Full wwPDB X-ray Structure Validation Report (i)

Nov 21, 2023 – 04:17 AM JST

PDB ID	:	7EYS
Title	:	Complex structure of SptF with Fe, alpha-ketoglutarate, and andiconin D
Authors	:	Tao, H.; Mori, T.; Abe, I.
Deposited on	:	2021-06-01
Resolution	:	1.95 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 (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:

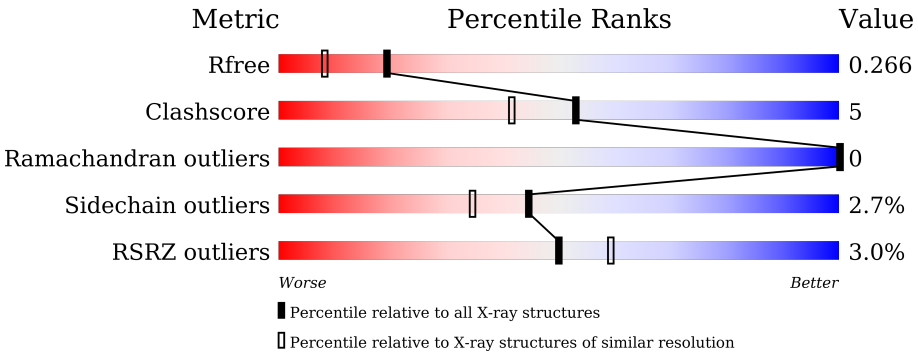
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (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 1.95 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	2580 (1.96-1.96)
Clashscore	141614	2705 (1.96-1.96)
Ramachandran outliers	138981	2678 (1.96-1.96)
Sidechain outliers	138945	2678 (1.96-1.96)
RSRZ outliers	127900	2539 (1.96-1.96)

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	296	4% 70% 13% • 16%
1	B	296	2% 80% 14% • 6%
1	C	296	3% 82% 11% 6%
1	D	296	2% 77% 14% 9%

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 9118 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 2-oxoglutarate/Fe(II)-dependent dioxygenase SptF.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	C	277	Total	C	N	O	S	0	0	0
			2214	1401	389	409	15			
1	A	248	Total	C	N	O	S	0	0	0
			1999	1267	348	369	15			
1	B	278	Total	C	N	O	S	0	1	0
			2232	1410	393	414	15			
1	D	268	Total	C	N	O	S	0	0	0
			2151	1360	376	400	15			

There are 56 discrepancies between the modelled and reference sequences:

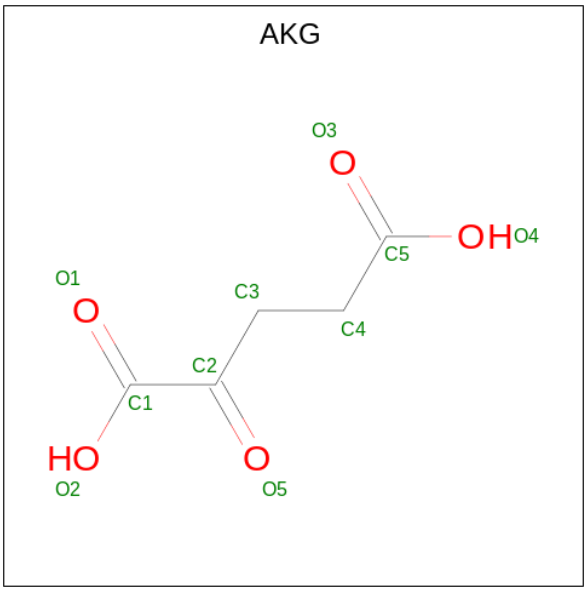
Chain	Residue	Modelled	Actual	Comment	Reference
C	3	MET	-	initiating methionine	UNP A0A6J4CX17
C	286	LYS	-	expression tag	UNP A0A6J4CX17
C	287	LEU	-	expression tag	UNP A0A6J4CX17
C	288	ALA	-	expression tag	UNP A0A6J4CX17
C	289	ALA	-	expression tag	UNP A0A6J4CX17
C	290	ALA	-	expression tag	UNP A0A6J4CX17
C	291	LEU	-	expression tag	UNP A0A6J4CX17
C	292	GLU	-	expression tag	UNP A0A6J4CX17
C	293	HIS	-	expression tag	UNP A0A6J4CX17
C	294	HIS	-	expression tag	UNP A0A6J4CX17
C	295	HIS	-	expression tag	UNP A0A6J4CX17
C	296	HIS	-	expression tag	UNP A0A6J4CX17
C	297	HIS	-	expression tag	UNP A0A6J4CX17
C	298	HIS	-	expression tag	UNP A0A6J4CX17
A	3	MET	-	initiating methionine	UNP A0A6J4CX17
A	286	LYS	-	expression tag	UNP A0A6J4CX17
A	287	LEU	-	expression tag	UNP A0A6J4CX17
A	288	ALA	-	expression tag	UNP A0A6J4CX17
A	289	ALA	-	expression tag	UNP A0A6J4CX17
A	290	ALA	-	expression tag	UNP A0A6J4CX17
A	291	LEU	-	expression tag	UNP A0A6J4CX17

Continued on next page...

Continued from previous page...

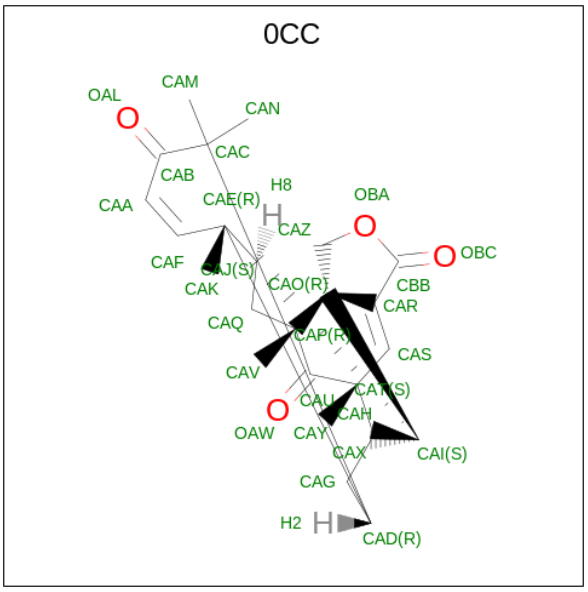
Chain	Residue	Modelled	Actual	Comment	Reference
A	292	GLU	-	expression tag	UNP A0A6J4CX17
A	293	HIS	-	expression tag	UNP A0A6J4CX17
A	294	HIS	-	expression tag	UNP A0A6J4CX17
A	295	HIS	-	expression tag	UNP A0A6J4CX17
A	296	HIS	-	expression tag	UNP A0A6J4CX17
A	297	HIS	-	expression tag	UNP A0A6J4CX17
A	298	HIS	-	expression tag	UNP A0A6J4CX17
B	3	MET	-	initiating methionine	UNP A0A6J4CX17
B	286	LYS	-	expression tag	UNP A0A6J4CX17
B	287	LEU	-	expression tag	UNP A0A6J4CX17
B	288	ALA	-	expression tag	UNP A0A6J4CX17
B	289	ALA	-	expression tag	UNP A0A6J4CX17
B	290	ALA	-	expression tag	UNP A0A6J4CX17
B	291	LEU	-	expression tag	UNP A0A6J4CX17
B	292	GLU	-	expression tag	UNP A0A6J4CX17
B	293	HIS	-	expression tag	UNP A0A6J4CX17
B	294	HIS	-	expression tag	UNP A0A6J4CX17
B	295	HIS	-	expression tag	UNP A0A6J4CX17
B	296	HIS	-	expression tag	UNP A0A6J4CX17
B	297	HIS	-	expression tag	UNP A0A6J4CX17
B	298	HIS	-	expression tag	UNP A0A6J4CX17
D	3	MET	-	initiating methionine	UNP A0A6J4CX17
D	286	LYS	-	expression tag	UNP A0A6J4CX17
D	287	LEU	-	expression tag	UNP A0A6J4CX17
D	288	ALA	-	expression tag	UNP A0A6J4CX17
D	289	ALA	-	expression tag	UNP A0A6J4CX17
D	290	ALA	-	expression tag	UNP A0A6J4CX17
D	291	LEU	-	expression tag	UNP A0A6J4CX17
D	292	GLU	-	expression tag	UNP A0A6J4CX17
D	293	HIS	-	expression tag	UNP A0A6J4CX17
D	294	HIS	-	expression tag	UNP A0A6J4CX17
D	295	HIS	-	expression tag	UNP A0A6J4CX17
D	296	HIS	-	expression tag	UNP A0A6J4CX17
D	297	HIS	-	expression tag	UNP A0A6J4CX17
D	298	HIS	-	expression tag	UNP A0A6J4CX17

- Molecule 2 is 2-OXOGLUTARIC ACID (three-letter code: AKG) (formula: $C_5H_6O_5$) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	C	1	Total C O 10 5 5	0	0

- Molecule 3 is Andiconin D (three-letter code: 0CC) (formula: $C_{25}H_{30}O_4$) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	C	1	Total C O 29 25 4	0	0

- Molecule 4 is FE (II) ION (three-letter code: FE2) (formula: Fe) (labeled as "Ligand of Interest" by depositor).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	C	1	Total Fe 1 1	0	0

- Molecule 5 is water.

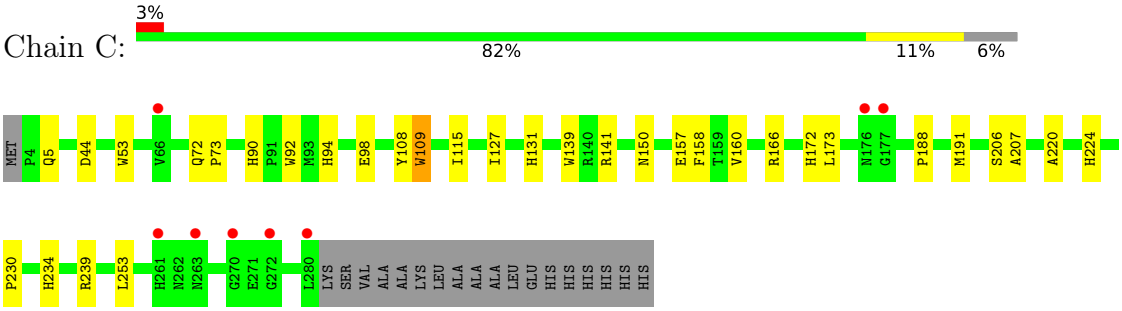
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	C	131	Total O 131 131	0	0
5	A	62	Total O 62 62	0	0
5	B	161	Total O 161 161	0	0
5	D	128	Total O 128 128	0	0

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.

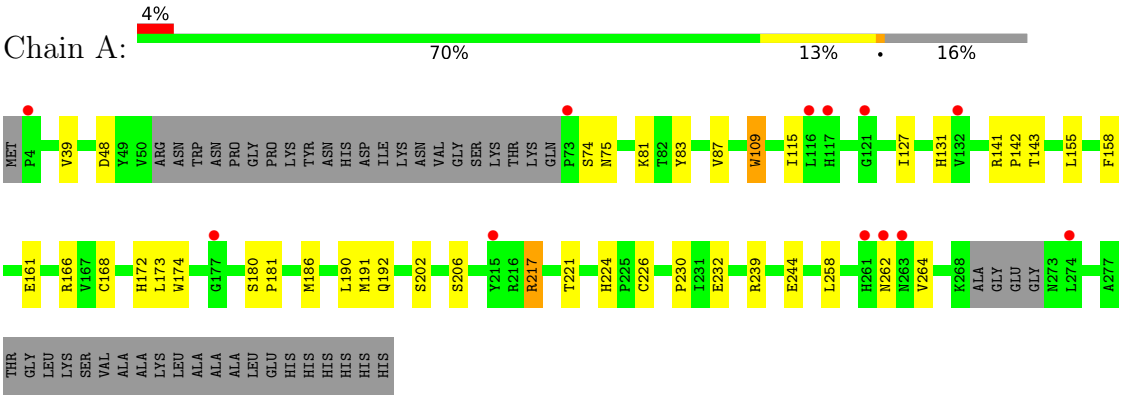
• Molecule 1: 2-oxoglutarate/Fe(II)-dependent dioxygenase SptF

Chain C:



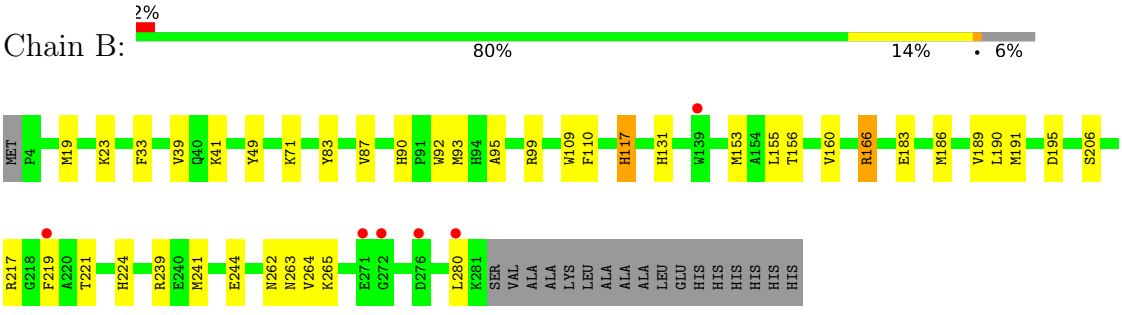
• Molecule 1: 2-oxoglutarate/Fe(II)-dependent dioxygenase SptF

Chain A:



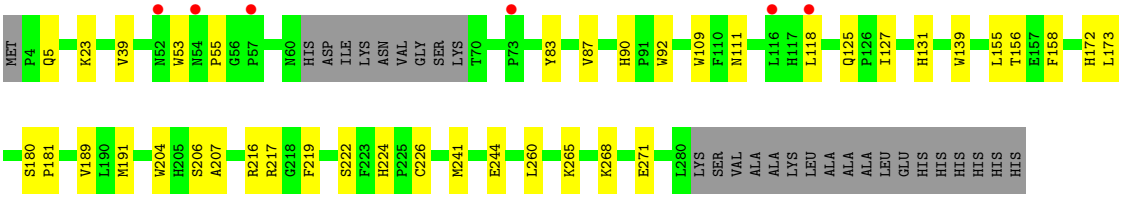
• Molecule 1: 2-oxoglutarate/Fe(II)-dependent dioxygenase SptF

Chain B:



• Molecule 1: 2-oxoglutarate/Fe(II)-dependent dioxygenase SptF

Chain D:


MET P4 Q5 K23 V39 N52 W53 N54 P55 G56 P57 N60 HIS ASP ILE LYS ASN VAL GLY SER LYS T70 P73 Y83 V87 H90 P91 W92 W109 F110 N111 L116 H117 L118 Q125 P126 I127 H131 W139 L155 T156 E157 F158 H172 L173
S180 P181 V189 L190 M191 W204 H205 S206 A207 R216 R217 G218 F219 S222 P223 H224 P225 C226 M241 E244 L260 K265 K268 E271 L280 LYS SER VAL ALA ALA LYS LEU ALA ALA ALA 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	49.37Å 72.23Å 77.61Å 81.58° 85.58° 70.24°	Depositor
Resolution (Å)	47.70 – 1.95 47.70 – 1.95	Depositor EDS
% Data completeness (in resolution range)	96.9 (47.70-1.95) 97.0 (47.70-1.95)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.62 (at 1.95Å)	Xtriage
Refinement program	PHENIX 1.19.2_4158	Depositor
R, R_{free}	0.213 , 0.268 0.212 , 0.266	Depositor DCC
R_{free} test set	1999 reflections (2.83%)	wwPDB-VP
Wilson B-factor (Å^2)	21.9	Xtriage
Anisotropy	0.725	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 44.4	EDS
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	9118	wwPDB-VP
Average B, all atoms (Å^2)	29.0	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 43.64 % 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 1.7129e-04. 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

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: FE2, 0CC, AKG

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	0.25	0/2060	0.44	0/2804
1	B	0.27	0/2301	0.46	0/3135
1	C	0.30	0/2283	0.46	0/3110
1	D	0.28	0/2218	0.44	0/3023
All	All	0.28	0/8862	0.45	0/12072

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts

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

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	1999	0	1895	22	0
1	B	2232	0	2118	28	0
1	C	2214	0	2098	19	0
1	D	2151	0	2031	23	0
2	C	10	0	4	1	0
3	C	29	0	0	0	0
4	C	1	0	0	0	0
5	A	62	0	0	1	0
5	B	161	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
5	C	131	0	0	2	0
5	D	128	0	0	0	0
All	All	9118	0	8146	90	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 5.

All (90) 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:74:SER:OG	1:A:75:ASN:N	2.30	0.65
1:C:94:HIS:O	1:C:98:GLU:HG3	1.99	0.63
1:B:39:VAL:HG23	1:B:219:PHE:CE2	2.34	0.63
1:A:262:ASN:O	1:A:264:VAL:N	2.31	0.62
1:B:110:PHE:HE1	1:B:221[B]:THR:HG23	1.65	0.62
1:D:118:LEU:HD22	1:D:216:ARG:HH21	1.65	0.62
1:D:39:VAL:HG23	1:D:219:PHE:CE2	2.36	0.60
1:D:265:LYS:HB2	1:D:268:LYS:HD3	1.84	0.60
1:C:98:GLU:HG2	1:C:108:TYR:HE1	1.67	0.60
1:B:39:VAL:HG23	1:B:219:PHE:HE2	1.67	0.59
1:C:158:PHE:HB2	1:C:191:MET:HB2	1.87	0.57
1:B:117:HIS:NE2	1:B:217:ARG:HD2	2.19	0.57
1:A:226:CYS:HB2	1:A:258:LEU:HD13	1.87	0.56
1:D:156:THR:HG23	1:D:217:ARG:HB2	1.88	0.56
1:A:262:ASN:O	1:A:264:VAL:HG23	2.05	0.55
1:B:160:VAL:HG23	1:B:190:LEU:HD21	1.89	0.55
1:B:110:PHE:CE1	1:B:221[B]:THR:HG23	2.41	0.54
1:B:241:MET:O	1:B:244:GLU:HG2	2.09	0.53
1:C:73:PRO:HD2	1:C:115:ILE:HB	1.89	0.52
1:C:5:GLN:NE2	5:C:404:HOH:O	2.30	0.51
1:D:39:VAL:HG22	1:D:155:LEU:HD13	1.93	0.50
1:A:232:GLU:OE2	5:A:301:HOH:O	2.20	0.50
1:A:190:LEU:HD11	1:B:41:LYS:HG3	1.94	0.50
1:B:39:VAL:HG22	1:B:155:LEU:HD13	1.93	0.49
1:C:150:ASN:HD21	1:C:220:ALA:HB1	1.78	0.49
1:B:166:ARG:NH1	1:B:183:GLU:OE2	2.45	0.49
1:A:168:CYS:HB2	1:A:186:MET:HE2	1.95	0.48
1:A:158:PHE:HB2	1:A:191:MET:HB2	1.95	0.48
1:D:268:LYS:HE2	1:D:271:GLU:HA	1.94	0.48
1:D:118:LEU:HD22	1:D:216:ARG:NH2	2.28	0.48
1:D:125:GLN:HB3	1:D:207:ALA:H	1.78	0.47

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:191:MET:HB3	1:B:195:ASP:HB2	1.97	0.47
1:D:111:ASN:HB3	1:D:222:SER:OG	2.14	0.47
1:D:158:PHE:HB2	1:D:191:MET:HB2	1.96	0.47
1:B:239:ARG:HE	1:B:239:ARG:HB2	1.57	0.47
1:B:83:TYR:OH	1:B:221[A]:THR:HG22	2.15	0.47
1:B:166:ARG:HB3	1:B:186:MET:HB3	1.97	0.46
1:C:53:TRP:CD2	1:C:73:PRO:HB3	2.51	0.46
1:B:90:HIS:HB3	1:B:93:MET:HE2	1.97	0.46
1:A:161:GLU:H	1:A:161:GLU:CD	2.18	0.46
1:D:172:HIS:CE1	1:D:173:LEU:HG	2.51	0.46
1:D:139:TRP:N	1:D:139:TRP:CD1	2.83	0.45
1:D:189:VAL:HG12	1:D:191:MET:HG3	1.98	0.45
1:D:226:CYS:SG	1:D:260:LEU:HD12	2.57	0.45
1:B:33:PHE:HE2	1:B:153:MET:HE3	1.81	0.45
1:A:39:VAL:HG13	1:A:155:LEU:HD23	1.98	0.45
1:A:83:TYR:OH	1:A:221:THR:HG22	2.17	0.45
1:D:83:TYR:HA	1:D:87:VAL:HB	1.97	0.45
1:D:127:ILE:HG21	1:D:181:PRO:HG2	1.99	0.45
1:C:72:GLN:HA	1:C:115:ILE:O	2.16	0.45
1:B:49:TYR:O	1:B:71:LYS:HD3	2.17	0.45
1:B:117:HIS:CE1	1:B:217:ARG:HD2	2.52	0.45
1:D:118:LEU:HB3	1:D:216:ARG:HB3	1.98	0.45
1:D:127:ILE:HG22	1:D:204:TRP:HE3	1.83	0.44
1:A:74:SER:HG	1:A:75:ASN:H	1.62	0.43
1:B:95:ALA:O	1:B:99:ARG:HG3	2.19	0.43
1:A:109:TRP:CZ2	1:A:230:PRO:HD3	2.53	0.43
1:B:262:ASN:OD1	1:B:264:VAL:HB	2.18	0.43
1:B:265:LYS:HE2	1:B:265:LYS:HB2	1.78	0.43
1:C:160:VAL:HG22	1:C:188:PRO:HB2	2.01	0.43
1:A:172:HIS:HB2	1:A:202:SER:OG	2.19	0.43
1:A:83:TYR:HA	1:A:87:VAL:HB	2.00	0.43
1:B:280:LEU:HD23	1:B:280:LEU:HA	1.91	0.43
1:C:44:ASP:OD1	5:C:401:HOH:O	2.21	0.43
1:D:127:ILE:HA	1:D:206:SER:HB3	2.01	0.43
1:A:262:ASN:C	1:A:264:VAL:H	2.20	0.42
1:C:109:TRP:CZ2	1:C:230:PRO:HD3	2.55	0.42
1:C:127:ILE:HA	1:C:206:SER:HB3	2.01	0.42
1:A:115:ILE:HG21	1:A:217:ARG:HE	1.84	0.42
1:D:53:TRP:CZ2	1:D:55:PRO:HB3	2.55	0.42
1:C:157:GLU:H	1:D:5:GLN:HE22	1.68	0.42
1:C:172:HIS:CE1	1:C:173:LEU:HG	2.55	0.42

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:141:ARG:HA	1:C:141:ARG:HD2	1.85	0.41
1:B:90:HIS:CB	1:B:93:MET:HE2	2.50	0.41
1:C:207:ALA:HB2	2:C:301:AKG:C5	2.50	0.41
1:D:241:MET:O	1:D:244:GLU:HG2	2.20	0.41
1:C:139:TRP:N	1:C:139:TRP:CD1	2.88	0.41
1:B:189:VAL:HG12	1:B:191:MET:HG3	2.03	0.41
1:C:234:HIS:CE1	1:C:253:LEU:HD22	2.56	0.41
1:A:174:TRP:CZ3	1:A:181:PRO:HG3	2.56	0.41
1:A:127:ILE:HA	1:A:206:SER:HB3	2.03	0.41
1:B:19:MET:O	1:B:23:LYS:HG2	2.20	0.41
1:B:83:TYR:HA	1:B:87:VAL:HB	2.03	0.41
1:B:90:HIS:CE1	1:B:92:TRP:HB2	2.56	0.41
1:C:90:HIS:CE1	1:C:92:TRP:HB2	2.56	0.41
1:A:142:PRO:HB3	1:A:173:LEU:HA	2.02	0.41
1:D:90:HIS:CE1	1:D:92:TRP:HB2	2.56	0.41
1:A:141:ARG:NH1	1:A:143:THR:OG1	2.54	0.40
1:A:81:LYS:NZ	1:A:244:GLU:OE2	2.54	0.40
1:B:156:THR:HG23	1:B:217:ARG:HB2	2.04	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	242/296 (82%)	238 (98%)	4 (2%)	0	100	100
1	B	277/296 (94%)	270 (98%)	7 (2%)	0	100	100
1	C	275/296 (93%)	266 (97%)	9 (3%)	0	100	100
1	D	264/296 (89%)	260 (98%)	4 (2%)	0	100	100
All	All	1058/1184 (89%)	1034 (98%)	24 (2%)	0	100	100

There are no Ramachandran outliers to report.

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	222/261 (85%)	213 (96%)	9 (4%)	30	18
1	B	246/261 (94%)	239 (97%)	7 (3%)	43	33
1	C	242/261 (93%)	237 (98%)	5 (2%)	53	46
1	D	237/261 (91%)	232 (98%)	5 (2%)	53	46
All	All	947/1044 (91%)	921 (97%)	26 (3%)	44	34

All (26) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	109	TRP
1	C	131	HIS
1	C	166	ARG
1	C	224	HIS
1	C	239	ARG
1	A	48	ASP
1	A	109	TRP
1	A	131	HIS
1	A	166	ARG
1	A	180	SER
1	A	192	GLN
1	A	217	ARG
1	A	224	HIS
1	A	239	ARG
1	B	109	TRP
1	B	117	HIS
1	B	131	HIS
1	B	166	ARG
1	B	206	SER
1	B	224	HIS
1	B	263	ASN
1	D	23	LYS
1	D	109	TRP
1	D	131	HIS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	D	180	SER
1	D	224	HIS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

5.3.3 RNA

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates

There are no monosaccharides in this entry.

5.6 Ligand geometry

Of 3 ligands modelled in this entry, 1 is monoatomic - leaving 2 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	AKG	C	301	4	9,9,9	5.88	2 (22%)	11,11,11	1.63	2 (18%)
3	0CC	C	302	-	31,34,34	4.39	10 (32%)	39,63,63	2.21	7 (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	AKG	C	301	4	-	0/9/9/9	-

All (12) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	C	301	AKG	C2-C1	-17.35	1.30	1.53
3	C	302	0CC	CAC-CAB	-13.17	1.40	1.53
3	C	302	0CC	CAT-CAU	-12.13	1.39	1.53
3	C	302	0CC	CAS-CAR	8.93	1.53	1.33
3	C	302	0CC	CAE-CAF	-8.74	1.39	1.51
3	C	302	0CC	CBB-CAR	-5.63	1.40	1.47
3	C	302	0CC	CAP-CAU	-4.94	1.40	1.53
3	C	302	0CC	CAA-CAF	4.23	1.39	1.33
3	C	302	0CC	CAA-CAB	-3.87	1.40	1.46
3	C	302	0CC	CAT-CAS	3.14	1.56	1.50
2	C	301	AKG	O2-C1	-2.42	1.23	1.30
3	C	302	0CC	CAX-CAT	-2.41	1.51	1.55

All (9) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	C	302	0CC	OBC-CBB-CAR	-7.91	121.88	130.45
3	C	302	0CC	OBA-CBB-CAR	5.63	113.51	107.88
3	C	302	0CC	OBA-CAZ-CAO	-5.40	102.22	106.56
3	C	302	0CC	CAZ-OBA-CBB	-4.59	105.06	110.33
2	C	301	AKG	C3-C2-C1	3.55	122.56	115.97
3	C	302	0CC	OBA-CBB-OBC	2.80	124.61	121.07
3	C	302	0CC	CAC-CAB-CAA	2.75	121.70	118.85
2	C	301	AKG	O1-C1-C2	-2.63	118.20	121.72
3	C	302	0CC	CAG-CAD-CAC	-2.52	110.53	113.61

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

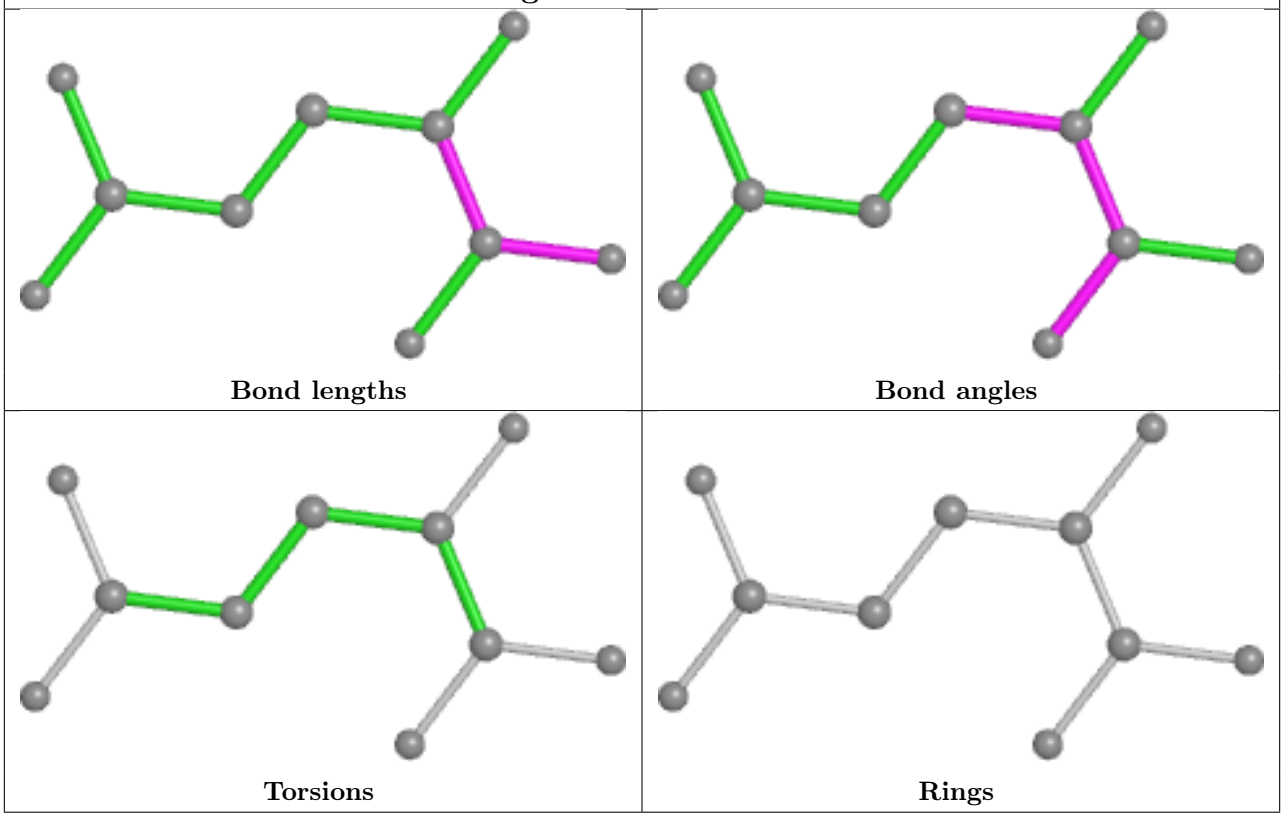
1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	C	301	AKG	1	0

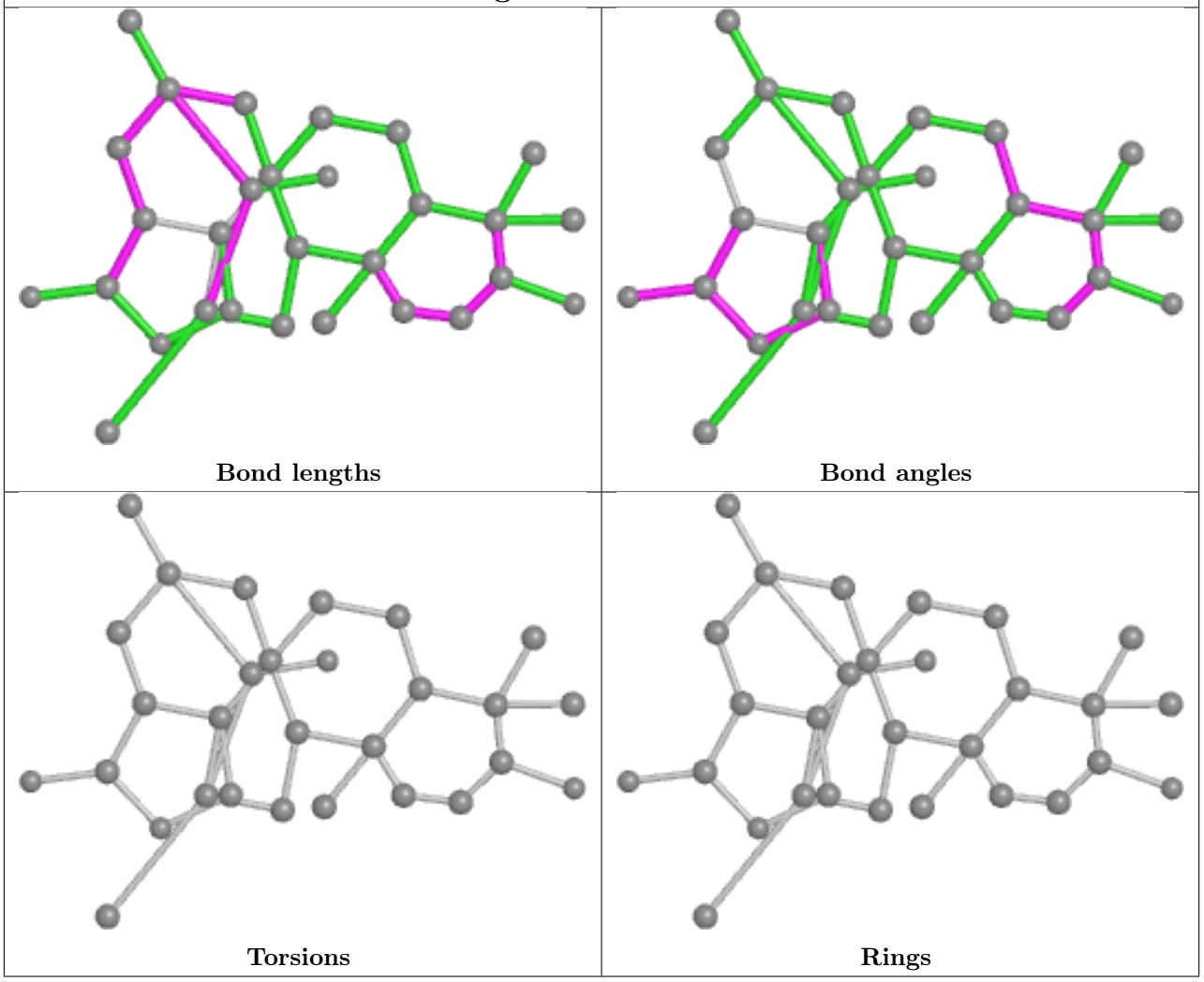
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will

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 AKG C 301



Ligand 0CC C 302



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	248/296 (83%)	0.48	12 (4%)	30	40	18, 35, 50, 66	0
1	B	278/296 (93%)	0.04	6 (2%)	62	70	17, 25, 41, 55	0
1	C	277/296 (93%)	0.16	8 (2%)	51	60	15, 27, 43, 56	0
1	D	268/296 (90%)	0.11	6 (2%)	62	70	17, 26, 42, 54	0
All	All	1071/1184 (90%)	0.19	32 (2%)	50	59	15, 27, 45, 66	0

All (32) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	272	GLY	4.1
1	A	177	GLY	3.8
1	C	177	GLY	3.8
1	A	274	LEU	3.8
1	A	261	HIS	3.8
1	A	262	ASN	3.7
1	A	121	GLY	3.6
1	C	66	VAL	3.6
1	C	263	ASN	3.5
1	A	263	ASN	3.4
1	B	272	GLY	3.1
1	C	261	HIS	3.0
1	A	73	PRO	3.0
1	C	270	GLY	2.9
1	A	117	HIS	2.8
1	B	271	GLU	2.7
1	B	280	LEU	2.7
1	A	116	LEU	2.6
1	C	280	LEU	2.6
1	B	219	PHE	2.6
1	A	215	TYR	2.5

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	139	TRP	2.4
1	D	118	LEU	2.4
1	D	73	PRO	2.3
1	C	176	ASN	2.3
1	D	54	ASN	2.2
1	A	4	PRO	2.2
1	A	132	VAL	2.2
1	B	276	ASP	2.1
1	D	52	ASN	2.1
1	D	116	LEU	2.1
1	D	57	PRO	2.0

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

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

6.3 Carbohydrates (i)

There are no monosaccharides in this entry.

6.4 Ligands (i)

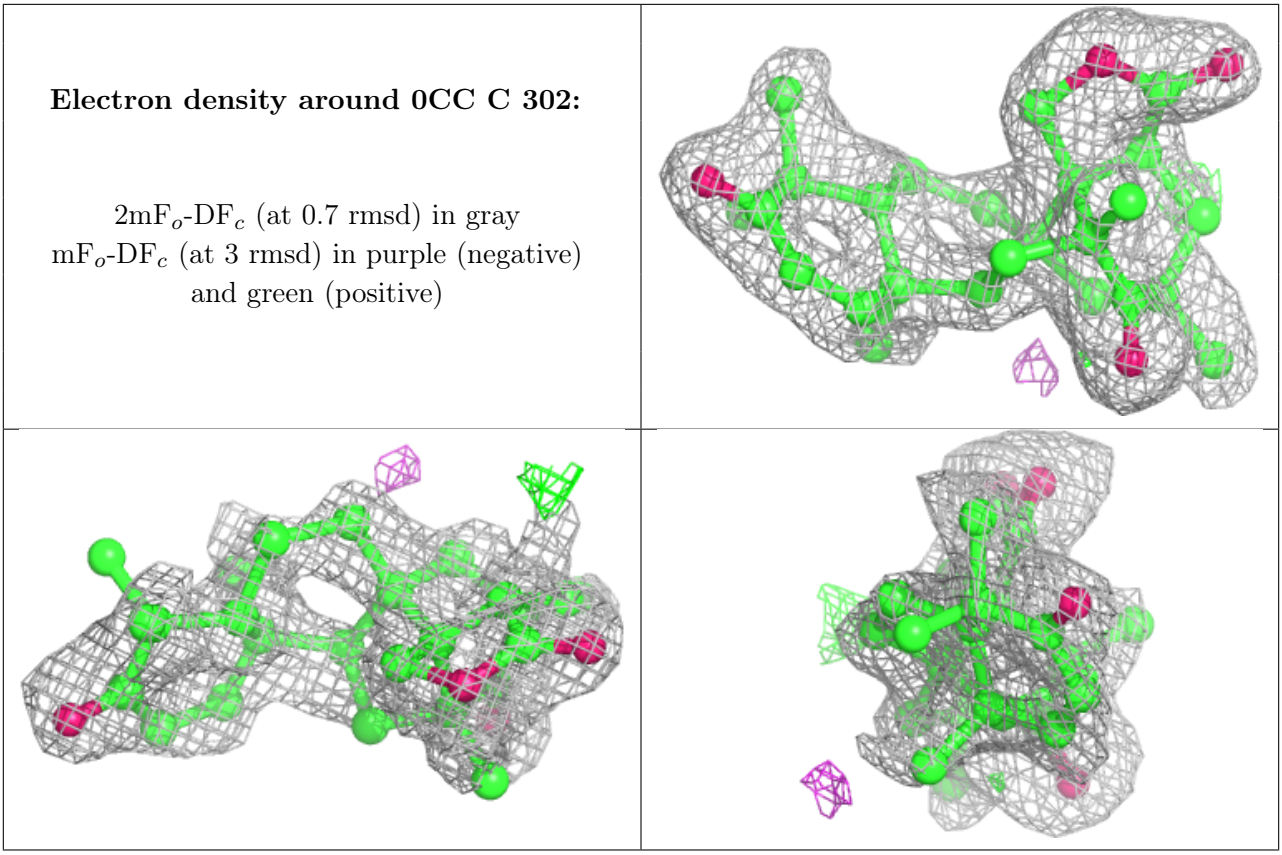
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' 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	0CC	C	302	29/29	0.83	0.20	25,28,33,34	29
2	AKG	C	301	10/10	0.85	0.15	22,29,32,33	10
4	FE2	C	303	1/1	0.86	0.10	27,27,27,27	1

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.

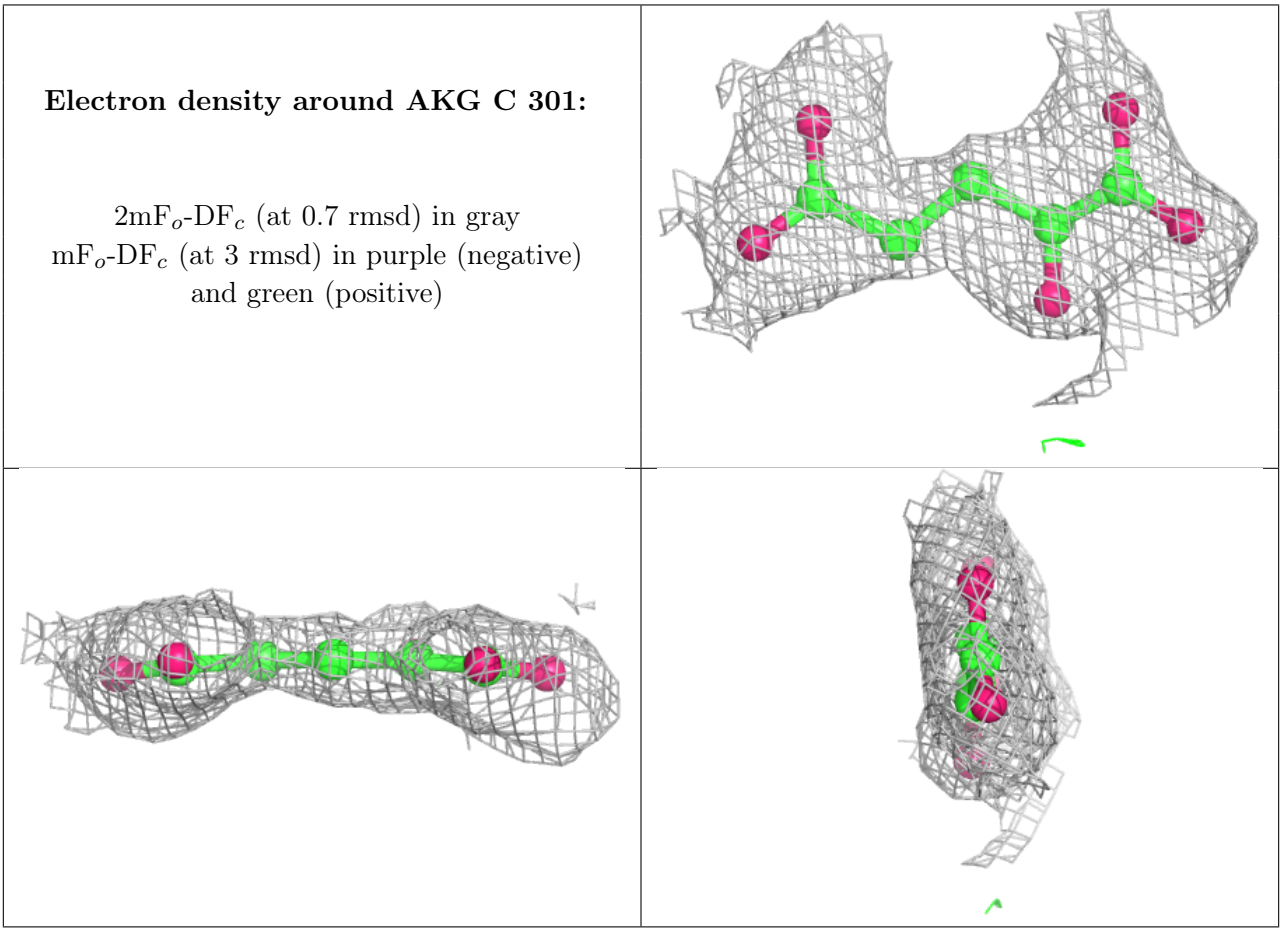
Electron density around 0CC C 302:

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



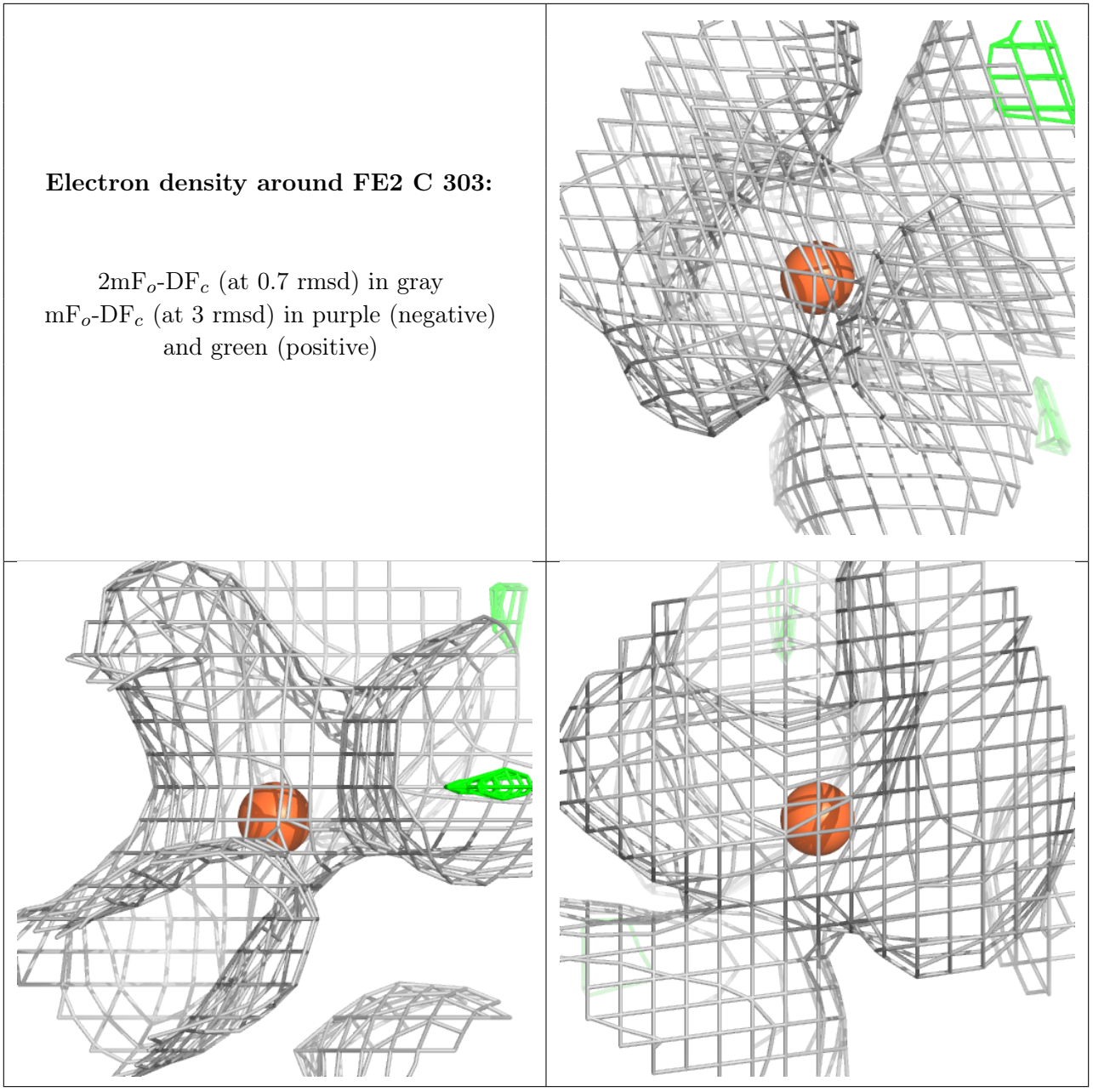
Electron density around AKG C 301:

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



Electron density around FE2 C 303:

$2mF_o$-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.