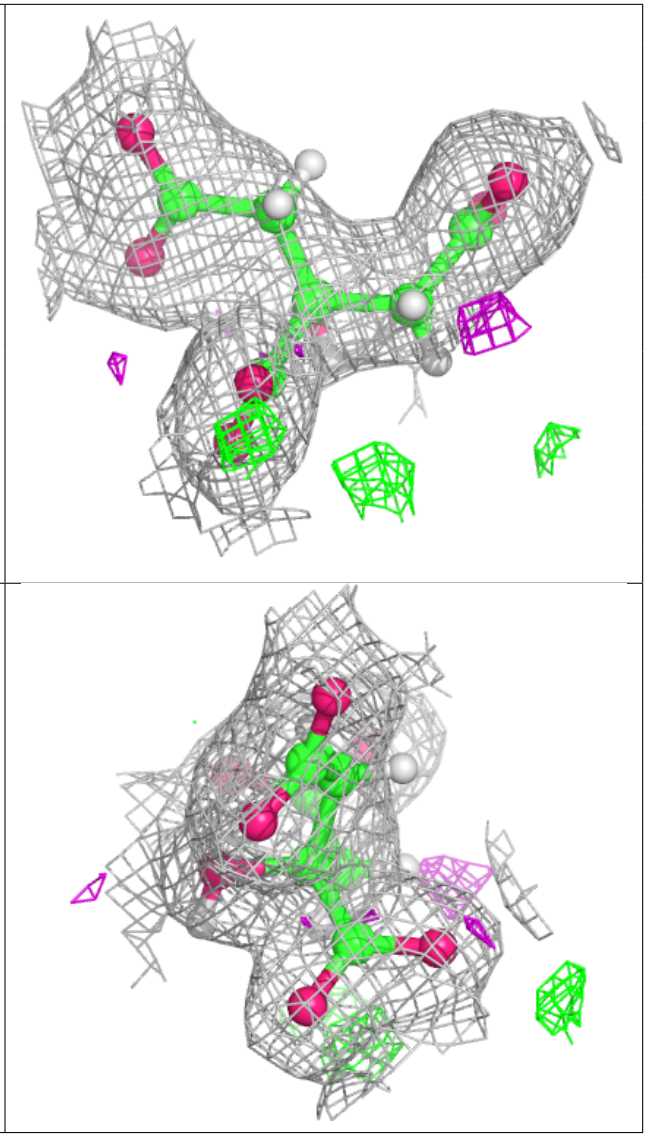
Electron density around FLC C 301:

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



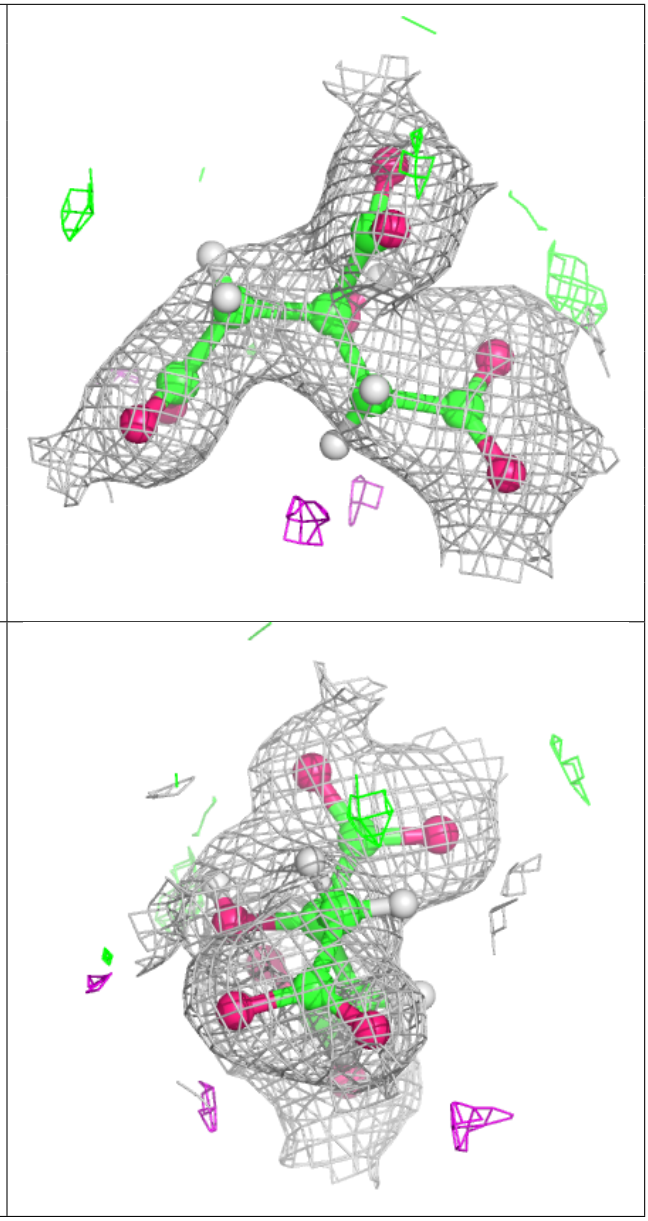
Electron density around FLC D 301:

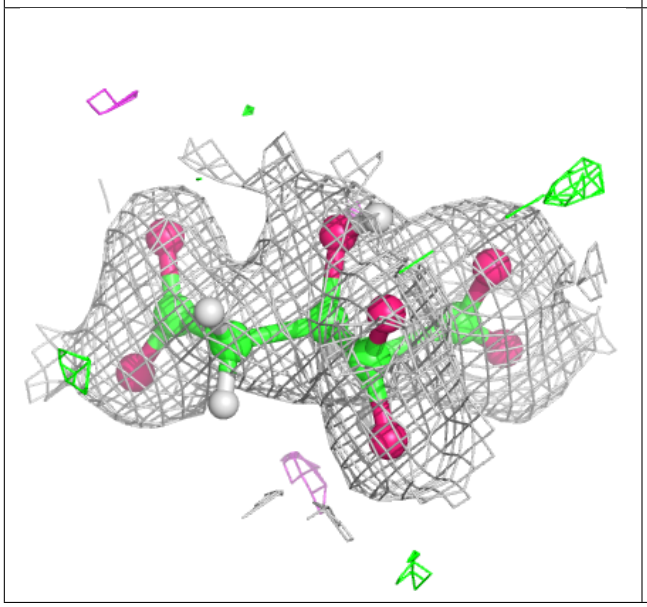
$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)



Electron density around FLC E 301:

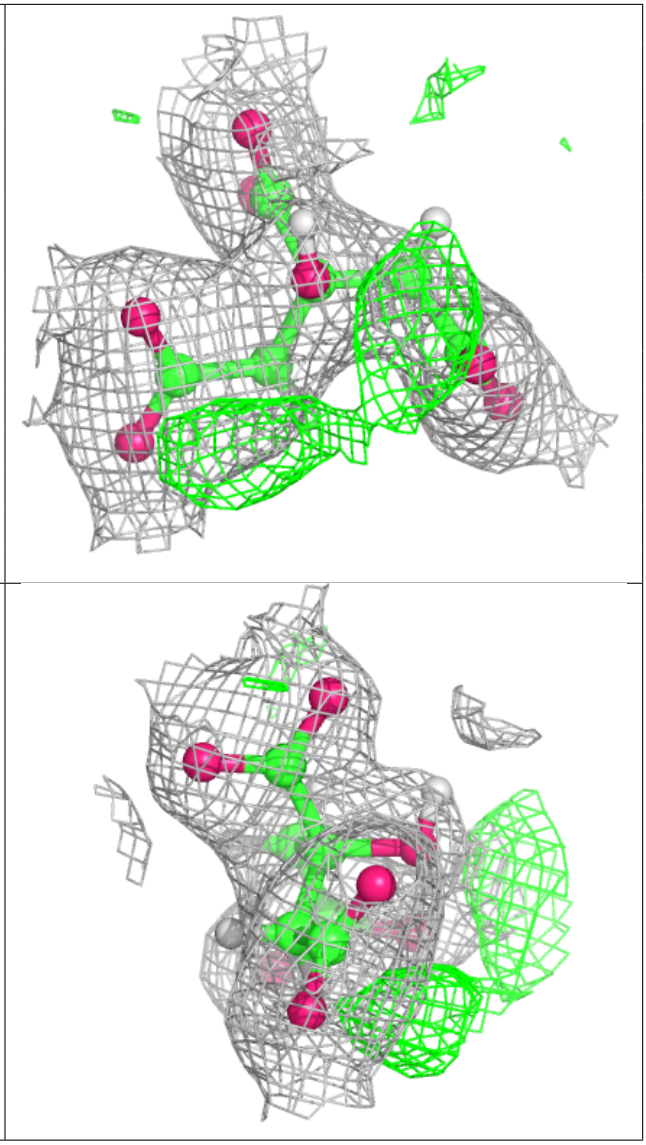
$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)





Electron density around FLC F 301:

$2mF_o$-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)

Unable to reproduce the depositors R factor - this section is therefore empty.