



# Full wwPDB X-ray Structure Validation Report i

Jan 26, 2026 – 10:07 AM EST

PDB ID : 9N33 / pdb\_00009n33  
Title : Glutarate L-2-hydroxylase N187C mutant-5'-Mal-C6-AGCT DNA conjugate  
Authors : Han, Z.; Mirkin, C.A.  
Deposited on : 2025-01-29  
Resolution : 2.61 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](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 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	:	4.5-2 with Phenix2.0
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	2.0
EDS	:	3.0
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.010 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.47

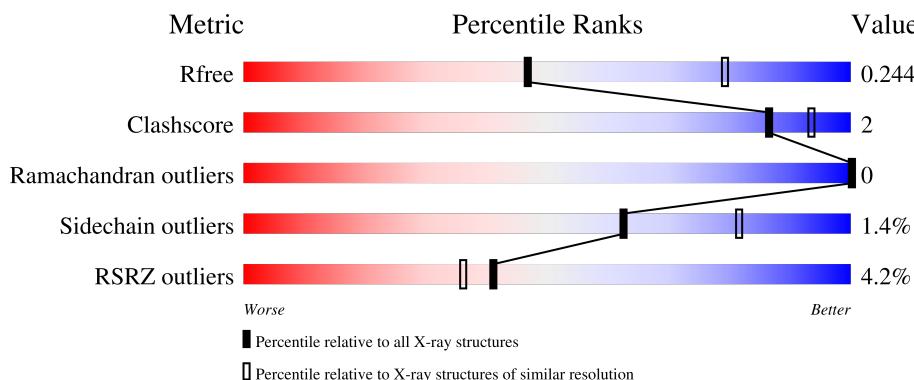
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## X-RAY DIFFRACTION

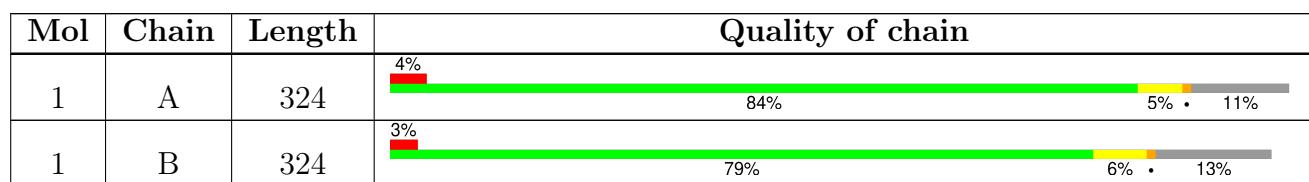
The reported resolution of this entry is 2.61 Å.

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))
$R_{free}$	164625	4623 (2.64-2.60)
Clashscore	180529	5071 (2.64-2.60)
Ramachandran outliers	177936	5006 (2.64-2.60)
Sidechain outliers	177891	5006 (2.64-2.60)
RSRZ outliers	164620	4622 (2.64-2.60)

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  $\geq 3$ , 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.



## 2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 9197 atoms, of which 4502 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.

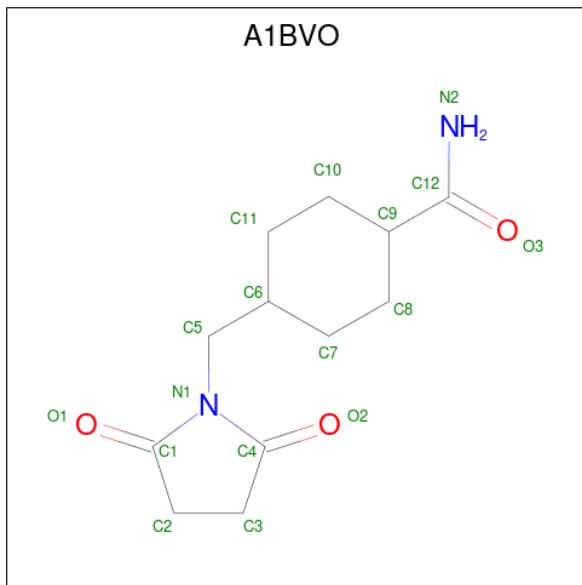
- Molecule 1 is a protein called Glutarate 2-hydroxylase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	288	Total	C	H	N	O	S	0	2	0
			4615	1506	2265	407	425	12			
1	B	282	Total	C	H	N	O	S	0	2	0
			4520	1477	2220	394	417	12			

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	24	THR	ILE	conflict	UNP B1IVJ9
A	187	CYS	ASN	engineered mutation	UNP B1IVJ9
A	321	HIS	-	expression tag	UNP B1IVJ9
A	322	HIS	-	expression tag	UNP B1IVJ9
A	323	HIS	-	expression tag	UNP B1IVJ9
A	324	HIS	-	expression tag	UNP B1IVJ9
B	24	THR	ILE	conflict	UNP B1IVJ9
B	187	CYS	ASN	engineered mutation	UNP B1IVJ9
B	321	HIS	-	expression tag	UNP B1IVJ9
B	322	HIS	-	expression tag	UNP B1IVJ9
B	323	HIS	-	expression tag	UNP B1IVJ9
B	324	HIS	-	expression tag	UNP B1IVJ9

- Molecule 2 is (1r,4r)-4-[(2,5-dioxopyrrolidin-1-yl)methyl]cyclohexane-1-carboxamide (CCD ID: A1BVO) (formula: C<sub>12</sub>H<sub>18</sub>N<sub>2</sub>O<sub>3</sub>) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	1	Total C H N O					0	0
			27	11	13	1	2		
2	B	1	Total C H N O					0	1
			9	4	2	1	2		
2	B	1	Total C H N O					0	1
			10	5	2	1	2		

- Molecule 3 is FE (II) ION (CCD ID: FE2) (formula: Fe) (labeled as "Ligand of Interest" by depositor).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	1	Total Fe		0	0
			1	1		
3	B	1	Total Fe		0	0
			1	1		

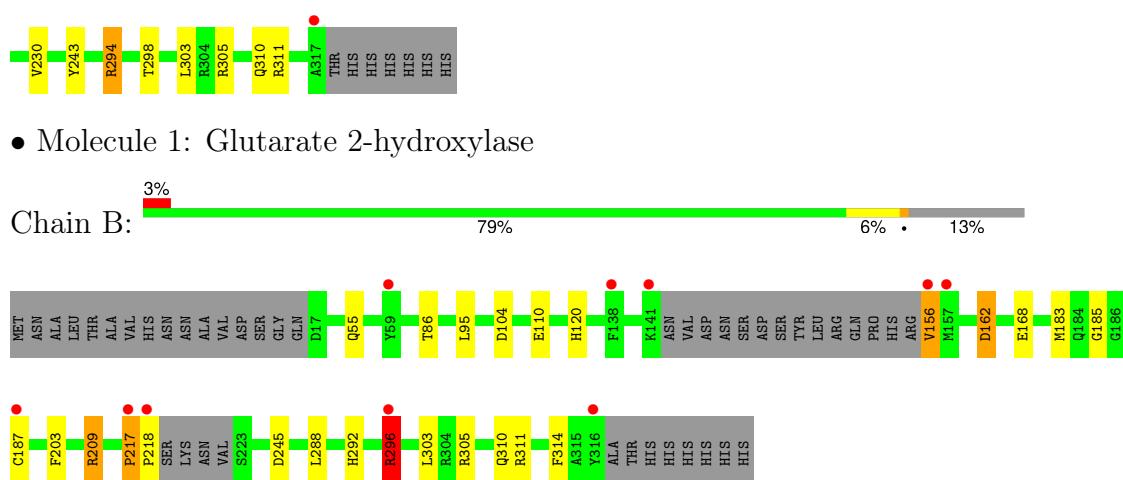
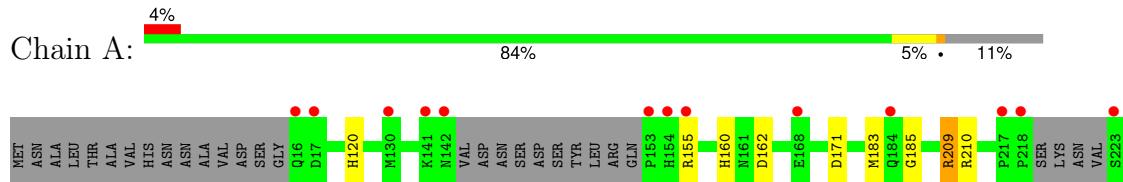
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	8	Total O		0	0
			8	8		
4	B	6	Total O		0	0
			6	6		

### 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: Glutarate 2-hydroxylase



## 4 Data and refinement statistics i

Property	Value	Source
Space group	P 4 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	124.94Å 124.94Å 127.29Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	127.29 – 2.61 127.29 – 2.61	Depositor EDS
% Data completeness (in resolution range)	99.9 (127.29-2.61) 99.9 (127.29-2.61)	Depositor EDS
$R_{merge}$	0.37	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle^1$	1.64 (at 2.62Å)	Xtriage
Refinement program	REFMAC 5.8.0430 (refmacat 0.4.88)	Depositor
$R$ , $R_{free}$	0.196, 0.240 0.200, 0.244	Depositor DCC
$R_{free}$ test set	1530 reflections (4.89%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	50.7	Xtriage
Anisotropy	0.302	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.40, 32.5	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.34$	Xtriage
Estimated twinning fraction	0.007 for -h,l,k 0.000 for -l,-k,-h	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	9197	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	55.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.67% 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: FE2, A1BVO

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	0.65	0/2414	1.15	5/3270 (0.2%)
1	B	0.60	0/2362	1.17	8/3200 (0.2%)
All	All	0.63	0/4776	1.16	13/6470 (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	4
1	B	0	1
All	All	0	5

There are no bond length outliers.

All (13) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed( $^{\circ}$ )	Ideal( $^{\circ}$ )
1	B	217	PRO	N-CA-C	10.65	123.69	110.70
1	A	162	ASP	CA-CB-CG	7.89	120.49	112.60
1	B	162	ASP	CA-CB-CG	7.35	119.95	112.60
1	B	120	HIS	CA-CB-CG	-5.81	107.99	113.80
1	B	55	GLN	N-CA-CB	5.74	118.56	110.12
1	A	298	THR	CA-CB-OG1	-5.63	101.16	109.60
1	B	104	ASP	CA-CB-CG	5.35	117.95	112.60
1	A	120	HIS	CA-CB-CG	-5.32	108.48	113.80
1	B	314	PHE	CA-CB-CG	5.17	118.97	113.80
1	A	171	ASP	CA-CB-CG	5.11	117.71	112.60
1	B	245	ASP	CA-CB-CG	5.09	117.69	112.60
1	B	110	GLU	CB-CG-CD	5.08	121.23	112.60

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	160	HIS	CA-CB-CG	5.05	118.85	113.80

There are no chirality outliers.

All (5) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	155	ARG	Sidechain
1	A	209	ARG	Sidechain
1	A	210	ARG	Sidechain
1	A	294	ARG	Sidechain
1	B	296	ARG	Sidechain

## 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	2350	2265	2303	6	0
1	B	2300	2220	2255	12	0
2	A	14	13	0	1	0
2	B	15	4	0	2	0
3	A	1	0	0	0	0
3	B	1	0	0	0	0
4	A	8	0	0	0	0
4	B	6	0	0	0	0
All	All	4695	4502	4558	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 2.

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:B:187[A]:CYS:HA	2:B:401[A]:A1BVO:C2	1.95	0.96
1:A:310:GLN:C	1:A:311:ARG:HD3	2.15	0.70
1:B:310:GLN:C	1:B:311:ARG:HD3	2.18	0.69

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:185:GLY:HA2	2:A:401:A1BVO:O2	1.94	0.67
1:B:203:PHE:O	1:B:209:ARG:NH2	2.28	0.67
1:B:185:GLY:HA2	2:B:402[B]:A1BVO:C2	2.34	0.58
1:B:183:MET:HE2	1:B:305:ARG:CZ	2.40	0.52
1:B:162:ASP:OD1	1:B:292:HIS:CE1	2.64	0.51
1:B:156:VAL:HG23	1:B:296:ARG:HG2	1.94	0.48
1:B:217:PRO:HG2	1:B:218:PRO:HD3	1.94	0.48
1:A:311:ARG:HD3	1:A:311:ARG:N	2.27	0.48
1:B:303:LEU:HD23	1:B:303:LEU:C	2.38	0.48
1:A:303:LEU:C	1:A:303:LEU:HD23	2.39	0.47
1:B:311:ARG:HD3	1:B:311:ARG:N	2.30	0.46
1:A:183:MET:HE2	1:A:305:ARG:HH11	1.81	0.46
1:B:86:THR:HG21	1:B:95:LEU:HD21	2.02	0.40
1:B:296:ARG:H	1:B:296:ARG:HG3	1.72	0.40
1:A:230:VAL:HG13	1:A:243:TYR:HB2	2.03	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	284/324 (88%)	280 (99%)	4 (1%)	0	100 100
1	B	278/324 (86%)	273 (98%)	5 (2%)	0	100 100
All	All	562/648 (87%)	553 (98%)	9 (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	255/285 (90%)	253 (99%)	2 (1%)	79 91
1	B	250/285 (88%)	245 (98%)	5 (2%)	50 73
All	All	505/570 (89%)	498 (99%)	7 (1%)	62 81

All (7) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	209	ARG
1	A	294	ARG
1	B	156	VAL
1	B	168	GLU
1	B	209	ARG
1	B	288	LEU
1	B	296	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (6) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	77	ASN
1	A	98	ASN
1	A	154	HIS
1	A	235	GLN
1	B	126	ASN
1	B	235	GLN

### 5.3.3 RNA [\(i\)](#)

There are no RNA molecules in this entry.

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

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 5 ligands modelled in this entry, 2 are monoatomic - leaving 3 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
2	A1BVO	A	401	1	15,15,18	0.35	0	18,20,25	1.25	3 (16%)
2	A1BVO	B	401[A]	1	7,7,18	0.59	0	9,9,25	1.10	0
2	A1BVO	B	402[B]	1	8,8,18	0.54	0	11,11,25	1.50	2 (18%)

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
2	A1BVO	A	401	1	-	2/4/25/31	0/2/2/2
2	A1BVO	B	401[A]	1	-	-	0/1/1/2
2	A1BVO	B	402[B]	1	-	-	0/1/1/2

There are no bond length outliers.

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	402[B]	A1BVO	O2-C4-N1	3.89	126.74	123.96
2	A	401	A1BVO	C1-N1-C4	-3.47	111.42	112.94
2	A	401	A1BVO	O2-C4-N1	2.74	126.87	123.94
2	A	401	A1BVO	C2-C3-C4	-2.45	102.73	105.24
2	B	402[B]	A1BVO	C1-N1-C4	-2.05	111.97	113.08

There are no chirality outliers.

All (2) torsion outliers are listed below:

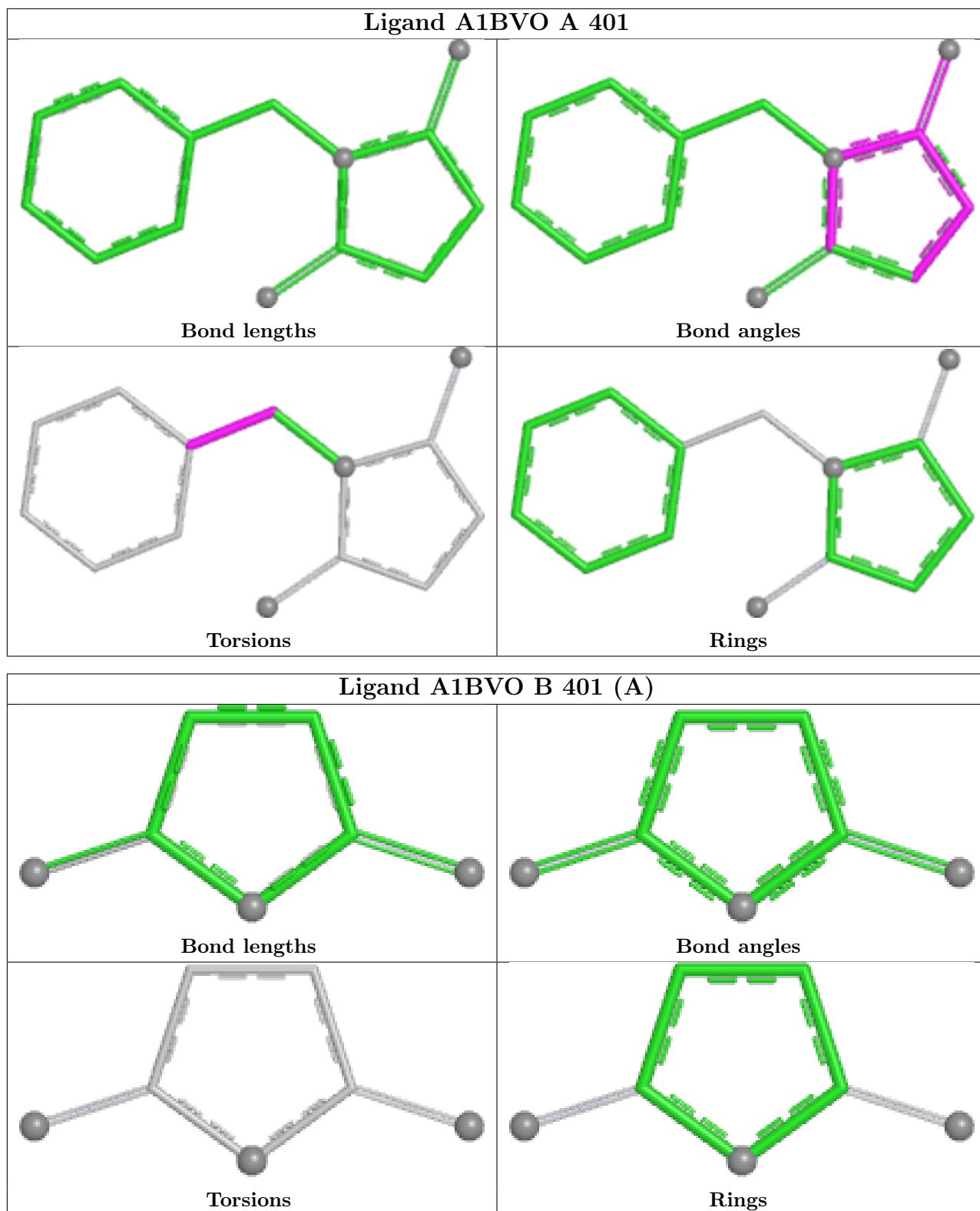
Mol	Chain	Res	Type	Atoms
2	A	401	A1BVO	N1-C5-C6-C11
2	A	401	A1BVO	N1-C5-C6-C7

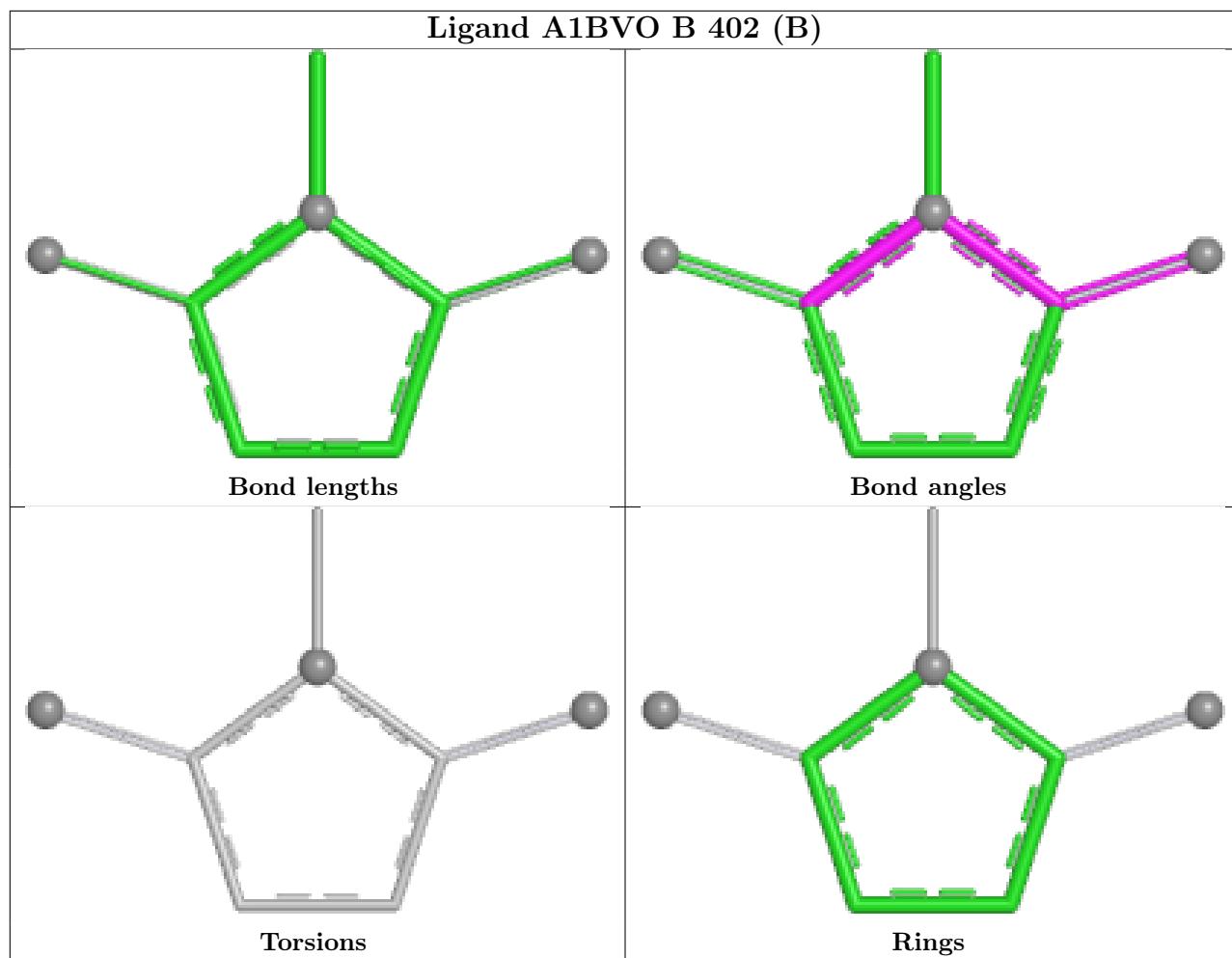
There are no ring outliers.

3 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	401	A1BVO	1	0
2	B	401[A]	A1BVO	1	0
2	B	402[B]	A1BVO	1	0

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less than 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.





## 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, 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	288/324 (88%)	-0.26	14 (4%) 36 31	31, 49, 82, 138	1 (0%)
1	B	282/324 (87%)	-0.30	10 (3%) 47 42	36, 52, 88, 118	1 (0%)
All	All	570/648 (87%)	-0.28	24 (4%) 41 36	31, 51, 87, 138	2 (0%)

All (24) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	153	PRO	7.8
1	A	142	ASN	6.6
1	B	218	PRO	5.6
1	A	154	HIS	5.2
1	B	141	LYS	5.1
1	A	218	PRO	5.1
1	B	156	VAL	4.8
1	B	316	TYR	4.6
1	A	155	ARG	4.0
1	B	217	PRO	3.9
1	A	317	ALA	3.8
1	A	217	PRO	3.5
1	A	16	GLN	3.4
1	A	223	SER	3.3
1	B	187[A]	CYS	2.9
1	B	59	TYR	2.8
1	A	130	MET	2.8
1	A	141	LYS	2.6
1	A	184	GLN	2.4
1	A	168	GLU	2.3
1	B	296	ARG	2.3
1	B	157	MET	2.2
1	A	17	ASP	2.2
1	B	138	PHE	2.2

## 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 [\(i\)](#)

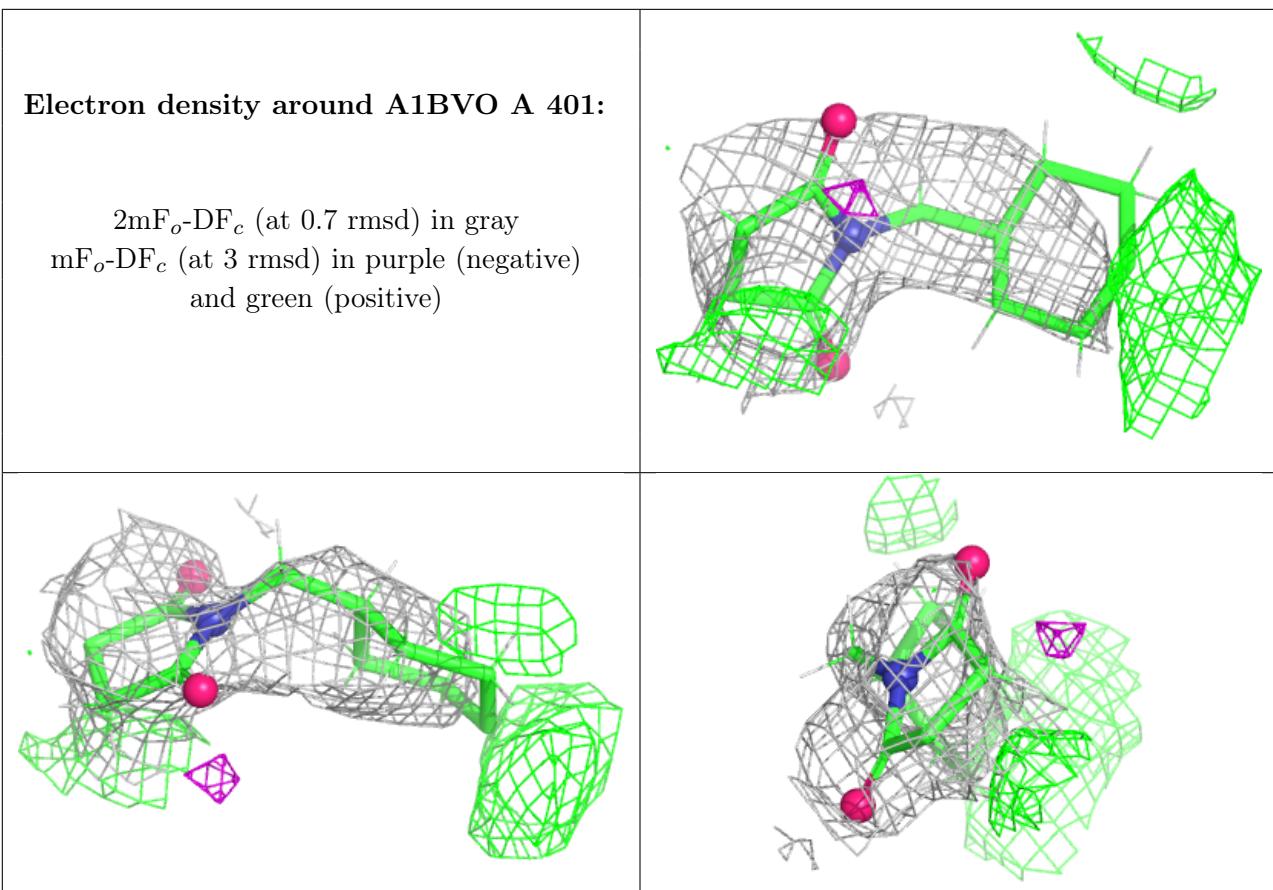
There are no oligosaccharides in this entry.

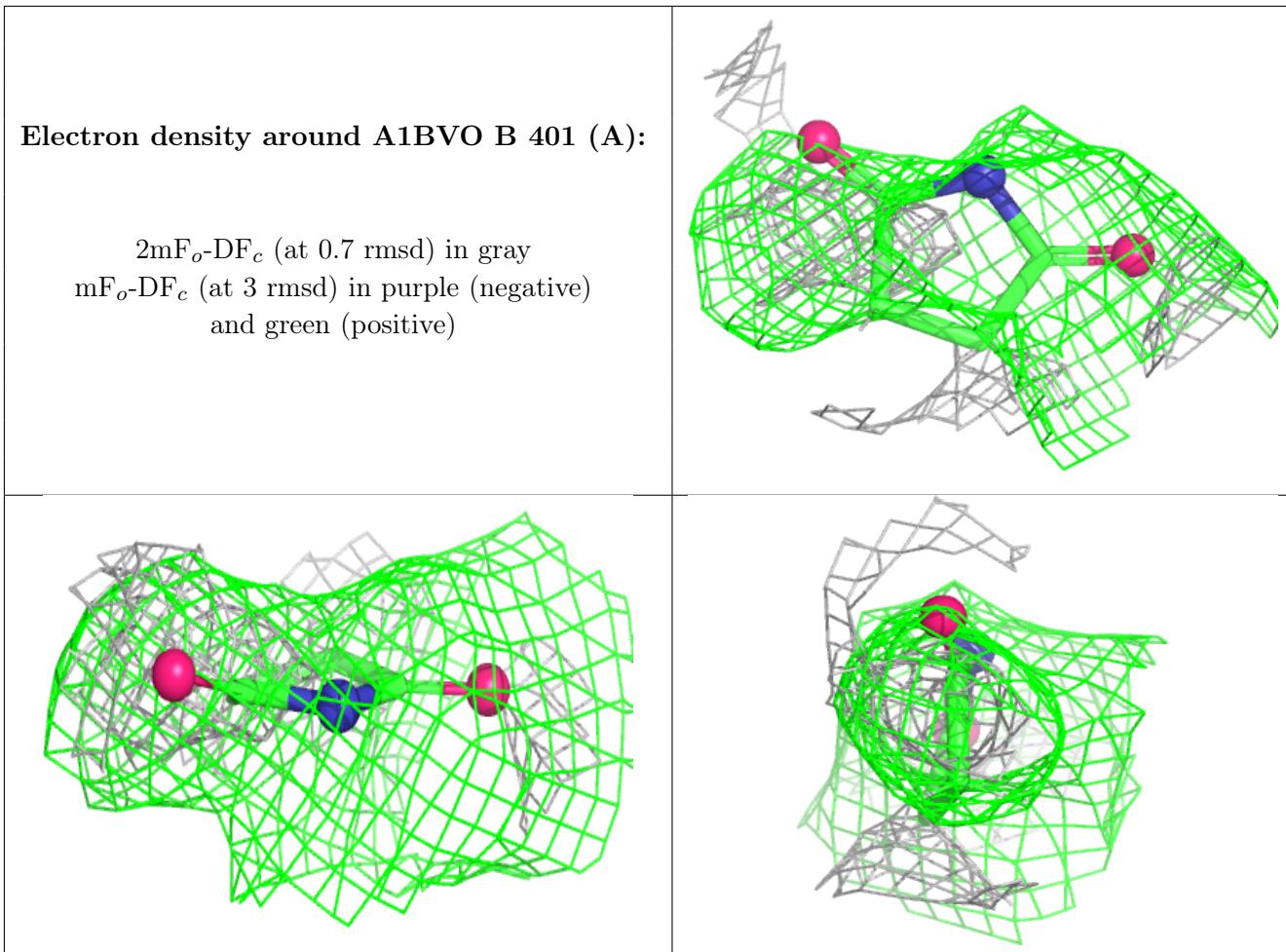
## 6.4 Ligands [\(i\)](#)

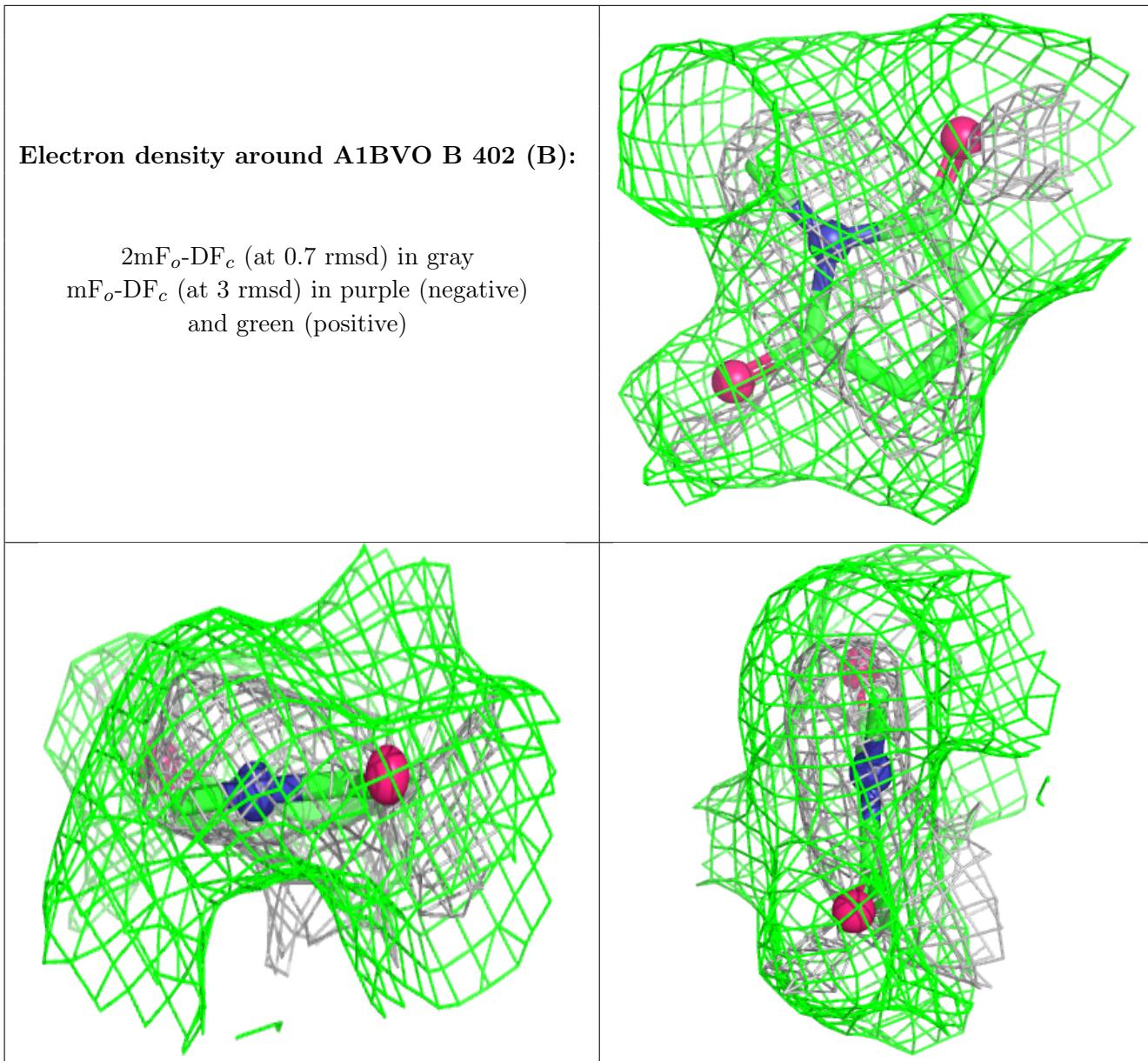
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q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

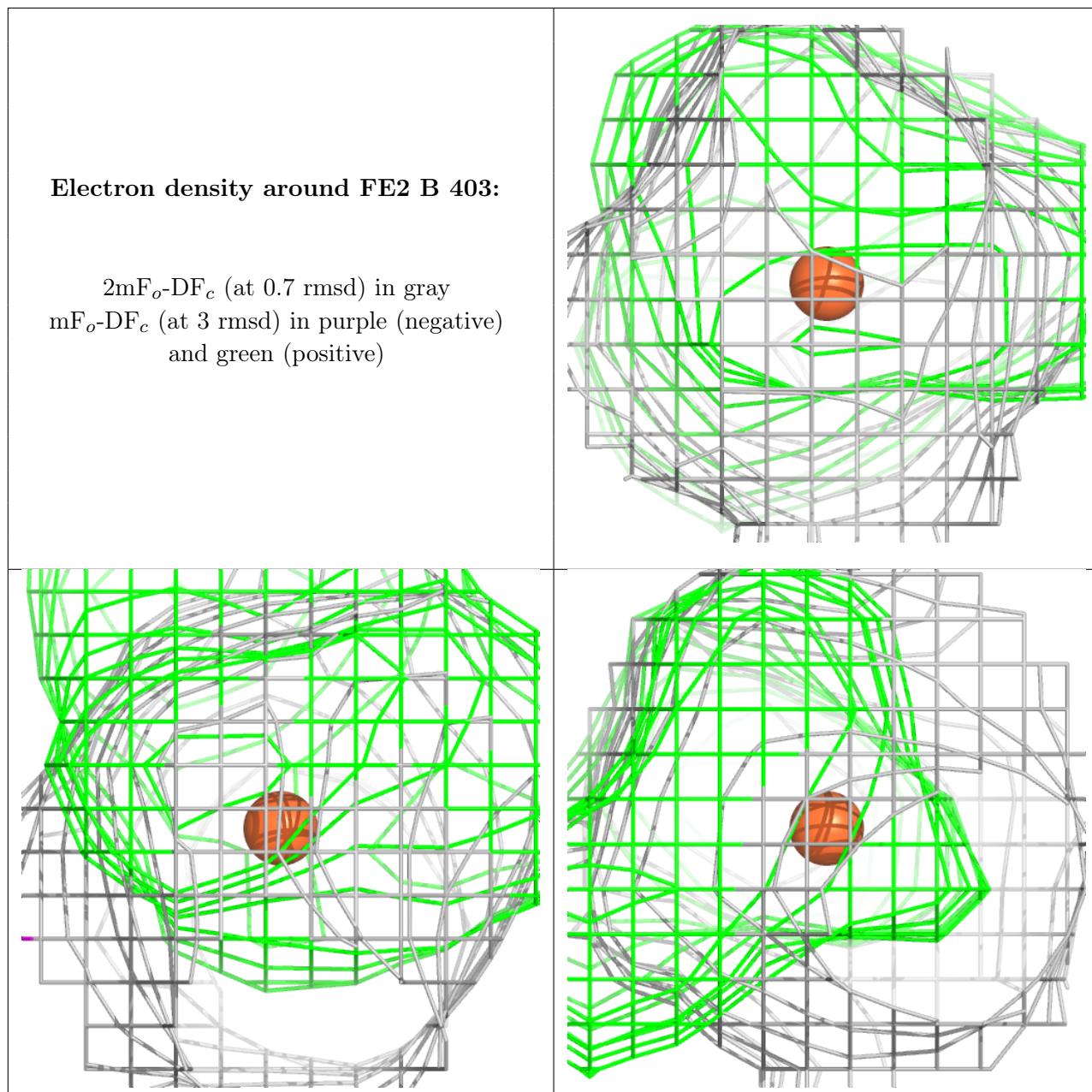
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
2	A1BVO	A	401	14/17	0.68	0.33	98,124,151,170	0
2	A1BVO	B	401[A]	7/17	0.84	0.52	121,166,170,174	9
2	A1BVO	B	402[B]	8/17	0.90	0.39	74,146,168,172	10
3	FE2	B	403	1/1	0.98	0.10	40,40,40,40	0
3	FE2	A	402	1/1	0.99	0.09	35,35,35,35	0

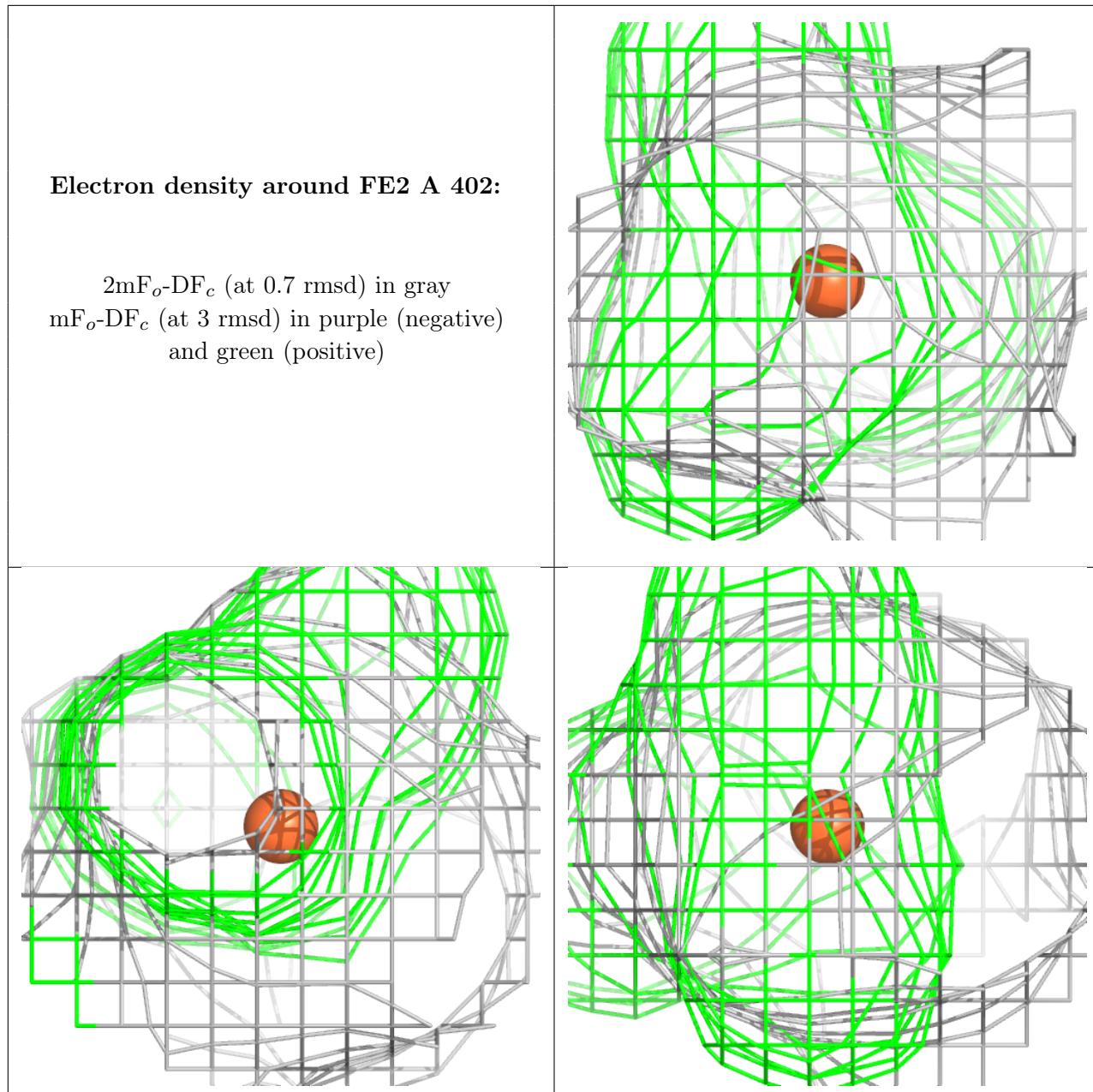
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation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











## 6.5 Other polymers [\(i\)](#)

There are no such residues in this entry.