



wwPDB X-ray Structure Validation Summary Report ⓘ

May 13, 2020 – 06:42 am BST

PDB ID : 5Z7H
Title : Crystal structure of CcpE regulatory domain in citrate-bound form from *Staphylococcus aureus*
Authors : Chen, J.; Wang, L.; Shang, F.; Xu, Y.
Deposited on : 2018-01-28
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 ⓘ symbol.

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

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

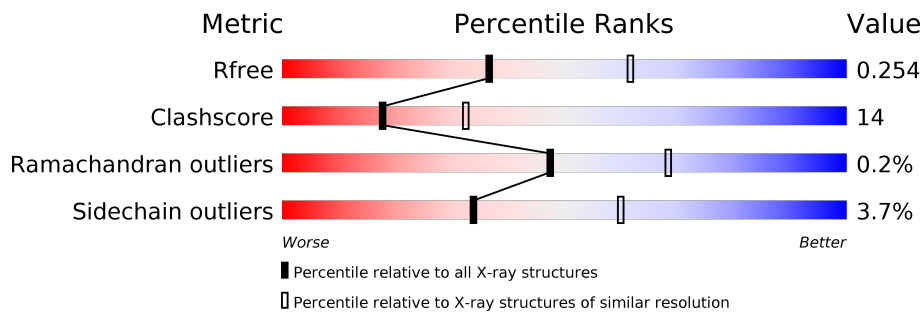
1 Overall quality at a glance

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 on the lower bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions $\leq 5\%$

Mol	Chain	Length	Quality of chain
1	A	288	
1	B	288	
1	C	288	
1	D	288	
1	E	288	
1	F	288	

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 crite-

ria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
2	FLC	A	301	-	-	X	-
2	FLC	D	301	-	-	X	-

2 Entry composition [i](#)

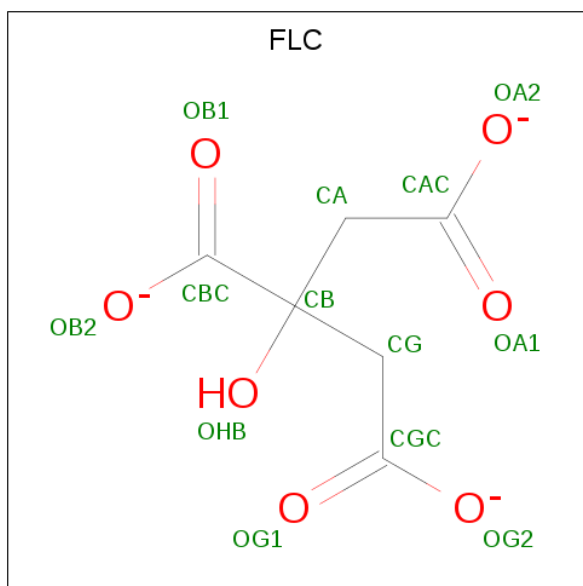
There are 3 unique types of molecules in this entry. The entry contains 19674 atoms, of which 9070 are hydrogens and 0 are deuteriums.

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	193	Total 3037	C 988	H 1498	N 246	O 295	S 10	0	0	0
1	B	196	Total 3078	C 1001	H 1516	N 254	O 297	S 10	0	0	0
1	C	193	Total 3038	C 989	H 1496	N 245	O 298	S 10	0	0	0
1	D	197	Total 3116	C 1010	H 1538	N 257	O 301	S 10	0	0	0
1	E	194	Total 3031	C 988	H 1492	N 246	O 295	S 10	0	0	0
1	F	194	Total 3047	C 992	H 1500	N 246	O 299	S 10	0	0	0

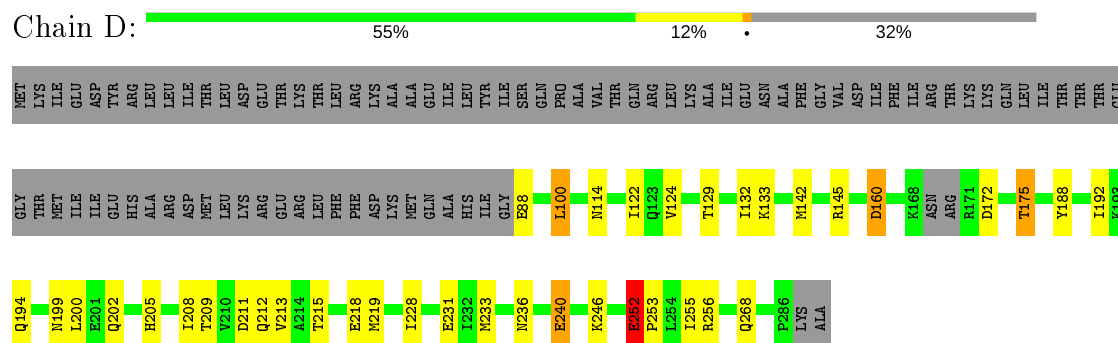
- Molecule 2 is CITRATE ANION (three-letter code: FLC) (formula: $C_6H_5O_7$) (labeled as "Ligand of Interest" by author).



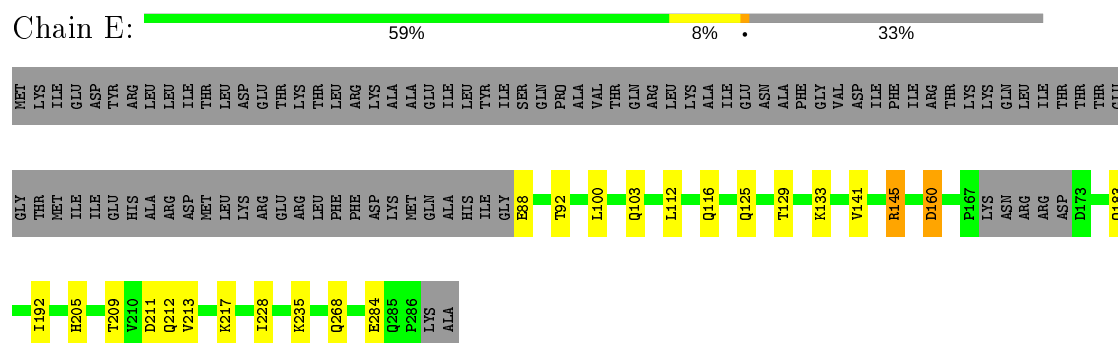
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total	C	H	O	0	0
			18	6	5	7		
2	B	1	Total	C	H	O	0	0
			18	6	5	7		
2	C	1	Total	C	H	O	0	0
			18	6	5	7		
2	D	1	Total	C	H	O	0	0
			18	6	5	7		
2	E	1	Total	C	H	O	0	0
			18	6	5	7		
2	F	1	Total	C	H	O	0	0
			18	6	5	7		

- Molecule 3 is water.

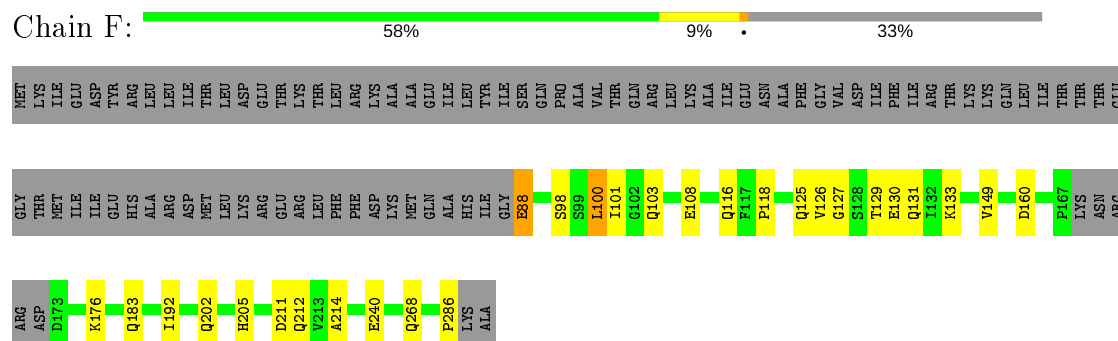
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	206	Total	O	0	0
			206	206		
3	B	207	Total	O	0	0
			207	207		
3	C	230	Total	O	0	0
			230	230		
3	D	182	Total	O	0	0
			182	182		
3	E	190	Total	O	0	0
			190	190		
3	F	204	Total	O	0	0
			204	204		



- Molecule 1: LysR family transcriptional regulator



- Molecule 1: LysR family transcriptional regulator



4 Data and refinement statistics

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
$\langle I/\sigma(I) \rangle$ ¹	2.42 (at 1.95Å)	Xtrriage
Refinement program	PHENIX 1.8.4_1496	Depositor
R, R_{free}	0.196 , 0.247 0.203 , 0.254	Depositor DCC
R_{free} test set	2000 reflections (1.70%)	wwPDB-VP
Wilson B-factor (Å ²)	28.5	Xtrriage
Anisotropy	0.042	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.46 , 74.2	EDS
L-test for twinning ²	$\langle L \rangle = 0.47$, $\langle L^2 \rangle = 0.30$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.96	EDS
Total number of atoms	19674	wwPDB-VP
Average B, all atoms (Å ²)	43.0	wwPDB-VP

Xtrriage'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.*

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 [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	0.66	1/1572 (0.1%)	0.75	3/2137 (0.1%)
1	B	0.64	0/1594	0.73	1/2164 (0.0%)
1	C	0.69	0/1575	0.70	0/2141
1	D	0.74	3/1611 (0.2%)	0.71	0/2187
1	E	0.65	0/1572	0.70	1/2138 (0.0%)
1	F	0.65	0/1580	0.68	0/2148
All	All	0.67	4/9504 (0.0%)	0.71	5/12915 (0.0%)

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	D	252	GLU	CD-OE2	6.67	1.32	1.25
1	D	252	GLU	CD-OE1	6.64	1.32	1.25
1	D	240	GLU	CD-OE2	5.38	1.31	1.25
1	A	211	ASP	CB-CG	-5.05	1.41	1.51

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	145	ARG	NE-CZ-NH2	-5.58	117.51	120.30
1	E	145	ARG	NE-CZ-NH2	-5.54	117.53	120.30
1	A	211	ASP	CB-CG-OD2	-5.15	113.67	118.30
1	B	145	ARG	NE-CZ-NH2	-5.11	117.75	120.30
1	A	211	ASP	CB-CA-C	-5.05	100.30	110.40

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	1539	1498	1503	48	0
1	B	1562	1516	1524	57	0
1	C	1542	1496	1503	44	1
1	D	1578	1538	1546	39	0
1	E	1539	1492	1497	30	0
1	F	1547	1500	1505	33	0
2	A	13	5	5	5	0
2	B	13	5	5	3	0
2	C	13	5	5	0	0
2	D	13	5	5	6	0
2	E	13	5	5	1	0
2	F	13	5	5	3	0
3	A	206	0	0	39	4
3	B	207	0	0	36	8
3	C	230	0	0	31	7
3	D	182	0	0	25	1
3	E	190	0	0	23	3
3	F	204	0	0	25	6
All	All	10604	9070	9108	252	16

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 14.

The worst 5 of 252 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:B:116:GLN:NE2	3:B:401:HOH:O	1.69	1.19
1:A:209:THR:O	3:A:401:HOH:O	1.66	1.12
1:E:116:GLN:NE2	3:E:401:HOH:O	1.96	0.98
1:F:129:THR:O	3:F:401:HOH:O	1.82	0.97
1:B:251:ASN:CA	1:B:252:GLU:HB2	1.97	0.94

The worst 5 of 16 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:C:582:HOH:O	3:E:573:HOH:O[2_755]	1.88	0.32
3:B:580:HOH:O	3:C:590:HOH:O[1_655]	1.90	0.30
3:B:445:HOH:O	3:C:497:HOH:O[1_655]	1.93	0.27
3:B:605:HOH:O	3:B:607:HOH:O[2_856]	1.94	0.26
3:B:510:HOH:O	3:F:530:HOH:O[4_755]	1.97	0.23

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	189/288 (66%)	184 (97%)	5 (3%)	0	100	100
1	B	190/288 (66%)	183 (96%)	6 (3%)	1 (0%)	29	48
1	C	189/288 (66%)	183 (97%)	5 (3%)	1 (0%)	29	48
1	D	193/288 (67%)	188 (97%)	5 (3%)	0	100	100
1	E	190/288 (66%)	183 (96%)	7 (4%)	0	100	100
1	F	190/288 (66%)	187 (98%)	3 (2%)	0	100	100
All	All	1141/1728 (66%)	1108 (97%)	31 (3%)	2 (0%)	47	68

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	202	GLN
1	B	252	GLU

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	176/263 (67%)	168 (96%)	8 (4%)	27	51
1	B	177/263 (67%)	168 (95%)	9 (5%)	24	45
1	C	177/263 (67%)	171 (97%)	6 (3%)	37	63
1	D	180/263 (68%)	172 (96%)	8 (4%)	28	52
1	E	175/263 (66%)	174 (99%)	1 (1%)	86	95
1	F	177/263 (67%)	170 (96%)	7 (4%)	31	56
All	All	1062/1578 (67%)	1023 (96%)	39 (4%)	34	60

5 of 39 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	88	GLU
1	C	211	ASP
1	F	160	ASP
1	C	126	VAL
1	C	160	ASP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 11 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	212	GLN
1	D	125	GLN
1	E	125	GLN
1	C	125	GLN
1	E	103	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 carbohydrates in this entry.

5.6 Ligand geometry

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	B	301	-	3,12,12	1.70	1 (33%)	3,17,17	1.70	1 (33%)
2	FLC	D	301	-	3,12,12	1.85	1 (33%)	3,17,17	0.83	0
2	FLC	E	301	-	3,12,12	1.65	0	3,17,17	1.39	0
2	FLC	C	301	-	3,12,12	1.26	0	3,17,17	1.79	1 (33%)
2	FLC	F	301	-	3,12,12	1.45	0	3,17,17	1.61	1 (33%)
2	FLC	A	301	-	3,12,12	2.17	2 (66%)	3,17,17	2.40	1 (33%)

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	B	301	-	-	4/6/16/16	-
2	FLC	D	301	-	-	1/6/16/16	-
2	FLC	E	301	-	-	0/6/16/16	-
2	FLC	C	301	-	-	0/6/16/16	-
2	FLC	F	301	-	-	3/6/16/16	-
2	FLC	A	301	-	-	3/6/16/16	-

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	D	301	FLC	CG-CB	-2.82	1.50	1.54
2	A	301	FLC	CA-CB	-2.36	1.51	1.54
2	A	301	FLC	CG-CB	-2.31	1.51	1.54
2	B	301	FLC	CA-CB	-2.31	1.51	1.54

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed($^{\circ}$)	Ideal($^{\circ}$)
2	A	301	FLC	CB-CA-CAC	-3.77	108.94	114.98
2	C	301	FLC	CB-CA-CAC	-2.73	110.62	114.98
2	B	301	FLC	CB-CA-CAC	-2.65	110.74	114.98
2	F	301	FLC	CB-CG-CGC	-2.42	111.11	114.98

There are no chirality outliers.

5 of 11 torsion outliers are listed below:

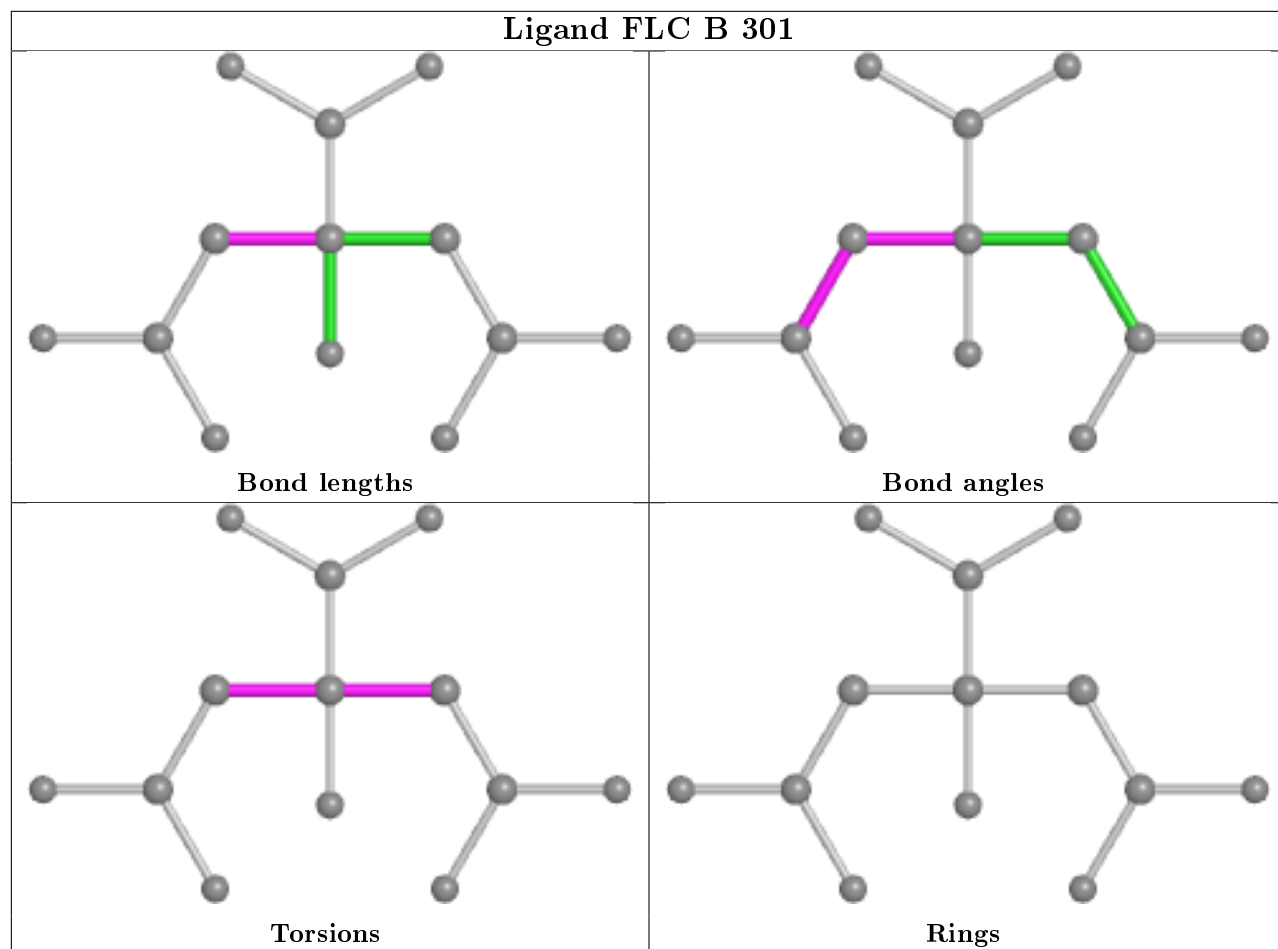
Mol	Chain	Res	Type	Atoms
2	F	301	FLC	CA-CB-CG-CGC
2	F	301	FLC	CBC-CB-CG-CGC
2	F	301	FLC	OHB-CB-CG-CGC
2	B	301	FLC	OHB-CB-CG-CGC
2	B	301	FLC	CAC-CA-CB-CBC

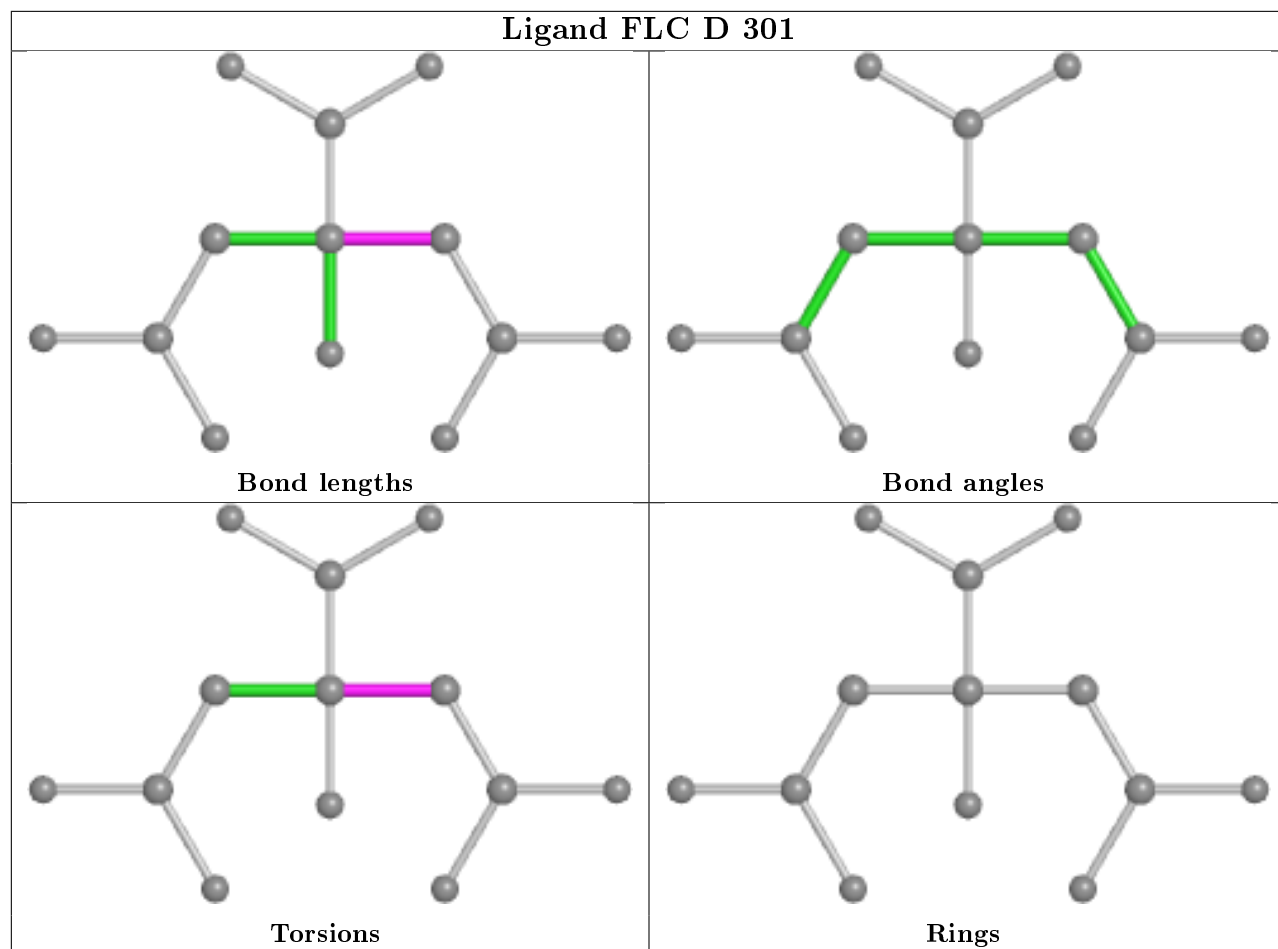
There are no ring outliers.

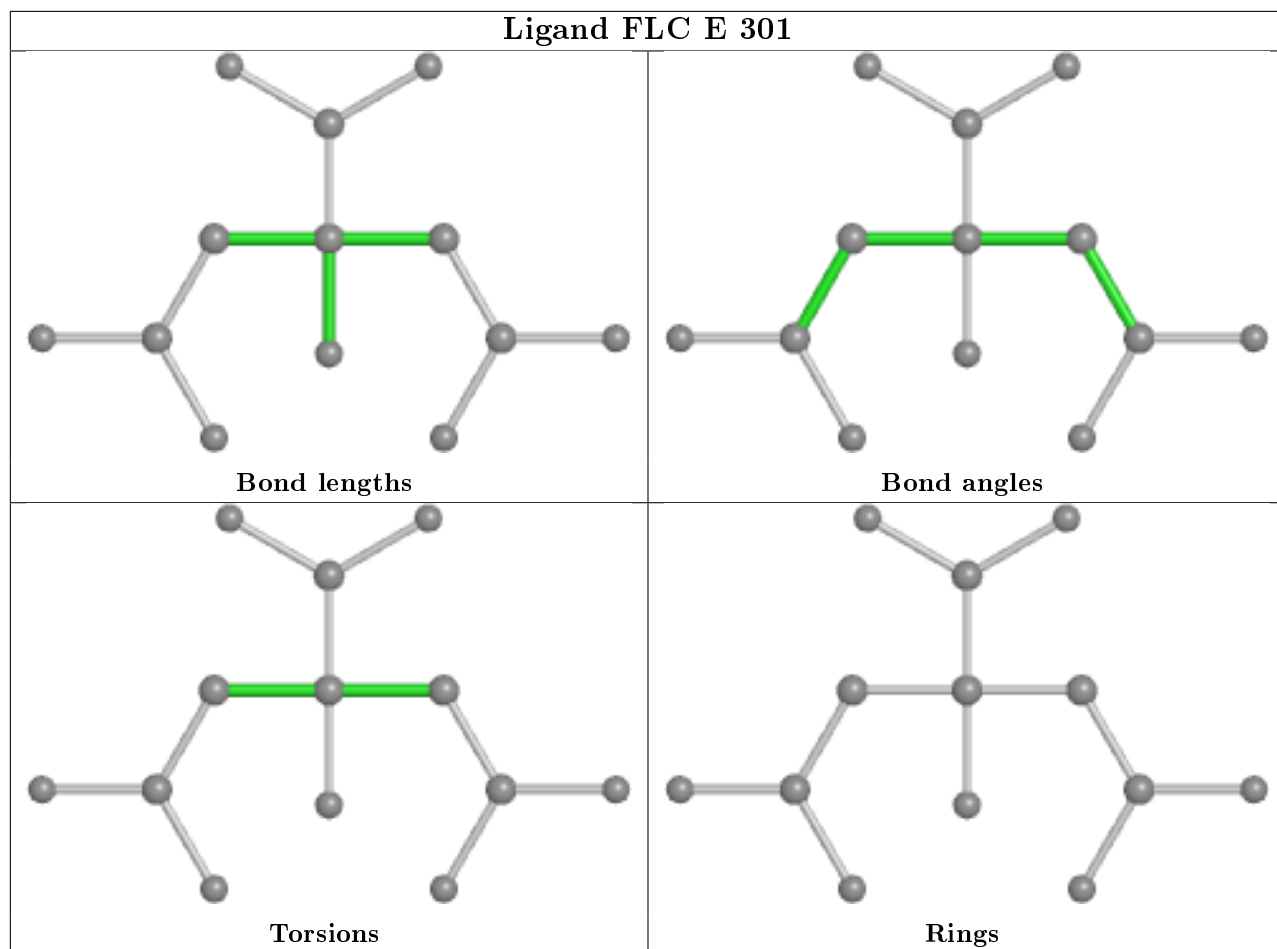
5 monomers are involved in 18 short contacts:

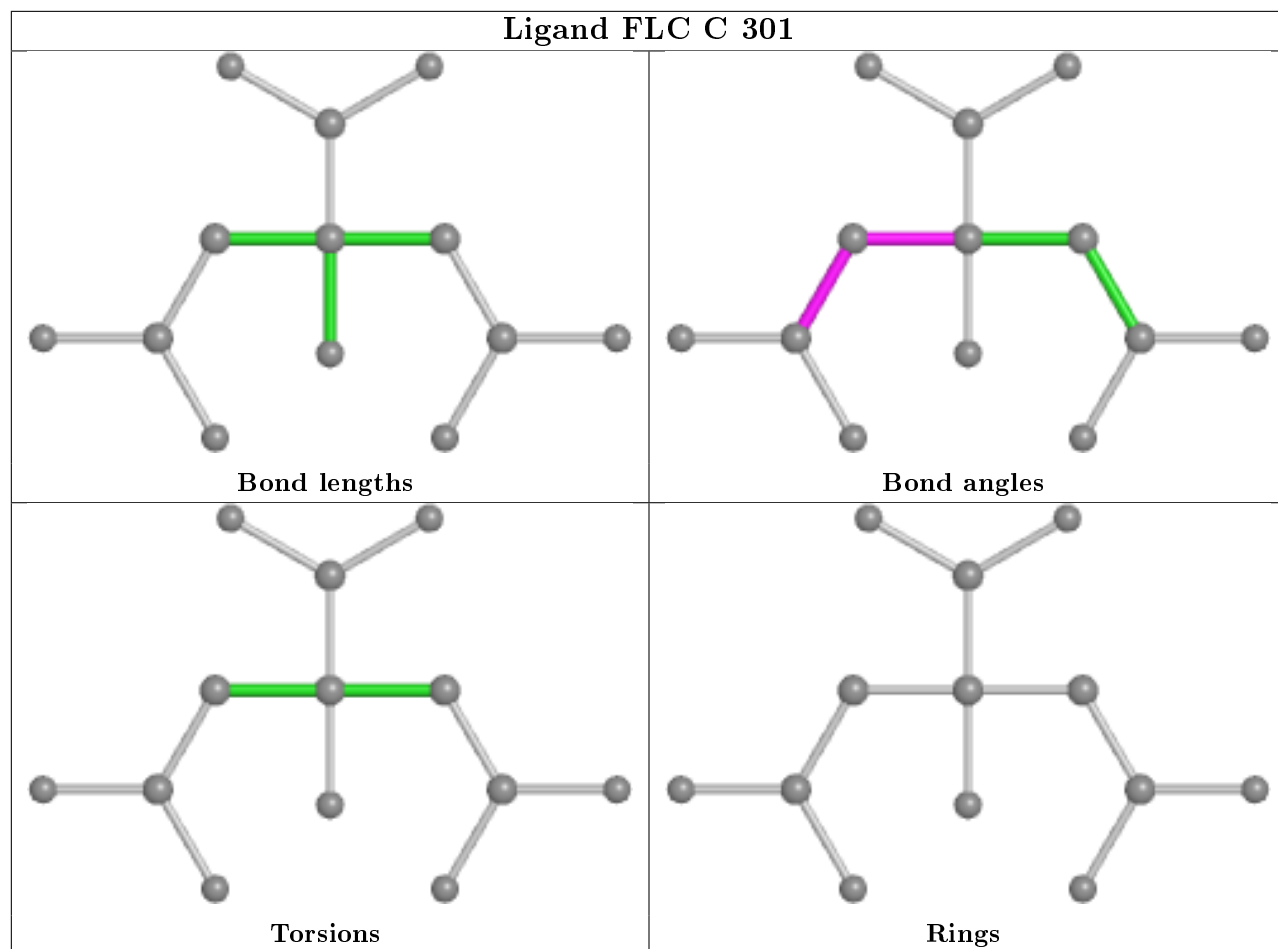
Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	B	301	FLC	3	0
2	D	301	FLC	6	0
2	E	301	FLC	1	0
2	F	301	FLC	3	0
2	A	301	FLC	5	0

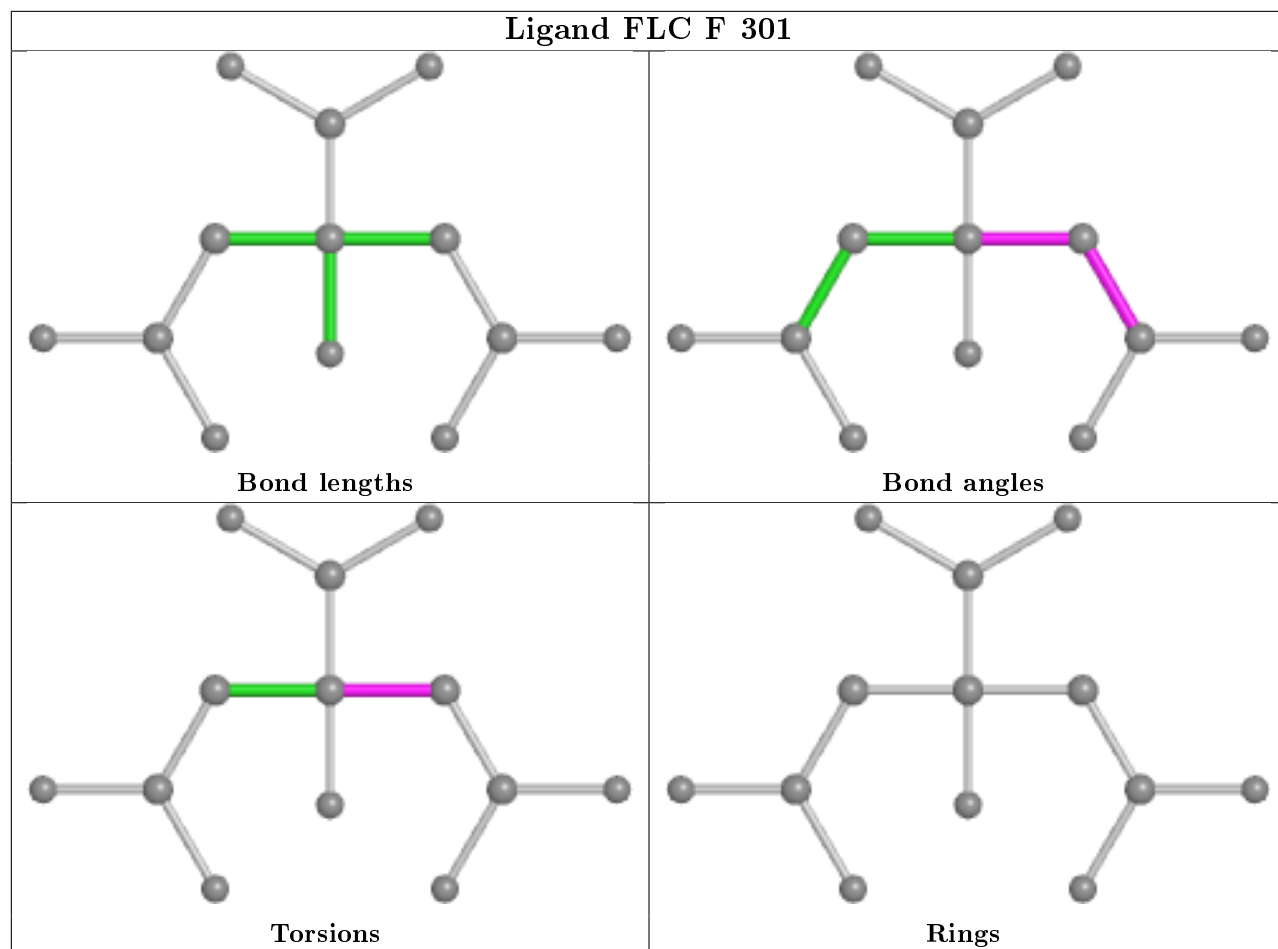
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included.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

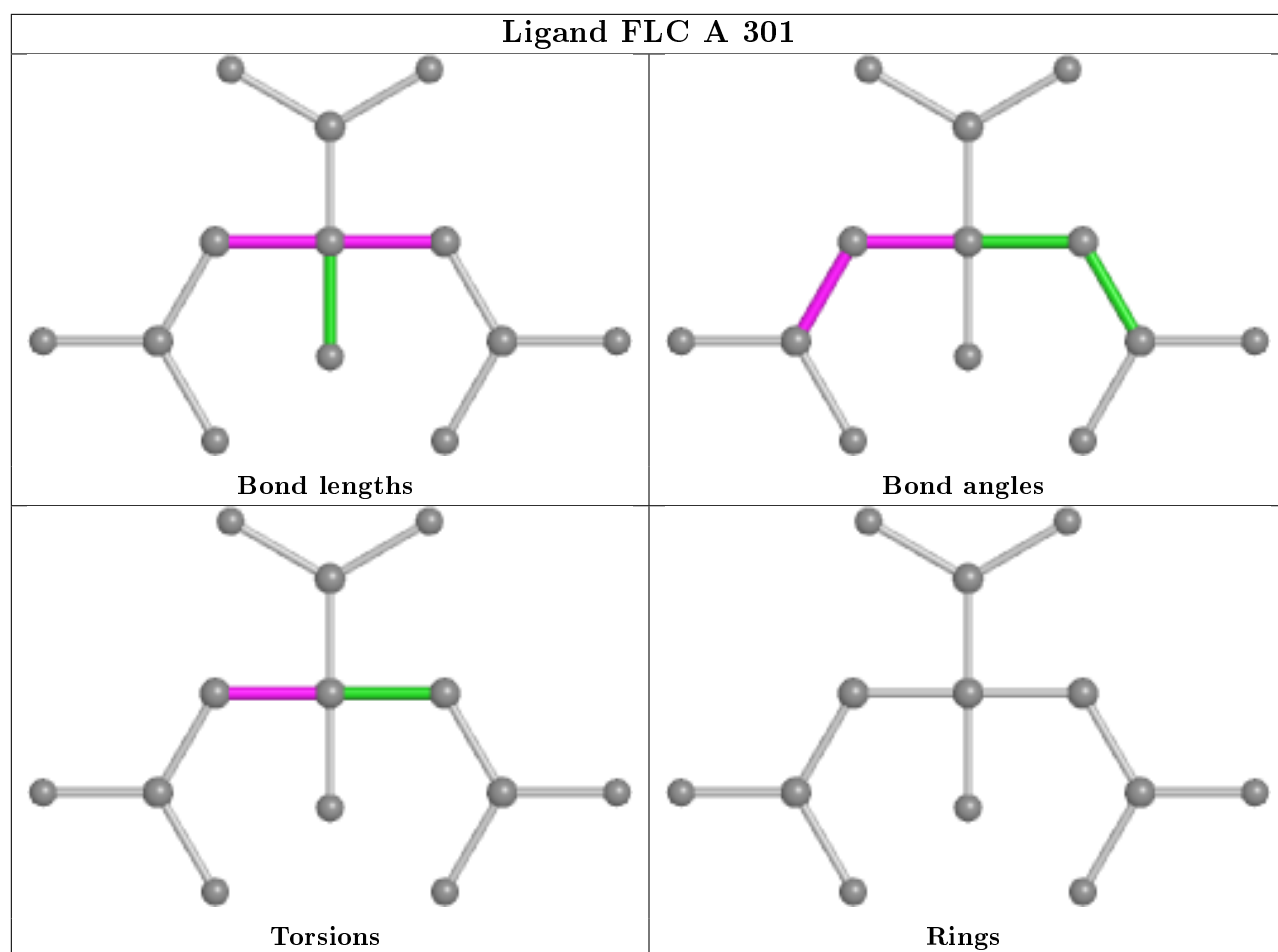












5.7 Other polymers [i](#)

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

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

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

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

6.3 Carbohydrates

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

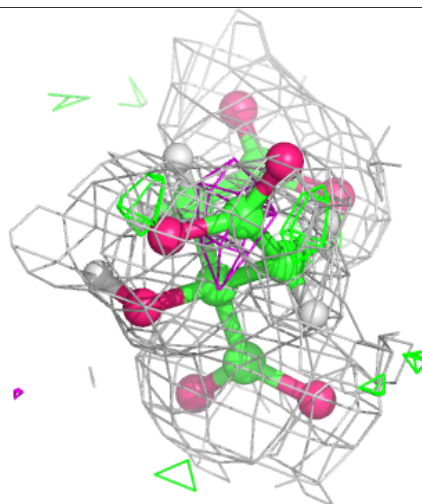
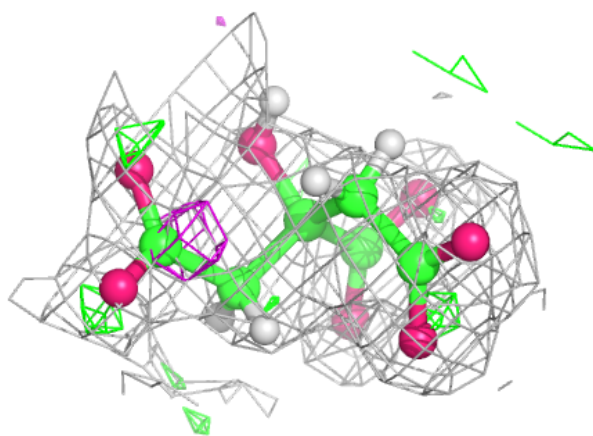
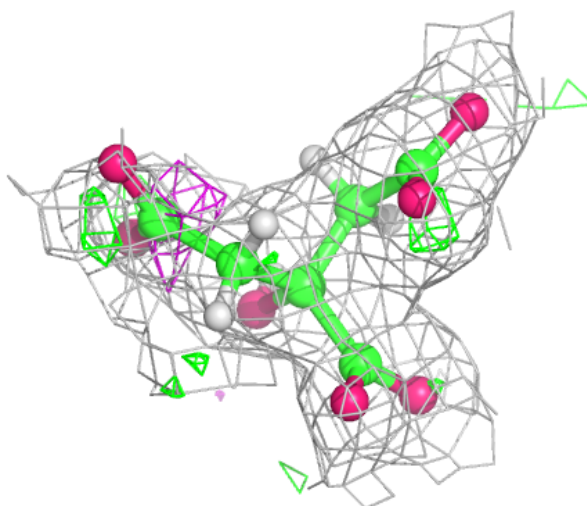
6.4 Ligands

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.

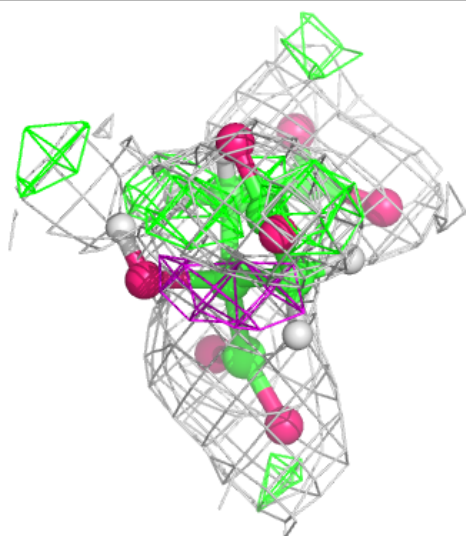
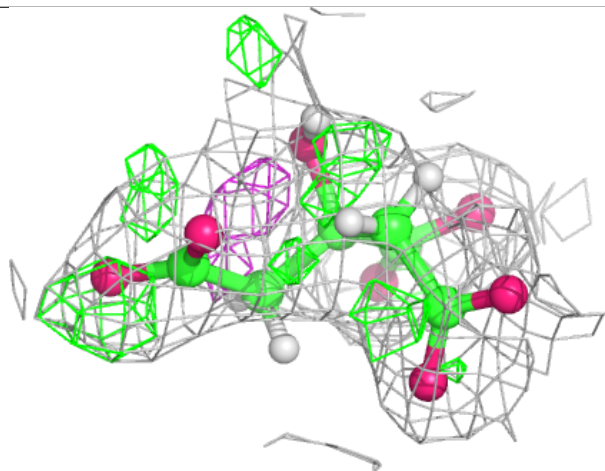
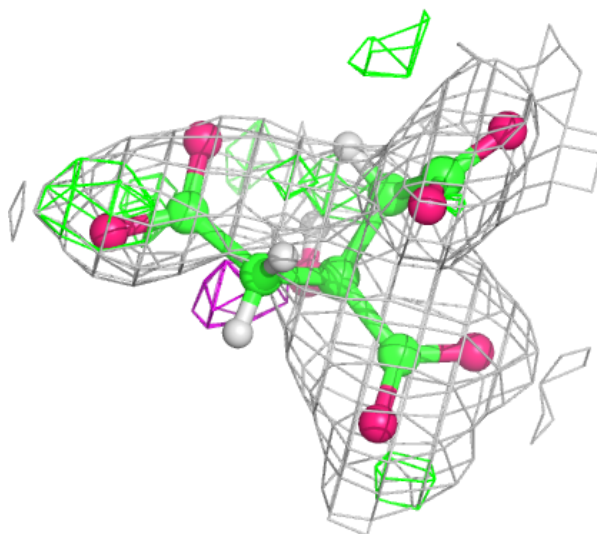
Electron density around FLC B 301:

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



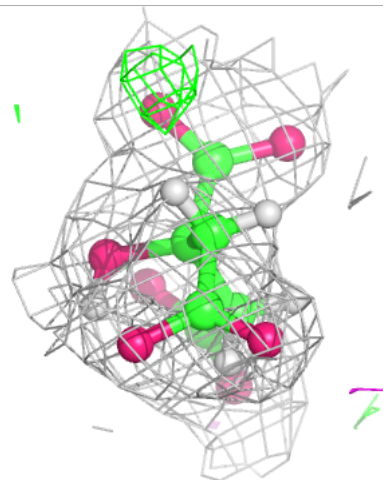
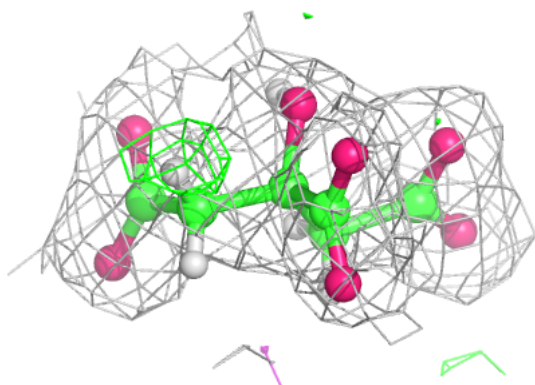
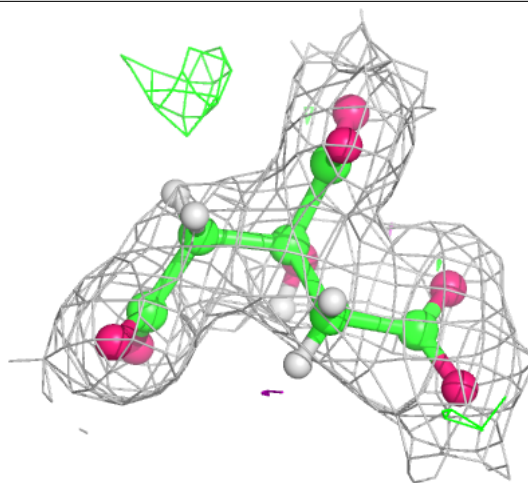
Electron density around FLC D 301:

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



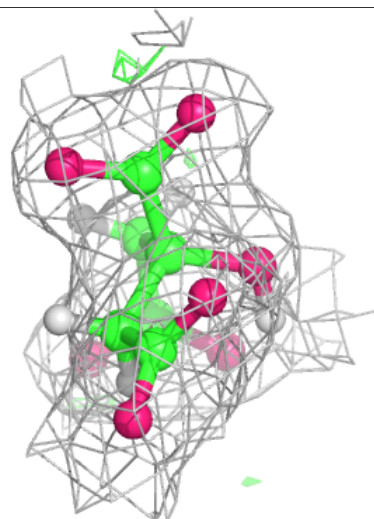
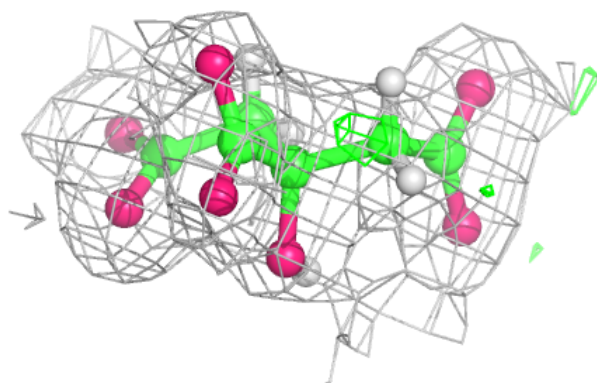
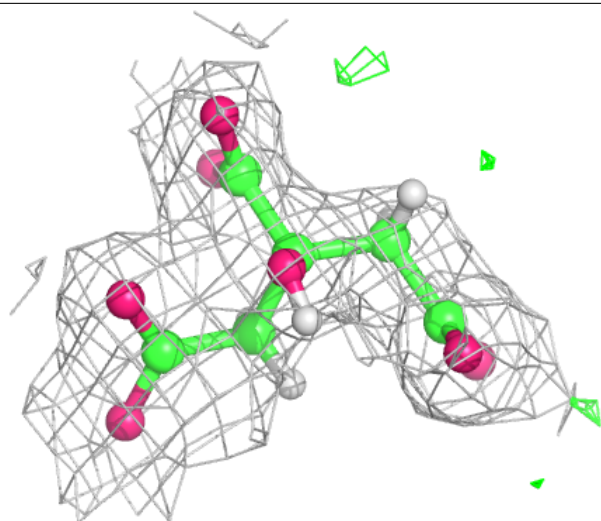
Electron density around FLC E 301:

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



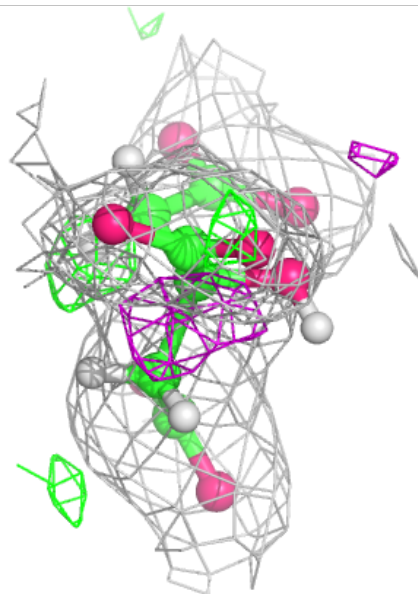
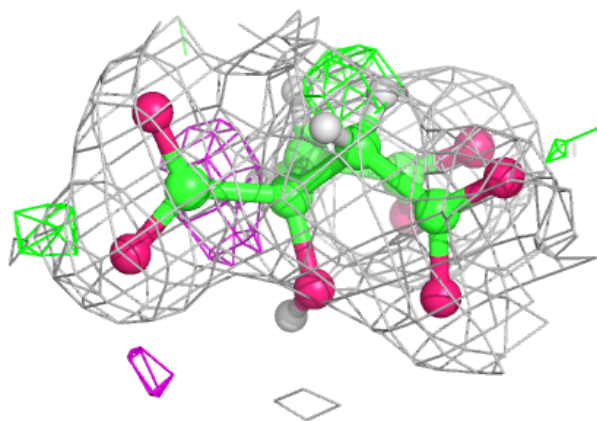
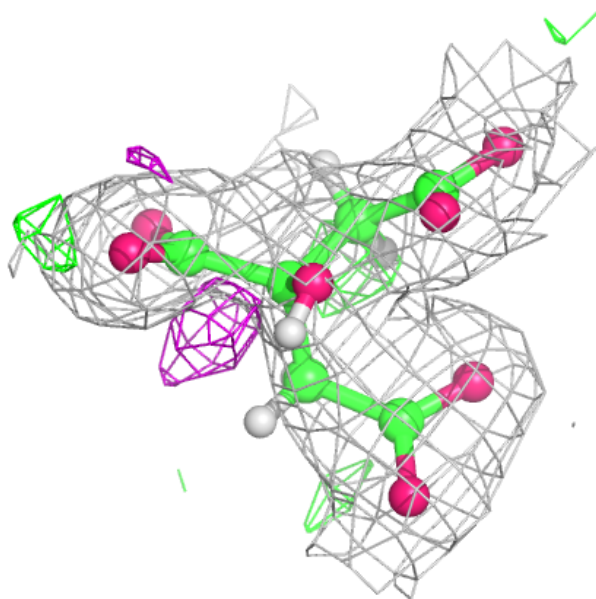
Electron density around FLC C 301:

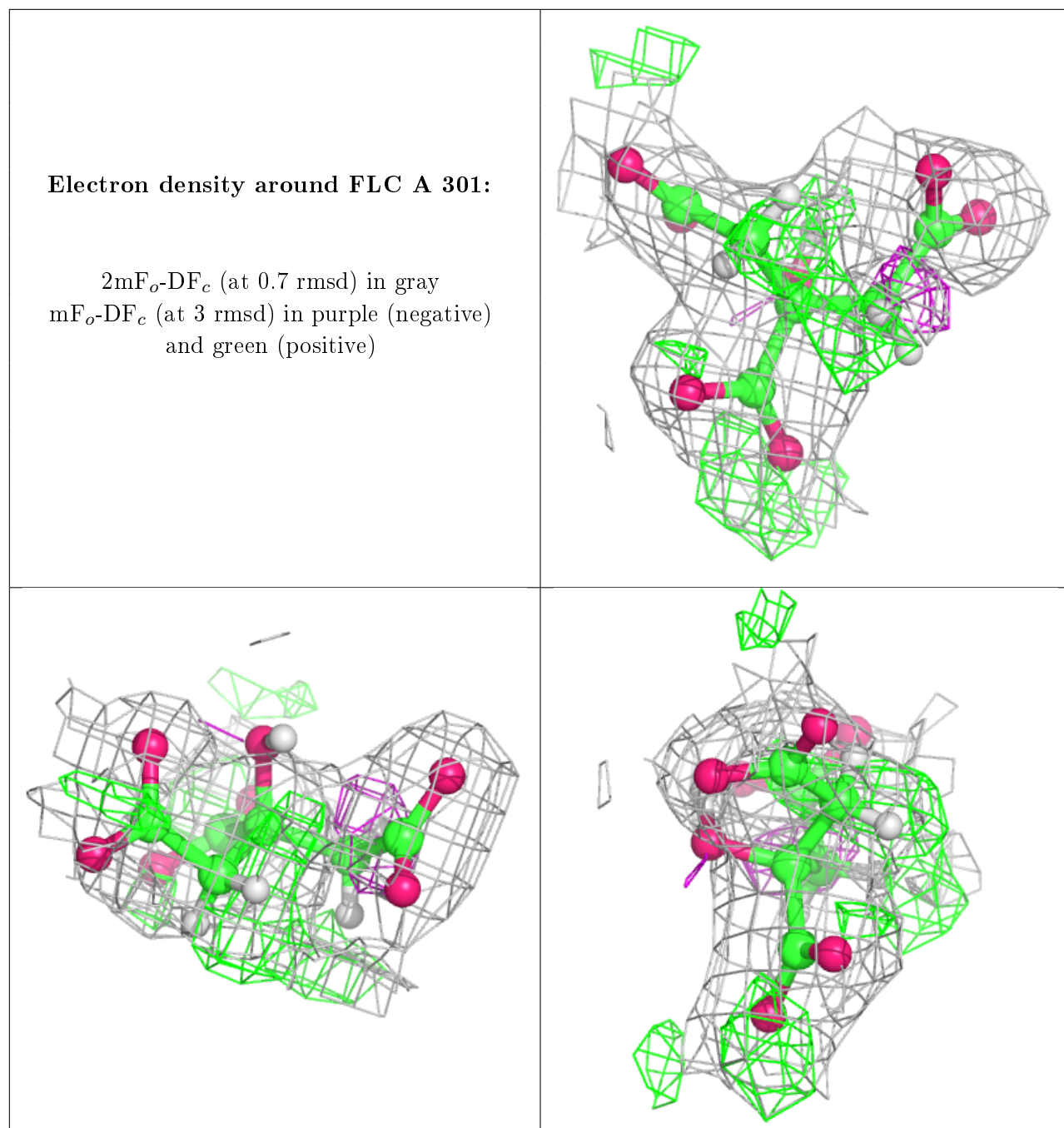
$2mF_o-DF_c$ (at 0.7 rmsd) in gray
 mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)



Electron density around FLC 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.