



Full wwPDB X-ray Structure Validation Report ⓘ

Sep 4, 2025 – 04:07 pm BST

PDB ID : 9GCB / pdb_00009gcb
Title : DUF4198 protein from Ideonella sakaiensis with Ni bound
Authors : Franco Cairo, J.P.L.; Correa, T.L.R.; Offen, W.A.; Nairn, A.; Walton, J.;
Davies, G.J.; Walton, P.H.; Sweeney, S.T.
Deposited on : 2024-08-01
Resolution : 1.44 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

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	:	4-5-2 with Phenix2.0rc1
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	2.0rc1
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.006 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.45.1

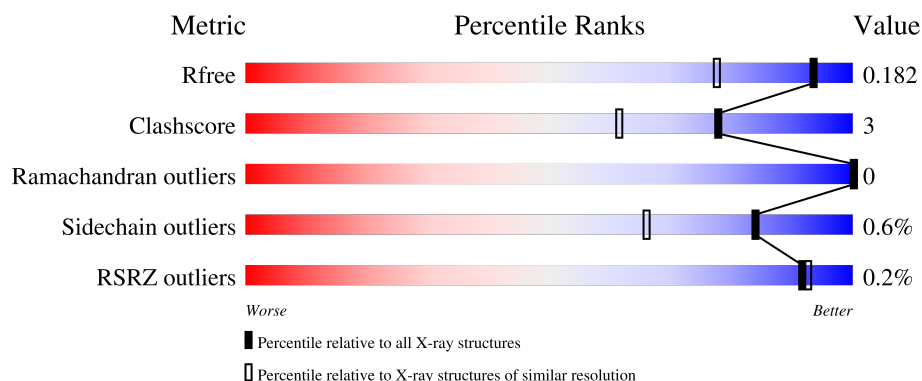
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.44 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	2809 (1.46-1.42)
Clashscore	180529	3008 (1.46-1.42)
Ramachandran outliers	177936	2971 (1.46-1.42)
Sidechain outliers	177891	2971 (1.46-1.42)
RSRZ outliers	164620	2809 (1.46-1.42)

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 $\leq 5\%$. 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	213	
1	C	213	

2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 3807 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.

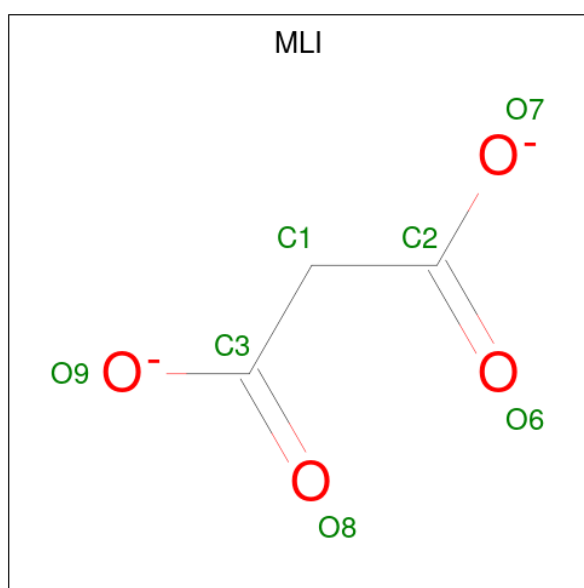
- Molecule 1 is a protein called Nickel ECF transporter, additional periplasmic component NikK.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	213	Total	C	N	O	S	0	8	0
			1615	1009	305	298	3			
1	C	213	Total	C	N	O	S	0	14	0
			1680	1045	319	313	3			

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

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	1	Total	Ni	0	0
			1	1		
2	C	1	Total	Ni	0	0
			1	1		

- Molecule 3 is MALONATE ION (CCD ID: MLI) (formula: C₃H₂O₄).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total C O 7 3 4	0	0
3	A	1	Total C O 7 3 4	0	0
3	C	1	Total C O 7 3 4	0	0
3	C	1	Total C O 7 3 4	0	0

- Molecule 4 is water.

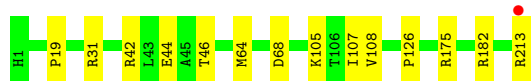
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	234	Total O 235 235	0	1
4	C	246	Total O 247 247	0	1

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.

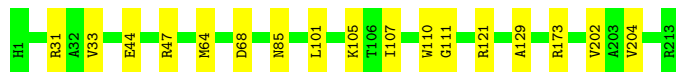
- Molecule 1: Nickel ECF transporter, additional periplasmic component NikK

Chain A:  93% 7%



- Molecule 1: Nickel ECF transporter, additional periplasmic component NikK

Chain C:  92% 8%



4 Data and refinement statistics

Property	Value
Space group	P 1
Cell constants a, b, c, α , β , γ	43.53Å 47.47Å 49.45Å 81.42° 81.50° 83.60°
Resolution (Å)	48.46 – 1.44 48.46 – 1.44
% Data completeness (in resolution range)	64.1 (48.46-1.44) 64.1 (48.46-1.44)
R_{merge}	0.03
R_{sym}	(Not available)
$\langle I/\sigma(I) \rangle$ ¹	1.59 (at 1.44Å)
Refinement program	REFMAC 5.8.0430 (refmacat 0.4.77), REFMAC 5.8.0430 (refmacat 0.4.77)
R, R_{free}	0.130 , 0.182 0.130 , 0.182
R_{free} test set	2199 reflections (4.92%)
Wilson B-factor (Å ²)	12.0
Anisotropy	0.533
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.37 , 38.9
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.33$
Estimated twinning fraction	0.003 for -h,-l,-k
F_o, F_c correlation	0.98
Total number of atoms	3807
Average B, all atoms (Å ²)	15.0

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.64% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical 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: MLI, 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	0.62	0/1657	1.12	3/2268 (0.1%)
1	C	0.62	0/1725	1.06	4/2359 (0.2%)
All	All	0.62	0/3382	1.09	7/4627 (0.2%)

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	3
1	C	0	1
All	All	0	4

There are no bond length outliers.

All (7) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	44	GLU	CB-CG-CD	6.69	123.97	112.60
1	C	44	GLU	CB-CG-CD	6.21	123.15	112.60
1	A	46	THR	CA-CB-OG1	-5.82	100.86	109.60
1	C	47	ARG	N-CA-CB	-5.34	101.90	109.85
1	C	68	ASP	CA-CB-CG	5.31	117.91	112.60
1	C	85	ASN	CA-CB-CG	5.08	117.69	112.60
1	A	68	ASP	CA-CB-CG	5.03	117.62	112.60

There are no chirality outliers.

All (4) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	175	ARG	Sidechain
1	A	182[A]	ARG	Sidechain
1	A	31	ARG	Sidechain
1	C	31	ARG	Sidechain

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	1615	0	1622	10	0
1	C	1680	0	1686	8	0
2	A	1	0	0	0	0
2	C	1	0	0	0	0
3	A	14	0	4	0	0
3	C	14	0	4	0	0
4	A	235	0	0	1	0
4	C	247	0	0	2	0
All	All	3807	0	3316	18	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 3.

All (18) 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:64[A]:MET:HE1	1:A:105[A]:LYS:HD3	1.38	1.05
1:A:64[A]:MET:CE	1:A:105[A]:LYS:HD3	2.09	0.82
1:A:64[A]:MET:HE2	1:A:105[A]:LYS:CG	2.15	0.76
1:A:64[A]:MET:HE2	1:A:105[A]:LYS:HG3	1.69	0.73
1:A:64[A]:MET:HE1	1:A:105[A]:LYS:CD	2.18	0.71
1:C:101:LEU:HB2	1:C:204[B]:VAL:HG12	1.75	0.68
1:A:107[A]:ILE:HD11	1:A:126:PRO:HG3	1.77	0.66
1:C:121:ARG:NH1	4:C:405:HOH:O	2.29	0.64
1:A:108[A]:VAL:HG12	4:A:612:HOH:O	1.98	0.62
1:A:64[A]:MET:CE	1:A:105[A]:LYS:CD	2.79	0.59
1:C:33:VAL:O	4:C:403:HOH:O	2.17	0.59
1:C:64[B]:MET:SD	1:C:105[B]:LYS:HD3	2.43	0.58

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:64[A]:MET:CE	1:A:105[A]:LYS:CG	2.85	0.54
1:C:111:GLY:C	1:C:129:ALA:HB1	2.39	0.47
1:C:202:VAL:HG22	1:C:204[B]:VAL:HG13	2.00	0.43
1:C:107[B]:ILE:HG21	1:C:110:TRP:CE2	2.53	0.42
1:A:42[B]:ARG:HE	1:A:42[B]:ARG:HB2	1.65	0.42
1:C:173[A]:ARG:HH11	1:C:173[A]:ARG:HD2	1.73	0.40

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	219/213 (103%)	215 (98%)	4 (2%)	0	100	100
1	C	226/213 (106%)	222 (98%)	4 (2%)	0	100	100
All	All	445/426 (104%)	437 (98%)	8 (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	163/158 (103%)	161 (99%)	2 (1%)	67	38
1	C	172/158 (109%)	172 (100%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
All	All	335/316 (106%)	333 (99%)	2 (1%)	84	66

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

Mol	Chain	Res	Type
1	A	19	PRO
1	A	213	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	118	GLN
1	A	207	ASN
1	C	207	ASN

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 6 ligands modelled in this entry, 2 are monoatomic - leaving 4 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	MLI	C	303	-	6,6,6	1.53	2 (33%)	7,7,7	1.12	0
3	MLI	C	302	2	6,6,6	1.36	1 (16%)	7,7,7	1.08	0
3	MLI	A	303	2	6,6,6	1.38	0	7,7,7	1.20	0
3	MLI	A	302	-	6,6,6	1.30	0	7,7,7	0.99	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	MLI	C	303	-	-	0/4/4/4	-
3	MLI	C	302	2	-	2/4/4/4	-
3	MLI	A	303	2	-	2/4/4/4	-
3	MLI	A	302	-	-	4/4/4/4	-

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	C	303	MLI	O7-C2	-2.16	1.23	1.30
3	C	303	MLI	O9-C3	-2.13	1.23	1.30
3	C	302	MLI	O7-C2	-2.08	1.23	1.30

There are no bond angle outliers.

There are no chirality outliers.

All (8) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	302	MLI	C2-C1-C3-O9
3	A	303	MLI	C3-C1-C2-O6
3	A	303	MLI	C3-C1-C2-O7
3	C	302	MLI	C2-C1-C3-O9
3	C	302	MLI	C2-C1-C3-O8
3	A	302	MLI	C2-C1-C3-O8
3	A	302	MLI	C3-C1-C2-O6
3	A	302	MLI	C3-C1-C2-O7

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, 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(Å ²)	Q<0.9
1	A	213/213 (100%)	-0.54	1 (0%) 87 89	5, 13, 23, 31	17 (7%)
1	C	213/213 (100%)	-0.71	0 100 100	3, 12, 19, 30	20 (9%)
All	All	426/426 (100%)	-0.62	1 (0%) 92 92	3, 12, 21, 31	37 (8%)

All (1) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	213	ARG	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 oligosaccharides 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, 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.

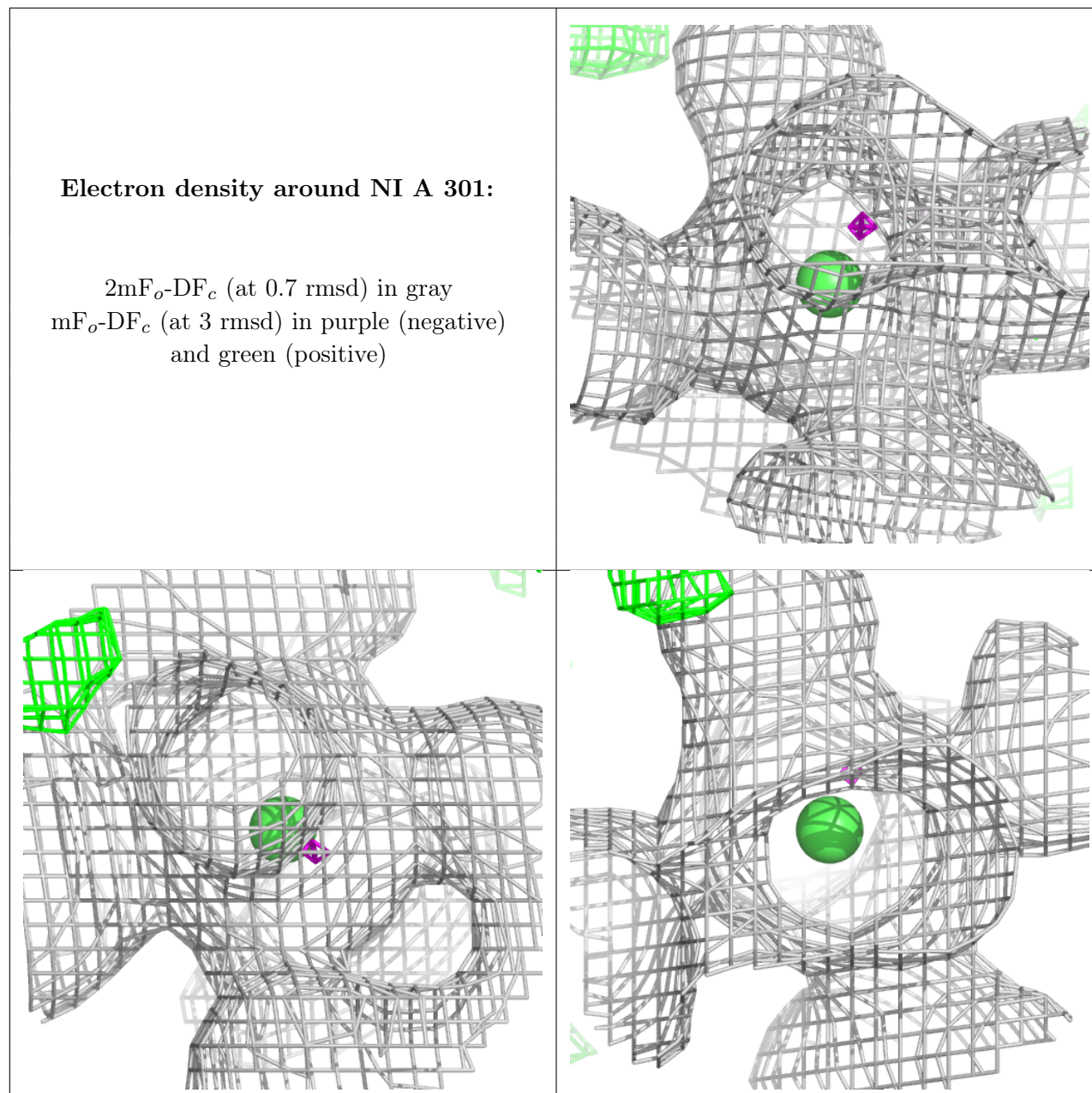
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å ²)	Q<0.9
3	MLI	C	302	7/7	0.92	0.10	17,20,21,22	7
3	MLI	C	303	7/7	0.92	0.08	18,20,22,22	7
3	MLI	A	303	7/7	0.93	0.09	23,27,28,32	7
3	MLI	A	302	7/7	0.95	0.06	19,21,22,25	0

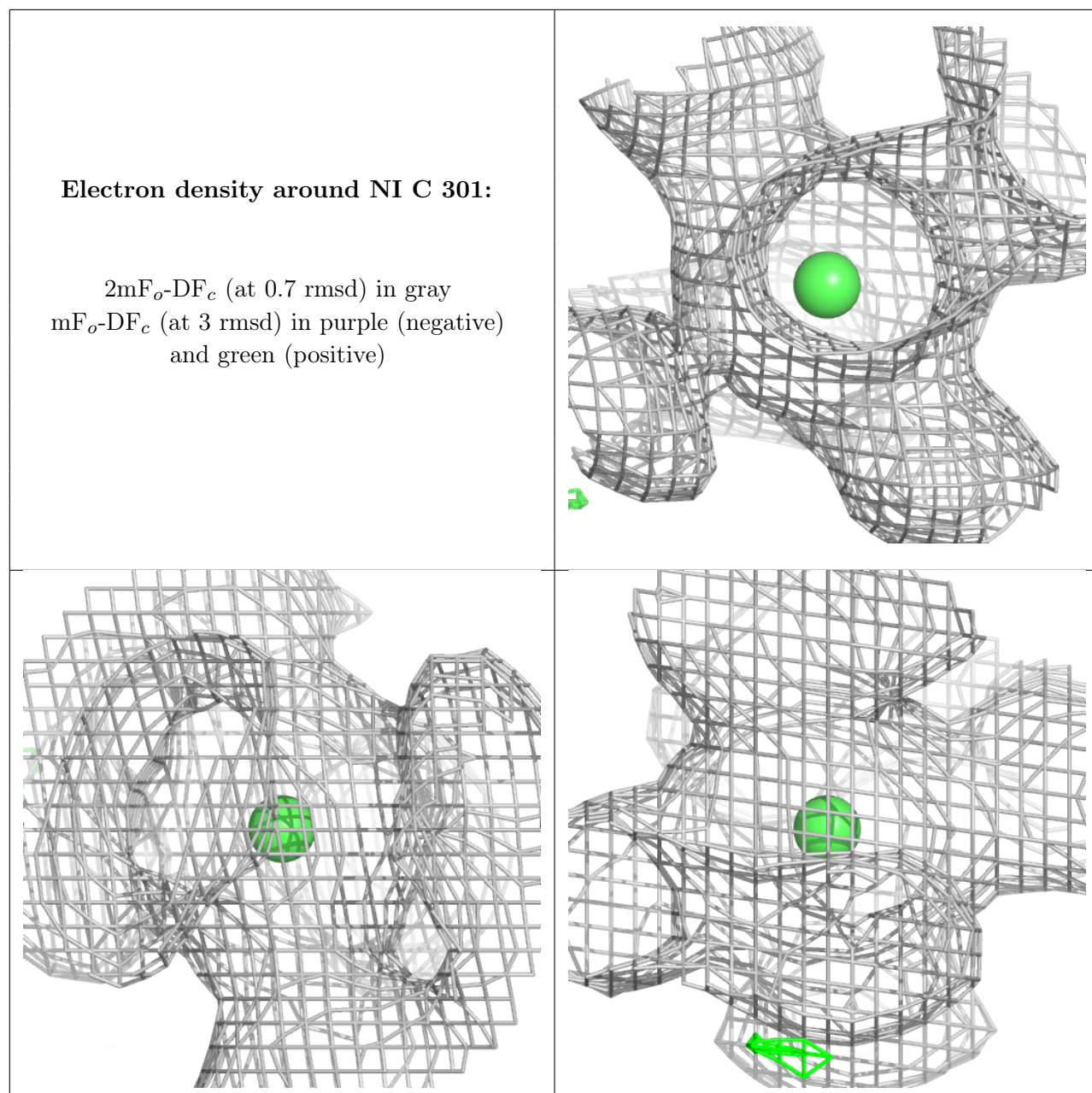
Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(\AA^2)	Q<0.9
2	NI	A	301	1/1	1.00	0.03	23,23,23,23	0
2	NI	C	301	1/1	1.00	0.01	22,22,22,22	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.





6.5 Other polymers ⓘ

There are no such residues in this entry.