



# wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 3, 2025 – 03:03 pm GMT

PDB ID : 8S71  
Title : Oxidoreductase B from Phytophthora sojae  
Authors : Banerjee, S.; Turella, S.; Morth, J.P.; Abou Hachem, M.  
Deposited on : 2024-02-28  
Resolution : 2.05 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 types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

---

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

MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.003 (Gargrove)
Density-Fitness	:	1.0.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.41

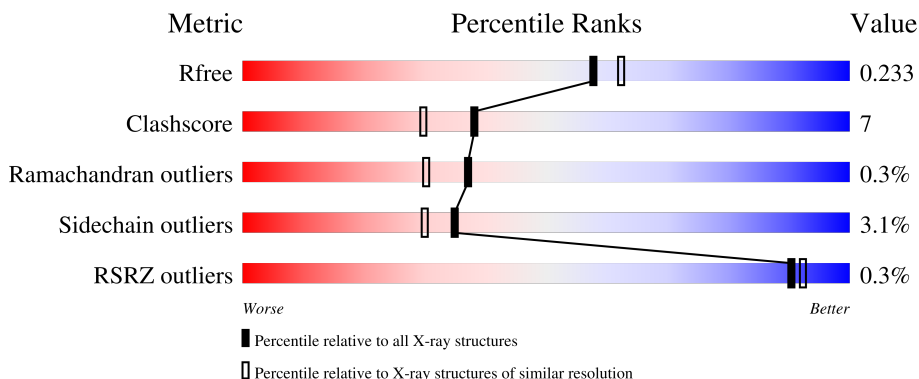
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*



The reported resolution of this entry is 2.05 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))
$R_{free}$	164625	2096 (2.04-2.04)
Clashscore	180529	2229 (2.04-2.04)
Ramachandran outliers	177936	2217 (2.04-2.04)
Sidechain outliers	177891	2217 (2.04-2.04)
RSRZ outliers	164620	2096 (2.04-2.04)

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	A	185	 79% 12% 7%
1	B	185	 78% 16% 6%

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 3057 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 Mucin-like protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	172	Total	C	N	O	S	0	6	0
			1392	874	228	279	11			
1	B	174	Total	C	N	O	S	0	6	0
			1409	885	231	280	13			

There are 18 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	200	LEU	-	expression tag	UNP G4YSM7
A	201	TRP	-	expression tag	UNP G4YSM7
A	202	SER	-	expression tag	UNP G4YSM7
A	203	HIS	-	expression tag	UNP G4YSM7
A	204	PRO	-	expression tag	UNP G4YSM7
A	205	GLN	-	expression tag	UNP G4YSM7
A	206	PHE	-	expression tag	UNP G4YSM7
A	207	GLU	-	expression tag	UNP G4YSM7
A	208	LYS	-	expression tag	UNP G4YSM7
B	200	LEU	-	expression tag	UNP G4YSM7
B	201	TRP	-	expression tag	UNP G4YSM7
B	202	SER	-	expression tag	UNP G4YSM7
B	203	HIS	-	expression tag	UNP G4YSM7
B	204	PRO	-	expression tag	UNP G4YSM7
B	205	GLN	-	expression tag	UNP G4YSM7
B	206	PHE	-	expression tag	UNP G4YSM7
B	207	GLU	-	expression tag	UNP G4YSM7
B	208	LYS	-	expression tag	UNP G4YSM7

- 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	A	1	Total 1	Cu 1	0	0
2	B	1	Total 1	Cu 1	0	0

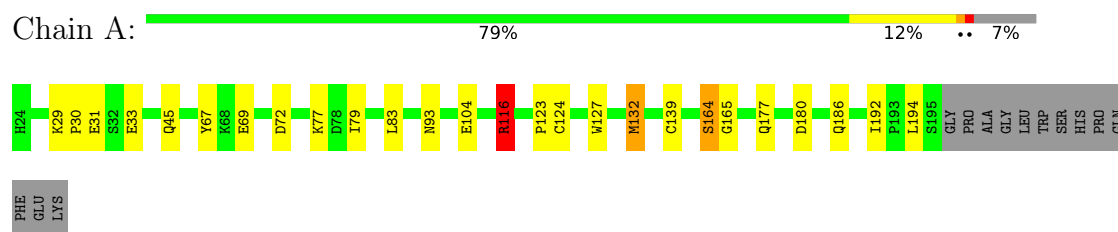
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	121	Total 122	O 122	0	1
3	B	132	Total 132	O 132	0	0

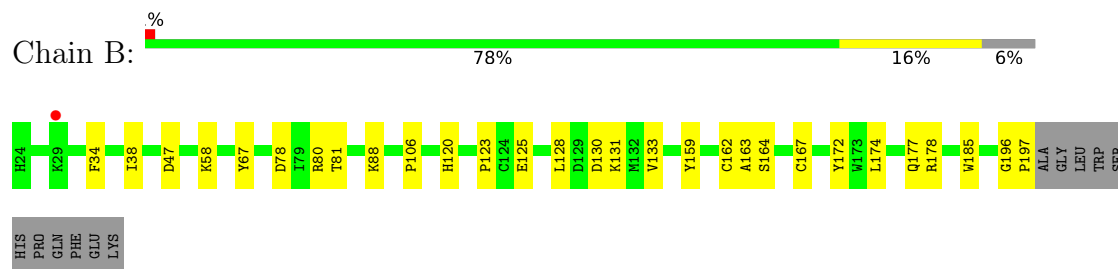
### 3 Residue-property plots [i](#)

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: Mucin-like protein



- Molecule 1: Mucin-like protein



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	40.07Å 106.90Å 44.81Å 90.00° 114.10° 90.00°	Depositor
Resolution (Å)	53.42 – 2.05 53.42 – 2.05	Depositor EDS
% Data completeness (in resolution range)	99.2 (53.42-2.05) 99.2 (53.42-2.05)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.42 (at 1.90Å)	Xtriage
Refinement program	REFMAC 5.8.0419	Depositor
R, $R_{free}$	0.168 , 0.232 0.177 , 0.233	Depositor DCC
$R_{free}$ test set	1026 reflections (4.78%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	25.9	Xtriage
Anisotropy	0.293	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 33.8	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	0.036 for h,-k,-h-l	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	3057	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	29.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.07% 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

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: 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	A	0.49	1/1428 (0.1%)	0.76	2/1941 (0.1%)
1	B	0.44	0/1446	0.73	0/1963
All	All	0.47	1/2874 (0.0%)	0.75	2/3904 (0.1%)

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	1
1	B	0	1
All	All	0	2

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	31	GLU	CD-OE2	8.11	1.34	1.25

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	116	ARG	NE-CZ-NH1	-6.37	117.11	120.30
1	A	116	ARG	NE-CZ-NH2	5.35	122.98	120.30

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	116	ARG	Sidechain
1	B	196	GLY	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	A	1392	0	1278	20	0
1	B	1409	0	1308	17	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
3	A	122	0	0	3	0
3	B	132	0	0	4	0
All	All	3057	0	2586	36	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 7.

The worst 5 of 36 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:124[B]:CYS:SG	1:A:139[B]:CYS:HB3	1.96	1.05
1:A:33:GLU:OE1	3:A:401:HOH:O	1.92	0.88
1:A:124[B]:CYS:HG	1:A:139[B]:CYS:CB	1.89	0.85
1:A:124[B]:CYS:SG	1:A:139[B]:CYS:CB	2.66	0.83
1:A:124[B]:CYS:HB3	1:A:139[B]:CYS:HG	1.54	0.73

There are no symmetry-related clashes.

## 5.3 Torsion angles [i](#)

### 5.3.1 Protein backbone [i](#)

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.



The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	176/185 (95%)	170 (97%)	5 (3%)	1 (1%)	22	13
1	B	178/185 (96%)	175 (98%)	3 (2%)	0	100	100
All	All	354/370 (96%)	345 (98%)	8 (2%)	1 (0%)	37	30

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	164	SER

### 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	150/157 (96%)	143 (95%)	7 (5%)	22	16
1	B	154/157 (98%)	151 (98%)	3 (2%)	52	50
All	All	304/314 (97%)	294 (97%)	10 (3%)	35	27

5 of 10 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	58	LYS
1	B	88	LYS
1	B	164	SER
1	A	116	ARG
1	A	132[A]	MET

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	140	GLN
1	A	177	GLN

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	B	140	GLN
1	B	148	GLN
1	B	177	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 oligosaccharides in this entry.

## 5.6 Ligand geometry ⓘ

Of 2 ligands modelled in this entry, 2 are monoatomic - leaving 0 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	A	172/185 (92%)	-0.23	0 <b>100</b> <b>100</b>	10, 27, 49, 59	6 (3%)
1	B	174/185 (94%)	-0.29	1 (0%) <b>85</b> <b>88</b>	10, 27, 48, 60	6 (3%)
All	All	346/370 (93%)	-0.26	1 (0%) <b>90</b> <b>92</b>	10, 27, 49, 60	12 (3%)

All (1) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	29	LYS	3.0

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

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

### 6.3 Carbohydrates [i](#)

There are no monosaccharides in this entry.

### 6.4 Ligands [i](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<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.

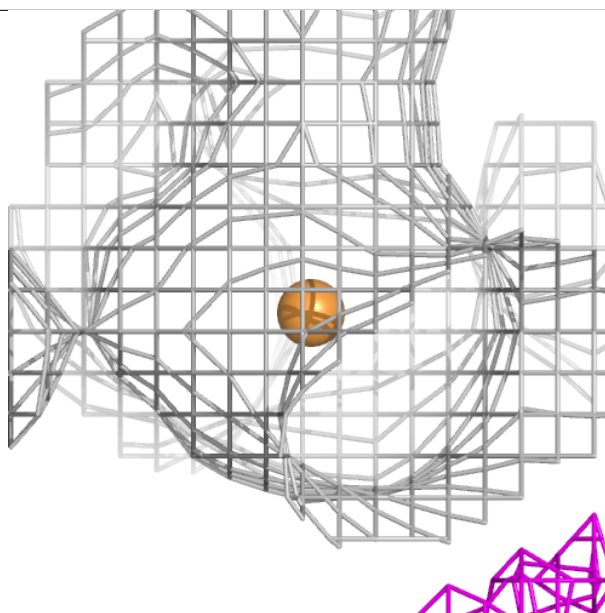
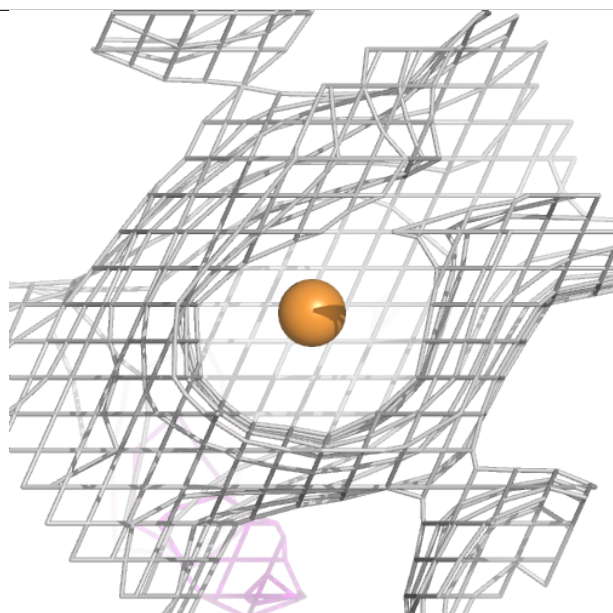
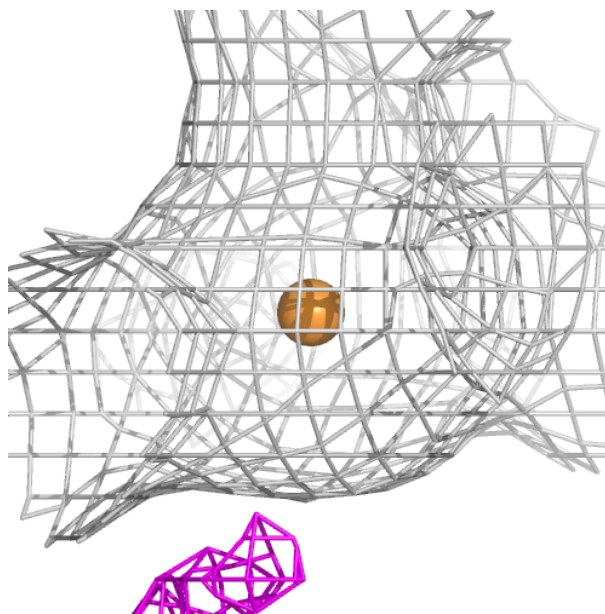
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
2	CU	A	301	1/1	1.00	0.01	28,28,28,28	0
2	CU	B	301	1/1	1.00	0.02	27,27,27,27	0

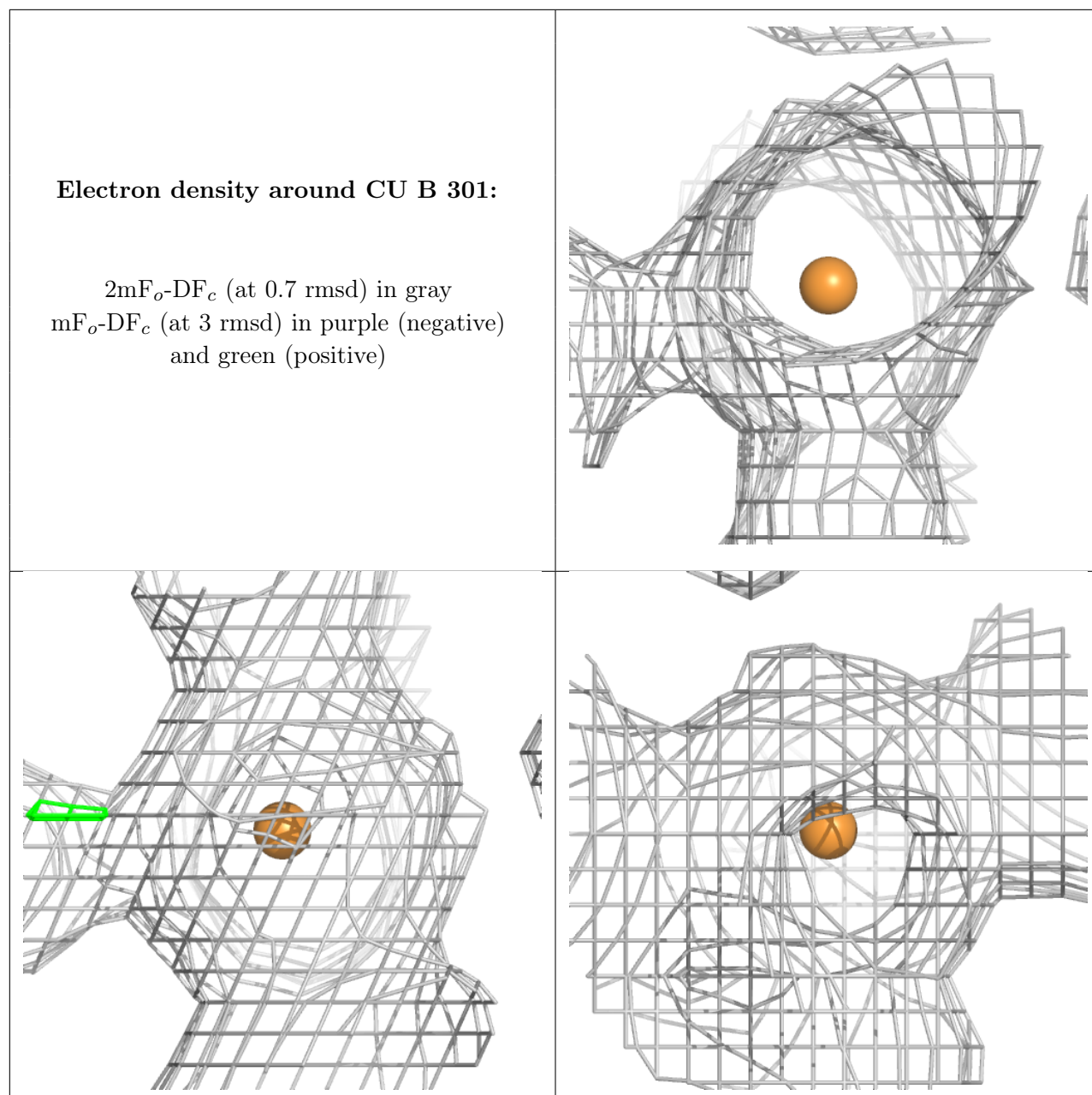
The following is a graphical depiction of the model fit to experimental electron density of all

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 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)





## 6.5 Other polymers ⓘ

There are no such residues in this entry.