WORLDWIDE PDB PROTEIN DATA BANK

Full wwPDB X-ray Structure Validation Report (i)

Oct 25, 2022 – 09:15 pm BST

PDB ID	:	8AUT
Title	:	WelO5* L221A bound to Zn(II), Cl, 2-oxoglutarate, and 12-epi-hapalindole C
Authors	:	Buller, R.; Hueppi, S.; Voss, M.; Schaub, D.
Deposited on	:	2022-08-25
Resolution	:	2.69 Å(reported)

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

|		
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.31.2
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0267
CCP4	:	7.1.010 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.31.2

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 2.69 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.
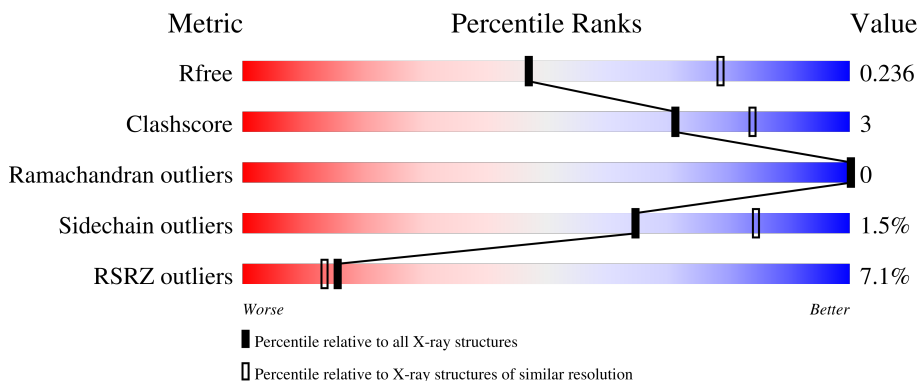
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	3863 (2.70-2.66)
Clashscore	141614	4210 (2.70-2.66)
Ramachandran outliers	138981	4141 (2.70-2.66)
Sidechain outliers	138945	4141 (2.70-2.66)
RSRZ outliers	127900	3780 (2.70-2.66)

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	310	4% / 79% 8% · 11%
1	B	310	10% / 79% 10% 11%
1	C	310	6% / 83% 6% 11%
1	D	310	6% / 77% 11% 12%

2 Entry composition ⓘ

There are 8 unique types of molecules in this entry. The entry contains 17580 atoms, of which 8619 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 Oxidoreductase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	275	Total	C	H	N	O	S	135	0	0
			4297	1384	2124	367	414	8			
1	B	275	Total	C	H	N	O	S	135	0	0
			4292	1382	2123	367	412	8			
1	C	277	Total	C	H	N	O	S	136	0	0
			4318	1390	2134	369	417	8			
1	D	274	Total	C	H	N	O	S	135	0	0
			4282	1379	2118	366	411	8			

There are 84 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-19	MET	-	initiating methionine	UNP A0A075X7C6
A	-18	GLY	-	expression tag	UNP A0A075X7C6
A	-17	SER	-	expression tag	UNP A0A075X7C6
A	-16	SER	-	expression tag	UNP A0A075X7C6
A	-15	HIS	-	expression tag	UNP A0A075X7C6
A	-14	HIS	-	expression tag	UNP A0A075X7C6
A	-13	HIS	-	expression tag	UNP A0A075X7C6
A	-12	HIS	-	expression tag	UNP A0A075X7C6
A	-11	HIS	-	expression tag	UNP A0A075X7C6
A	-10	HIS	-	expression tag	UNP A0A075X7C6
A	-9	SER	-	expression tag	UNP A0A075X7C6
A	-8	SER	-	expression tag	UNP A0A075X7C6
A	-7	GLY	-	expression tag	UNP A0A075X7C6
A	-6	LEU	-	expression tag	UNP A0A075X7C6
A	-5	VAL	-	expression tag	UNP A0A075X7C6
A	-4	PRO	-	expression tag	UNP A0A075X7C6
A	-3	ARG	-	expression tag	UNP A0A075X7C6
A	-2	GLY	-	expression tag	UNP A0A075X7C6
A	-1	SER	-	expression tag	UNP A0A075X7C6
A	0	HIS	-	expression tag	UNP A0A075X7C6
A	221	ALA	LEU	engineered mutation	UNP A0A075X7C6

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	-19	MET	-	initiating methionine	UNP A0A075X7C6
B	-18	GLY	-	expression tag	UNP A0A075X7C6
B	-17	SER	-	expression tag	UNP A0A075X7C6
B	-16	SER	-	expression tag	UNP A0A075X7C6
B	-15	HIS	-	expression tag	UNP A0A075X7C6
B	-14	HIS	-	expression tag	UNP A0A075X7C6
B	-13	HIS	-	expression tag	UNP A0A075X7C6
B	-12	HIS	-	expression tag	UNP A0A075X7C6
B	-11	HIS	-	expression tag	UNP A0A075X7C6
B	-10	HIS	-	expression tag	UNP A0A075X7C6
B	-9	SER	-	expression tag	UNP A0A075X7C6
B	-8	SER	-	expression tag	UNP A0A075X7C6
B	-7	GLY	-	expression tag	UNP A0A075X7C6
B	-6	LEU	-	expression tag	UNP A0A075X7C6
B	-5	VAL	-	expression tag	UNP A0A075X7C6
B	-4	PRO	-	expression tag	UNP A0A075X7C6
B	-3	ARG	-	expression tag	UNP A0A075X7C6
B	-2	GLY	-	expression tag	UNP A0A075X7C6
B	-1	SER	-	expression tag	UNP A0A075X7C6
B	0	HIS	-	expression tag	UNP A0A075X7C6
B	221	ALA	LEU	engineered mutation	UNP A0A075X7C6
C	-19	MET	-	initiating methionine	UNP A0A075X7C6
C	-18	GLY	-	expression tag	UNP A0A075X7C6
C	-17	SER	-	expression tag	UNP A0A075X7C6
C	-16	SER	-	expression tag	UNP A0A075X7C6
C	-15	HIS	-	expression tag	UNP A0A075X7C6
C	-14	HIS	-	expression tag	UNP A0A075X7C6
C	-13	HIS	-	expression tag	UNP A0A075X7C6
C	-12	HIS	-	expression tag	UNP A0A075X7C6
C	-11	HIS	-	expression tag	UNP A0A075X7C6
C	-10	HIS	-	expression tag	UNP A0A075X7C6
C	-9	SER	-	expression tag	UNP A0A075X7C6
C	-8	SER	-	expression tag	UNP A0A075X7C6
C	-7	GLY	-	expression tag	UNP A0A075X7C6
C	-6	LEU	-	expression tag	UNP A0A075X7C6
C	-5	VAL	-	expression tag	UNP A0A075X7C6
C	-4	PRO	-	expression tag	UNP A0A075X7C6
C	-3	ARG	-	expression tag	UNP A0A075X7C6
C	-2	GLY	-	expression tag	UNP A0A075X7C6
C	-1	SER	-	expression tag	UNP A0A075X7C6
C	0	HIS	-	expression tag	UNP A0A075X7C6
C	221	ALA	LEU	engineered mutation	UNP A0A075X7C6

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	-19	MET	-	initiating methionine	UNP A0A075X7C6
D	-18	GLY	-	expression tag	UNP A0A075X7C6
D	-17	SER	-	expression tag	UNP A0A075X7C6
D	-16	SER	-	expression tag	UNP A0A075X7C6
D	-15	HIS	-	expression tag	UNP A0A075X7C6
D	-14	HIS	-	expression tag	UNP A0A075X7C6
D	-13	HIS	-	expression tag	UNP A0A075X7C6
D	-12	HIS	-	expression tag	UNP A0A075X7C6
D	-11	HIS	-	expression tag	UNP A0A075X7C6
D	-10	HIS	-	expression tag	UNP A0A075X7C6
D	-9	SER	-	expression tag	UNP A0A075X7C6
D	-8	SER	-	expression tag	UNP A0A075X7C6
D	-7	GLY	-	expression tag	UNP A0A075X7C6
D	-6	LEU	-	expression tag	UNP A0A075X7C6
D	-5	VAL	-	expression tag	UNP A0A075X7C6
D	-4	PRO	-	expression tag	UNP A0A075X7C6
D	-3	ARG	-	expression tag	UNP A0A075X7C6
D	-2	GLY	-	expression tag	UNP A0A075X7C6
D	-1	SER	-	expression tag	UNP A0A075X7C6
D	0	HIS	-	expression tag	UNP A0A075X7C6
D	221	ALA	LEU	engineered mutation	UNP A0A075X7C6

- Molecule 2 is ZINC ION (three-letter code: ZN) (formula: Zn) (labeled as "Ligand of Interest" by depositor).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	1	Total 1	Zn 1	0	0
2	B	1	Total 1	Zn 1	0	0
2	C	1	Total 1	Zn 1	0	0
2	D	1	Total 1	Zn 1	0	0

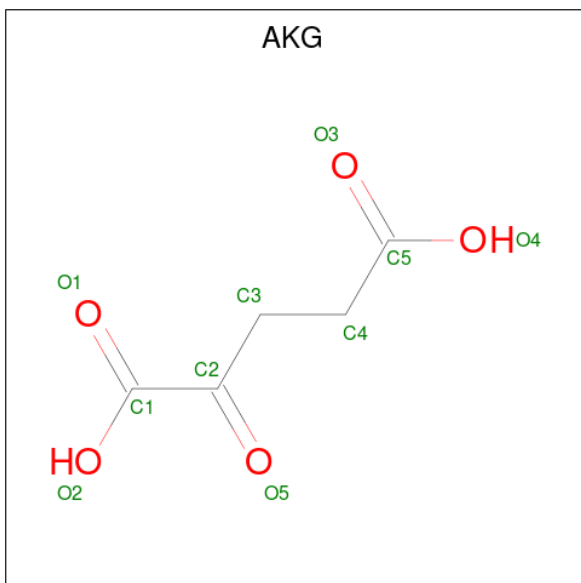- Molecule 3 is CHLORIDE ION (three-letter code: CL) (formula: Cl) (labeled as "Ligand of Interest" by depositor).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	1	Total 1	Cl 1	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	B	1	Total Cl 1 1		0	0
3	C	1	Total Cl 1 1		0	0
3	D	1	Total Cl 1 1		0	0

- Molecule 4 is 2-OXOGLUTARIC ACID (three-letter code: AKG) (formula: $C_5H_6O_5$) (labeled as "Ligand of Interest" by depositor).
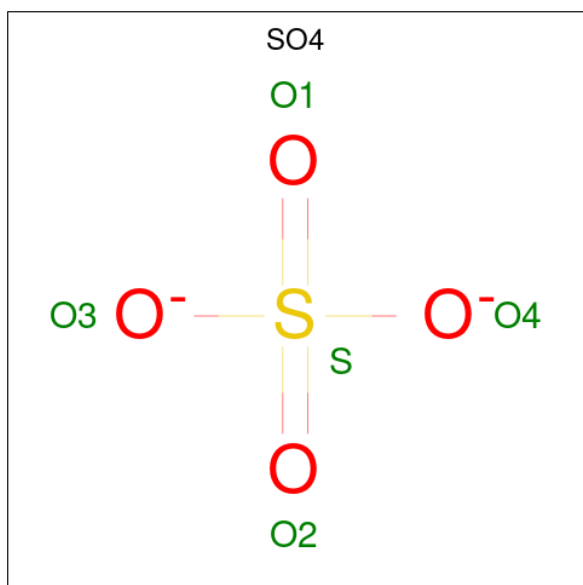

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	A	1	Total C H O 14 5 4 5				0	0
4	B	1	Total C H O 14 5 4 5				0	0
4	C	1	Total C H O 14 5 4 5				0	0
4	D	1	Total C H O 14 5 4 5				0	0

- Molecule 5 is 3-[(1 {S},2 {R},3 {S},6 {S})-3-ethenyl-2-isocyano-3-methyl-6-prop-1-en-2-yl-cyclohexyl]-1 {H}-indole (three-letter code: OAU) (formula: $C_{21}H_{24}N_2$) (labeled as "Ligand of Interest" by depositor).
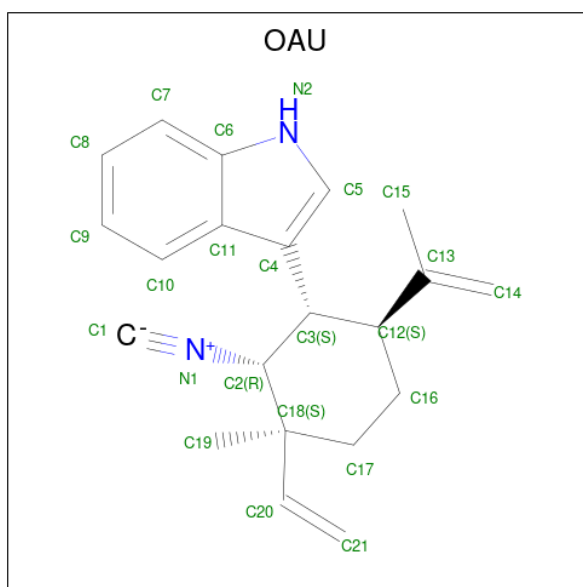
OAU

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
5	A	1	Total 47	C 21	H 24	N 2	0	0
5	B	1	Total 47	C 21	H 24	N 2	0	0
5	C	1	Total 47	C 21	H 24	N 2	0	0
5	D	1	Total 47	C 21	H 24	N 2	0	0

- Molecule 6 is SULFATE ION (three-letter code: SO4) (formula: O_4S).

SO4

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
6	A	1	Total 5	O 4	S 1	0	0
6	B	1	Total 5	O 4	S 1	0	0
6	B	1	Total 5	O 4	S 1	0	0

- Molecule 7 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).
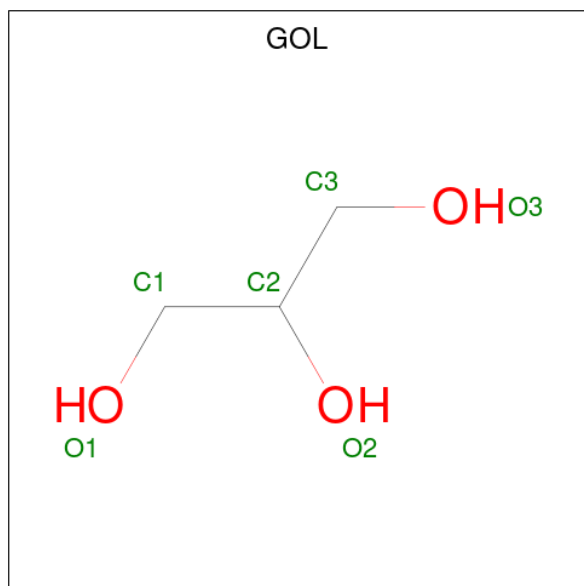
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
7	B	1	Total 14	C 3	H 8	O 3	2	0

- Molecule 8 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	A	25	Total 25	O 25	0	0
8	B	30	Total 30	O 30	0	0
8	C	23	Total 23	O 23	0	0
8	D	32	Total 32	O 32	0	0

3 Residue-property plots ⓘ

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.

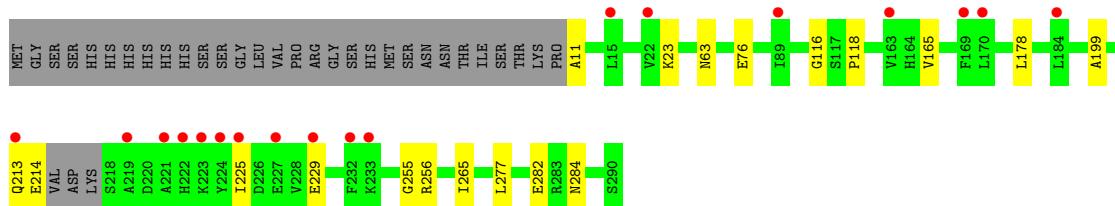• Molecule 1: Oxidoreductase

Chain A:
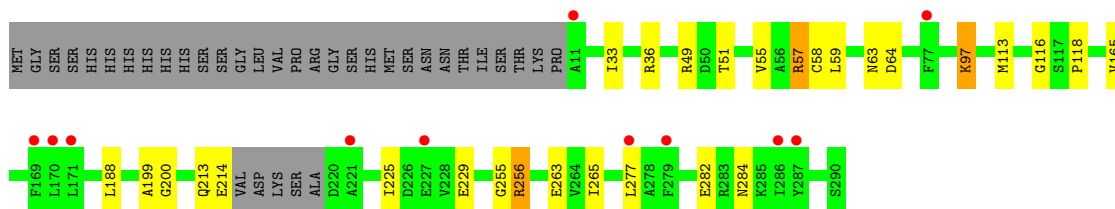
• Molecule 1: Oxidoreductase

Chain B:
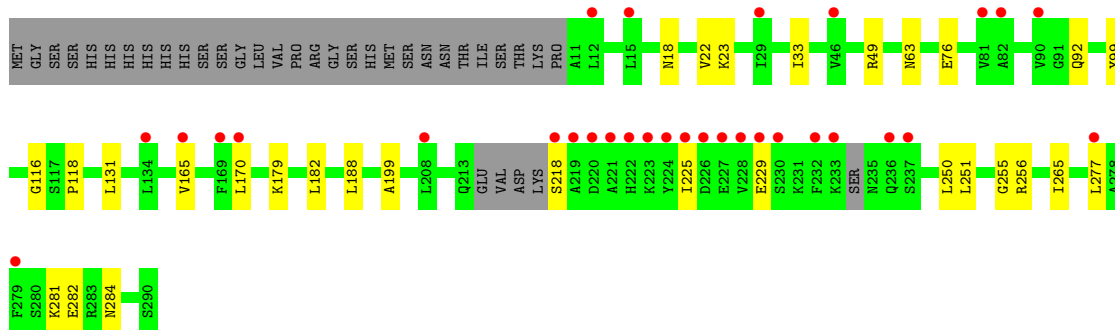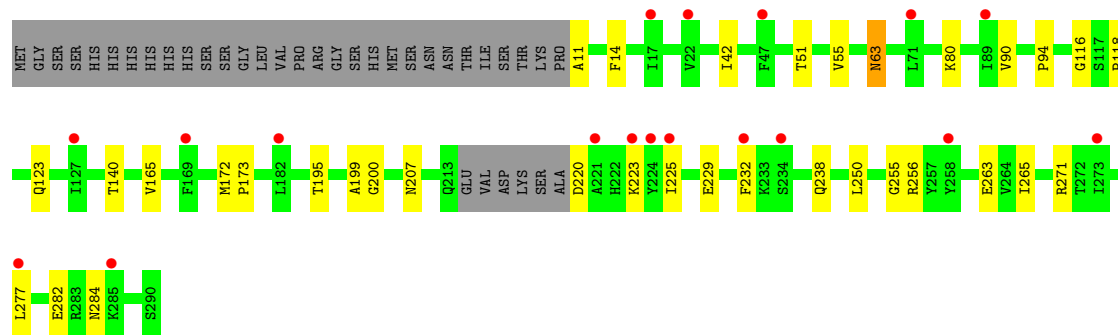
• Molecule 1: Oxidoreductase

Chain C:

• Molecule 1: Oxidoreductase

Chain D:

MET GLY SER SER HIS HIS HIS HIS HIS HIS SER SER GLY LEU VAL PRO ARG GLY SER HIS MET SER ASN ASN THR ILE SER THR LYS PRO A11 F14 I17 V22 I42 F47 T51 V55 N63 L71 K80 I89 V90 P94 G116 S117 P118

Q123 I127 T140 V165 F169 M172 P173 L182 T195 A199 G200 N207 Q213 GLU VAL ASP LYS SER ALA D220 A221 H222 K223 Y224 I225 E229 F232 K233 S234 Q238 L250 G255 R256 Y257 Y258 E263 V264 I265 R271 T272 I273

L277 E282 R283 N284 K285 S290

4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 61 2 2	Depositor
Cell constants a, b, c, α, β, γ	106.46Å 106.46Å 385.22Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	49.17 – 2.69 49.17 – 2.68	Depositor EDS
% Data completeness (in resolution range)	99.8 (49.17-2.69) 99.9 (49.17-2.68)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.61 (at 2.69Å)	Xtriage
Refinement program	REFMAC 5.8.0352	Depositor
R, R_{free}	0.186 , 0.232 0.192 , 0.236	Depositor DCC
R_{free} test set	1862 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å^2)	77.3	Xtriage
Anisotropy	0.438	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	(Not available) , (Not available)	EDS
L-test for twinning[2]	$< |L| > = 0.49, < L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.97	EDS
Total number of atoms	17580	wwPDB-VP
Average B, all atoms (Å^2)	92.0	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 45.10 % 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 1.3869e-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: SO4, OAU, AKG, CL, GOL, ZN

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 5$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	#$|Z|$>5	RMSZ	#$|Z|$>5
1	A	0.45	0/2224	0.77	0/3010
1	B	0.42	0/2219	0.74	0/3002
1	C	0.43	0/2235	0.74	0/3025
1	D	0.46	0/2215	0.76	0/2998
All	All	0.44	0/8893	0.75	0/12035

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	3
1	B	0	2
1	C	0	2
1	D	0	2
All	All	0	9

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (9) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	256	ARG	Sidechain
1	A	57	ARG	Sidechain
1	A	63	ASN	Peptide
1	B	256	ARG	Sidechain
1	B	63	ASN	Peptide

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Group
1	C	256	ARG	Sidechain
1	C	63	ASN	Peptide
1	D	256	ARG	Sidechain
1	D	63	ASN	Peptide

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	2173	2124	2115	17	0
1	B	2169	2123	2113	14	0
1	C	2184	2134	2125	7	0
1	D	2164	2118	2109	22	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
2	C	1	0	0	0	0
2	D	1	0	0	0	0
3	A	1	0	0	0	0
3	B	1	0	0	0	0
3	C	1	0	0	0	0
3	D	1	0	0	0	0
4	A	10	4	4	0	0
4	B	10	4	4	0	0
4	C	10	4	4	0	0
4	D	10	4	4	0	0
5	A	23	24	0	0	0
5	B	23	24	0	0	0
5	C	23	24	0	0	0
5	D	23	24	0	0	0
6	A	5	0	0	0	0
6	B	10	0	0	0	0
7	B	6	8	8	0	0
8	A	25	0	0	0	0
8	B	30	0	0	1	0
8	C	23	0	0	1	0
8	D	32	0	0	4	0
All	All	8961	8619	8486	60	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 3.

All (60) 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:C:282:GLU:HB2	1:C:284:ASN:OD1	1.71	0.89
1:D:11:ALA:HB3	8:D:409:HOH:O	1.75	0.85
1:D:282:GLU:HB3	1:D:284:ASN:OD1	1.85	0.75
1:A:282:GLU:HB2	1:A:284:ASN:OD1	1.85	0.75
1:D:94:PRO:HG3	1:D:173:PRO:HB2	1.70	0.74
1:C:225:ILE:O	1:C:229:GLU:HG2	1.90	0.72
1:D:225:ILE:O	1:D:229:GLU:HG2	1.90	0.71
1:B:225:ILE:O	1:B:229:GLU:HG2	1.90	0.70
1:A:213:GLN:O	1:A:214:GLU:HB2	1.92	0.69
1:A:225:ILE:O	1:A:229:GLU:HG2	1.92	0.69
1:B:170:LEU:O	1:B:179:LYS:HE3	1.92	0.68
1:D:140:THR:HG22	8:D:417:HOH:O	1.93	0.67
1:B:22:VAL:HG23	1:B:131:LEU:HD21	1.81	0.62
1:C:11:ALA:HA	8:C:409:HOH:O	2.00	0.60
1:B:92:GLN:HG2	1:B:99:TYR:HB2	1.84	0.59
1:D:11:ALA:CB	8:D:409:HOH:O	2.43	0.59
1:D:80:LYS:HB2	8:D:408:HOH:O	2.04	0.57
1:D:282:GLU:CB	1:D:284:ASN:OD1	2.52	0.56
1:A:49:ARG:HH11	1:A:49:ARG:HG2	1.70	0.54
1:B:182:LEU:O	1:B:281:LYS:HE2	2.07	0.54
1:D:118:PRO:HG3	1:D:123:GLN:HE21	1.74	0.53
1:A:58:CYS:HB3	1:A:113:MET:HE3	1.92	0.52
1:B:49:ARG:HH11	1:B:49:ARG:HG2	1.77	0.50
1:B:18:ASN:HB3	8:B:424:HOH:O	2.13	0.49
1:A:64:ASP:HA	1:A:113:MET:HE3	1.95	0.48
1:D:116:GLY:C	1:D:118:PRO:HA	2.33	0.48
1:C:116:GLY:C	1:C:118:PRO:HA	2.34	0.48
1:A:116:GLY:C	1:A:118:PRO:HA	2.34	0.47
1:B:116:GLY:C	1:B:118:PRO:HA	2.34	0.47
1:D:165:VAL:HG22	1:D:255:GLY:O	2.15	0.47
1:C:213:GLN:O	1:C:214:GLU:CB	2.63	0.46
1:B:282:GLU:HB3	1:B:284:ASN:ND2	2.31	0.46
1:A:49:ARG:HG2	1:A:49:ARG:NH1	2.31	0.45
1:A:213:GLN:HA	1:A:213:GLN:OE1	2.16	0.45
1:D:42:ILE:HG12	1:D:250:LEU:HD13	1.98	0.45
1:D:172:MET:HE2	1:D:172:MET:HB2	1.89	0.44
1:D:14:PHE:HE2	1:D:250:LEU:HD11	1.82	0.44

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:199:ALA:HB3	1:C:265:ILE:HB	2.00	0.44
1:D:90:VAL:HA	1:D:172:MET:HE1	1.98	0.43
1:A:199:ALA:HB3	1:A:265:ILE:HB	1.99	0.43
1:B:165:VAL:HG22	1:B:255:GLY:O	2.18	0.43
1:A:165:VAL:HG22	1:A:255:GLY:O	2.17	0.43
1:B:199:ALA:HB3	1:B:265:ILE:HB	2.00	0.43
1:C:165:VAL:HG22	1:C:255:GLY:O	2.19	0.43
1:A:51:THR:O	1:A:55:VAL:HG23	2.19	0.42
1:D:199:ALA:HB3	1:D:265:ILE:HB	2.02	0.42
1:A:200:GLY:HA3	1:A:263:GLU:O	2.20	0.42
1:D:207:ASN:ND2	1:D:238:GLN:HE21	2.17	0.42
1:A:97:LYS:HD3	1:A:97:LYS:C	2.40	0.42
1:A:33:ILE:HD11	1:A:188:LEU:HG	2.02	0.41
1:A:59:LEU:HG	1:A:113:MET:HE2	2.03	0.41
1:D:51:THR:O	1:D:55:VAL:HG23	2.20	0.41
1:D:94:PRO:CG	1:D:173:PRO:HB2	2.46	0.41
1:B:49:ARG:HG2	1:B:49:ARG:NH1	2.36	0.41
1:B:250:LEU:HD23	1:B:251:LEU:N	2.36	0.41
1:B:33:ILE:HD11	1:B:188:LEU:HG	2.03	0.41
1:D:220:ASP:OD2	1:D:223:LYS:HG3	2.21	0.40
1:A:36:ARG:NH1	1:A:256:ARG:HD3	2.37	0.40
1:D:195:THR:HG21	1:D:271:ARG:NH1	2.36	0.40
1:D:200:GLY:HA3	1:D:263:GLU:O	2.21	0.40

There are no symmetry-related clashes.

5.3 Torsion angles ⓘ

5.3.1 Protein backbone ⓘ

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	271/310 (87%)	261 (96%)	10 (4%)	0	100	100
1	B	269/310 (87%)	262 (97%)	7 (3%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	C	273/310 (88%)	266 (97%)	7 (3%)	0	100	100
1	D	270/310 (87%)	263 (97%)	7 (3%)	0	100	100
All	All	1083/1240 (87%)	1052 (97%)	31 (3%)	0	100	100

There are no Ramachandran outliers to report.

5.3.2 Protein sidechains ⓘ

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	234/265 (88%)	231 (99%)	3 (1%)	69	86
1	B	233/265 (88%)	229 (98%)	4 (2%)	60	82
1	C	235/265 (89%)	231 (98%)	4 (2%)	60	82
1	D	233/265 (88%)	230 (99%)	3 (1%)	69	86
All	All	935/1060 (88%)	921 (98%)	14 (2%)	65	84

All (14) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	57	ARG
1	A	97	LYS
1	A	277	LEU
1	B	23	LYS
1	B	76	GLU
1	B	218	SER
1	B	277	LEU
1	C	23	LYS
1	C	76	GLU
1	C	178	LEU
1	C	277	LEU
1	D	63	ASN
1	D	232	PHE
1	D	277	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (9) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	63	ASN
1	A	115	GLN
1	A	267	ASN
1	B	238	GLN
1	C	63	ASN
1	D	63	ASN
1	D	123	GLN
1	D	213	GLN
1	D	238	GLN

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 20 ligands modelled in this entry, 8 are monoatomic - leaving 12 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	$\#|Z| > 2$
4	AKG	B	303	2	9,9,9	1.06	1 (11%)	11,11,11	1.11	1 (9%)
4	AKG	D	303	2	9,9,9	1.50	2 (22%)	11,11,11	2.00	3 (27%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	#|Z| > 2	Counts	RMSZ	#|Z| > 2
4	AKG	C	303	2	9,9,9	2.11	4 (44%)	11,11,11	1.91	3 (27%)
5	OAU	C	304	-	22,25,25	0.97	1 (4%)	25,37,37	1.42	6 (24%)
7	GOL	B	305	-	5,5,5	0.24	0	5,5,5	0.50	0
6	SO4	B	307	-	4,4,4	0.23	0	6,6,6	0.40	0
6	SO4	B	306	-	4,4,4	0.27	0	6,6,6	0.18	0
6	SO4	A	305	-	4,4,4	0.24	0	6,6,6	0.58	0
5	OAU	A	304	-	22,25,25	0.80	1 (4%)	25,37,37	1.02	3 (12%)
5	OAU	B	304	-	22,25,25	0.94	1 (4%)	25,37,37	1.32	5 (20%)
5	OAU	D	304	-	22,25,25	0.86	1 (4%)	25,37,37	1.35	5 (20%)
4	AKG	A	303	2	9,9,9	0.59	0	11,11,11	1.55	1 (9%)

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	AKG	B	303	2	-	4/9/9/9	-
4	AKG	D	303	2	-	0/9/9/9	-
4	AKG	C	303	2	-	4/9/9/9	-
5	OAU	C	304	-	-	3/7/33/33	0/3/3/3
7	GOL	B	305	-	-	4/4/4/4	-
5	OAU	A	304	-	-	4/7/33/33	0/3/3/3
5	OAU	B	304	-	-	3/7/33/33	0/3/3/3
5	OAU	D	304	-	-	3/7/33/33	0/3/3/3
4	AKG	A	303	2	-	3/9/9/9	-

All (11) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	C	303	AKG	C2-C1	4.41	1.59	1.53
5	C	304	OAU	C4-C11	3.91	1.45	1.40
5	B	304	OAU	C4-C11	3.53	1.44	1.40
4	D	303	AKG	O1-C1	3.40	1.31	1.22
5	D	304	OAU	C4-C11	3.28	1.44	1.40
4	C	303	AKG	O3-C5	2.85	1.31	1.22
5	A	304	OAU	C4-C11	2.82	1.43	1.40
4	D	303	AKG	O2-C1	-2.55	1.23	1.30
4	C	303	AKG	O4-C5	-2.35	1.22	1.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	B	303	AKG	O3-C5	2.32	1.29	1.22
4	C	303	AKG	O5-C2	2.06	1.27	1.23

All (27) **bond angle outliers** are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	A	303	AKG	O5-C2-C1	4.66	126.18	119.43
4	D	303	AKG	O1-C1-C2	-4.26	116.04	121.72
4	C	303	AKG	O1-C1-C2	4.23	127.37	121.72
5	D	304	OAU	C19-C18-C20	-3.33	102.81	109.74
5	C	304	OAU	C19-C18-C20	-3.23	103.01	109.74
4	D	303	AKG	O5-C2-C1	3.10	123.92	119.43
4	C	303	AKG	O5-C2-C1	2.96	123.72	119.43
5	B	304	OAU	C19-C18-C20	-2.88	103.75	109.74
5	D	304	OAU	C17-C18-C2	2.83	113.69	108.13
4	D	303	AKG	O2-C1-C2	2.80	121.64	113.97
5	C	304	OAU	C17-C18-C2	2.74	113.52	108.13
5	B	304	OAU	C4-C3-C12	2.61	118.63	114.16
4	B	303	AKG	O5-C2-C1	2.57	123.15	119.43
5	B	304	OAU	C17-C18-C2	2.52	113.10	108.13
5	C	304	OAU	C3-C12-C13	2.51	117.93	113.48
5	A	304	OAU	C17-C18-C20	-2.49	104.38	110.25
5	C	304	OAU	C4-C3-C12	2.48	118.40	114.16
5	B	304	OAU	C16-C17-C18	2.33	116.22	112.64
5	C	304	OAU	C16-C17-C18	2.33	116.22	112.64
5	A	304	OAU	C17-C18-C2	2.23	112.51	108.13
5	D	304	OAU	C16-C17-C18	2.23	116.06	112.64
5	A	304	OAU	C8-C7-C6	-2.16	116.97	120.08
5	B	304	OAU	C4-C3-C2	-2.11	107.97	111.89
5	D	304	OAU	C8-C7-C6	-2.07	117.11	120.08
4	C	303	AKG	C3-C4-C5	2.05	118.01	113.60
5	D	304	OAU	C3-C12-C13	2.05	117.12	113.48
5	C	304	OAU	C4-C3-C2	-2.03	108.12	111.89

There are no chirality outliers.

All (28) **torsion outliers** are listed below:

Mol	Chain	Res	Type	Atoms
4	B	303	AKG	O2-C1-C2-C3
5	A	304	OAU	C2-C18-C20-C21
7	B	305	GOL	O1-C1-C2-C3
7	B	305	GOL	C1-C2-C3-O3

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms
7	B	305	GOL	O1-C1-C2-O2
7	B	305	GOL	O2-C2-C3-O3
5	A	304	OAU	C17-C18-C20-C21
4	B	303	AKG	C1-C2-C3-C4
4	B	303	AKG	O2-C1-C2-O5
4	C	303	AKG	C2-C3-C4-C5
5	A	304	OAU	C3-C12-C13-C14
5	C	304	OAU	C3-C12-C13-C14
5	C	304	OAU	C3-C12-C13-C15
5	A	304	OAU	C16-C12-C13-C14
5	B	304	OAU	C16-C12-C13-C15
5	C	304	OAU	C19-C18-C20-C21
5	D	304	OAU	C16-C12-C13-C14
5	B	304	OAU	C3-C12-C13-C14
5	D	304	OAU	C3-C12-C13-C14
4	C	303	AKG	C3-C4-C5-O4
4	C	303	AKG	C3-C4-C5-O3
4	A	303	AKG	C3-C4-C5-O3
4	A	303	AKG	C1-C2-C3-C4
4	C	303	AKG	C1-C2-C3-C4
5	B	304	OAU	C16-C12-C13-C14
5	D	304	OAU	C16-C12-C13-C15
4	B	303	AKG	O1-C1-C2-C3
4	A	303	AKG	C3-C4-C5-O4

There are no ring outliers.

No monomer is involved in short contacts.

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 B 303

Bond lengths

Bond angles

Torsions

Rings

Ligand AKG D 303

Bond lengths

Bond angles

Torsions

Rings

Full wwPDB X-ray Structure Validation Report

Ligand AKG C 303

Bond lengths

Bond angles

Torsions

Rings

Ligand OAU C 304

Bond lengths

Bond angles

Torsions

Rings

Ligand OAU A 304

Bond lengths

Bond angles

Torsions

Rings

Ligand OAU B 304

Bond lengths

Bond angles

Torsions

Rings

Ligand OAU D 304

Bond lengths

Bond angles

Torsions

Rings

Ligand AKG A 303

Bond lengths

Bond angles

Torsions

Rings

5.7 Other polymers ⓘ

There are no such residues in this entry.

5.8 Polymer linkage issues ⓘ

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, 95th 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	275/310 (88%)	0.63	11 (4%)	38	36	63, 91, 129, 149	0
1	B	275/310 (88%)	0.84	31 (11%)	5	4	66, 86, 146, 173	0
1	C	277/310 (89%)	0.66	18 (6%)	18	16	65, 90, 147, 179	0
1	D	274/310 (88%)	0.62	18 (6%)	18	16	61, 90, 132, 150	0
All	All	1101/1240 (88%)	0.69	78 (7%)	16	13	61, 89, 134, 179	0

All (78) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	221	ALA	8.5
1	B	225	ILE	6.9
1	B	169	PHE	5.9
1	B	224	TYR	5.6
1	B	228	VAL	5.2
1	C	219	ALA	4.9
1	B	230	SER	4.6
1	D	169	PHE	4.5
1	C	224	TYR	4.5
1	B	222	HIS	4.4
1	C	223	LYS	4.4
1	C	221	ALA	4.3
1	B	219	ALA	4.2
1	C	225	ILE	4.0
1	B	220	ASP	3.9
1	B	229	GLU	3.9
1	B	170	LEU	3.7
1	C	169	PHE	3.7
1	B	29	ILE	3.7
1	C	232	PHE	3.5
1	B	227	GLU	3.5

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	D	127	ILE	3.4
1	C	213	GLN	3.4
1	C	89	ILE	3.4
1	C	227	GLU	3.2
1	B	226	ASP	3.2
1	C	229	GLU	3.2
1	B	81	VAL	3.1
1	D	223	LYS	3.0
1	B	232	PHE	3.0
1	D	225	ILE	3.0
1	B	134	LEU	2.9
1	B	223	LYS	2.9
1	B	237	SER	2.9
1	A	279	PHE	2.8
1	A	171	LEU	2.8
1	D	258	TYR	2.8
1	B	236	GLN	2.8
1	A	169	PHE	2.8
1	A	286	ILE	2.8
1	B	233	LYS	2.8
1	B	279	PHE	2.7
1	D	232	PHE	2.7
1	A	170	LEU	2.7
1	C	170	LEU	2.7
1	A	287	TYR	2.7
1	B	218	SER	2.6
1	C	163	VAL	2.6
1	C	233	LYS	2.6
1	A	221	ALA	2.6
1	D	221	ALA	2.5
1	B	277	LEU	2.5
1	D	22	VAL	2.5
1	B	15	LEU	2.5
1	B	82	ALA	2.5
1	B	90	VAL	2.4
1	D	234	SER	2.4
1	D	182	LEU	2.3
1	B	12	LEU	2.3
1	C	184	LEU	2.3
1	A	77	PHE	2.2
1	B	208	LEU	2.2
1	A	277	LEU	2.2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	227	GLU	2.2
1	C	222	HIS	2.1
1	D	273	ILE	2.1
1	D	89	ILE	2.1
1	D	277	LEU	2.1
1	A	11	ALA	2.1
1	D	224	TYR	2.1
1	C	15	LEU	2.1
1	C	22	VAL	2.1
1	D	285	LYS	2.0
1	B	46	VAL	2.0
1	B	165	VAL	2.0
1	D	17	ILE	2.0
1	D	47	PHE	2.0
1	D	71	LEU	2.0

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, 95th 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
7	GOL	B	305	6/6	0.82	0.20	53,103,111,114	2
6	SO4	B	306	5/5	0.88	0.16	96,98,122,125	0
5	OAU	B	304	23/23	0.91	0.60	87,120,153,163	0
6	SO4	B	307	5/5	0.95	0.17	89,93,114,115	0
5	OAU	C	304	23/23	0.95	0.49	90,114,134,143	0
4	AKG	B	303	10/10	0.97	0.35	82,95,106,110	0
5	OAU	A	304	23/23	0.97	0.53	88,103,116,120	0
6	SO4	A	305	5/5	0.97	0.23	83,84,107,108	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å²)	Q<0.9
3	CL	B	302	1/1	0.98	0.38	129,129,129,129	0
5	OAU	D	304	23/23	0.98	0.43	88,98,109,110	0
4	AKG	C	303	10/10	0.98	0.28	91,100,118,127	0
4	AKG	A	303	10/10	0.99	0.30	76,93,130,164	0
2	ZN	B	301	1/1	0.99	0.18	87,87,87,87	0
2	ZN	C	301	1/1	0.99	0.18	90,90,90,90	0
4	AKG	D	303	10/10	0.99	0.33	74,84,99,100	0
2	ZN	A	301	1/1	0.99	0.21	81,81,81,81	0
3	CL	C	302	1/1	0.99	0.36	112,112,112,112	0
2	ZN	D	301	1/1	1.00	0.20	81,81,81,81	0
3	CL	A	302	1/1	1.00	0.18	97,97,97,97	0
3	CL	D	302	1/1	1.00	0.25	90,90,90,90	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.
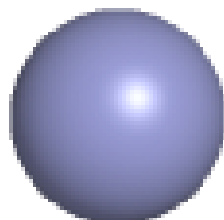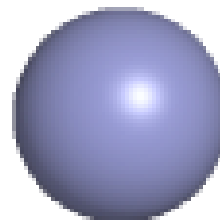
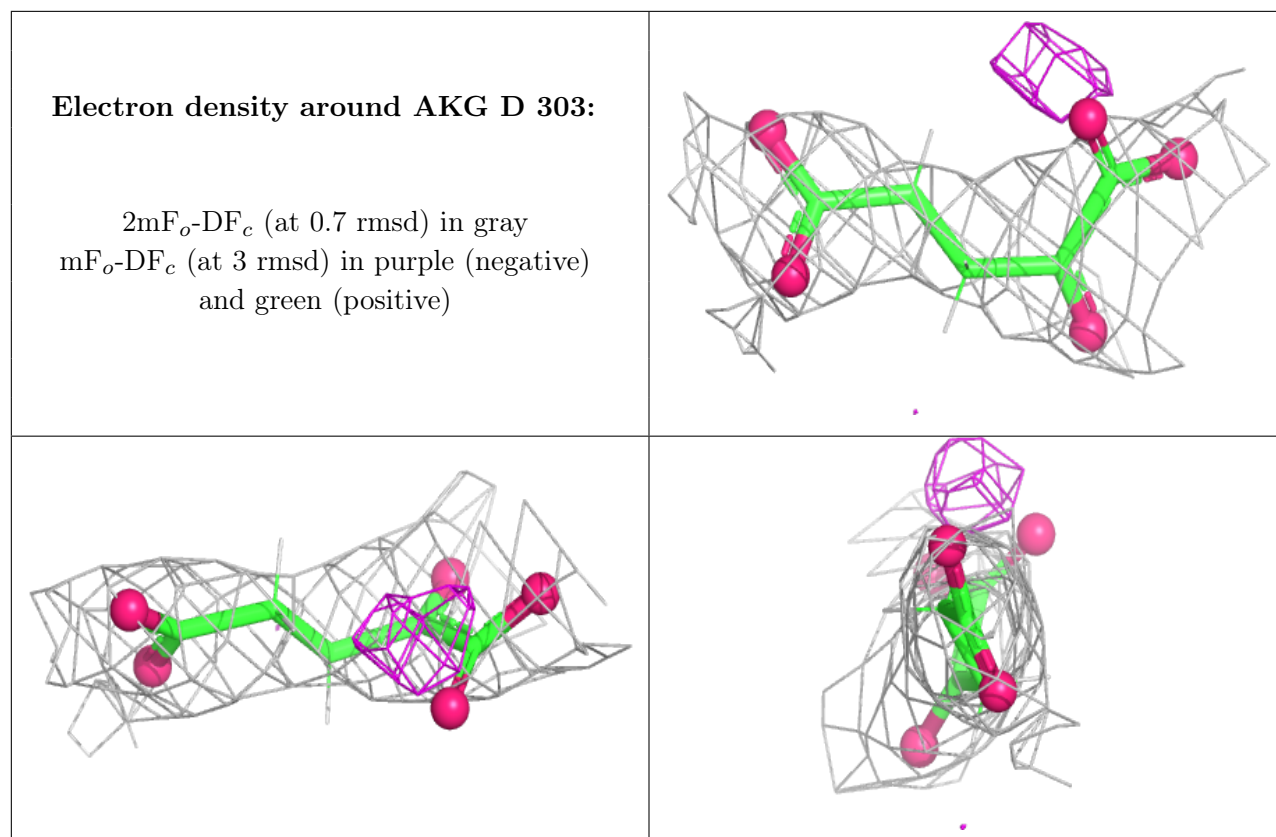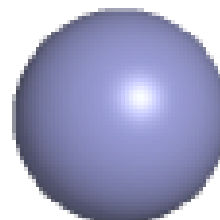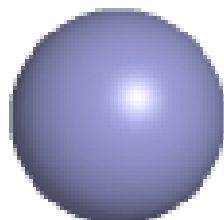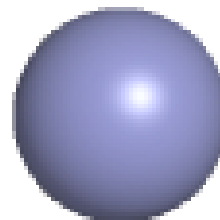**Electron density around OAU B 304:**

$2mF_o\text{-}DF_c$ (at 0.7 rmsd) in gray
$mF_o\text{-}DF_c$ (at 3 rmsd) in purple (negative)
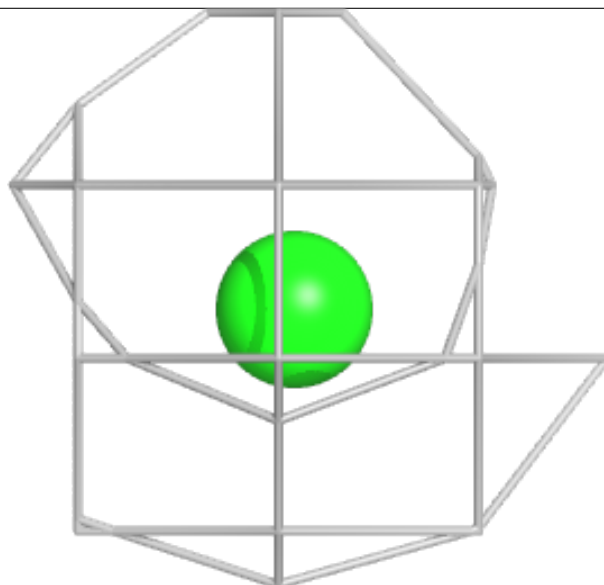and green (positive)
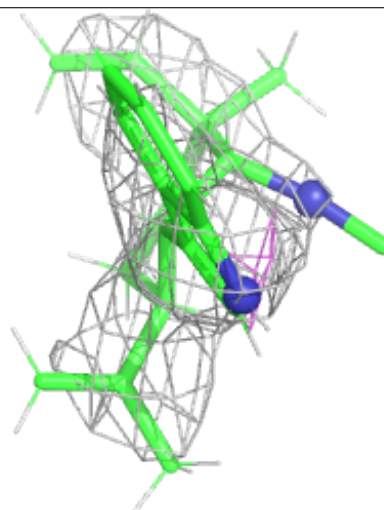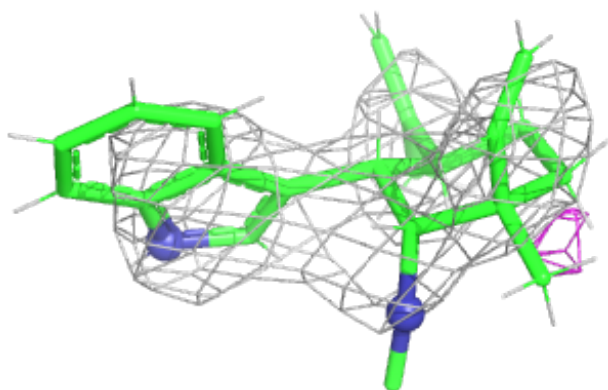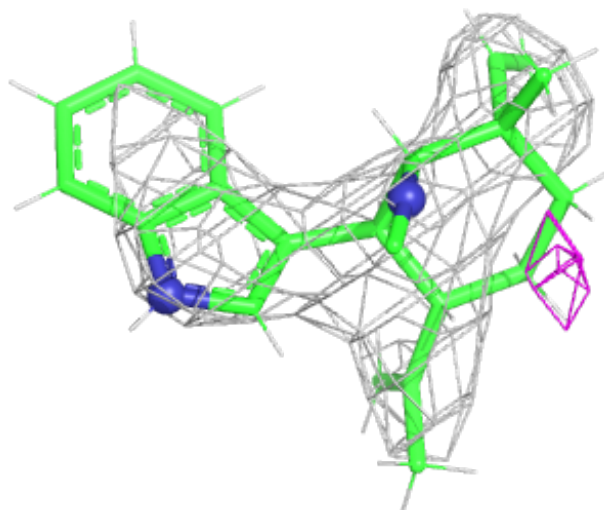
Electron density around OAU C 304:

$2mF_o\text{-}DF_c$ (at 0.7 rmsd) in gray
$mF_o\text{-}DF_c$ (at 3 rmsd) in purple (negative)
and green (positive)
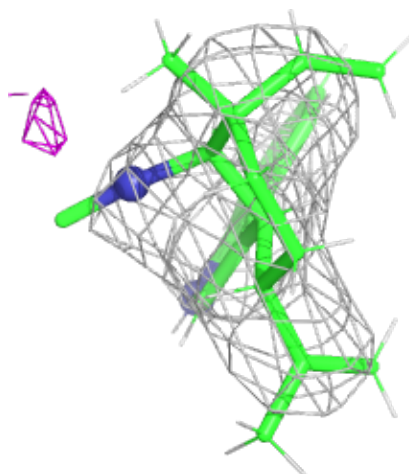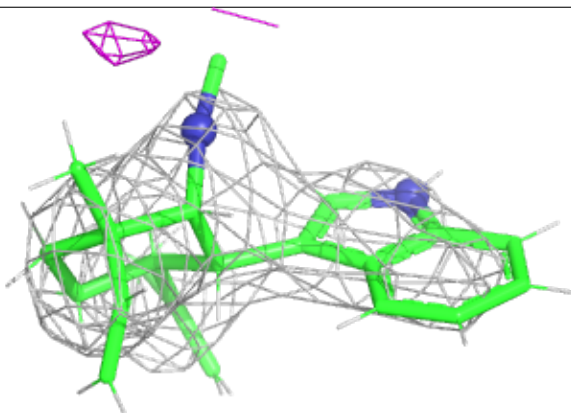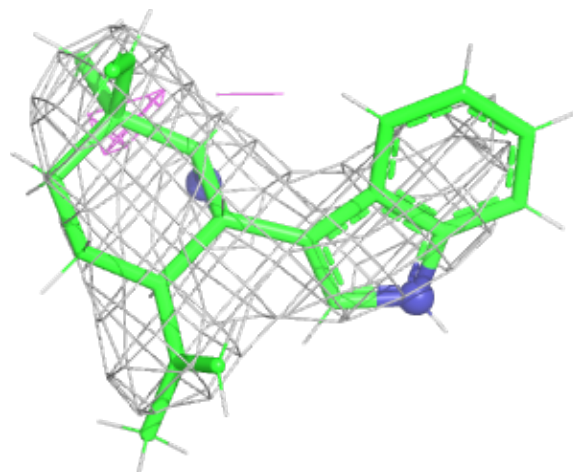
Electron density around AKG B 303:

$2mF_o\text{-}DF_c$ (at 0.7 rmsd) in gray
$mF_o\text{-}DF_c$ (at 3 rmsd) in purple (negative)
and green (positive)
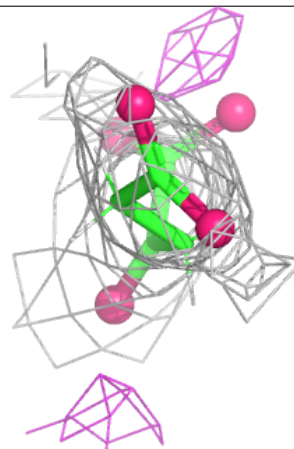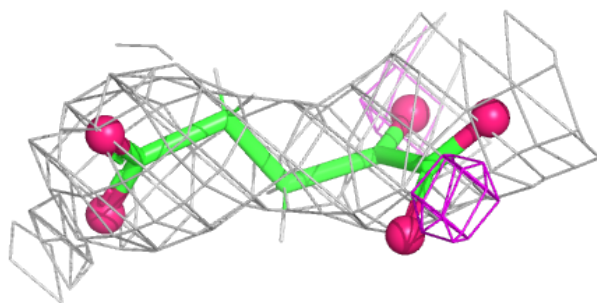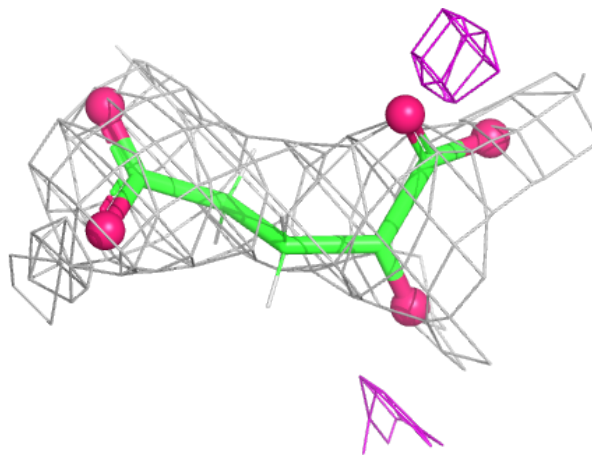
Electron density around OAU A 304:

$2mF_o\text{-}DF_c$ (at 0.7 rmsd) in gray
$mF_o\text{-}DF_c$ (at 3 rmsd) in purple (negative)
and green (positive)
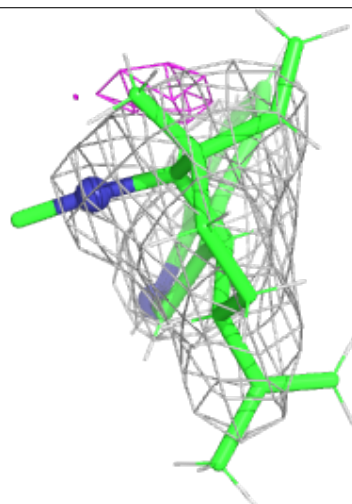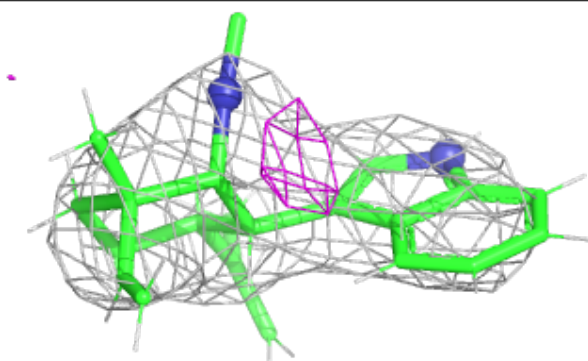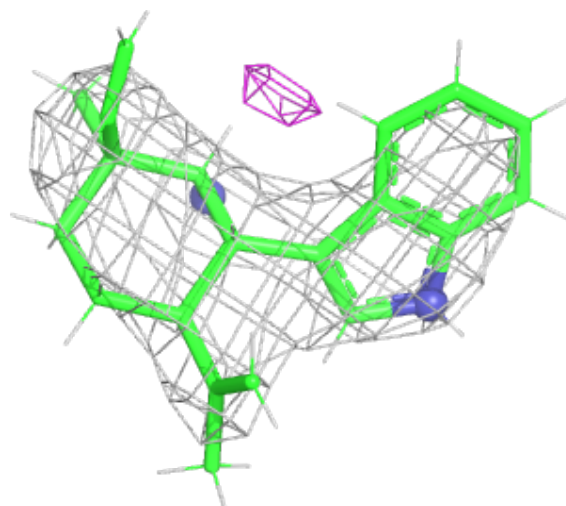
Electron density around CL B 302:

$2mF_o-DF_c$ (at 0.7 rmsd) in gray
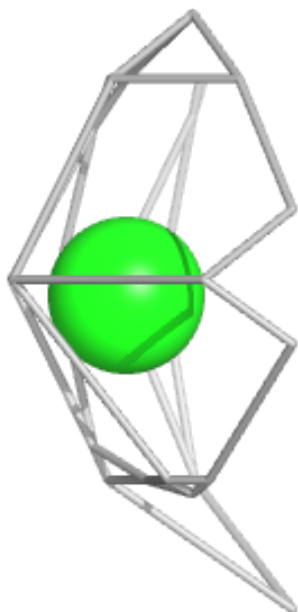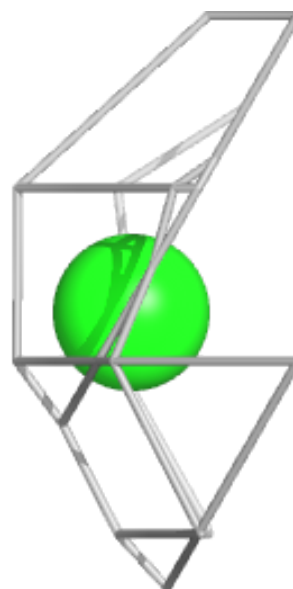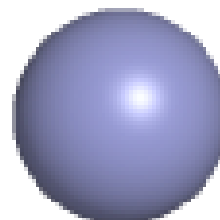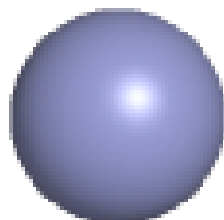mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)
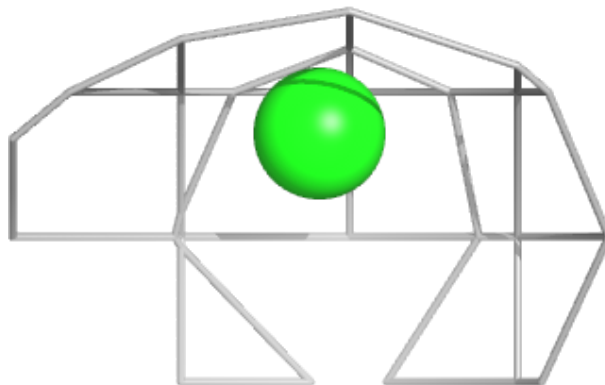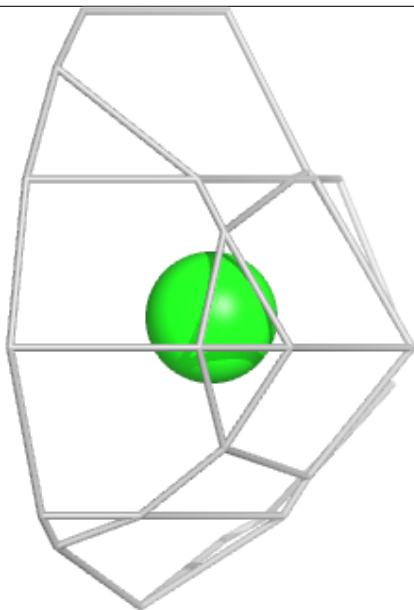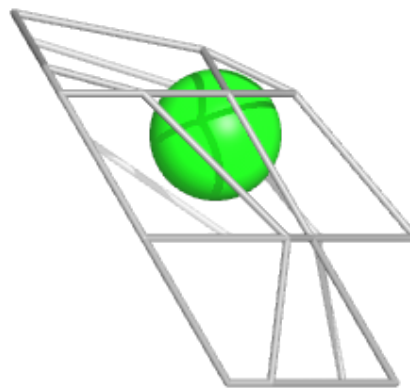
Electron density around OAU D 304:

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

Electron density around AKG 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)

Electron density around AKG A 303:

$2mF_o\text{-}DF_c$ (at 0.7 rmsd) in gray
$mF_o\text{-}DF_c$ (at 3 rmsd) in purple (negative)
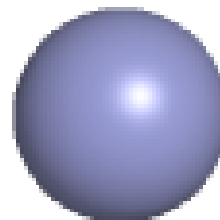and green (positive)
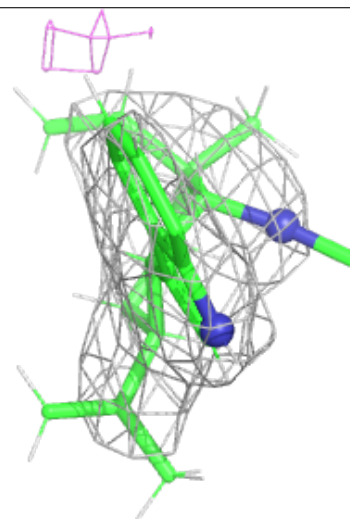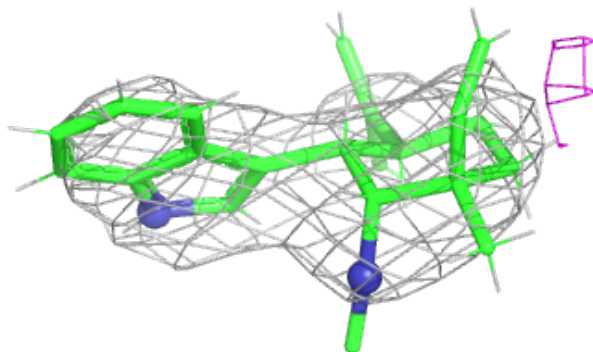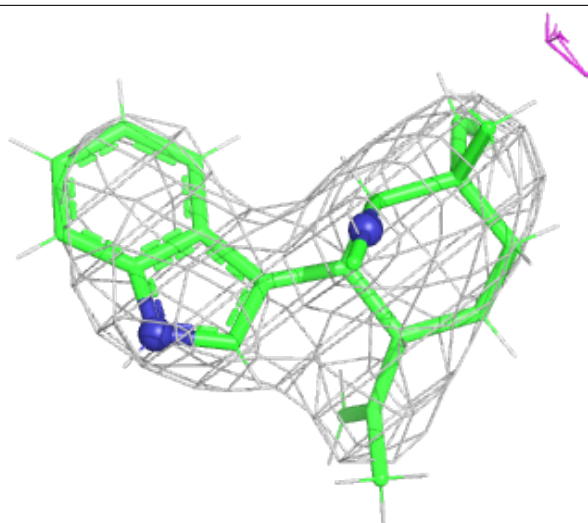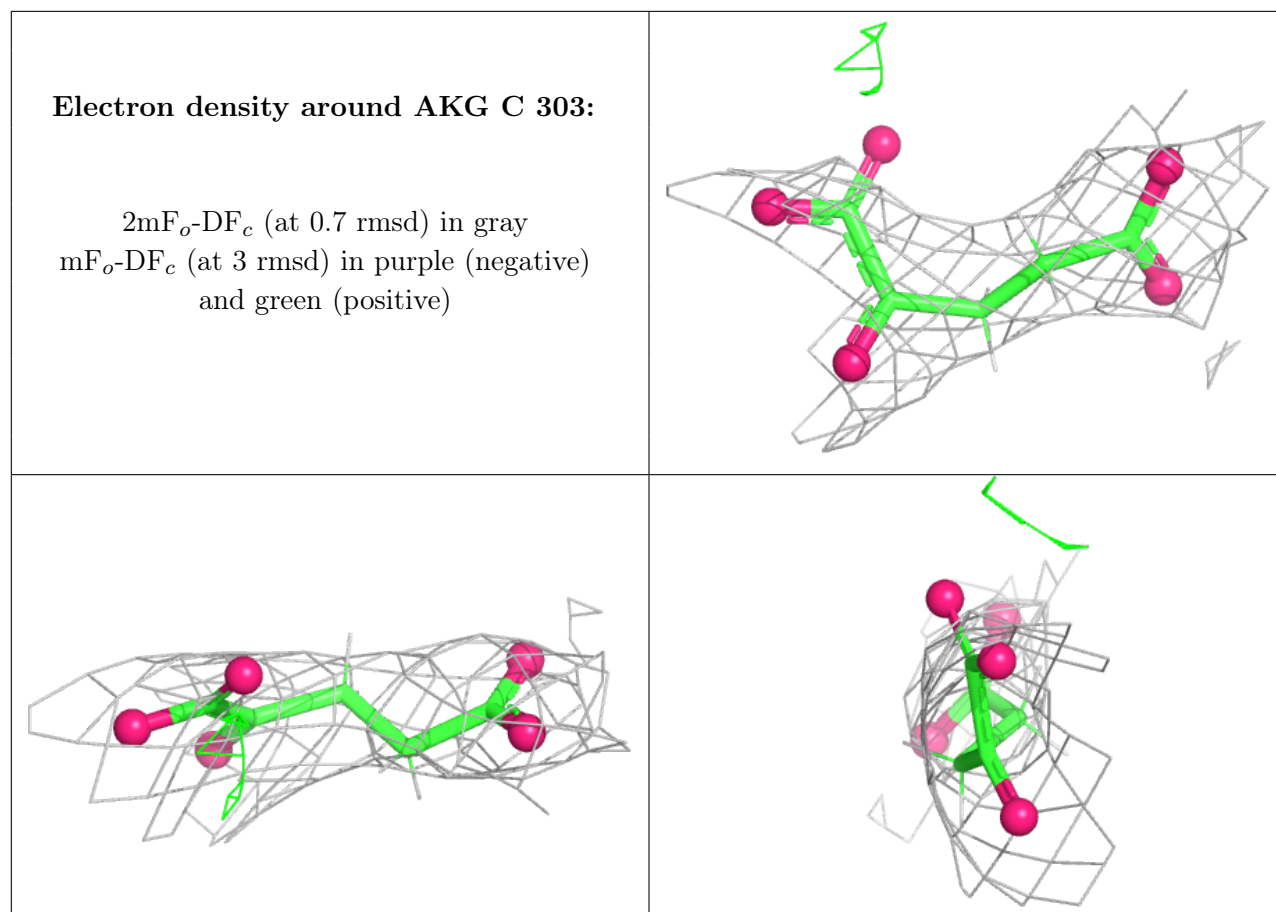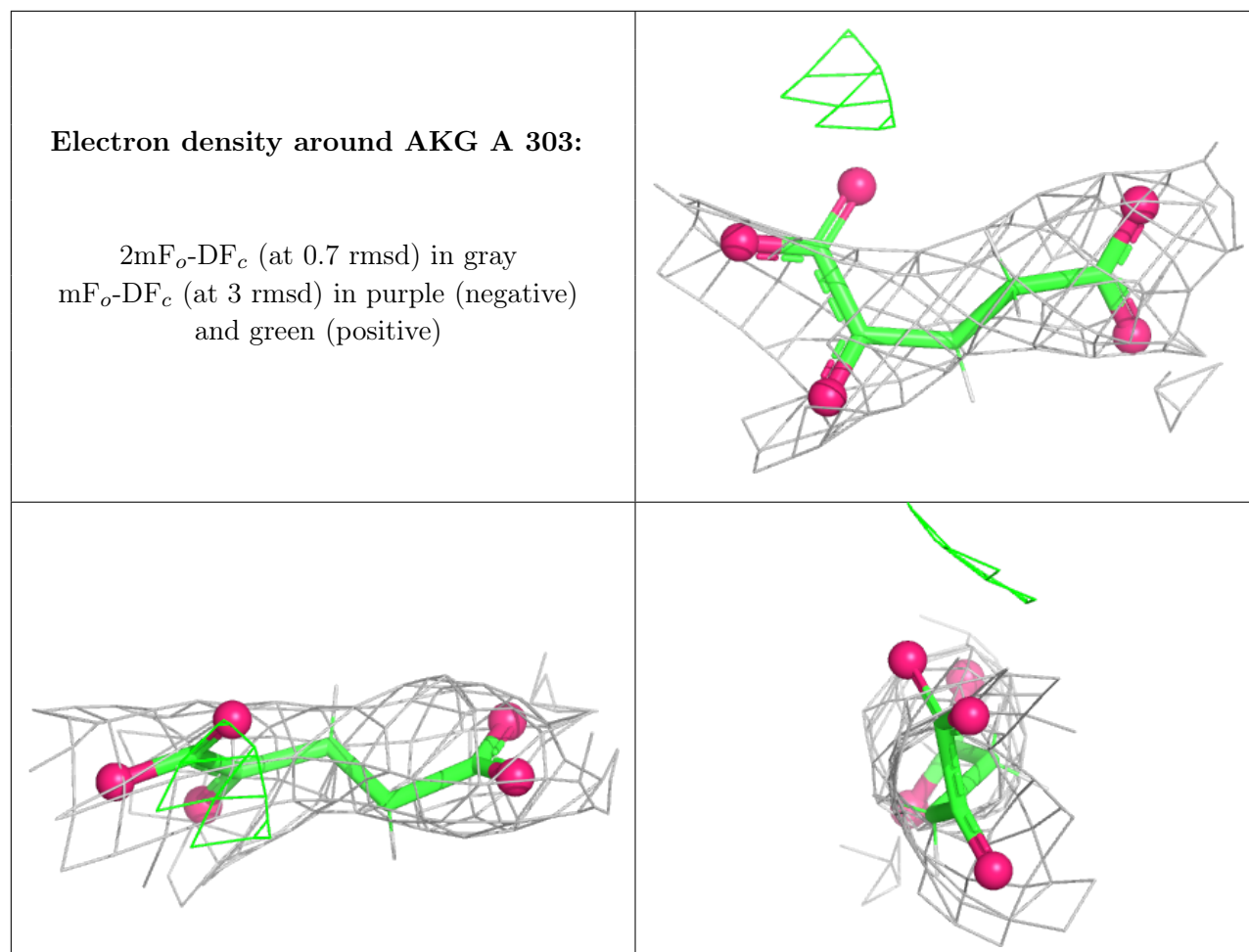
Electron density around ZN B 301:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)
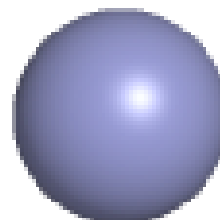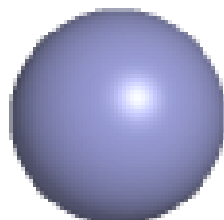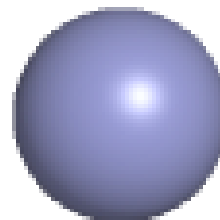
Electron density around ZN C 301:

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 AKG D 303:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative)
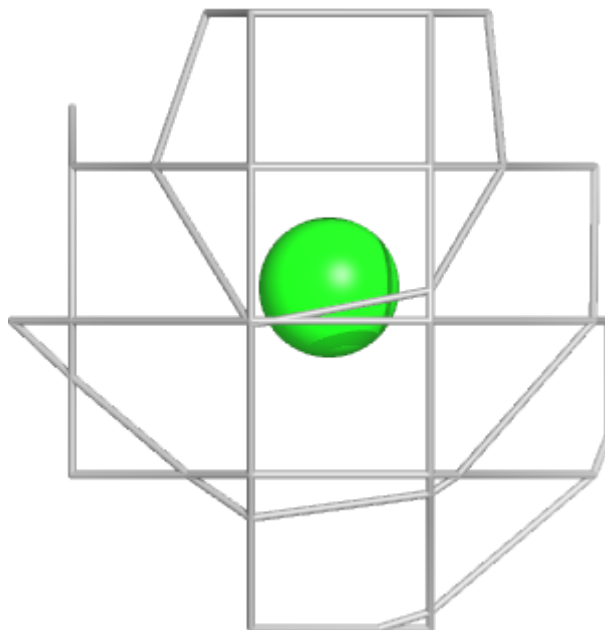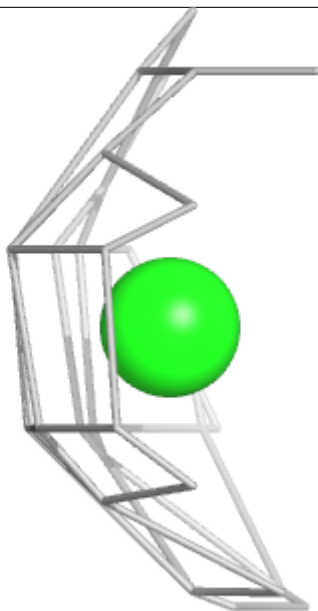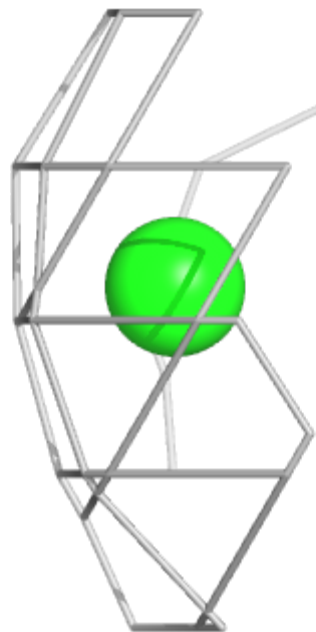and green (positive)
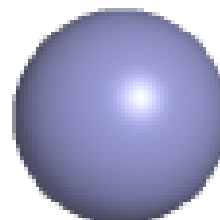
Electron density around ZN A 301:

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

Electron density around CL C 302:

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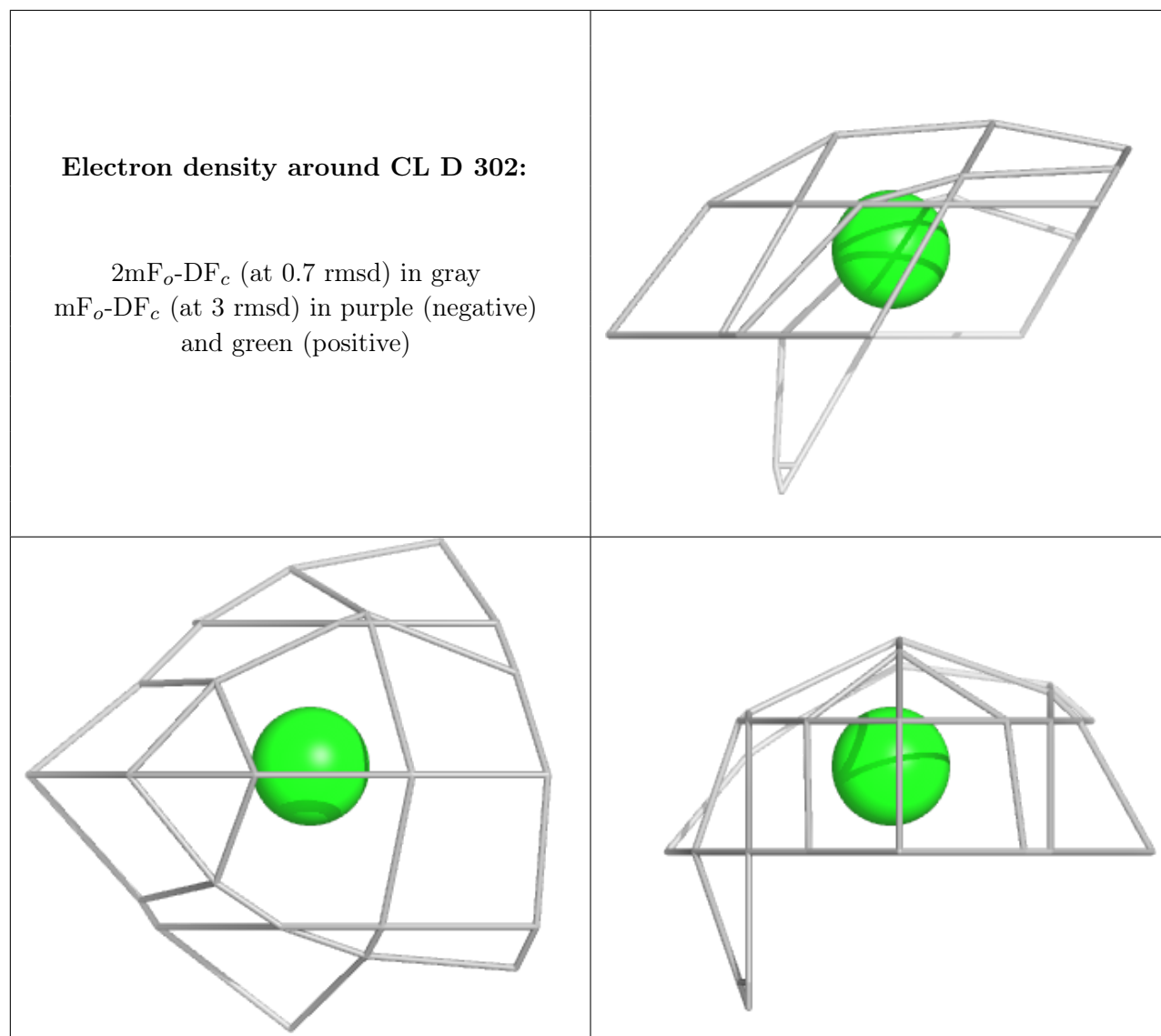and green (positive)

Electron density around ZN D 301:

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

Electron density around CL A 302:

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

Electron density around CL D 302:

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

6.5 Other polymers ⓘ

There are no such residues in this entry.