Full wwPDB X-ray Structure Validation Report ⓘ

Feb 1, 2024 – 12:08 PM JST

PDB ID	:	8HB2
Title	:	Crystal structure of Caenorhabditis elegans NMAD-1 in complex with ligand II
Authors	:	Shang, G.; Chen, Z.
Deposited on	:	2022-10-27
Resolution	:	3.06 Å(reported)

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

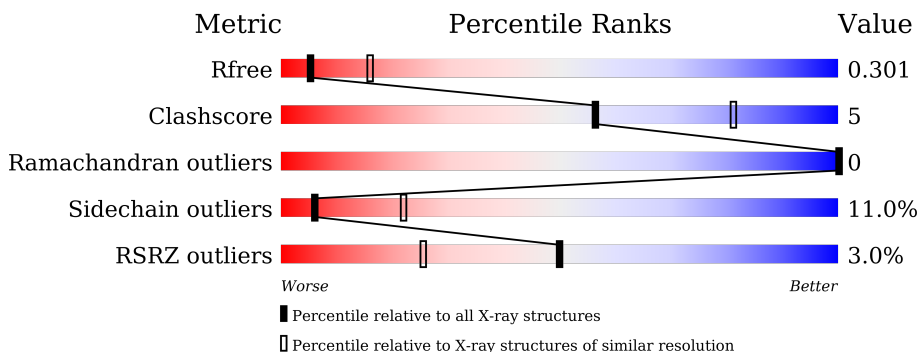|		
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh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

1 Overall quality at a glance ⓘ

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

The reported resolution of this entry is 3.06 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	Percentile Ranks	Value
Rfree		0.301
Clashscore		5
Ramachandran outliers		0
Sidechain outliers		11.0%
RSRZ outliers		3.0%

Worse — *Better*

▮ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	1754 (3.10-3.02)
Clashscore	141614	1864 (3.10-3.02)
Ramachandran outliers	138981	1794 (3.10-3.02)
Sidechain outliers	138945	1793 (3.10-3.02)
RSRZ outliers	127900	1713 (3.10-3.02)

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	305	% 65% 17% • 17%
1	B	305	% 69% 15% 16%
1	C	305	5% 67% 14% • 17%
1	D	305	3% 64% 19% • 16%

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
3	AKG	B	401	-	X	-	-

2 Entry composition ⓘ

There are 4 unique types of molecules in this entry. The entry contains 7972 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 DNA N6-methyl adenine demethylase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	254	Total	C	N	O	S	0	0	0
			1966	1236	337	377	16			
1	B	255	Total	C	N	O	S	0	0	0
			2010	1265	346	383	16			
1	C	254	Total	C	N	O	S	0	0	0
			1967	1235	336	380	16			
1	D	255	Total	C	N	O	S	0	0	0
			1976	1243	340	377	16			

There are 68 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-13	HIS	-	expression tag	UNP Q8MNT9
A	-12	HIS	-	expression tag	UNP Q8MNT9
A	-11	HIS	-	expression tag	UNP Q8MNT9
A	-10	HIS	-	expression tag	UNP Q8MNT9
A	-9	HIS	-	expression tag	UNP Q8MNT9
A	-8	HIS	-	expression tag	UNP Q8MNT9
A	-7	GLU	-	expression tag	UNP Q8MNT9
A	-6	ASN	-	expression tag	UNP Q8MNT9
A	-5	LEU	-	expression tag	UNP Q8MNT9
A	-4	TYR	-	expression tag	UNP Q8MNT9
A	-3	PHE	-	expression tag	UNP Q8MNT9
A	-2	GLN	-	expression tag	UNP Q8MNT9
A	-1	GLY	-	expression tag	UNP Q8MNT9
A	0	SER	-	expression tag	UNP Q8MNT9
A	109	LYS	GLU	engineered mutation	UNP Q8MNT9
A	112	LYS	GLN	engineered mutation	UNP Q8MNT9
A	114	LYS	GLN	engineered mutation	UNP Q8MNT9
B	-13	HIS	-	expression tag	UNP Q8MNT9
B	-12	HIS	-	expression tag	UNP Q8MNT9
B	-11	HIS	-	expression tag	UNP Q8MNT9
B	-10	HIS	-	expression tag	UNP Q8MNT9

Continued on next page...

W O R L D W I D E
wwPDB
PROTEIN DATA BANK

Continued from previous page...

| Chain	Residue	Modelled	Actual	Comment	Reference
B	-9	HIS	-	expression tag	UNP Q8MNT9
B	-8	HIS	-	expression tag	UNP Q8MNT9
B	-7	GLU	-	expression tag	UNP Q8MNT9
B	-6	ASN	-	expression tag	UNP Q8MNT9
B	-5	LEU	-	expression tag	UNP Q8MNT9
B	-4	TYR	-	expression tag	UNP Q8MNT9
B	-3	PHE	-	expression tag	UNP Q8MNT9
B	-2	GLN	-	expression tag	UNP Q8MNT9
B	-1	GLY	-	expression tag	UNP Q8MNT9
B	0	SER	-	expression tag	UNP Q8MNT9
B	109	LYS	GLU	engineered mutation	UNP Q8MNT9
B	112	LYS	GLN	engineered mutation	UNP Q8MNT9
B	114	LYS	GLN	engineered mutation	UNP Q8MNT9
C	-13	HIS	-	expression tag	UNP Q8MNT9
C	-12	HIS	-	expression tag	UNP Q8MNT9
C	-11	HIS	-	expression tag	UNP Q8MNT9
C	-10	HIS	-	expression tag	UNP Q8MNT9
C	-9	HIS	-	expression tag	UNP Q8MNT9
C	-8	HIS	-	expression tag	UNP Q8MNT9
C	-7	GLU	-	expression tag	UNP Q8MNT9
C	-6	ASN	-	expression tag	UNP Q8MNT9
C	-5	LEU	-	expression tag	UNP Q8MNT9
C	-4	TYR	-	expression tag	UNP Q8MNT9
C	-3	PHE	-	expression tag	UNP Q8MNT9
C	-2	GLN	-	expression tag	UNP Q8MNT9
C	-1	GLY	-	expression tag	UNP Q8MNT9
C	0	SER	-	expression tag	UNP Q8MNT9
C	109	LYS	GLU	engineered mutation	UNP Q8MNT9
C	112	LYS	GLN	engineered mutation	UNP Q8MNT9
C	114	LYS	GLN	engineered mutation	UNP Q8MNT9
D	-13	HIS	-	expression tag	UNP Q8MNT9
D	-12	HIS	-	expression tag	UNP Q8MNT9
D	-11	HIS	-	expression tag	UNP Q8MNT9
D	-10	HIS	-	expression tag	UNP Q8MNT9
D	-9	HIS	-	expression tag	UNP Q8MNT9
D	-8	HIS	-	expression tag	UNP Q8MNT9
D	-7	GLU	-	expression tag	UNP Q8MNT9
D	-6	ASN	-	expression tag	UNP Q8MNT9
D	-5	LEU	-	expression tag	UNP Q8MNT9
D	-4	TYR	-	expression tag	UNP Q8MNT9
D	-3	PHE	-	expression tag	UNP Q8MNT9
D	-2	GLN	-	expression tag	UNP Q8MNT9

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	-1	GLY	-	expression tag	UNP Q8MNT9
D	0	SER	-	expression tag	UNP Q8MNT9
D	109	LYS	GLU	engineered mutation	UNP Q8MNT9
D	112	LYS	GLN	engineered mutation	UNP Q8MNT9
D	114	LYS	GLN	engineered mutation	UNP Q8MNT9

- Molecule 2 is MANGANESE (II) ION (three-letter code: MN) (formula: Mn) (labeled as "Ligand of Interest" by depositor).

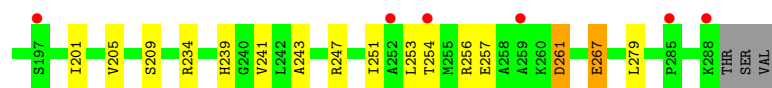| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	1	Total 1	Mn 1	0	0
2	B	1	Total 1	Mn 1	0	0
2	C	1	Total 1	Mn 1	0	0
2	D	1	Total 1	Mn 1	0	0

- Molecule 3 is 2-OXOGLUTARIC ACID (three-letter code: AKG) (formula: $C_5H_6O_5$) (labeled as "Ligand of Interest" by depositor).
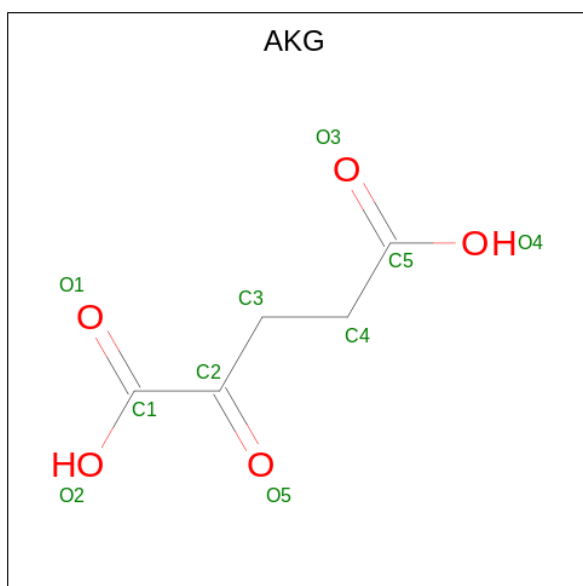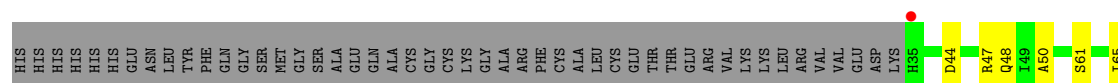

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	B	1	Total 10	C 5	O 5	0	0
3	C	1	Total 10	C 5	O 5	0	0

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	10	Total 10	O 10	0	0
4	B	16	Total 16	O 16	0	0
4	C	3	Total 3	O 3	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: DNA N6-methyl adenine demethylase

- Molecule 1: DNA N6-methyl adenine demethylase

- Molecule 1: DNA N6-methyl adenine demethylase

S197 I201 V205 S209 R234 H239 G240 V241 L242 A243 R247 I251 A252 L253 T254 M255 R256 E257 A258 A259 K260 D261 E267 L279 P285 K288 THR SER VAL

• Molecule 1: DNA N6-methyl adenine demethylase

Chain D:

3%

64% 19% • 16%

HIS HIS HIS HIS HIS HIS GLU ASN LEU TYR PHE GLN GLY SER MET GLY SER ALA GLU GLN ALA CYS GLY CYS LYS GLY ALA ARG PHE CYS ALA LEU CYS GLU THR THR GLU ARG VAL LYS LYS LEU ARG VAL VAL GLU ASP LYS H35 D44 R47 Q48 I49 A50 S61 I65

Q73 S74 V75 I81 E82 I83 D84 S96 K100 M104 I105 V108 S113 R118 G123 V126 N127 F128 V133 D136 T137 F138 V139 G140 M141 N150 S153 D156 V157 K158 K159 L160 F166 E167 M168 C169 V176 S179 Q185 D186

G192 I196 L200 I201 N202 M206 T207 L208 S214 V219 P222 H223 R224 S225 R234 K238 L242 A243 H244 R247 L253 E257 A263 Y269 E270 A274 R278 L279 R283 V284 P285 K288 T289 SER VAL

4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	180.24Å 75.39Å 117.22Å 90.00° 112.53° 90.00°	Depositor
Resolution (Å)	30.19 – 3.06 30.19 – 3.06	Depositor EDS
% Data completeness (in resolution range)	95.7 (30.19-3.06) 95.7 (30.19-3.06)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	3.09 (at 3.06Å)	Xtriage
Refinement program	PHENIX (1.18.2_3874: ???)	Depositor
R, R_{free}	0.275 , 0.296 0.288 , 0.301	Depositor DCC
R_{free} test set	1337 reflections (5.04%)	wwPDB-VP
Wilson B-factor (Å^2)	67.4	Xtriage
Anisotropy	0.231	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.31 , 36.3	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	0.86	EDS
Total number of atoms	7972	wwPDB-VP
Average B, all atoms (Å^2)	64.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.34% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[1]Intensities estimated from amplitudes.
[2]Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality ⓘ

5.1 Standard geometry ⓘ

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

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	0.57	0/2005	0.70	0/2717
1	B	0.36	0/2050	0.51	0/2770
1	C	0.57	0/2005	0.69	0/2715
1	D	0.48	0/2016	0.63	0/2733
All	All	0.50	0/8076	0.64	0/10935

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	1966	0	1836	17	4
1	B	2010	0	1923	24	3
1	C	1967	0	1850	12	1
1	D	1976	0	1859	24	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
2	C	1	0	0	1	0
2	D	1	0	0	0	0
3	B	10	0	4	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	C	10	0	4	1	0
4	A	10	0	0	0	0
4	B	16	0	0	2	0
4	C	3	0	0	0	0
All	All	7972	0	7476	77	4

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 (77) 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:D:200:LEU:HD11	1:D:253:LEU:HG	1.68	0.76
1:A:233:CYS:HA	1:A:237:TRP:HB2	1.79	0.65
1:D:206:MET:HB3	1:D:219:VAL:HG13	1.79	0.63
1:B:117:ARG:NH2	4:B:501:HOH:O	2.34	0.61
1:A:193:ASN:HD21	1:A:231:ASP:H	1.50	0.59
1:A:113:SER:HB2	1:A:119:LYS:HG2	1.84	0.58
1:D:47:ARG:NH2	1:D:73:GLN:O	2.37	0.58
1:B:98:GLU:HG3	1:B:200:LEU:HD22	1.86	0.58
1:B:114:LYS:NZ	4:B:509:HOH:O	2.27	0.58
1:C:138:PHE:HB3	1:C:279:LEU:HD13	1.87	0.56
1:D:47:ARG:NH1	1:D:75:VAL:O	2.33	0.56
1:B:57:ASN:H	1:B:60:SER:HB2	1.71	0.54
1:D:61:SER:O	1:D:65:ILE:HG13	2.06	0.54
1:D:50:ALA:HB3	1:D:81:ILE:HG13	1.90	0.53
1:A:192:GLY:HA3	1:A:257:GLU:HG3	1.91	0.52
1:A:156:ASP:HB3	1:A:159:LYS:HB2	1.91	0.52
1:C:180:ALA:HB2	1:C:243:ALA:HB2	1.91	0.52
1:D:208:LEU:HA	1:D:238:LYS:O	2.09	0.52
1:D:126:VAL:HG22	1:D:133:VAL:HG13	1.93	0.51
1:B:86:LEU:HD21	1:B:219:VAL:HG21	1.93	0.50
1:C:103:ASN:O	1:C:107:THR:HG23	2.11	0.50
1:B:120:GLN:HB2	1:B:170:ASN:HB2	1.93	0.50
1:D:44:ASP:O	1:D:48:GLN:N	2.45	0.50
1:D:138:PHE:CE2	1:D:279:LEU:HB3	2.47	0.50
1:B:89:ILE:HD12	1:B:226:LEU:HD22	1.93	0.49
1:A:62:LEU:HD12	1:A:205:VAL:HG23	1.95	0.49
1:B:185:GLN:HG3	1:B:238:LYS:HG2	1.95	0.49
1:A:77:THR:O	1:A:80:SER:HB3	2.13	0.48
1:D:141:MET:HB2	1:D:283:ARG:HH21	1.78	0.48

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:261:ASP:HB3	1:B:268:LEU:HD12	1.96	0.47
1:B:193:ASN:O	1:B:234:ARG:HD3	2.13	0.47
1:B:202:ASN:O	1:B:250:ARG:HG3	2.14	0.47
1:B:126:VAL:HG22	1:B:133:VAL:HG22	1.96	0.47
1:A:181:ILE:HG23	1:A:241:VAL:H	1.79	0.47
2:C:400:MN:MN	3:C:401:AKG:O5	1.56	0.47
1:D:185:GLN:HG3	1:D:238:LYS:HG3	1.97	0.47
1:A:126:VAL:HG13	1:A:133:VAL:HG22	1.97	0.46
1:B:156:ASP:HB3	1:B:159:LYS:HB2	1.97	0.46
1:C:62:LEU:HD12	1:C:205:VAL:HG23	1.98	0.46
1:C:156:ASP:HB3	1:C:159:LYS:HB2	1.97	0.46
1:C:102:LEU:HD23	1:C:102:LEU:HA	1.81	0.46
1:D:192:GLY:HA3	1:D:257:GLU:HG3	1.97	0.46
1:B:62:LEU:O	1:B:66:ILE:HG13	2.16	0.45
1:A:223:HIS:CD2	1:A:224:ARG:HG3	2.52	0.45
1:A:93:LEU:HD23	1:A:151:LYS:HE2	1.98	0.45
1:B:200:LEU:HB2	1:B:251:ILE:HB	1.99	0.45
1:C:192:GLY:HA3	1:C:257:GLU:HG3	1.98	0.45
1:D:104:MET:O	1:D:108:VAL:HG23	2.17	0.45
1:D:123:GLY:HA3	1:D:167:GLU:HG3	1.99	0.44
1:A:104:MET:O	1:A:108:VAL:HG23	2.16	0.44
1:C:261:ASP:HB3	1:C:267:GLU:HG3	2.00	0.44
1:A:89:ILE:HD12	1:A:226:LEU:HD22	1.99	0.43
1:B:275:GLU:CD	1:B:278:ARG:HH21	2.21	0.43
1:C:119:LYS:HE3	1:C:119:LYS:HB2	1.76	0.43
1:D:242:LEU:HD23	1:D:244:HIS:CE1	2.54	0.43
1:D:263:ALA:O	1:D:269:TYR:HB2	2.18	0.43
1:C:253:LEU:HD12	1:C:253:LEU:HA	1.84	0.43
1:A:193:ASN:ND2	1:A:231:ASP:H	2.15	0.42
1:A:140:GLY:HA2	1:A:279:LEU:O	2.19	0.42
1:B:196:ILE:O	1:B:254:THR:HA	2.20	0.42
1:C:107:THR:HG22	1:D:202:ASN:HD21	1.84	0.42
1:C:126:VAL:HG13	1:C:133:VAL:HG22	2.01	0.42
1:B:148:LEU:HD11	1:B:253:LEU:HD21	2.01	0.42
1:D:274:ALA:O	1:D:278:ARG:HD3	2.20	0.42
1:D:269:TYR:CZ	1:D:274:ALA:HB2	2.55	0.42
1:B:208:LEU:HA	1:B:238:LYS:O	2.20	0.42
1:B:152:MET:HB3	1:B:160:LEU:HB3	2.01	0.42
1:A:255:MET:HE3	1:A:255:MET:HB3	1.70	0.41
1:B:86:LEU:HD11	1:B:227:LEU:HD11	2.01	0.41
1:B:44:ASP:O	1:B:48:GLN:N	2.54	0.41

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:195:LEU:HD11	1:B:234:ARG:HD2	2.02	0.41
1:D:222:PRO:HG2	1:D:225:SER:HB3	2.01	0.41
1:A:261:ASP:O	1:A:266:GLY:HA3	2.21	0.41
1:B:192:GLY:N	1:B:257:GLU:O	2.54	0.41
1:D:44:ASP:O	1:D:48:GLN:HA	2.21	0.41
1:D:247:ARG:HD3	1:D:247:ARG:HA	1.81	0.40
1:D:44:ASP:O	1:D:48:GLN:CA	2.69	0.40

All (4) 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)
1:A:236:ASP:OD1	1:B:270:GLU:OE2[1_545]	1.68	0.52
1:A:278:ARG:NH1	1:C:79:GLY:O[1_545]	2.07	0.13
1:A:236:ASP:CG	1:B:270:GLU:OE2[1_545]	2.10	0.10
1:A:236:ASP:OD2	1:B:270:GLU:OE2[1_545]	2.11	0.09

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	252/305 (83%)	248 (98%)	4 (2%)	0	100	100
1	B	253/305 (83%)	249 (98%)	4 (2%)	0	100	100
1	C	252/305 (83%)	244 (97%)	8 (3%)	0	100	100
1	D	253/305 (83%)	249 (98%)	4 (2%)	0	100	100
All	All	1010/1220 (83%)	990 (98%)	20 (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	204/272 (75%)	174 (85%)	30 (15%)	3	11
1	B	216/272 (79%)	214 (99%)	2 (1%)	78	90
1	C	208/272 (76%)	172 (83%)	36 (17%)	2	7
1	D	208/272 (76%)	184 (88%)	24 (12%)	5	20
All	All	836/1088 (77%)	744 (89%)	92 (11%)	6	22

All (92) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	36	VAL
1	A	51	ILE
1	A	53	THR
1	A	54	THR
1	A	58	SER
1	A	60	SER
1	A	62	LEU
1	A	69	SER
1	A	70	THR
1	A	77	THR
1	A	96	SER
1	A	114	LYS
1	A	131	LYS
1	A	141	MET
1	A	153	SER
1	A	157	VAL
1	A	160	LEU
1	A	166	PHE
1	A	176	VAL
1	A	181	ILE
1	A	185	GLN
1	A	186	ASP
1	A	188	MET
1	A	201	ILE

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	233	CYS
1	A	241	VAL
1	A	253	LEU
1	A	255	MET
1	A	256	ARG
1	A	268	LEU
1	B	201	ILE
1	B	233	CYS
1	C	40	VAL
1	C	54	THR
1	C	58	SER
1	C	60	SER
1	C	62	LEU
1	C	73	GLN
1	C	74	SER
1	C	100	LYS
1	C	103	ASN
1	C	109	LYS
1	C	115	SER
1	C	119	LYS
1	C	135	THR
1	C	136	ASP
1	C	138	PHE
1	C	139	VAL
1	C	143	GLU
1	C	153	SER
1	C	157	VAL
1	C	164	GLN
1	C	168	MET
1	C	170	ASN
1	C	171	LEU
1	C	176	VAL
1	C	179	SER
1	C	201	ILE
1	C	209	SER
1	C	234	ARG
1	C	239	HIS
1	C	241	VAL
1	C	247	ARG
1	C	251	ILE
1	C	254	THR
1	C	256	ARG

Continued from previous page...

Mol	Chain	Res	Type
1	C	261	ASP
1	C	267	GLU
1	D	82	GLU
1	D	84	ASP
1	D	96	SER
1	D	100	LYS
1	D	105	ILE
1	D	113	SER
1	D	118	ARG
1	D	139	VAL
1	D	150	ASN
1	D	153	SER
1	D	156	ASP
1	D	157	VAL
1	D	159	LYS
1	D	160	LEU
1	D	166	PHE
1	D	176	VAL
1	D	179	SER
1	D	186	ASP
1	D	196	ILE
1	D	200	LEU
1	D	214	SER
1	D	219	VAL
1	D	223	HIS
1	D	234	ARG

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

Mol	Chain	Res	Type
1	A	193	ASN
1	A	223	HIS
1	A	245	HIS

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 6 ligands modelled in this entry, 4 are 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	#$|Z| > 2$
3	AKG	C	401	2	9,9,9	6.12	2 (22%)	11,11,11	1.72	2 (18%)
3	AKG	B	401	2	9,9,9	1.37	1 (11%)	11,11,11	1.65	4 (36%)

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
3	AKG	C	401	2	-	6/9/9/9	-
3	AKG	B	401	2	-	8/9/9/9	-

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	C	401	AKG	C2-C1	-18.04	1.29	1.53
3	C	401	AKG	O2-C1	-2.69	1.22	1.30
3	B	401	AKG	O5-C2	-2.30	1.18	1.23

All (6) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	C	401	AKG	C3-C2-C1	4.02	123.43	115.97
3	C	401	AKG	O1-C1-C2	-2.43	118.47	121.72
3	B	401	AKG	O2-C1-C2	2.42	120.59	113.97
3	B	401	AKG	C3-C2-C1	2.30	120.24	115.97
3	B	401	AKG	C4-C3-C2	-2.05	109.18	113.03
3	B	401	AKG	O4-C5-C4	2.01	120.50	114.03

There are no chirality outliers.

All (14) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	B	401	AKG	O1-C1-C2-O5
3	B	401	AKG	O1-C1-C2-C3
3	B	401	AKG	O2-C1-C2-C3
3	B	401	AKG	C1-C2-C3-C4
3	B	401	AKG	C2-C3-C4-C5
3	C	401	AKG	O1-C1-C2-O5
3	C	401	AKG	O1-C1-C2-C3
3	C	401	AKG	O2-C1-C2-C3
3	C	401	AKG	C2-C3-C4-C5
3	B	401	AKG	O5-C2-C3-C4
3	C	401	AKG	C3-C4-C5-O4
3	C	401	AKG	C3-C4-C5-O3
3	B	401	AKG	C3-C4-C5-O4
3	B	401	AKG	C3-C4-C5-O3

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	C	401	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 401	
Bond lengths	Bond angles
Torsions	Rings

Ligand AKG B 401	
Bond lengths	Bond angles
Torsions	Rings
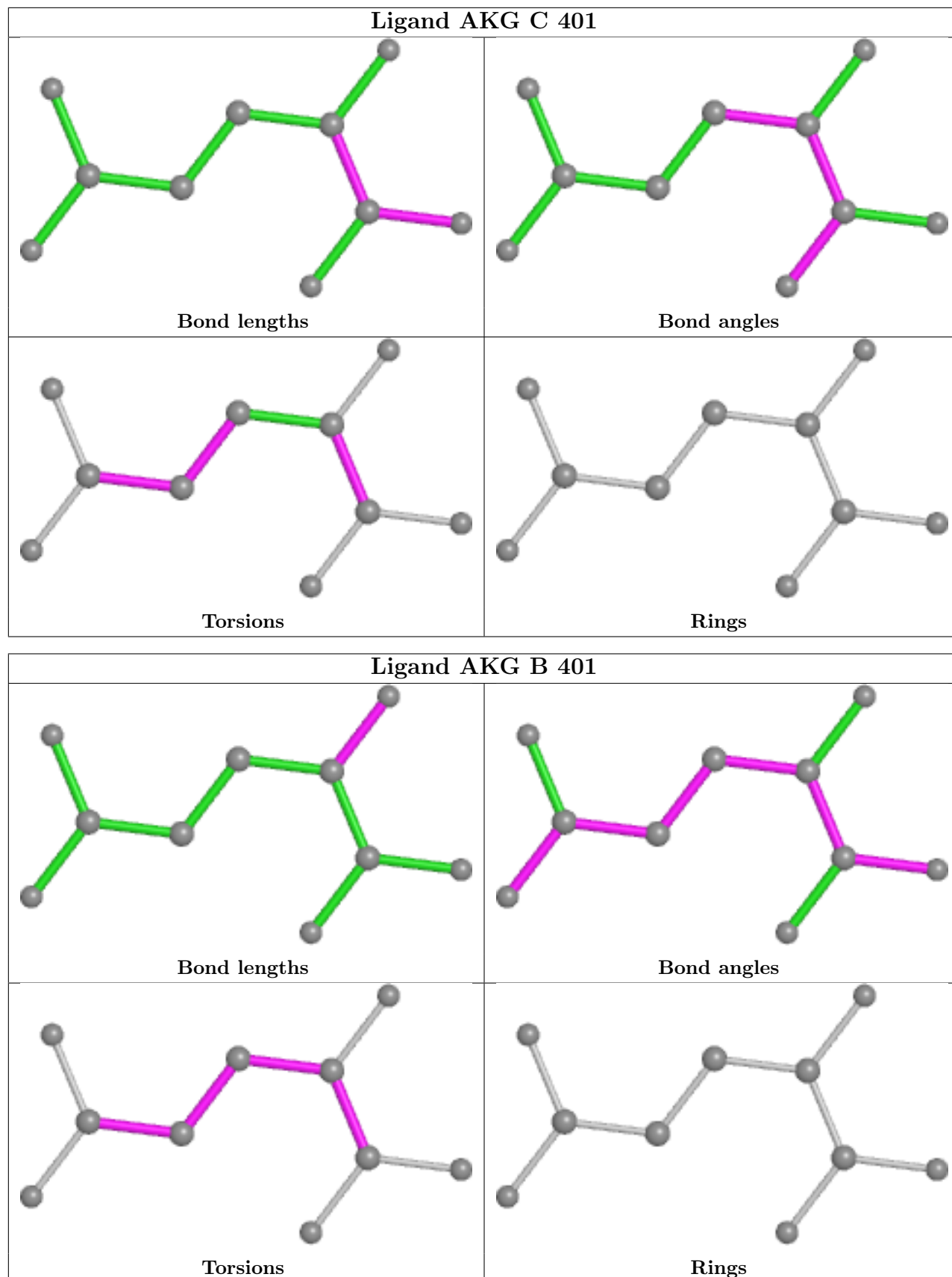
5.7 Other polymers ⓘ

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(Å²)	Q<0.9
1	A	254/305 (83%)	0.07	4 (1%)	72	49	46, 57, 71, 78	0
1	B	255/305 (83%)	-0.12	2 (0%)	86	70	36, 47, 60, 75	0
1	C	254/305 (83%)	0.32	15 (5%)	22	9	60, 74, 99, 107	0
1	D	255/305 (83%)	0.28	10 (3%)	39	19	59, 74, 97, 106	0
All	All	1018/1220 (83%)	0.14	31 (3%)	50	25	36, 65, 95, 107	0

All (31) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	130	HIS	3.6
1	C	113	SER	3.5
1	C	285	PRO	3.3
1	D	35	HIS	3.1
1	D	136	ASP	3.0
1	C	288	LYS	2.9
1	A	289	THR	2.9
1	C	36	VAL	2.8
1	B	35	HIS	2.8
1	D	270	GLU	2.7
1	C	197	SER	2.7
1	C	259	ALA	2.7
1	C	169	CYS	2.6
1	D	285	PRO	2.6
1	C	112	LYS	2.6
1	C	254	THR	2.6
1	C	111	ALA	2.6
1	D	128	PHE	2.5
1	A	36	VAL	2.4
1	C	188	MET	2.4
1	D	288	LYS	2.3

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	D	137	THR	2.2
1	B	178	LYS	2.2
1	C	135	THR	2.2
1	D	289	THR	2.2
1	D	169	CYS	2.2
1	A	236	ASP	2.1
1	D	185	GLN	2.1
1	C	252	ALA	2.1
1	A	254	THR	2.1
1	C	127	ASN	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
2	MN	C	400	1/1	0.90	0.12	78,78,78,78	0
2	MN	A	301	1/1	0.91	0.14	50,50,50,50	0
2	MN	D	301	1/1	0.94	0.25	76,76,76,76	0
3	AKG	C	401	10/10	0.96	0.24	73,73,73,73	0
2	MN	B	400	1/1	0.97	0.16	40,40,40,40	0
3	AKG	B	401	10/10	0.98	0.21	43,43,43,43	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.
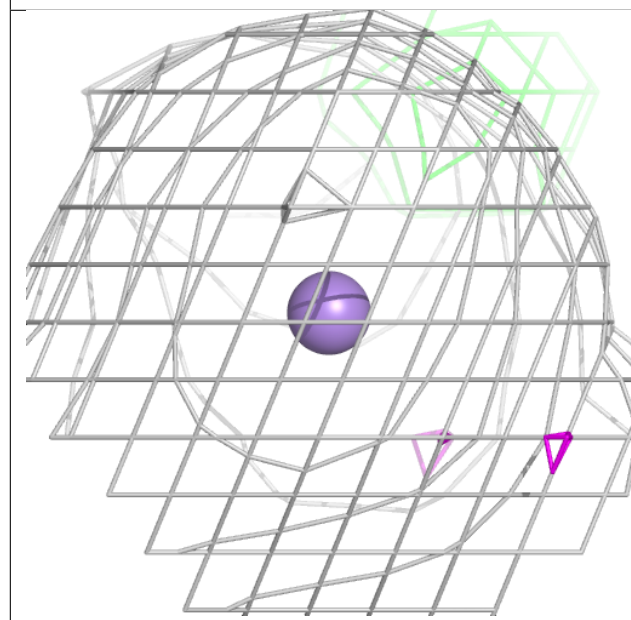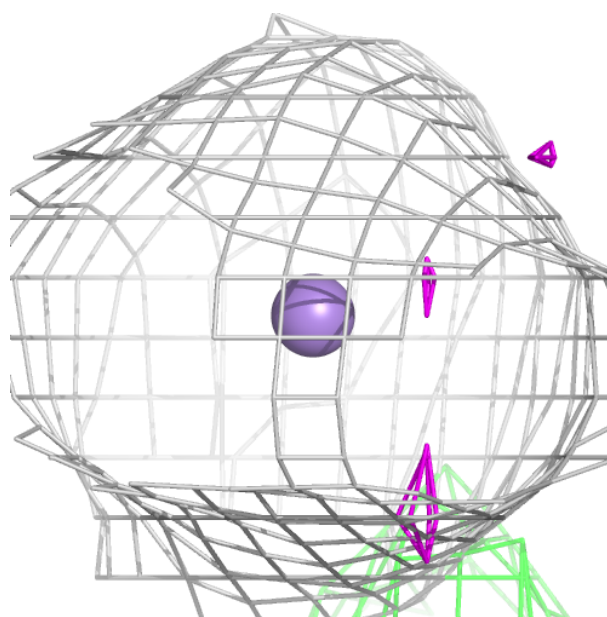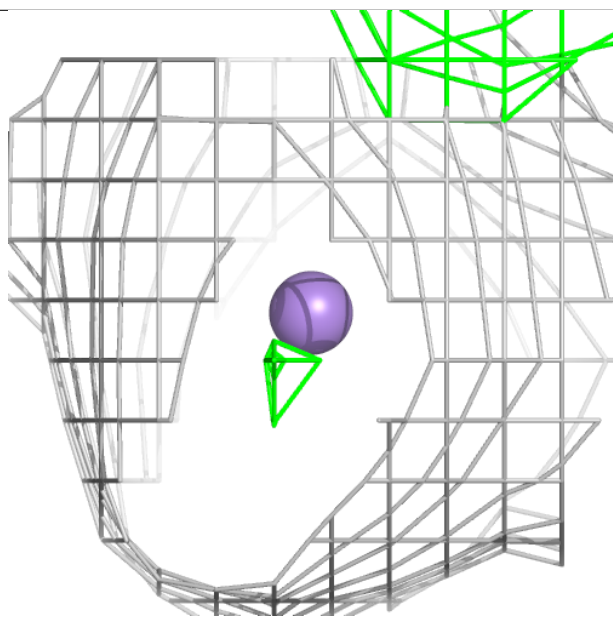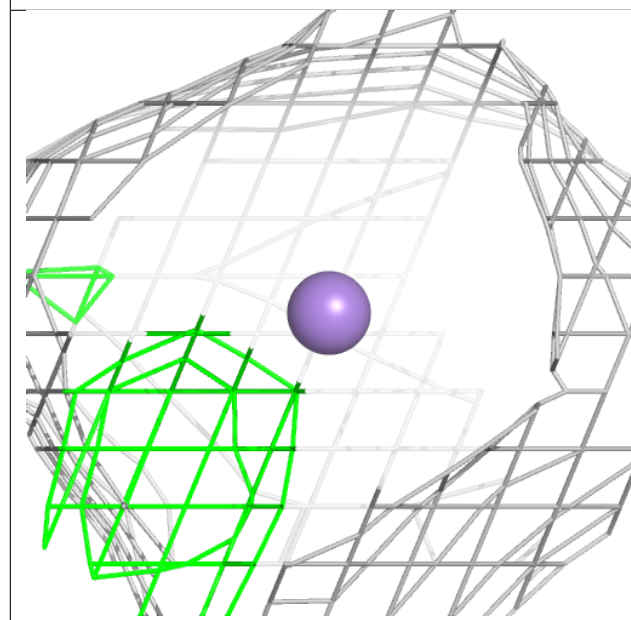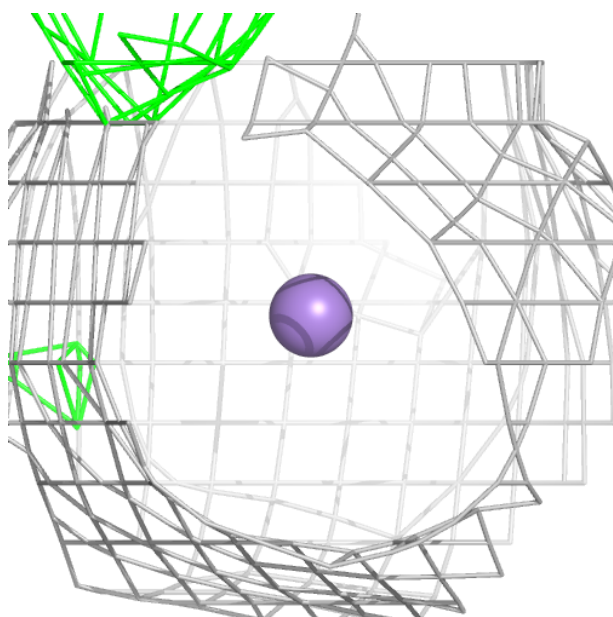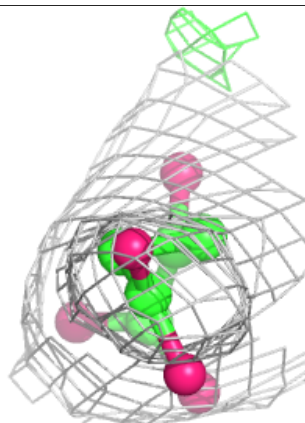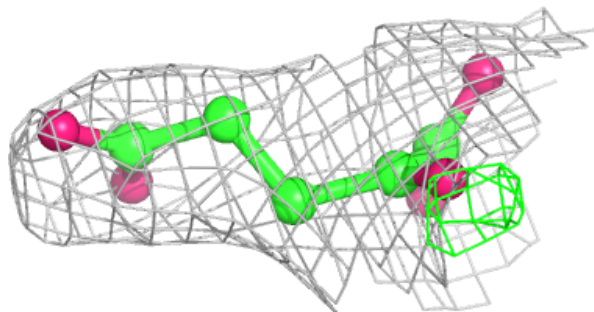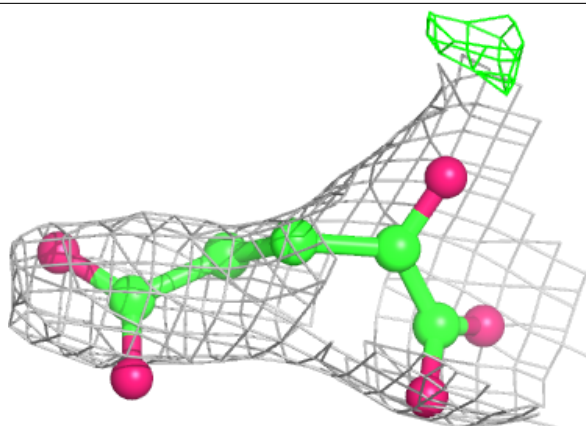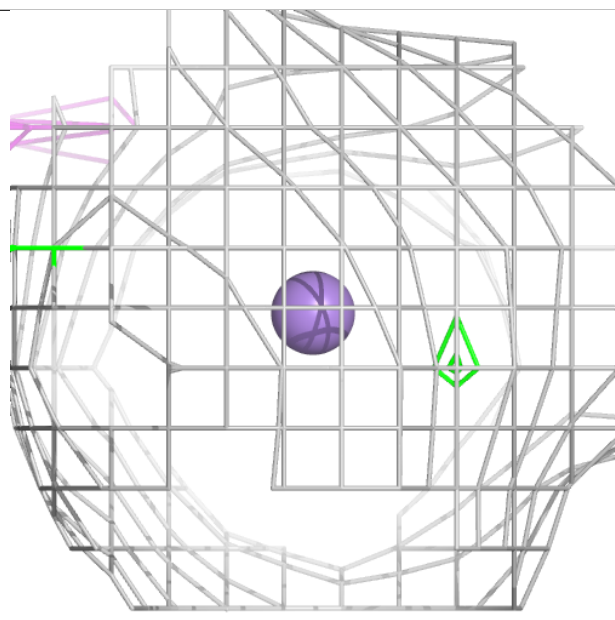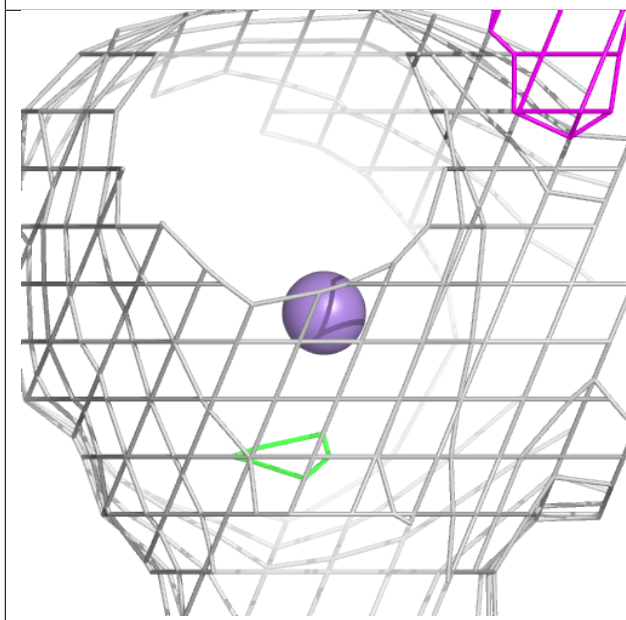
Electron density around MN C 400:

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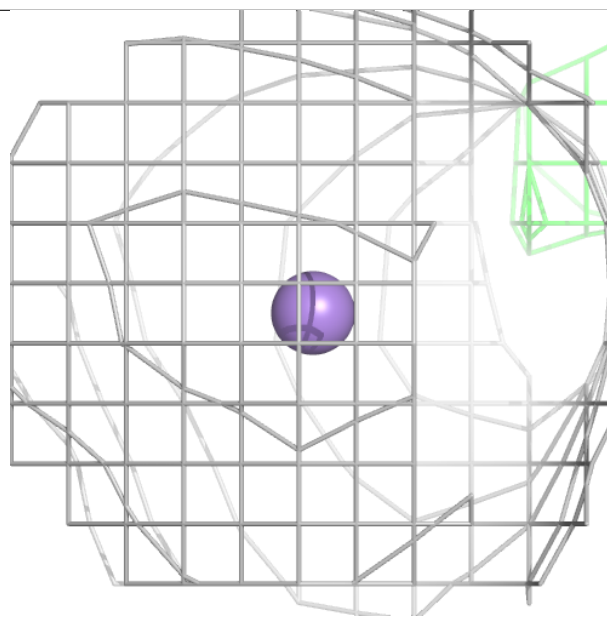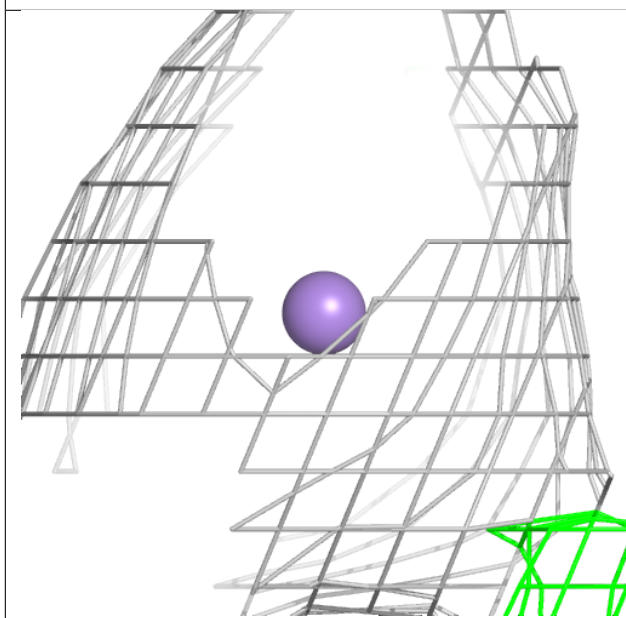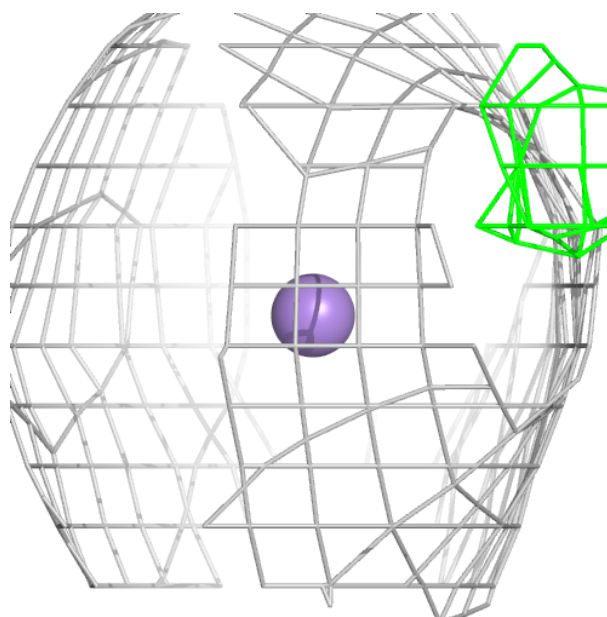
Electron density around MN A 301:

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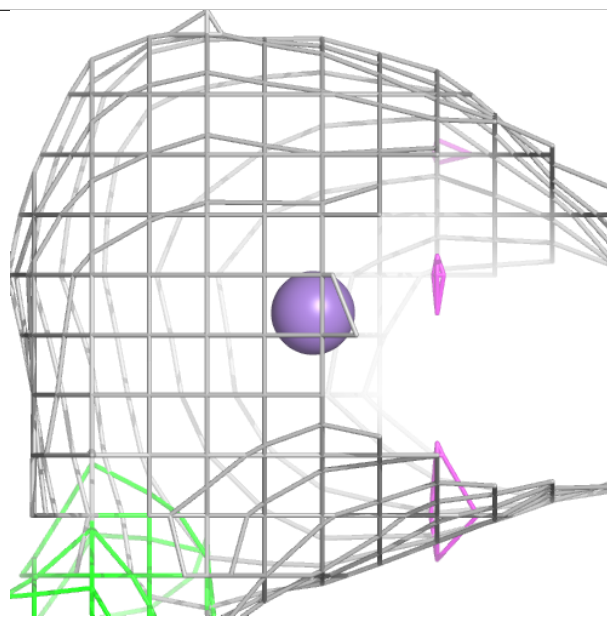
Electron density around MN D 301:

2mF$_o$-DF$_c$ (at 0.7 rmsd) in gray
mF$_o$-DF$_c$ (at 3 rmsd) in purple (negative)
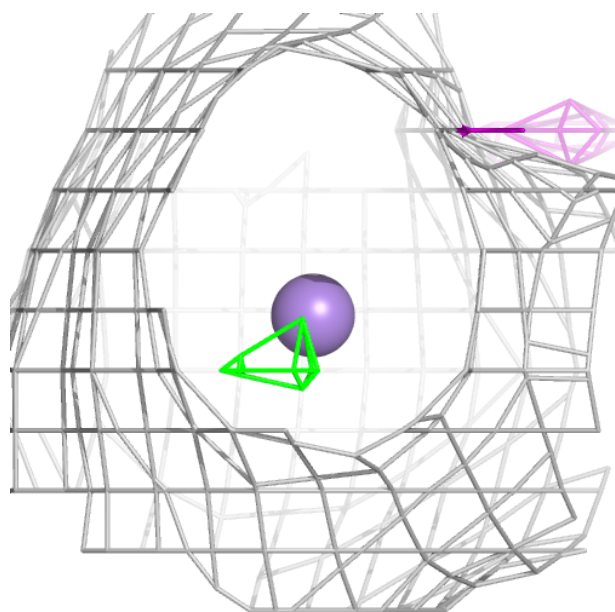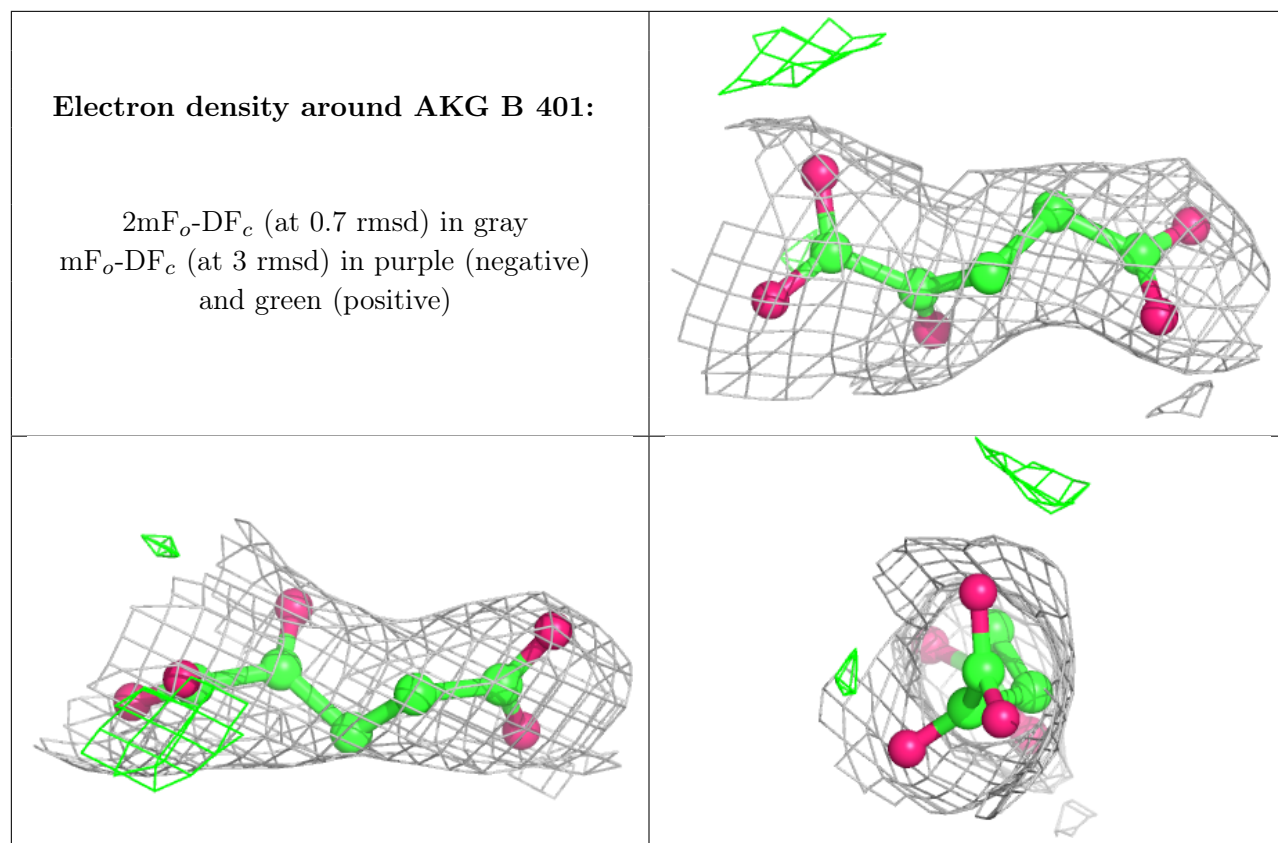and green (positive)

Electron density around AKG C 401:

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 MN B 400:

$2mF_o$-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 B 401:

$2mF_o$-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 ⓘ

There are no such residues in this entry.