



# wwPDB X-ray Structure Validation Summary Report ⓘ

Apr 28, 2022 – 10:22 pm BST

PDB ID : 7OVA  
Title : Crystal structure of an AA9 LPMO  
Authors : Males, A.; Correa, T.L.R.; Murakami, M.T.; Walton, P.H.; Davies, G.J.  
Deposited on : 2021-06-14  
Resolution : 1.75 Å(reported)

This is a wwPDB X-ray Structure Validation Summary Report for a publicly released PDB entry.

We welcome your comments at [validation@mail.wwpdb.org](mailto:validation@mail.wwpdb.org)

A user guide is available at

<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

---

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

MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.28.1
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)	:	2.28.1

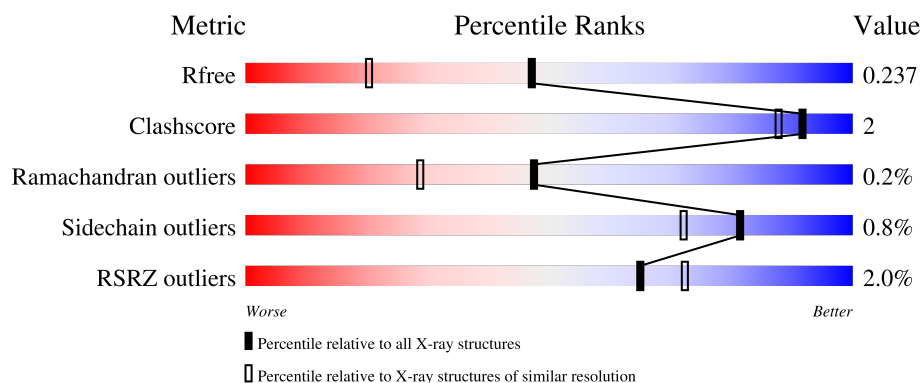
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.75 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	2340 (1.76-1.76)
Clashscore	141614	2466 (1.76-1.76)
Ramachandran outliers	138981	2437 (1.76-1.76)
Sidechain outliers	138945	2437 (1.76-1.76)
RSRZ outliers	127900	2298 (1.76-1.76)

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  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions  $\leq 5\%$ . 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	AAA	234	<div> <div>2%</div> <div> <div></div> <div>92%</div> <div>6%</div> </div> <div>.</div> </div>
1	BBB	234	<div> <div>%</div> <div> <div></div> <div>91%</div> <div>6%</div> </div> <div>.</div> </div>
1	CCC	234	<div> <div>%</div> <div> <div></div> <div>94%</div> <div>.</div> </div> <div>.</div> </div>
1	DDD	234	<div> <div>3%</div> <div> <div></div> <div>92%</div> <div>7%</div> </div> <div>.</div> </div>

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 14017 atoms, of which 6469 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 Endoglucanase, putative.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	AAA	229	Total	C	H	N	O	S	67	2	0
			3329	1078	1613	283	348	7			
1	BBB	226	Total	C	H	N	O	S	64	2	0
			3302	1068	1605	279	343	7			
1	CCC	228	Total	C	H	N	O	S	64	0	0
			3300	1067	1602	280	344	7			
1	DDD	231	Total	C	H	N	O	S	71	2	0
			3374	1091	1637	290	349	7			

There are 24 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
AAA	229	HIS	-	expression tag	UNP A1D2G7
AAA	230	HIS	-	expression tag	UNP A1D2G7
AAA	231	HIS	-	expression tag	UNP A1D2G7
AAA	232	HIS	-	expression tag	UNP A1D2G7
AAA	233	HIS	-	expression tag	UNP A1D2G7
AAA	234	HIS	-	expression tag	UNP A1D2G7
BBB	229	HIS	-	expression tag	UNP A1D2G7
BBB	230	HIS	-	expression tag	UNP A1D2G7
BBB	231	HIS	-	expression tag	UNP A1D2G7
BBB	232	HIS	-	expression tag	UNP A1D2G7
BBB	233	HIS	-	expression tag	UNP A1D2G7
BBB	234	HIS	-	expression tag	UNP A1D2G7
CCC	229	HIS	-	expression tag	UNP A1D2G7
CCC	230	HIS	-	expression tag	UNP A1D2G7
CCC	231	HIS	-	expression tag	UNP A1D2G7
CCC	232	HIS	-	expression tag	UNP A1D2G7
CCC	233	HIS	-	expression tag	UNP A1D2G7
CCC	234	HIS	-	expression tag	UNP A1D2G7
DDD	229	HIS	-	expression tag	UNP A1D2G7
DDD	230	HIS	-	expression tag	UNP A1D2G7
DDD	231	HIS	-	expression tag	UNP A1D2G7

*Continued on next page...*

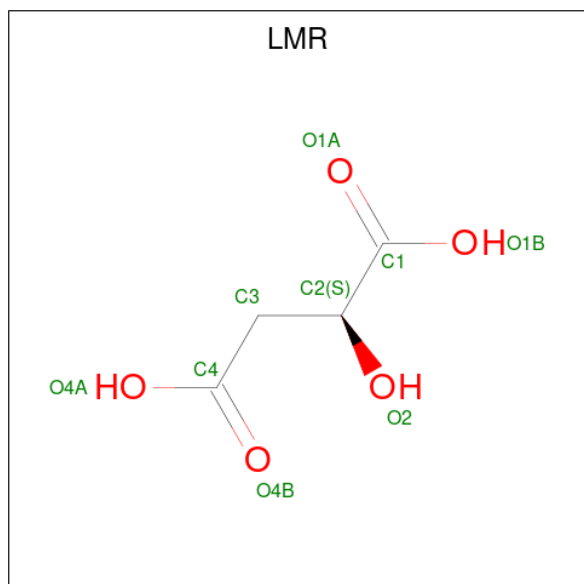
Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
DDD	232	HIS	-	expression tag	UNP A1D2G7
DDD	233	HIS	-	expression tag	UNP A1D2G7
DDD	234	HIS	-	expression tag	UNP A1D2G7

- Molecule 2 is COPPER (II) ION (three-letter code: CU) (formula: Cu) (labeled as "Ligand of Interest" by depositor).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	AAA	1	Total Cu 1 1	0	0
2	BBB	1	Total Cu 1 1	0	0
2	CCC	1	Total Cu 1 1	0	0
2	DDD	1	Total Cu 1 1	0	0

- Molecule 3 is (2S)-2-hydroxybutanedioic acid (three-letter code: LMR) (formula: C<sub>4</sub>H<sub>6</sub>O<sub>5</sub>).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	AAA	1	Total C H O 13 4 4 5	1	0
3	CCC	1	Total C H O 13 4 4 5	1	0
3	DDD	1	Total C H O 13 4 4 5	1	0

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	AAA	173	Total 175	O 175	0	2
4	BBB	166	Total 167	O 167	0	1
4	CCC	152	Total 153	O 153	0	1
4	DDD	172	Total 174	O 174	0	2

### 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.

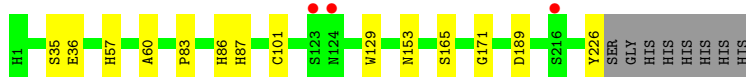
- Molecule 1: Endoglucanase, putative

Chain AAA: 



- Molecule 1: Endoglucanase, putative

Chain BBB: 



- Molecule 1: Endoglucanase, putative

Chain CCC: 



- Molecule 1: Endoglucanase, putative

Chain DDD: 



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	37.51Å 66.41Å 86.70Å 83.71° 80.83° 90.00°	Depositor
Resolution (Å)	49.55 – 1.75 49.55 – 1.75	Depositor EDS
% Data completeness (in resolution range)	97.4 (49.55-1.75) 94.9 (49.55-1.75)	Depositor EDS
$R_{merge}$	0.12	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.42 (at 1.75Å)	Xtriage
Refinement program	REFMAC 5.8.0267	Depositor
R, $R_{free}$	0.184 , 0.226 0.195 , 0.237	Depositor DCC
$R_{free}$ test set	3864 reflections (4.92%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	15.7	Xtriage
Anisotropy	0.368	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	(Not available) , (Not available)	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.51$ , $\langle L^2 \rangle = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	14017	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	18.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 13.18% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

## 5 Model quality [i](#)

### 5.1 Standard geometry [i](#)

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

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	AAA	0.72	1/1767 (0.1%)	0.84	0/2411
1	BBB	0.74	0/1747	0.83	0/2383
1	CCC	0.72	0/1742	0.81	0/2376
1	DDD	0.73	0/1790	0.83	0/2441
All	All	0.73	1/7046 (0.0%)	0.83	0/9611

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	AAA	30	GLU	CD-OE2	-5.38	1.19	1.25

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

### 5.2 Too-close contacts [i](#)

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

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	AAA	1716	1613	1606	5	0
1	BBB	1697	1605	1601	6	0
1	CCC	1698	1602	1596	3	0
1	DDD	1737	1637	1630	7	0
2	AAA	1	0	0	0	0
2	BBB	1	0	0	0	0

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	CCC	1	0	0	0	0
2	DDD	1	0	0	0	0
3	AAA	9	4	4	0	0
3	CCC	9	4	4	0	0
3	DDD	9	4	4	0	0
4	AAA	175	0	0	0	0
4	BBB	167	0	0	0	0
4	CCC	153	0	0	0	0
4	DDD	174	0	0	0	0
All	All	7548	6469	6445	21	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 2.

The worst 5 of 21 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:DDD:119:LEU:HD21	1:DDD:122:ASP:OD1	1.84	0.77
1:BBB:101:CYS:H	1:BBB:153:ASN:ND2	2.01	0.58
1:BBB:189:ASP:OD2	1:BBB:226:TYR:OH	2.30	0.49
1:BBB:165[A]:SER:OG	1:BBB:171:GLY:O	2.31	0.48
1:DDD:88:GLY:HA3	1:DDD:161:ILE:O	2.13	0.48

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	AAA	229/234 (98%)	218 (95%)	11 (5%)	0	100	100
1	BBB	226/234 (97%)	219 (97%)	6 (3%)	1 (0%)	34	17

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	CCC	226/234 (97%)	218 (96%)	8 (4%)	0	100	100
1	DDD	231/234 (99%)	227 (98%)	3 (1%)	1 (0%)	34	17
All	All	912/936 (97%)	882 (97%)	28 (3%)	2 (0%)	47	29

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	DDD	57	HIS
1	BBB	57	HIS

### 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	AAA	184/188 (98%)	182 (99%)	2 (1%)	73	60
1	BBB	183/188 (97%)	182 (100%)	1 (0%)	88	83
1	CCC	182/188 (97%)	180 (99%)	2 (1%)	73	60
1	DDD	187/188 (100%)	186 (100%)	1 (0%)	88	83
All	All	736/752 (98%)	730 (99%)	6 (1%)	81	72

5 of 6 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	CCC	36	GLU
1	CCC	57	HIS
1	DDD	36	GLU
1	AAA	57	HIS
1	AAA	36	GLU

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 7 ligands modelled in this entry, 4 are monoatomic - leaving 3 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	LMR	DDD	302	-	2,8,8	0.57	0	3,10,10	0.84	0
3	LMR	AAA	302	-	2,8,8	0.49	0	3,10,10	0.67	0
3	LMR	CCC	302	-	2,8,8	0.37	0	3,10,10	1.14	0

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	LMR	DDD	302	-	-	0/2/8/8	-
3	LMR	AAA	302	-	-	1/2/8/8	-
3	LMR	CCC	302	-	-	0/2/8/8	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (1) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	AAA	302	LMR	O2-C2-C3-C4

There are no ring outliers.

No monomer is involved in short contacts.

## 5.7 Other polymers [i](#)

There are no such residues in this entry.

## 5.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

## 6 Fit of model and data [i](#)

### 6.1 Protein, DNA and RNA chains [i](#)

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<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	AAA	229/234 (97%)	0.10	5 (2%) 62 69	11, 17, 29, 54	0
1	BBB	226/234 (96%)	0.08	3 (1%) 77 83	10, 16, 27, 52	0
1	CCC	228/234 (97%)	0.12	3 (1%) 77 83	11, 18, 30, 46	0
1	DDD	231/234 (98%)	0.14	7 (3%) 50 56	11, 17, 30, 76	0
All	All	914/936 (97%)	0.11	18 (1%) 65 72	10, 17, 29, 76	0

The worst 5 of 18 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	DDD	123	SER	4.5
1	AAA	229	HIS	4.3
1	DDD	230	HIS	3.9
1	BBB	216	SER	3.8
1	DDD	216	SER	3.7

### 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](#)

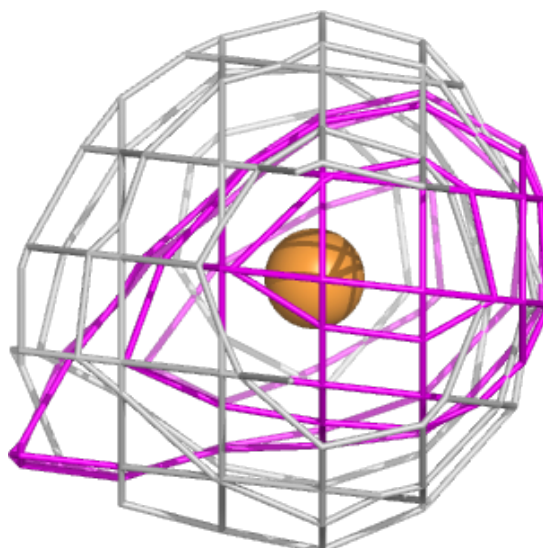
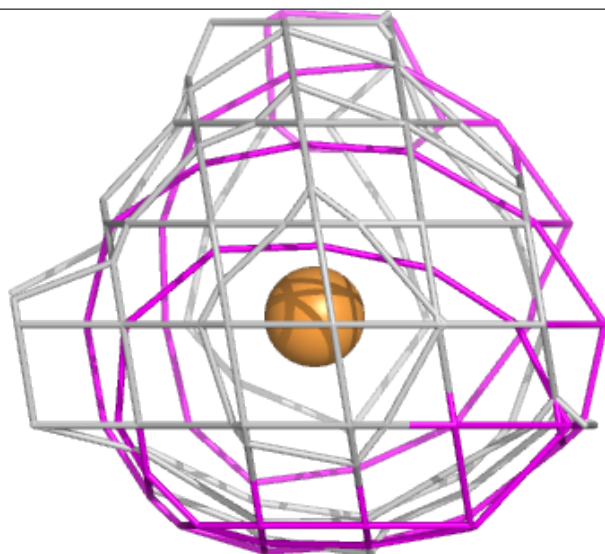
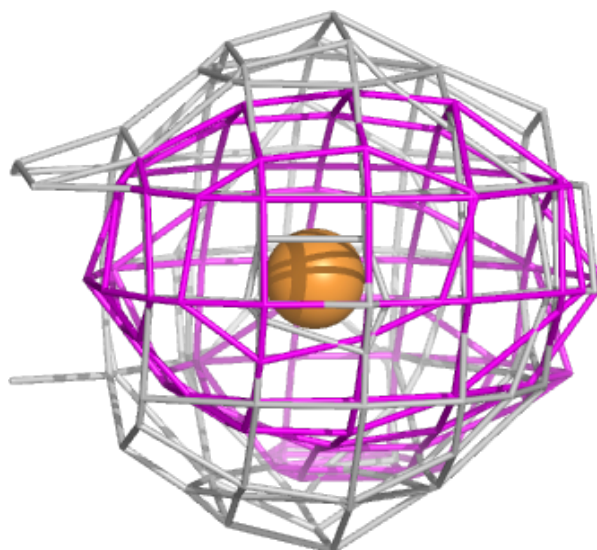
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled ‘Q< 0.9’ lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
3	LMR	AAA	302	9/9	0.70	0.20	28,41,48,53	1
3	LMR	DDD	302	9/9	0.89	0.14	24,27,32,32	1
3	LMR	CCC	302	9/9	0.90	0.15	22,27,32,34	1
2	CU	CCC	301	1/1	0.98	0.04	19,19,19,19	0
2	CU	AAA	301	1/1	0.98	0.07	19,19,19,19	0
2	CU	BBB	301	1/1	0.99	0.07	17,17,17,17	0
2	CU	DDD	301	1/1	1.00	0.06	17,17,17,17	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.

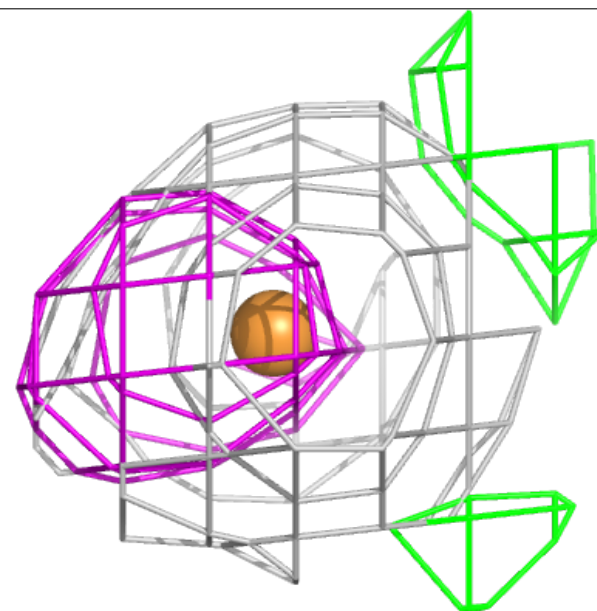
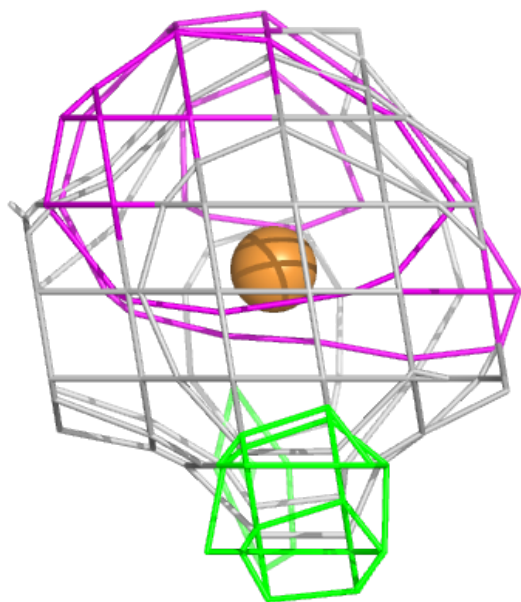
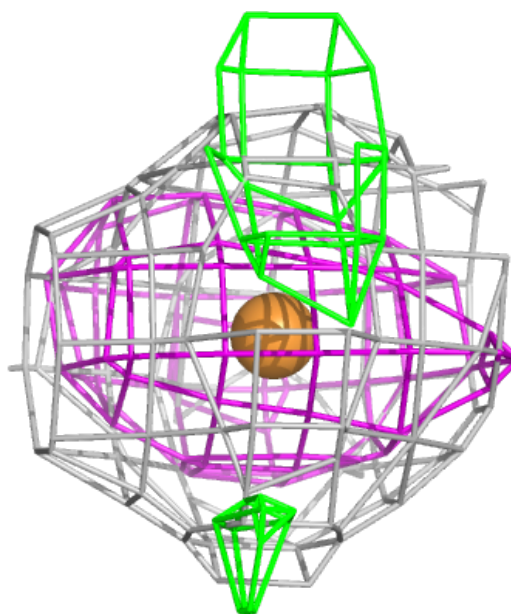
**Electron density around CU CCC 301:**

2mF<sub>o</sub>-DF<sub>c</sub> (at 0.7 rmsd) in gray  
 mF<sub>o</sub>-DF<sub>c</sub> (at 3 rmsd) in purple (negative)  
 and green (positive)



**Electron density around CU AAA 301:**

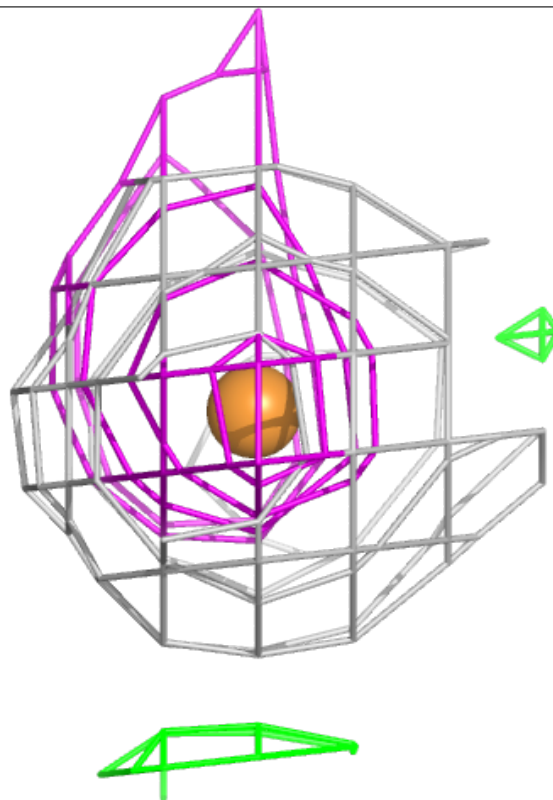
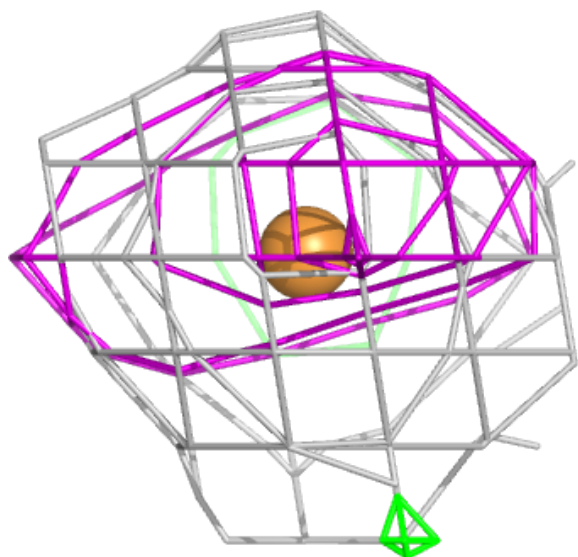
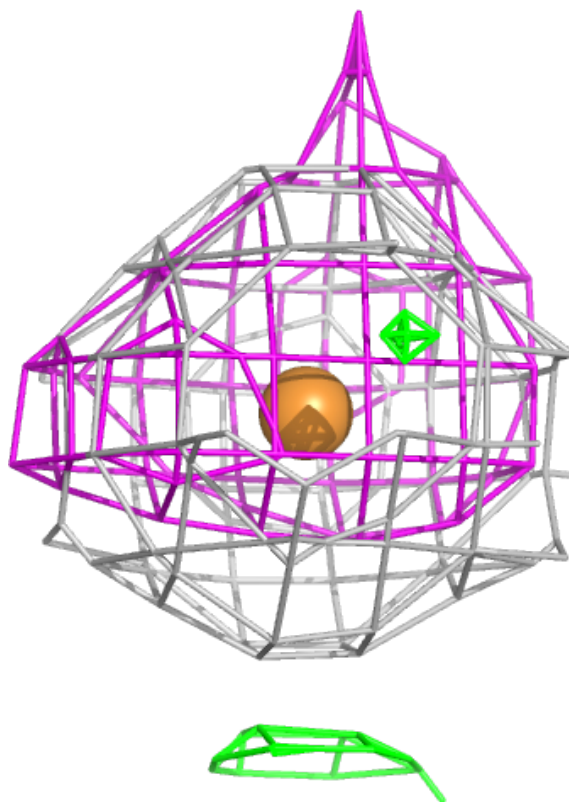
$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)





**Electron density around CU BBB 301:**

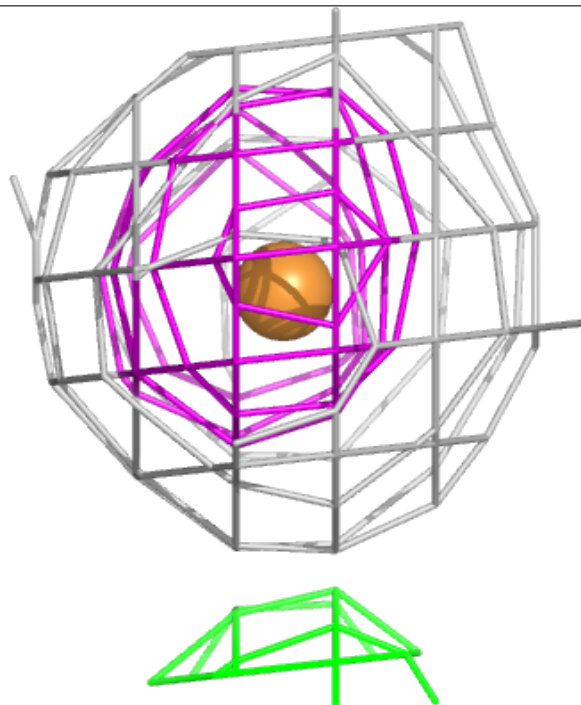
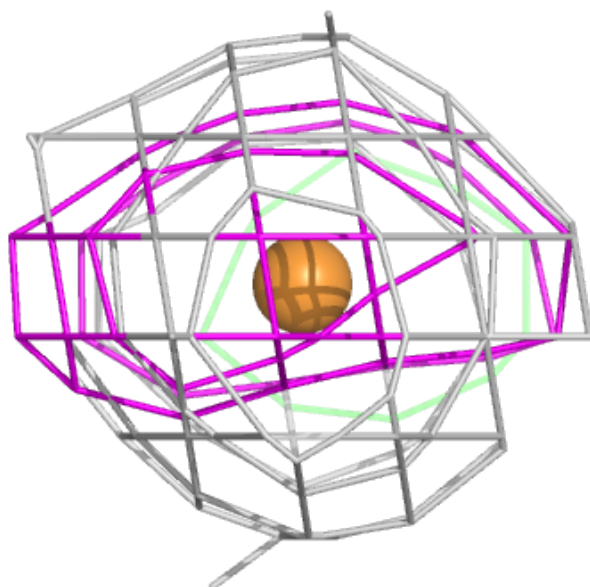
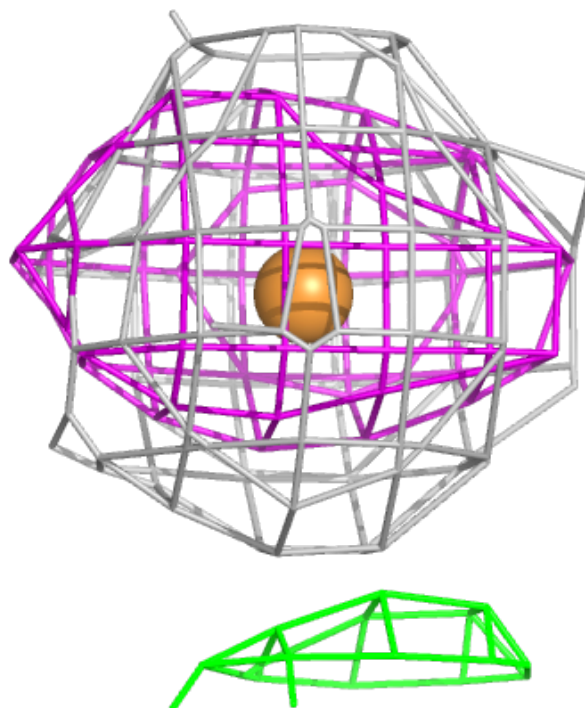
$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)





**Electron density around CU DDD 301:**

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



## 6.5 Other polymers [i](#)

There are no such residues in this entry.