



# Full wwPDB X-ray Structure Validation Report ⓘ

Mar 3, 2025 – 02:05 PM EST

PDB ID : 9BL1  
Title : Crystal structure of heme-binding protein from Populus trichocarpa  
Authors : Kumaran, D.; Grosjean, N.; Blaby, E.C.  
Deposited on : 2024-04-29  
Resolution : 1.80 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 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  
Mogul : 2022.3.0, CSD as543be (2022)  
Xtriage (Phenix) : 1.21  
EDS : 3.0  
Percentile statistics : 20231227.v01 (using entries in the PDB archive December 27th 2023)  
CCP4 : 9.0.004 (Gargrove)  
Density-Fitness : 1.0.11  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.41.4

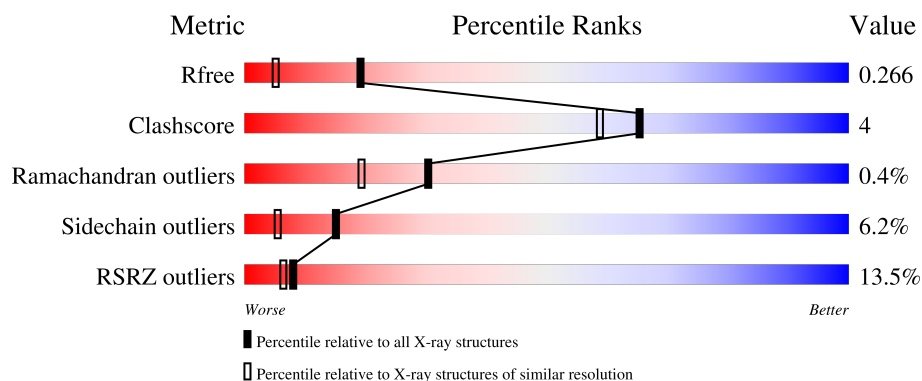
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.80 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	7108 (1.80-1.80)
Clashscore	180529	8162 (1.80-1.80)
Ramachandran outliers	177936	8077 (1.80-1.80)
Sidechain outliers	177891	8076 (1.80-1.80)
RSRZ outliers	164620	7108 (1.80-1.80)

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	291	

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 2161 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 Heme-binding protein, PtHOZ1.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	264	Total	C	N	O	S	Se	0	0	0
			2062	1305	356	394	2	5			

There are 14 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	MSE	-	initiating methionine	UNP A0A3N7FVG6
A	279	GLU	-	expression tag	UNP A0A3N7FVG6
A	280	ASN	-	expression tag	UNP A0A3N7FVG6
A	281	LEU	-	expression tag	UNP A0A3N7FVG6
A	282	TYR	-	expression tag	UNP A0A3N7FVG6
A	283	PHE	-	expression tag	UNP A0A3N7FVG6
A	284	GLN	-	expression tag	UNP A0A3N7FVG6
A	285	SER	-	expression tag	UNP A0A3N7FVG6
A	286	HIS	-	expression tag	UNP A0A3N7FVG6
A	287	HIS	-	expression tag	UNP A0A3N7FVG6
A	288	HIS	-	expression tag	UNP A0A3N7FVG6
A	289	HIS	-	expression tag	UNP A0A3N7FVG6
A	290	HIS	-	expression tag	UNP A0A3N7FVG6
A	291	HIS	-	expression tag	UNP A0A3N7FVG6

- Molecule 2 is NICKEL (II) ION (three-letter code: NI) (formula: Ni) (labeled as "Ligand of Interest" by depositor).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	2	Total	Ni	0	0
			2	2		

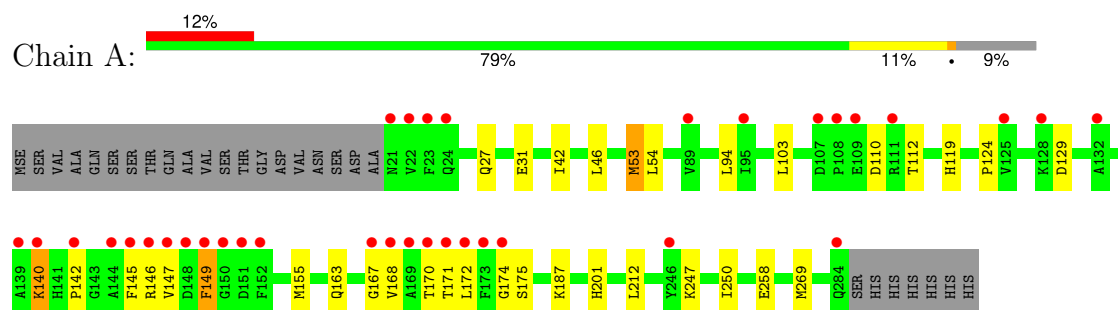
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	97	Total	O	0	0
			97	97		

### 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: Heme-binding protein, PtHOZ1



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 42 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	77.31Å 77.31Å 92.42Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	38.68 – 1.80 38.68 – 1.80	Depositor EDS
% Data completeness (in resolution range)	96.6 (38.68-1.80) 96.6 (38.68-1.80)	Depositor EDS
$R_{merge}$	0.06	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.63 (at 1.81Å)	Xtriage
Refinement program	REFMAC 5.8.0267	Depositor
R, $R_{free}$	0.206 , 0.255 0.213 , 0.266	Depositor DCC
$R_{free}$ test set	1309 reflections (4.92%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	35.6	Xtriage
Anisotropy	0.043	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 25.9	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	2161	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	45.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.23% 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: NI

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	0.71	0/2102	0.88	0/2843

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 [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	2062	0	2041	15	0
2	A	2	0	0	0	0
3	A	97	0	0	1	0
All	All	2161	0	2041	15	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 4.

All (15) 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:170:THR:HG23	1:A:174:GLY:HA2	1.76	0.67

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:124:PRO:HA	1:A:155:MSE:SE	2.55	0.56
1:A:145:PHE:HB2	1:A:149:PHE:CD2	2.40	0.56
1:A:27:GLN:O	1:A:31:GLU:HG3	2.09	0.52
1:A:170:THR:O	1:A:174:GLY:HA3	2.12	0.49
1:A:140:LYS:C	1:A:142:PRO:HD3	2.34	0.47
1:A:54:LEU:HD21	1:A:94:LEU:HG	1.96	0.47
1:A:53:MSE:HE2	1:A:53:MSE:HB3	1.99	0.44
1:A:250:ILE:HG22	1:A:269:MSE:HE1	1.99	0.44
1:A:119:HIS:NE2	1:A:163:GLN:NE2	2.66	0.43
1:A:112:THR:O	1:A:167:GLY:N	2.43	0.43
1:A:201:HIS:HE1	3:A:501:HOH:O	2.00	0.43
1:A:112:THR:OG1	1:A:168:VAL:HG23	2.19	0.43
1:A:103:LEU:C	1:A:103:LEU:HD23	2.39	0.42
1:A:46:LEU:HD23	1:A:46:LEU:HA	1.89	0.41

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	262/291 (90%)	250 (95%)	11 (4%)	1 (0%)	30 19

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	110	ASP

### 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	227/244 (93%)	213 (94%)	14 (6%)	15 5

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

Mol	Chain	Res	Type
1	A	42	ILE
1	A	53	MSE
1	A	129	ASP
1	A	140	LYS
1	A	146	ARG
1	A	147	VAL
1	A	149	PHE
1	A	171	THR
1	A	172	LEU
1	A	175	SER
1	A	187	LYS
1	A	212	LEU
1	A	247	LYS
1	A	258	GLU

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	27	GLN
1	A	141	HIS
1	A	163	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 [i](#)

There are no oligosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

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 [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 ⓘ

### 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, 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	259/291 (89%)	0.68	35 (13%) 8 6	25, 36, 96, 145	0

All (35) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	169	ALA	10.3
1	A	147	VAL	7.1
1	A	172	LEU	6.3
1	A	171	THR	5.5
1	A	168	VAL	5.5
1	A	173	PHE	5.3
1	A	170	THR	4.8
1	A	174	GLY	4.8
1	A	145	PHE	4.5
1	A	152	PHE	4.0
1	A	23	PHE	4.0
1	A	107	ASP	3.6
1	A	142	PRO	3.6
1	A	149	PHE	3.4
1	A	95	ILE	3.2
1	A	108	PRO	3.1
1	A	132	ALA	3.0
1	A	150	GLY	2.7
1	A	24	GLN	2.6
1	A	21	ASN	2.6
1	A	109	GLU	2.6
1	A	148	ASP	2.6
1	A	151	ASP	2.5
1	A	144	ALA	2.5
1	A	167	GLY	2.5
1	A	284	GLN	2.4
1	A	246	TYR	2.4

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	128	LYS	2.3
1	A	22	VAL	2.3
1	A	111	ARG	2.2
1	A	139	ALA	2.2
1	A	140	LYS	2.2
1	A	89	VAL	2.2
1	A	146	ARG	2.1
1	A	125	VAL	2.0

## 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](#)

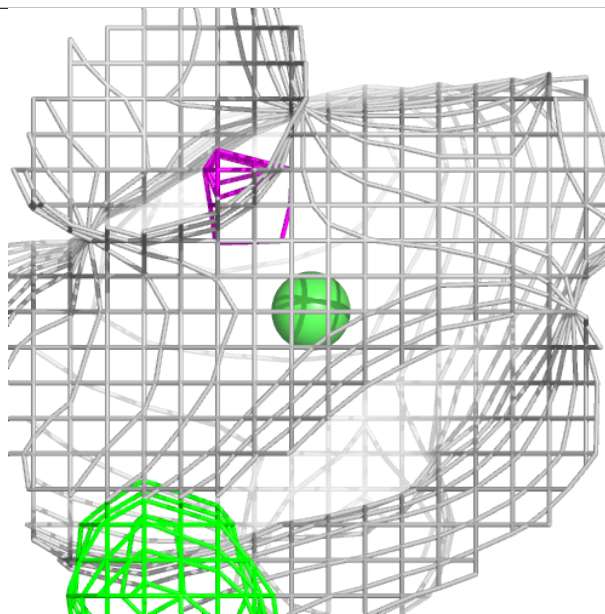
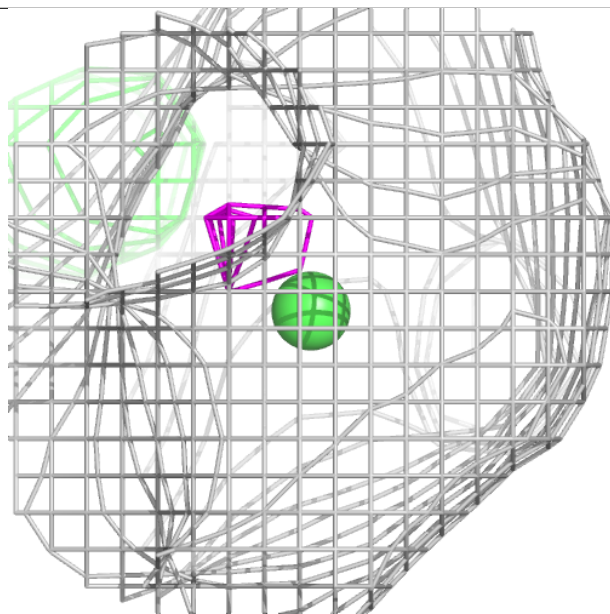
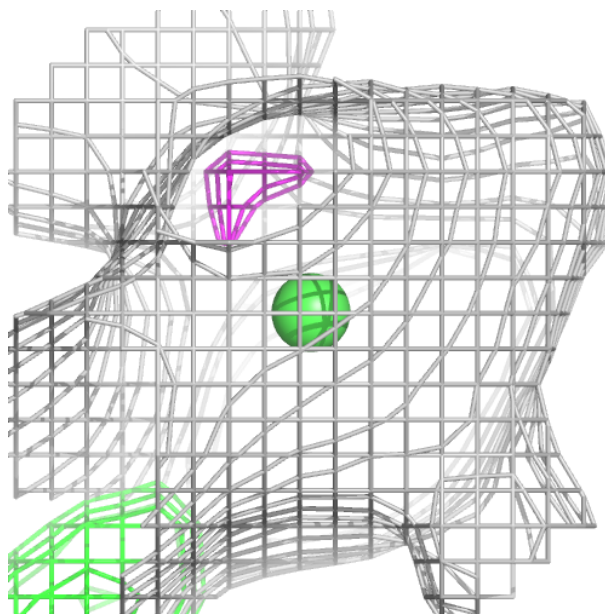
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q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

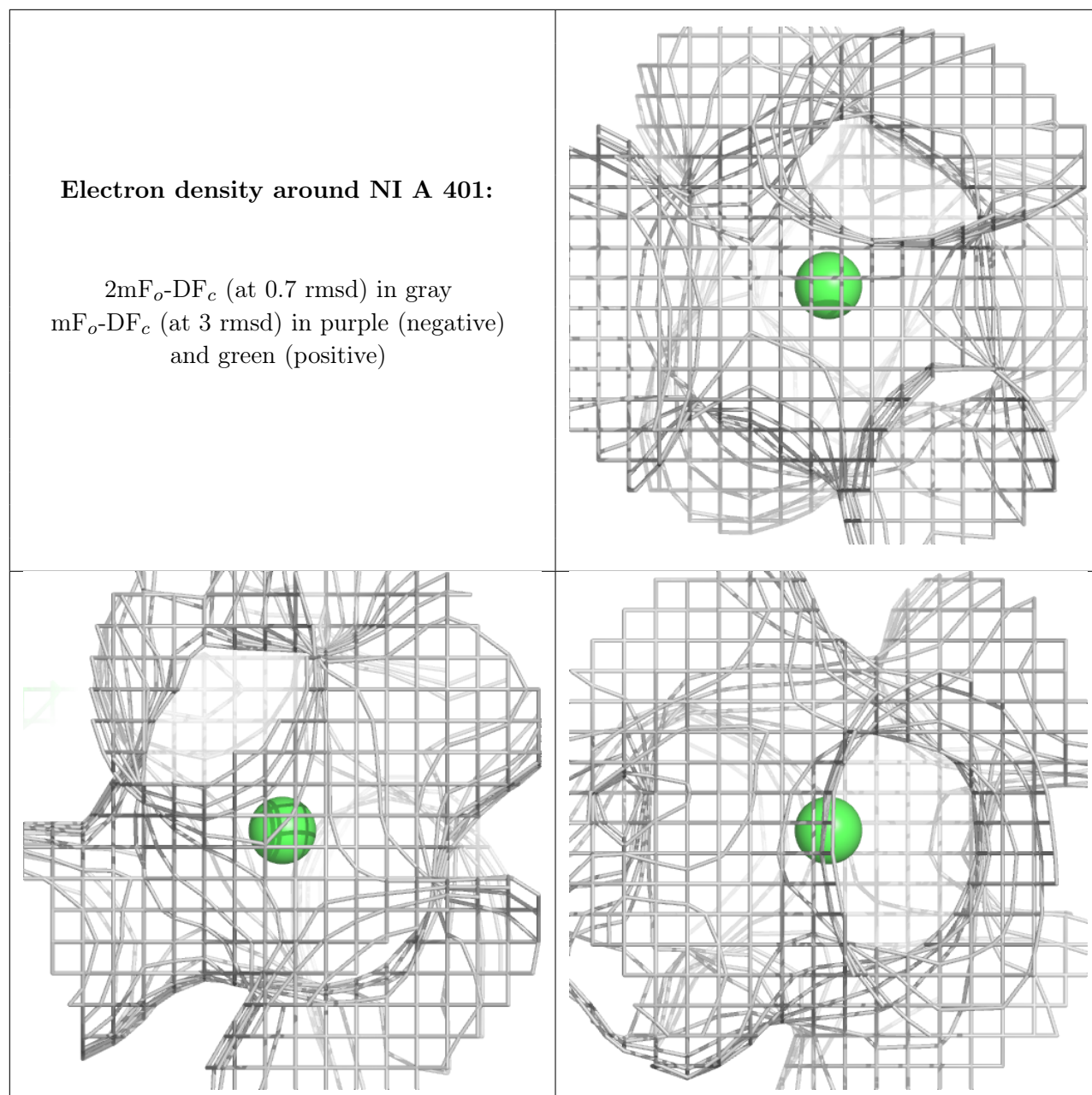
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
2	NI	A	402	1/1	0.98	0.05	53,53,53,53	0
2	NI	A	401	1/1	0.99	0.02	28,28,28,28	0

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 NI A 402:**

$2mF_o-DF_c$  (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.