Full wwPDB X-ray Structure Validation Report (i)

Aug 23, 2023 – 10:34 am BST

PDB ID	:	8BDD
Title	:	Crystal structure of Bacteroides ovatus CP926 PL17 alginate lyase
Authors	:	Roenne, M.E.; Tandrup, T.; Wilkens, C.
Deposited on	:	2022-10-19
Resolution	:	1.61 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 (i) 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 (1)) were used in the production of this report:

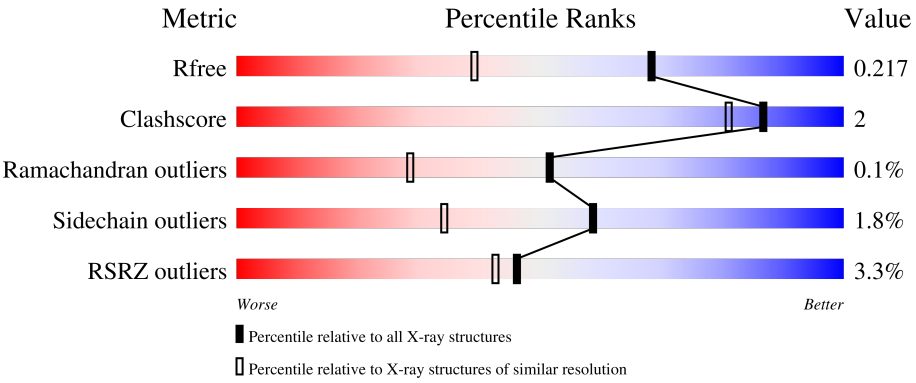
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.35
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

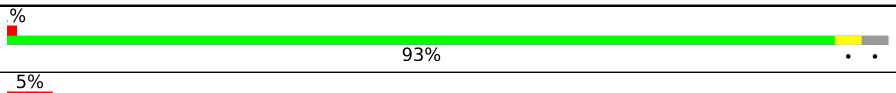
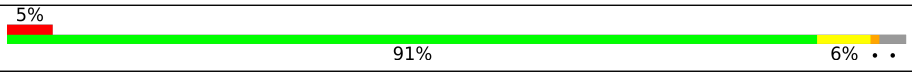
The reported resolution of this entry is 1.61 Å.

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	4693 (1.64-1.60)
Clashscore	141614	5002 (1.64-1.60)
Ramachandran outliers	138981	4888 (1.64-1.60)
Sidechain outliers	138945	4887 (1.64-1.60)
RSRZ outliers	127900	4609 (1.64-1.60)

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 <=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	738	
1	B	738	

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 12071 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 Alginate lyase family protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	715	Total	C	N	O	S	0	1	0
			5657	3602	962	1067	26			
1	B	715	Total	C	N	O	S	0	1	0
			5657	3602	962	1067	26			

There are 46 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	MET	-	initiating methionine	UNP A0A5M5BZE4
A	2	GLY	-	expression tag	UNP A0A5M5BZE4
A	3	SER	-	expression tag	UNP A0A5M5BZE4
A	4	SER	-	expression tag	UNP A0A5M5BZE4
A	5	HIS	-	expression tag	UNP A0A5M5BZE4
A	6	HIS	-	expression tag	UNP A0A5M5BZE4
A	7	HIS	-	expression tag	UNP A0A5M5BZE4
A	8	HIS	-	expression tag	UNP A0A5M5BZE4
A	9	HIS	-	expression tag	UNP A0A5M5BZE4
A	10	HIS	-	expression tag	UNP A0A5M5BZE4
A	11	SER	-	expression tag	UNP A0A5M5BZE4
A	12	SER	-	expression tag	UNP A0A5M5BZE4
A	13	GLY	-	expression tag	UNP A0A5M5BZE4
A	14	LEU	-	expression tag	UNP A0A5M5BZE4
A	15	VAL	-	expression tag	UNP A0A5M5BZE4
A	16	PRO	-	expression tag	UNP A0A5M5BZE4
A	17	ARG	-	expression tag	UNP A0A5M5BZE4
A	18	GLY	-	expression tag	UNP A0A5M5BZE4
A	19	SER	-	expression tag	UNP A0A5M5BZE4
A	20	HIS	-	expression tag	UNP A0A5M5BZE4
A	21	MET	-	expression tag	UNP A0A5M5BZE4
A	22	ALA	-	expression tag	UNP A0A5M5BZE4
A	23	SER	-	expression tag	UNP A0A5M5BZE4
B	1	MET	-	initiating methionine	UNP A0A5M5BZE4
B	2	GLY	-	expression tag	UNP A0A5M5BZE4

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	3	SER	-	expression tag	UNP A0A5M5BZE4
B	4	SER	-	expression tag	UNP A0A5M5BZE4
B	5	HIS	-	expression tag	UNP A0A5M5BZE4
B	6	HIS	-	expression tag	UNP A0A5M5BZE4
B	7	HIS	-	expression tag	UNP A0A5M5BZE4
B	8	HIS	-	expression tag	UNP A0A5M5BZE4
B	9	HIS	-	expression tag	UNP A0A5M5BZE4
B	10	HIS	-	expression tag	UNP A0A5M5BZE4
B	11	SER	-	expression tag	UNP A0A5M5BZE4
B	12	SER	-	expression tag	UNP A0A5M5BZE4
B	13	GLY	-	expression tag	UNP A0A5M5BZE4
B	14	LEU	-	expression tag	UNP A0A5M5BZE4
B	15	VAL	-	expression tag	UNP A0A5M5BZE4
B	16	PRO	-	expression tag	UNP A0A5M5BZE4
B	17	ARG	-	expression tag	UNP A0A5M5BZE4
B	18	GLY	-	expression tag	UNP A0A5M5BZE4
B	19	SER	-	expression tag	UNP A0A5M5BZE4
B	20	HIS	-	expression tag	UNP A0A5M5BZE4
B	21	MET	-	expression tag	UNP A0A5M5BZE4
B	22	ALA	-	expression tag	UNP A0A5M5BZE4
B	23	SER	-	expression tag	UNP A0A5M5BZE4

- 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	1	Total Ni 1 1	0	0
2	B	1	Total Ni 1 1	0	0

- Molecule 3 is water.

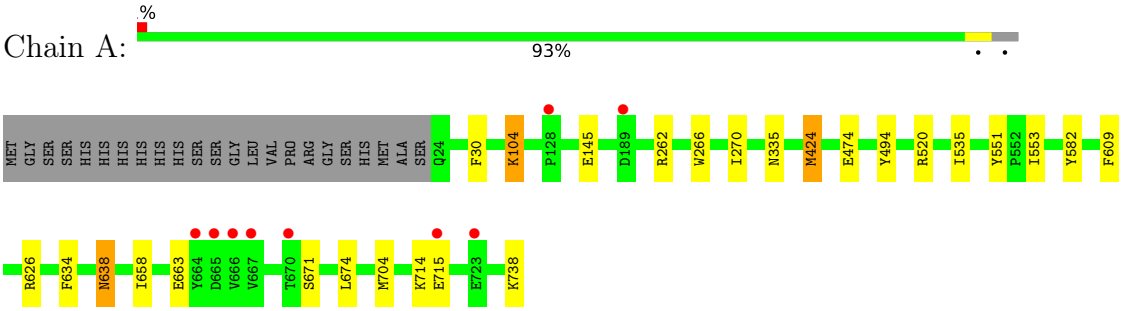
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	439	Total O 439 439	0	0
3	B	316	Total O 316 316	0	0

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.

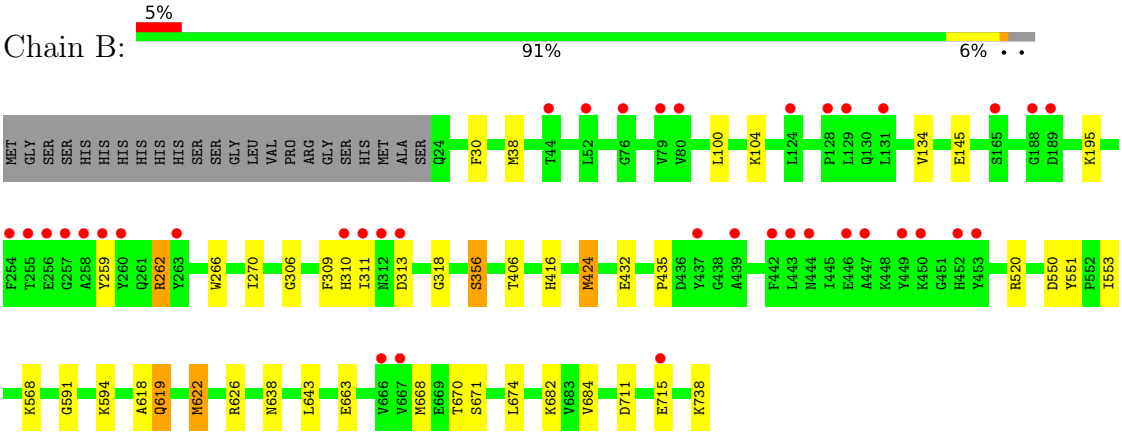
• Molecule 1: Alginate lyase family protein

Chain A:



• Molecule 1: Alginate lyase family protein

Chain B:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	89.59Å 120.95Å 90.24Å 90.00° 117.76° 90.00°	Depositor
Resolution (Å)	48.21 – 1.61 48.21 – 1.61	Depositor EDS
% Data completeness (in resolution range)	99.3 (48.21-1.61) 99.3 (48.21-1.61)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	0.96 (at 1.61Å)	Xtriage
Refinement program	REFMAC 8, PHENIX 1.19.2_4158	Depositor
R, R_{free}	0.190 , 0.212 0.191 , 0.217	Depositor DCC
R_{free} test set	2277 reflections (1.05%)	wwPDB-VP
Wilson B-factor (Å^2)	28.3	Xtriage
Anisotropy	0.198	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.36 , 39.9	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	0.000 for -h-l,k,h 0.000 for l,k,-h-l 0.011 for h,-k,-h-l 0.012 for -h-l,-k,l 0.011 for l,-k,h	Xtriage
F_o,F_c correlation	0.97	EDS
Total number of atoms	12071	wwPDB-VP
Average B, all atoms (Å^2)	32.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.26% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ 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: 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	0.35	0/5803	0.58	0/7866
1	B	0.35	0/5803	0.57	0/7866
All	All	0.35	0/11606	0.58	0/15732

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

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	5657	0	5475	22	0
1	B	5657	0	5475	33	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
3	A	439	0	0	2	1
3	B	316	0	0	4	1
All	All	12071	0	10950	47	1

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.

All (47) 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:B:266:TRP:NE1	1:B:270:ILE:HD11	2.06	0.70
1:B:682:LYS:HE3	1:B:684:VAL:HG12	1.73	0.69
1:B:38:MET:HE1	1:B:100:LEU:HD11	1.77	0.66
1:B:313:ASP:HB2	1:B:416:HIS:HB3	1.81	0.62
1:A:626:ARG:HH22	1:B:626:ARG:HH22	1.49	0.61
1:A:638:ASN:OD1	1:B:638:ASN:ND2	2.32	0.61
1:B:668:MET:HG3	1:B:670:THR:HG23	1.84	0.59
1:A:671:SER:HB2	1:A:674:LEU:HD11	1.85	0.58
1:A:663:GLU:O	1:A:671:SER:HA	2.04	0.58
1:B:551:TYR:HE2	1:B:553:ILE:HD11	1.69	0.58
1:A:424:MET:HG2	1:A:520:ARG:CZ	2.35	0.56
1:A:266:TRP:NE1	1:A:270:ILE:HD11	2.22	0.55
1:B:309:PHE:O	1:B:311:ILE:N	2.35	0.55
1:B:424:MET:HG2	1:B:520:ARG:CZ	2.37	0.54
1:B:551:TYR:CE2	1:B:553:ILE:HD11	2.42	0.53
1:B:594:LYS:HE2	3:B:985:HOH:O	2.09	0.53
1:B:259:TYR:O	1:B:262:ARG:HB3	2.10	0.51
1:A:335:ASN:HB2	3:A:1296:HOH:O	2.12	0.50
1:B:550:ASP:HB3	1:B:643:LEU:HD11	1.95	0.49
1:A:551:TYR:HE2	1:A:553:ILE:HD11	1.79	0.48
1:B:619:GLN:NE2	3:B:910:HOH:O	2.46	0.48
1:A:626:ARG:HH12	1:B:626:ARG:CZ	2.27	0.47
1:B:671:SER:HB2	1:B:674:LEU:HD11	1.96	0.47
3:A:1146:HOH:O	1:B:626:ARG:HD3	2.15	0.47
1:B:38:MET:CE	1:B:100:LEU:HD11	2.43	0.46
1:B:432:GLU:HG2	1:B:435:PRO:HB3	1.96	0.46
1:B:356:SER:OG	3:B:901:HOH:O	2.21	0.46
1:A:551:TYR:CE2	1:A:553:ILE:HD11	2.51	0.45
1:A:626:ARG:NH1	1:B:626:ARG:CZ	2.80	0.45
1:B:663:GLU:O	1:B:671:SER:HA	2.18	0.44
1:A:104:LYS:H	1:A:104:LYS:NZ	2.16	0.44
1:A:609:PHE:CE1	1:A:674:LEU:HD22	2.52	0.44
1:A:626:ARG:HH12	1:B:626:ARG:NH2	2.16	0.43
1:A:714:LYS:HB3	1:A:714:LYS:HE2	1.70	0.43
1:A:626:ARG:HD2	1:B:638:ASN:ND2	2.34	0.43
1:A:474:GLU:HG3	1:A:582:TYR:CD1	2.53	0.43
1:A:634:PHE:CZ	1:B:568:LYS:HE3	2.53	0.43
1:B:594:LYS:HD2	1:B:618:ALA:HB1	2.01	0.42
1:A:494:TYR:CE1	1:A:714:LYS:HD2	2.55	0.42
1:A:626:ARG:NH2	1:B:626:ARG:HH12	2.18	0.41
1:B:406:THR:O	1:B:424:MET:HA	2.20	0.41

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:711:ASP:OD2	3:B:902:HOH:O	2.22	0.41
1:A:494:TYR:CZ	1:A:714:LYS:HD2	2.56	0.41
1:B:306:GLY:O	1:B:318:GLY:HA3	2.21	0.41
1:B:591:GLY:O	1:B:622:MET:HG2	2.21	0.40
1:B:100:LEU:HA	1:B:100:LEU:HD13	1.91	0.40
1:A:535:ILE:HB	1:A:658:ILE:HB	2.04	0.40

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:A:1095:HOH:O	3:B:1096:HOH:O[1_454]	2.17	0.03

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	714/738 (97%)	693 (97%)	21 (3%)	0	100	100
1	B	714/738 (97%)	688 (96%)	25 (4%)	1 (0%)	51	28
All	All	1428/1476 (97%)	1381 (97%)	46 (3%)	1 (0%)	51	28

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	310	HIS

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	601/619 (97%)	592 (98%)	9 (2%)	65	43
1	B	601/619 (97%)	589 (98%)	12 (2%)	55	29
All	All	1202/1238 (97%)	1181 (98%)	21 (2%)	59	36

All (21) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	30	PHE
1	A	104	LYS
1	A	145	GLU
1	A	262	ARG
1	A	424	MET
1	A	638	ASN
1	A	704	MET
1	A	715	GLU
1	A	738	LYS
1	B	30	PHE
1	B	104	LYS
1	B	134	VAL
1	B	145	GLU
1	B	195	LYS
1	B	262	ARG
1	B	356	SER
1	B	424	MET
1	B	619	GLN
1	B	622	MET
1	B	715	GLU
1	B	738	LYS

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 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

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^{th} 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	715/738 (96%)	-0.29	9 (1%)	77	76	20, 29, 42, 105	0
1	B	715/738 (96%)	0.02	38 (5%)	26	23	19, 31, 56, 82	0
All	All	1430/1476 (96%)	-0.13	47 (3%)	46	42	19, 30, 50, 105	0

All (47) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	667	VAL	11.2
1	B	447	ALA	7.4
1	B	449	TYR	5.6
1	B	258	ALA	5.2
1	B	259	TYR	4.8
1	B	255	THR	4.7
1	B	443	LEU	4.4
1	B	667	VAL	4.3
1	B	254	PHE	4.2
1	B	189	ASP	4.2
1	B	444	ASN	4.1
1	B	312	ASN	3.9
1	B	311	ILE	3.9
1	B	437	TYR	3.7
1	B	442	PHE	3.6
1	B	76	GLY	3.6
1	B	44	THR	3.3
1	A	666	VAL	3.3
1	B	310	HIS	3.3
1	B	80	VAL	3.1
1	A	189	ASP	3.1
1	A	723	GLU	3.1
1	A	665	ASP	3.0
1	B	131	LEU	2.9

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	453	TYR	2.7
1	B	452	HIS	2.7
1	B	666	VAL	2.6
1	B	263	TYR	2.6
1	B	257	GLY	2.6
1	B	313	ASP	2.6
1	B	715	GLU	2.6
1	B	260	TYR	2.5
1	B	439	ALA	2.5
1	A	670	THR	2.4
1	B	124	LEU	2.4
1	B	79	VAL	2.3
1	B	128	PRO	2.3
1	B	446	GLU	2.2
1	B	450	LYS	2.2
1	B	129	LEU	2.2
1	A	128	PRO	2.1
1	B	52	LEU	2.1
1	B	188	GLY	2.1
1	A	715	GLU	2.1
1	B	256	GLU	2.1
1	B	165	SER	2.1
1	A	664	TYR	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^{th} 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.

Continued on next page...

Continued from previous page...

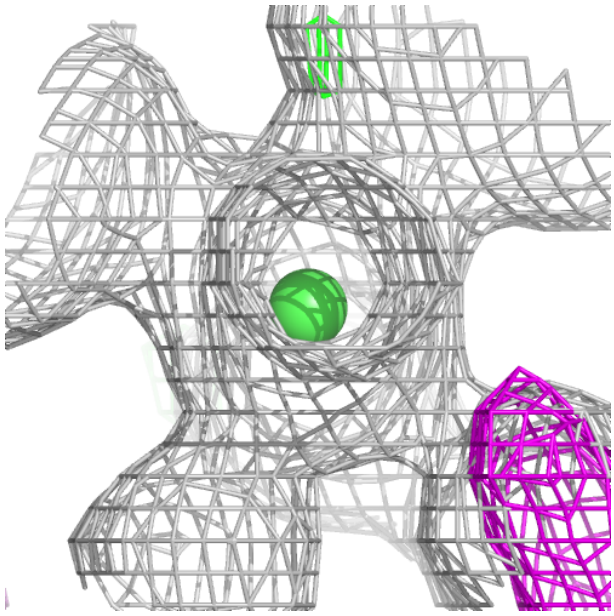
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9

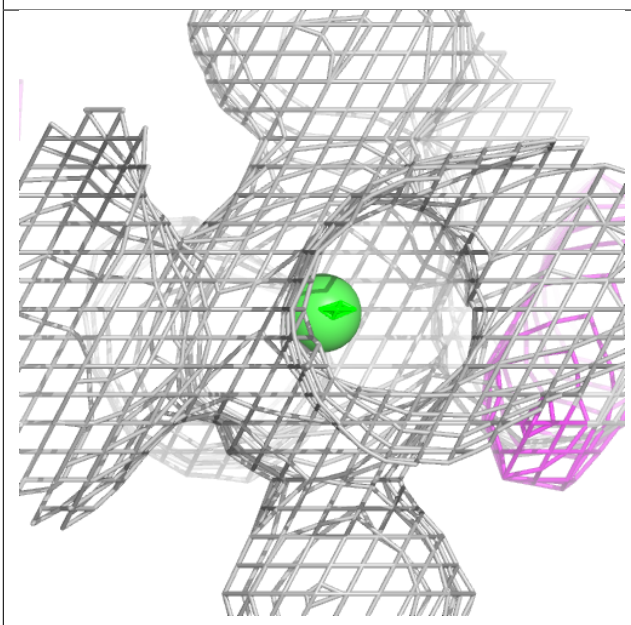
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
2	NI	A	801	1/1	0.99	0.07	19,19,19,19	1
2	NI	B	801	1/1	0.99	0.06	25,25,25,25	1

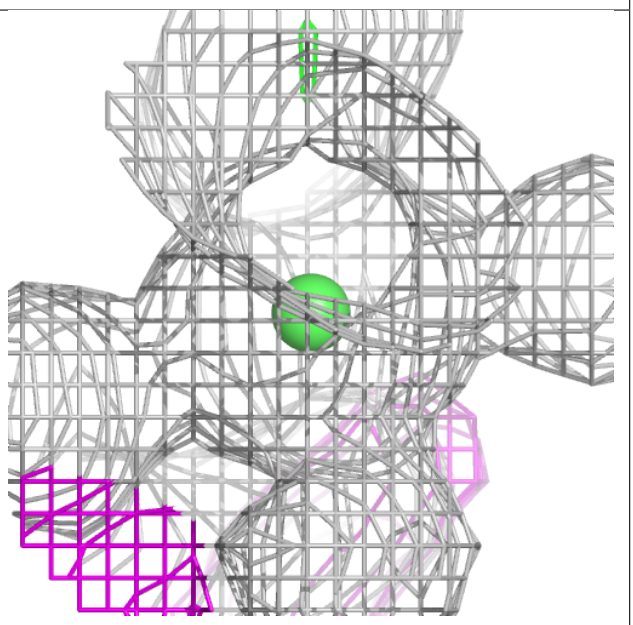
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation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

Electron density around NI A 801:

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

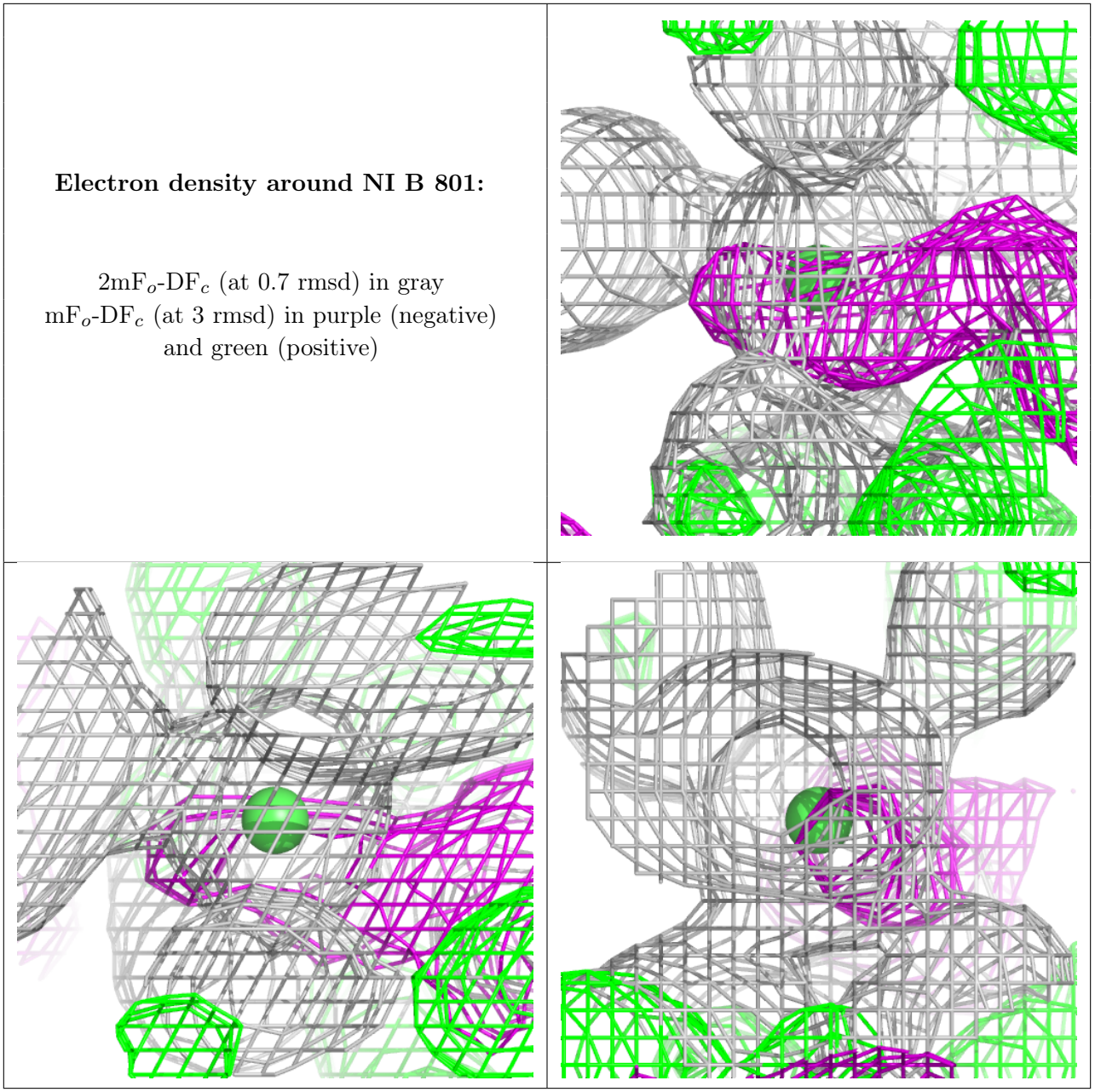






Electron density around NI B 801:

$2mF_o$-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.