



# Full wwPDB X-ray Structure Validation Report ⓘ

Nov 15, 2021 – 12:06 pm GMT

PDB ID : 7OS5  
Title : Crystal structure of Rhizobium etli inducible L-asparaginase ReAV (orthorhombic form OP)  
Authors : Loch, J.I.; Imiolczyk, B.; Gilski, M.; Jaskolski, M.  
Deposited on : 2021-06-07  
Resolution : 1.29 Å(reported)

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

We welcome your comments at [validation@mail.wwpdb.org](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 (270009), CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.23.2  
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) : 2.23.2

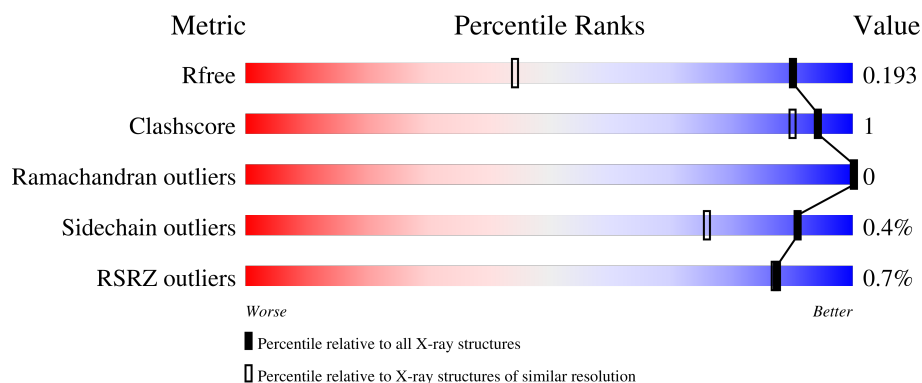
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 1.29 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	1058 (1.30-1.30)
Clashscore	141614	1101 (1.30-1.30)
Ramachandran outliers	138981	1058 (1.30-1.30)
Sidechain outliers	138945	1058 (1.30-1.30)
RSRZ outliers	127900	1029 (1.30-1.30)

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	373	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red, orange, yellow, green, grey);"></div> <div style="display: flex; justify-content: space-between; padding: 0 5px;"> <span>%</span> <span>95%</span> <span>• •</span> </div> </div>
1	BBB	373	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red, orange, yellow, green, grey);"></div> <div style="display: flex; justify-content: space-between; padding: 0 5px;"> <span>%</span> <span>92%</span> <span>• 6%</span> </div> </div>

## 2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 6073 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 L-asparaginase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	AAA	366	Total	C	N	O	S	0	11	0
			2767	1709	504	529	25			
1	BBB	349	Total	C	N	O	S	0	7	0
			2617	1617	479	498	23			

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
AAA	-5	GLY	-	expression tag	UNP Q2K0Z2
AAA	-4	ILE	-	expression tag	UNP Q2K0Z2
AAA	-3	ASP	-	expression tag	UNP Q2K0Z2
AAA	-2	PRO	-	expression tag	UNP Q2K0Z2
AAA	-1	PHE	-	expression tag	UNP Q2K0Z2
AAA	0	THR	-	expression tag	UNP Q2K0Z2
BBB	-5	GLY	-	expression tag	UNP Q2K0Z2
BBB	-4	ILE	-	expression tag	UNP Q2K0Z2
BBB	-3	ASP	-	expression tag	UNP Q2K0Z2
BBB	-2	PRO	-	expression tag	UNP Q2K0Z2
BBB	-1	PHE	-	expression tag	UNP Q2K0Z2
BBB	0	THR	-	expression tag	UNP Q2K0Z2

- 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	AAA	1	Total	Zn	0	0
			1	1		
2	BBB	1	Total	Zn	0	0
			1	1		

- Molecule 3 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: C<sub>2</sub>H<sub>6</sub>O<sub>2</sub>).

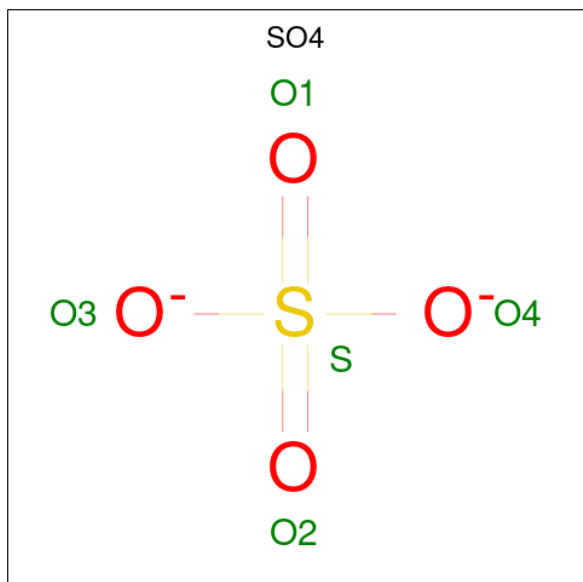


Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	AAA	1	Total	C	O	0	0
			4	2	2		
3	BBB	1	Total	C	O	0	0
			4	2	2		

- Molecule 4 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	AAA	1	Total	Cl	0	0
			1	1		

- Molecule 5 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	AAA	1	Total O S 5 4 1	0	0
5	AAA	1	Total O S 5 4 1	0	0
5	AAA	1	Total O S 5 4 1	0	0
5	AAA	1	Total O S 5 4 1	0	0
5	BBB	1	Total O S 5 4 1	0	0
5	BBB	1	Total O S 5 4 1	0	0
5	BBB	1	Total O S 5 4 1	0	0
5	BBB	1	Total O S 5 4 1	0	0
5	BBB	1	Total O S 5 4 1	0	0
5	BBB	1	Total O S 5 4 1	0	0

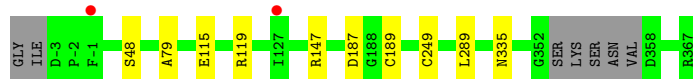
- Molecule 6 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	AAA	317	Total O 317 317	0	0
6	BBB	311	Total O 311 311	0	0

### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density ( $RSRZ > 2$ ). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

- Molecule 1: L-asparaginase



- Molecule 1: L-asparaginase



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	78.23Å 91.14Å 106.02Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	69.11 – 1.29 69.11 – 1.29	Depositor EDS
% Data completeness (in resolution range)	68.6 (69.11-1.29) 68.6 (69.11-1.29)	Depositor EDS
$R_{merge}$	0.17	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.80 (at 1.29Å)	Xtriage
Refinement program	REFMAC 5.8.0267	Depositor
R, $R_{free}$	0.155 , 0.194 0.155 , 0.193	Depositor DCC
$R_{free}$ test set	1000 reflections (0.77%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	6.4	Xtriage
Anisotropy	1.607	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>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	6073	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	16.0	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 31.05 % 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 1.1798e-03. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

<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: ZN, CL, EDO, SO4, CSO

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	0.78	0/2826	0.79	0/3816
1	BBB	0.77	0/2664	0.80	0/3596
All	All	0.78	0/5490	0.79	0/7412

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	AAA	0	1

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	AAA	79	ALA	Peptide

### 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	2767	0	2745	8	0
1	BBB	2617	0	2599	7	0
2	AAA	1	0	0	0	0
2	BBB	1	0	0	0	0
3	AAA	4	0	6	0	0
3	BBB	4	0	6	0	0
4	AAA	1	0	0	0	0
5	AAA	20	0	0	0	0
5	BBB	30	0	0	1	0
6	AAA	317	0	0	3	0
6	BBB	311	0	0	3	0
All	All	6073	0	5356	14	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 1.

All (14) 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:BBB:249[B]:CSO:OD	6:BBB:501:HOH:O	1.87	0.92
1:BBB:249[A]:CSO:OD	6:BBB:502:HOH:O	1.99	0.81
1:AAA:249[B]:CSO:OD	6:AAA:501:HOH:O	2.07	0.69
1:AAA:249[A]:CSO:OD	6:AAA:502:HOH:O	2.11	0.63
1:AAA:187[B]:ASP:OD1	1:AAA:189:CYS:SG	2.57	0.62
1:AAA:187[B]:ASP:OD2	1:AAA:189:CYS:SG	2.58	0.62
1:BBB:163[B]:MET:HE2	1:BBB:163[B]:MET:HA	1.85	0.58
1:AAA:187[B]:ASP:CG	1:AAA:189:CYS:SG	2.83	0.57
1:BBB:152:GLY:HA2	5:BBB:405:SO4:O3	2.14	0.48
1:AAA:48:SER:OG	6:AAA:503:HOH:O	2.20	0.47
1:AAA:115:GLU:O	1:AAA:119:ARG:HG3	2.14	0.47
1:BBB:37:LEU:HD11	1:BBB:310[B]:LEU:HD21	2.00	0.43
1:BBB:48:SER:OG	6:BBB:503:HOH:O	2.21	0.41
1:AAA:335:ASN:HB2	1:BBB:188:GLY:HA2	2.03	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	AAA	371/373 (100%)	363 (98%)	8 (2%)	0	100	100
1	BBB	352/373 (94%)	344 (98%)	8 (2%)	0	100	100
All	All	723/746 (97%)	707 (98%)	16 (2%)	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	AAA	286/282 (101%)	284 (99%)	2 (1%)	84	61
1	BBB	268/282 (95%)	268 (100%)	0	100	100
All	All	554/564 (98%)	552 (100%)	2 (0%)	91	76

All (2) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	AAA	147	ARG
1	AAA	289	LEU

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 ⓘ

4 non-standard protein/DNA/RNA residues are modelled in this entry.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

## 5.5 Carbohydrates ⓘ

There are no monosaccharides in this entry.

## 5.6 Ligand geometry ⓘ

Of 15 ligands modelled in this entry, 3 are monoatomic - leaving 12 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

## 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 [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	365/373 (97%)	-0.19	2 (0%) 91 91	4, 11, 31, 68	0
1	BBB	348/373 (93%)	-0.22	3 (0%) 84 85	5, 12, 29, 68	0
All	All	713/746 (95%)	-0.20	5 (0%) 87 87	4, 11, 31, 68	0

All (5) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	BBB	348	PHE	4.4
1	AAA	127	ILE	3.7
1	BBB	216	ASP	2.9
1	BBB	220	GLY	2.9
1	AAA	-1	PHE	2.6

### 6.2 Non-standard residues in protein, DNA, RNA chains [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(Å <sup>2</sup> )	Q<0.9
1	CSO	AAA	249[A]	7/8	0.98	0.07	8,8,14,14	4
1	CSO	AAA	249[B]	7/8	0.98	0.07	8,8,11,18	4
1	CSO	BBB	249[A]	7/8	0.98	0.07	8,8,10,15	4
1	CSO	BBB	249[B]	7/8	0.98	0.07	8,8,16,17	4

### 6.3 Carbohydrates [i](#)

There are no monosaccharides in this entry.

## 6.4 Ligands

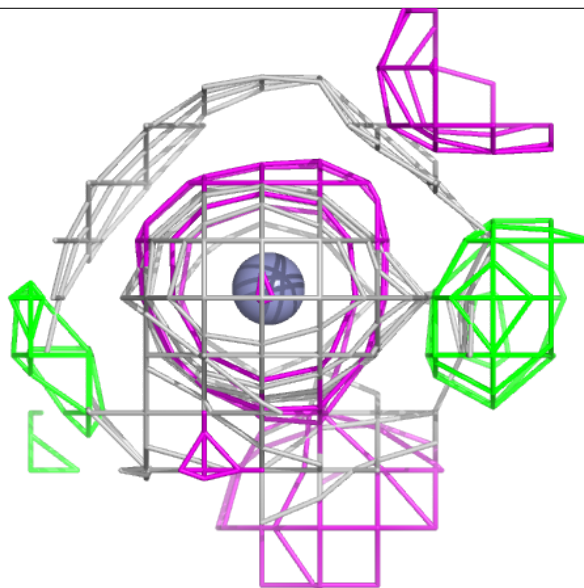
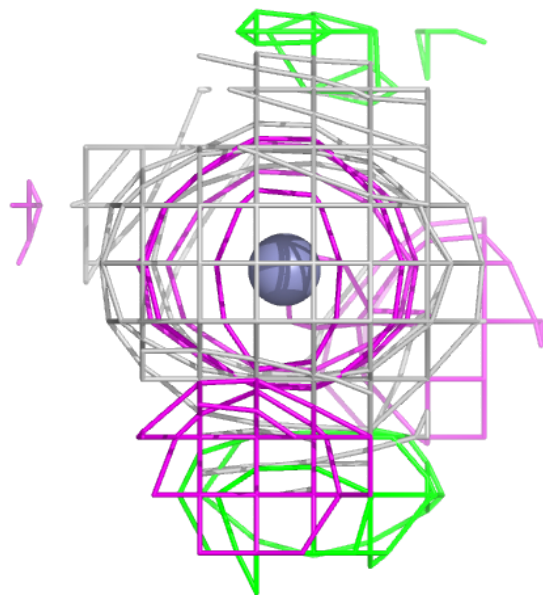
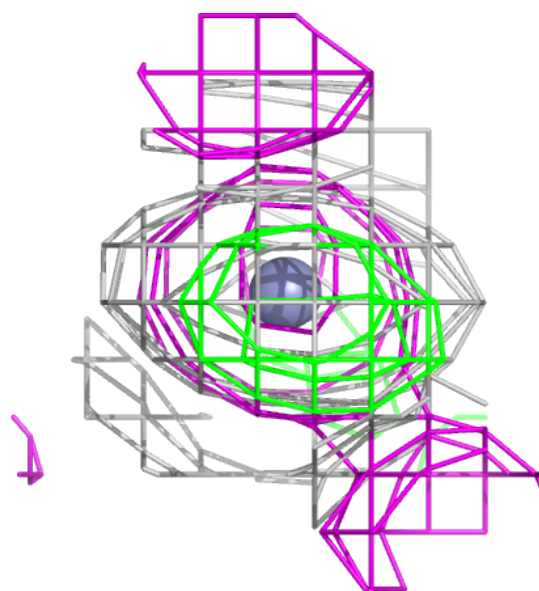
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q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

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
3	EDO	BBB	404	4/4	0.82	0.17	23,24,29,33	0
5	SO4	BBB	407	5/5	0.84	0.14	26,39,47,65	5
3	EDO	AAA	402	4/4	0.85	0.18	53,57,57,60	0
5	SO4	AAA	405	5/5	0.86	0.16	35,36,39,40	5
5	SO4	BBB	403	5/5	0.90	0.12	27,34,39,42	5
5	SO4	BBB	408	5/5	0.92	0.13	37,40,44,46	5
4	CL	AAA	403	1/1	0.93	0.06	46,46,46,46	0
5	SO4	BBB	406	5/5	0.93	0.11	19,26,28,31	5
5	SO4	AAA	406	5/5	0.93	0.13	25,27,28,36	5
5	SO4	AAA	407	5/5	0.93	0.12	13,23,32,35	5
5	SO4	AAA	404	5/5	0.94	0.12	16,24,27,28	5
5	SO4	BBB	405	5/5	0.98	0.16	30,35,42,43	5
5	SO4	BBB	402	5/5	0.98	0.07	11,12,15,15	5
2	ZN	BBB	401	1/1	1.00	0.03	8,8,8,8	1
2	ZN	AAA	401	1/1	1.00	0.03	8,8,8,8	1

The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation to approximate a three-dimensional view.

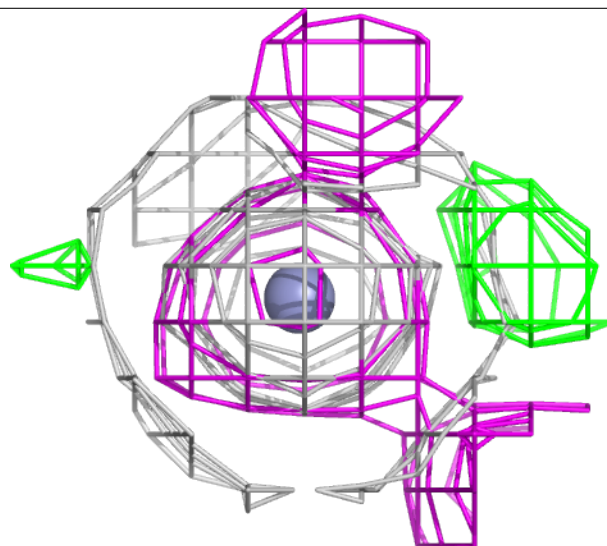
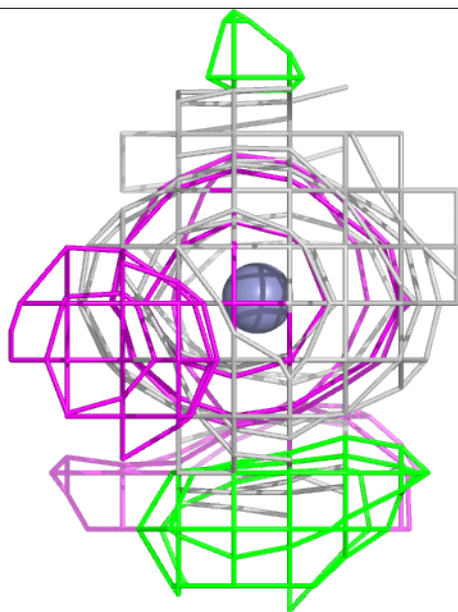
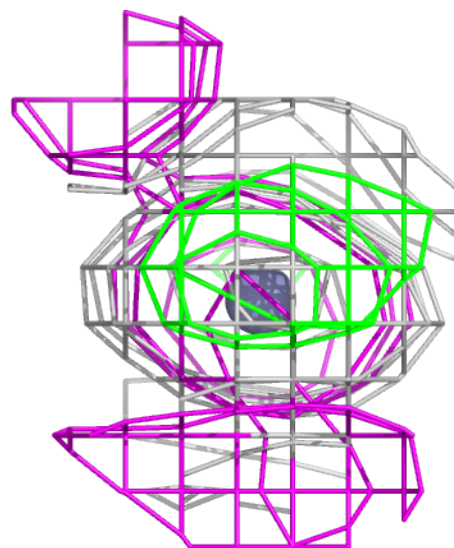
**Electron density around ZN BBB 401:**

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



**Electron density around ZN AAA 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.