



wwPDB X-ray Structure Validation Summary Report

Apr 25, 2026 – 03:02 PM EDT

PDB ID : 4OZO / pdb_00004ozo
Title : Crystal structure of an α -L-fucosidase GH29 from *Bacteroides thetaiotaomicron* (BT2192) in complex with α NPTG
Authors : Lafite, P.; Daniellou, R.; Guillotin, L.
Deposited on : 2014-02-17
Resolution : 2.60 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

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 Phenix2.0
Mogul : 2022.3.0, CSD as543be (2022)
Xtriage (Phenix) : 2.0
EDS : 3.0
Buster-report : wwPDB partial adaption of 1.1.7 (2018)
Percentile statistics : 20250101.v01 (using entries in the PDB archive January 1st 2025)
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) : 2.49

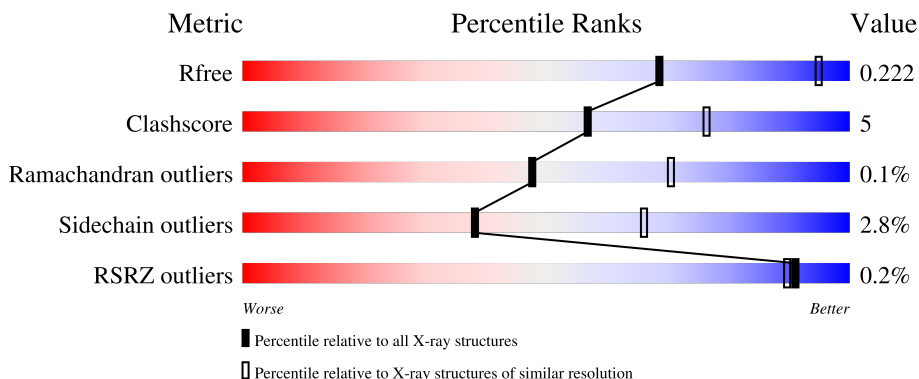
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.60 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}	180053	4008 (2.60-2.60)
Clashscore	190562	4347 (2.60-2.60)
Ramachandran outliers	187476	4277 (2.60-2.60)
Sidechain outliers	187428	4277 (2.60-2.60)
RSRZ outliers	180081	4008 (2.60-2.60)

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 ≥ 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	469	 85% 12% ..
1	B	469	 83% 14% .

2 Entry composition i

There are 4 unique types of molecules in this entry. The entry contains 7569 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 Putative lipoprotein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	459	3650	2332	618	684	16	2	0	0
1	B	459	3650	2332	618	684	16	3	0	0

There are 22 discrepancies between the modelled and reference sequences:

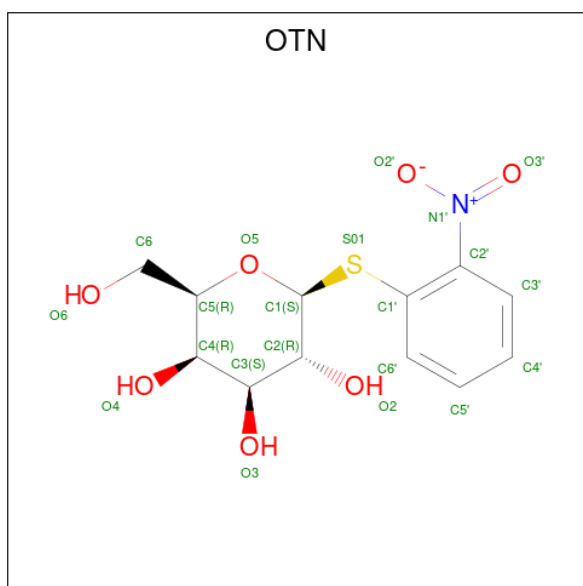
Chain	Residue	Modelled	Actual	Comment	Reference
A	23	MET	-	expression tag	UNP Q8A5P6
A	24	SER	-	expression tag	UNP Q8A5P6
A	25	LEU	-	expression tag	UNP Q8A5P6
A	484	GLU	-	expression tag	UNP Q8A5P6
A	485	GLY	-	expression tag	UNP Q8A5P6
A	486	HIS	-	expression tag	UNP Q8A5P6
A	487	HIS	-	expression tag	UNP Q8A5P6
A	488	HIS	-	expression tag	UNP Q8A5P6
A	489	HIS	-	expression tag	UNP Q8A5P6
A	490	HIS	-	expression tag	UNP Q8A5P6
A	491	HIS	-	expression tag	UNP Q8A5P6
B	23	MET	-	expression tag	UNP Q8A5P6
B	24	SER	-	expression tag	UNP Q8A5P6
B	25	LEU	-	expression tag	UNP Q8A5P6
B	484	GLU	-	expression tag	UNP Q8A5P6
B	485	GLY	-	expression tag	UNP Q8A5P6
B	486	HIS	-	expression tag	UNP Q8A5P6
B	487	HIS	-	expression tag	UNP Q8A5P6
B	488	HIS	-	expression tag	UNP Q8A5P6
B	489	HIS	-	expression tag	UNP Q8A5P6
B	490	HIS	-	expression tag	UNP Q8A5P6
B	491	HIS	-	expression tag	UNP Q8A5P6

- Molecule 2 is GLYCEROL (CCD ID: GOL) (formula: C₃H₈O₃).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	1	Total C O 6 3 3	0	0
2	A	1	Total C O 6 3 3	0	0
2	B	1	Total C O 6 3 3	0	0
2	B	1	Total C O 6 3 3	0	0

- Molecule 3 is 2-nitrophenyl 1-thio-beta-D-galactopyranoside (CCD ID: OTN) (formula: $C_{12}H_{15}NO_7S$).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	A	1	Total	C	N	O	S	0	0
			21	12	1	7	1		
3	B	1	Total	C	N	O	S	0	0
			21	12	1	7	1		

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	106	Total	O	0	0
			106	106		
4	B	97	Total	O	0	0
			97	97		

4 Data and refinement statistics

Property	Value	Source
Space group	I 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	125.30Å 122.17Å 159.65Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	49.00 – 2.60 49.00 – 2.60	Depositor EDS
% Data completeness (in resolution range)	100.0 (49.00-2.60) 100.0 (49.00-2.60)	Depositor EDS
R_{merge}	0.13	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.52 (at 2.61Å)	Xtrriage
Refinement program	PHENIX (phenix.refine: 1.8.1_1168)	Depositor
R, R_{free}	0.165 , 0.220 0.170 , 0.222	Depositor DCC
R_{free} test set	1901 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å ²)	37.8	Xtrriage
Anisotropy	0.879	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.37 , 36.7	EDS
L-test for twinning ²	$\langle L \rangle = 0.47$, $\langle L^2 \rangle = 0.31$	Xtrriage
Estimated twinning fraction	0.074 for -k,-h,-l	Xtrriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	7569	wwPDB-VP
Average B, all atoms (Å ²)	18.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.98% 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 [i](#)

5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: GOL, OTN

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	0.46	0/3753	0.80	5/5106 (0.1%)
1	B	0.48	0/3753	0.82	3/5106 (0.1%)
All	All	0.47	0/7506	0.81	8/10212 (0.1%)

There are no bond length outliers.

The worst 5 of 8 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	223	TYR	CA-C-N	-7.72	111.88	120.45
1	A	223	TYR	C-N-CA	-7.72	111.88	120.45
1	B	223	TYR	CA-C-N	-7.12	112.47	119.87
1	B	223	TYR	C-N-CA	-7.12	112.47	119.87
1	A	455	LYS	N-CA-C	-5.11	107.59	113.88

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	3650	0	3485	32	0
1	B	3650	0	3485	35	0
2	A	12	0	16	3	0
2	B	12	0	16	1	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	A	21	0	15	2	0
3	B	21	0	15	2	0
4	A	106	0	0	2	0
4	B	97	0	0	2	0
All	All	7569	0	7032	66	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 5.

The worst 5 of 66 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:208:ARG:NH2	1:B:214:CYS:O	2.22	0.71
1:A:339:GLN:NE2	1:B:343:GLU:OE2	2.25	0.67
1:A:57:GLU:OE2	4:A:650:HOH:O	2.15	0.64
1:A:310:ARG:HH21	2:A:502:GOL:H12	1.61	0.64
1:A:284:TRP:CD1	2:A:501:GOL:H11	2.34	0.62

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	457/469 (97%)	440 (96%)	16 (4%)	1 (0%)	43 66
1	B	457/469 (97%)	441 (96%)	16 (4%)	0	100 100
All	All	914/938 (97%)	881 (96%)	32 (4%)	1 (0%)	48 70

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	254	GLU

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	377/393 (96%)	367 (97%)	10 (3%)	39	67
1	B	377/393 (96%)	366 (97%)	11 (3%)	37	65
All	All	754/786 (96%)	733 (97%)	21 (3%)	38	66

5 of 21 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	186	TRP
1	B	389	ASP
1	B	449	LEU
1	B	391	ILE
1	B	373	ASN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 6 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	373	ASN
1	A	462	GLN
1	B	456	GLN
1	A	170	HIS
1	A	149	HIS

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](#)

6 ligands are modelled in this entry.

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	OTN	B	503	-	22,22,22	2.43	4 (18%)	28,31,31	1.38	4 (14%)
2	GOL	A	501	-	5,5,5	1.15	0	5,5,5	0.46	0
3	OTN	A	503	-	22,22,22	2.32	4 (18%)	28,31,31	1.46	5 (17%)
2	GOL	B	502	-	5,5,5	1.14	0	5,5,5	0.50	0
2	GOL	B	501	-	5,5,5	0.93	0	5,5,5	0.26	0
2	GOL	A	502	-	5,5,5	0.94	0	5,5,5	0.59	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	OTN	B	503	-	-	0/8/30/30	0/2/2/2
2	GOL	A	501	-	-	0/4/4/4	-
3	OTN	A	503	-	-	0/8/30/30	0/2/2/2
2	GOL	B	502	-	-	2/4/4/4	-
2	GOL	B	501	-	-	0/4/4/4	-
2	GOL	A	502	-	-	2/4/4/4	-

The worst 5 of 8 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	B	503	OTN	O3'-N1'	7.76	1.36	1.22
3	A	503	OTN	O3'-N1'	7.63	1.35	1.22
3	A	503	OTN	C2'-N1'	-5.85	1.35	1.45
3	B	503	OTN	C2'-N1'	-5.47	1.35	1.45
3	B	503	OTN	C1'-S01	-4.88	1.70	1.77

The worst 5 of 9 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	503	OTN	C2-C1-S01	-2.92	104.38	109.37
3	B	503	OTN	O3'-N1'-C2'	2.85	123.91	119.03
3	A	503	OTN	C2'-C1'-S01	2.83	124.74	122.11
3	B	503	OTN	C2-C1-S01	-2.68	104.80	109.37
3	A	503	OTN	O3'-N1'-C2'	2.54	123.37	119.03

There are no chirality outliers.

All (4) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	502	GOL	C1-C2-C3-O3
2	B	502	GOL	C1-C2-C3-O3
2	A	502	GOL	O2-C2-C3-O3
2	B	502	GOL	O2-C2-C3-O3

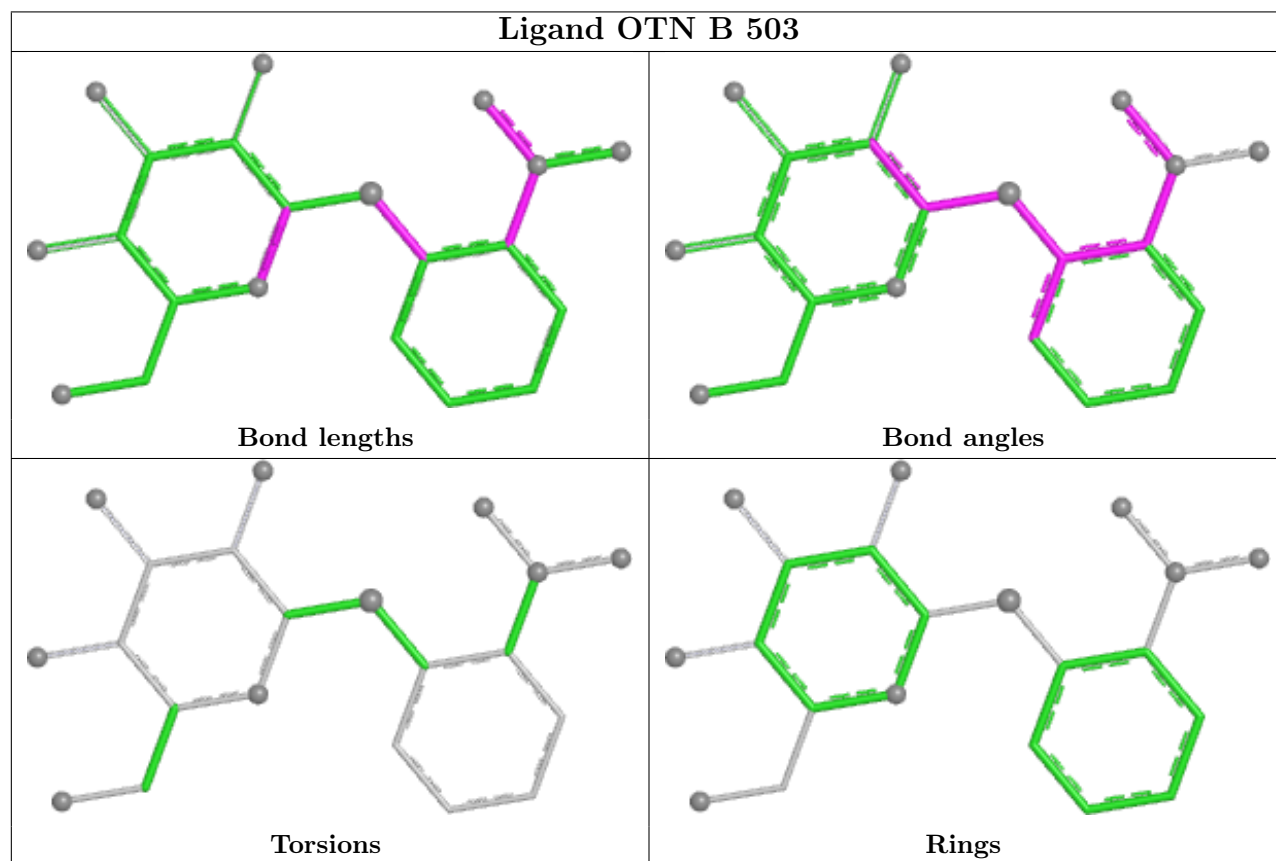
There are no ring outliers.

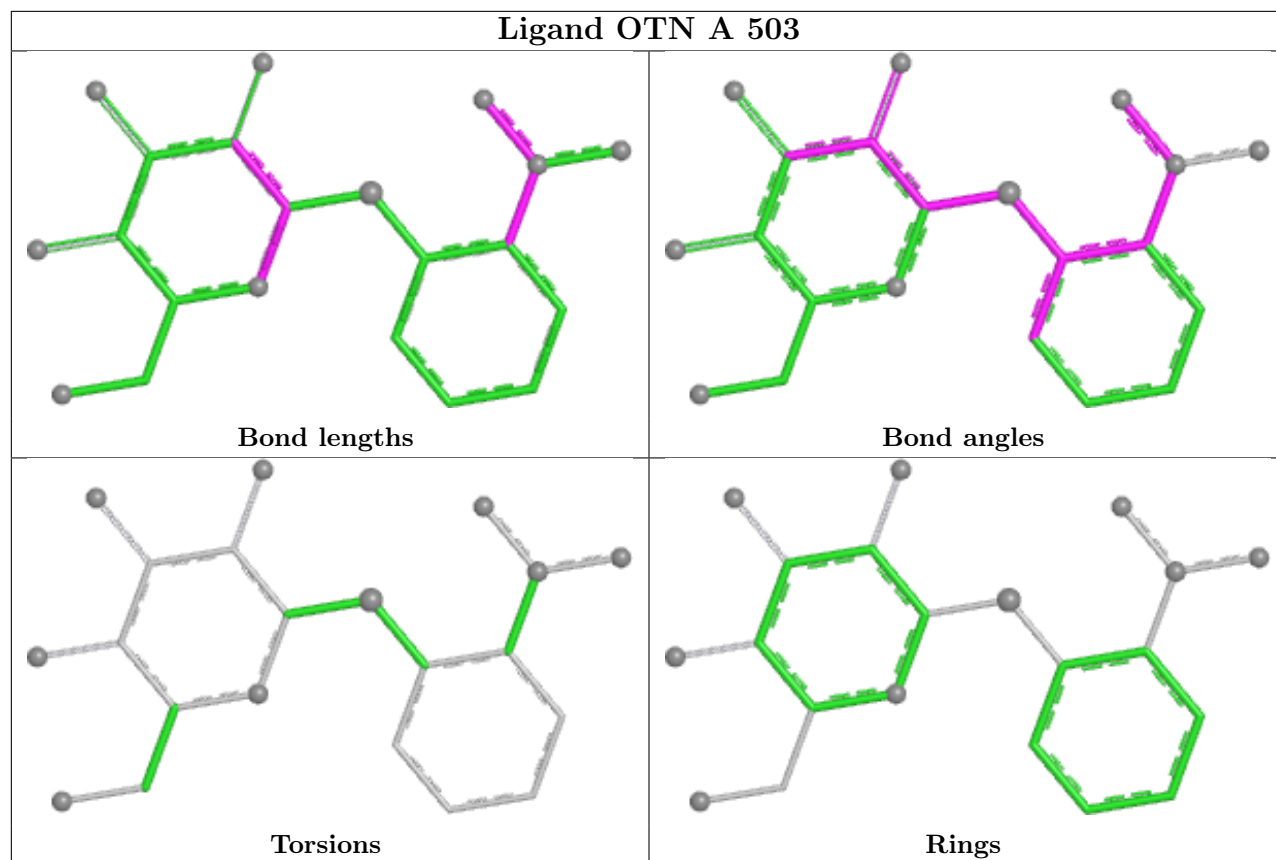
5 monomers are involved in 8 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	B	503	OTN	2	0
2	A	501	GOL	2	0
3	A	503	OTN	2	0
2	B	501	GOL	1	0
2	A	502	GOL	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, 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	459/469 (97%)	-0.76	1 (0%) 91 90	7, 17, 33, 59	3 (0%)
1	B	459/469 (97%)	-0.74	1 (0%) 91 90	7, 16, 34, 62	4 (0%)
All	All	918/938 (97%)	-0.75	2 (0%) 91 90	7, 17, 33, 62	7 (0%)

All (2) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	428	ASN	2.9
1	B	428	ASN	2.9

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.

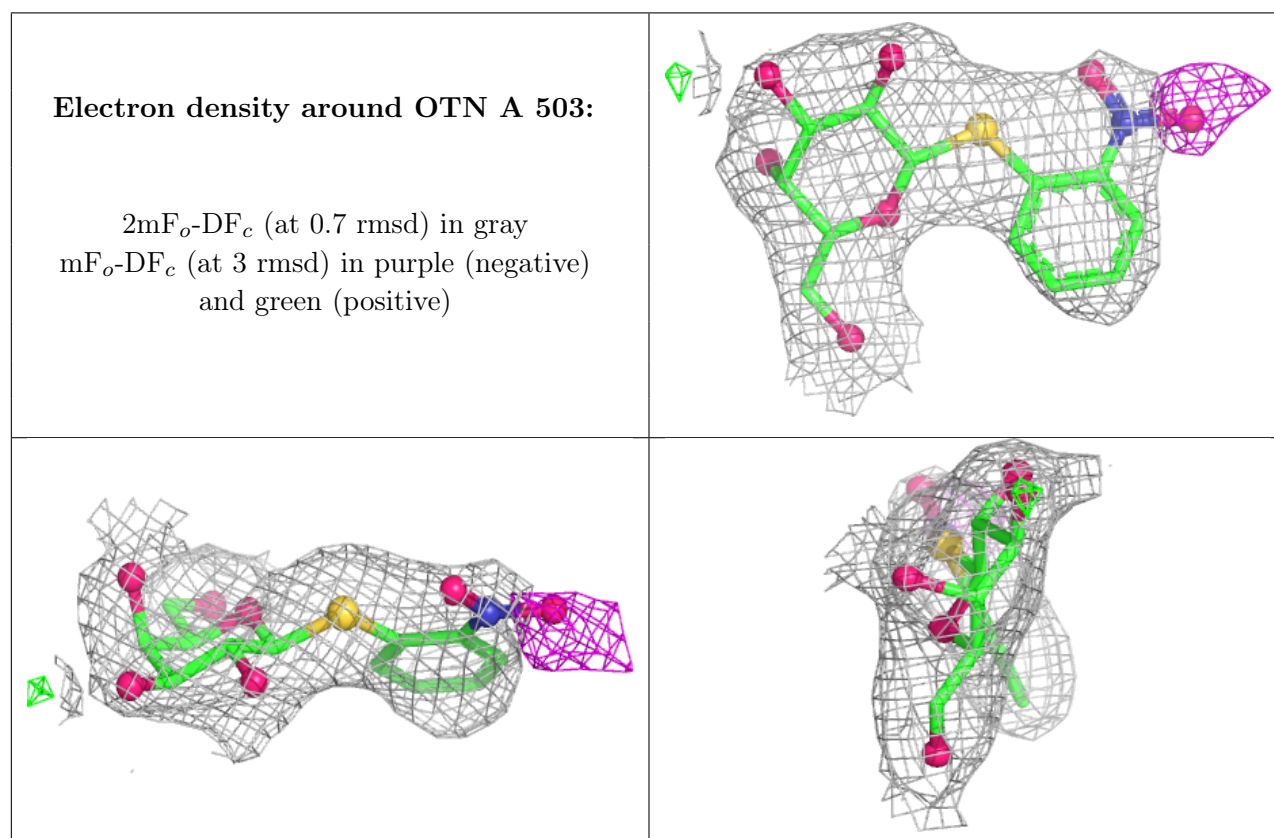
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å ²)	Q<0.9
2	GOL	A	502	6/6	0.82	0.12	22,24,30,36	0
2	GOL	B	501	6/6	0.83	0.14	25,28,33,34	0
3	OTN	A	503	21/21	0.94	0.08	16,25,31,54	0

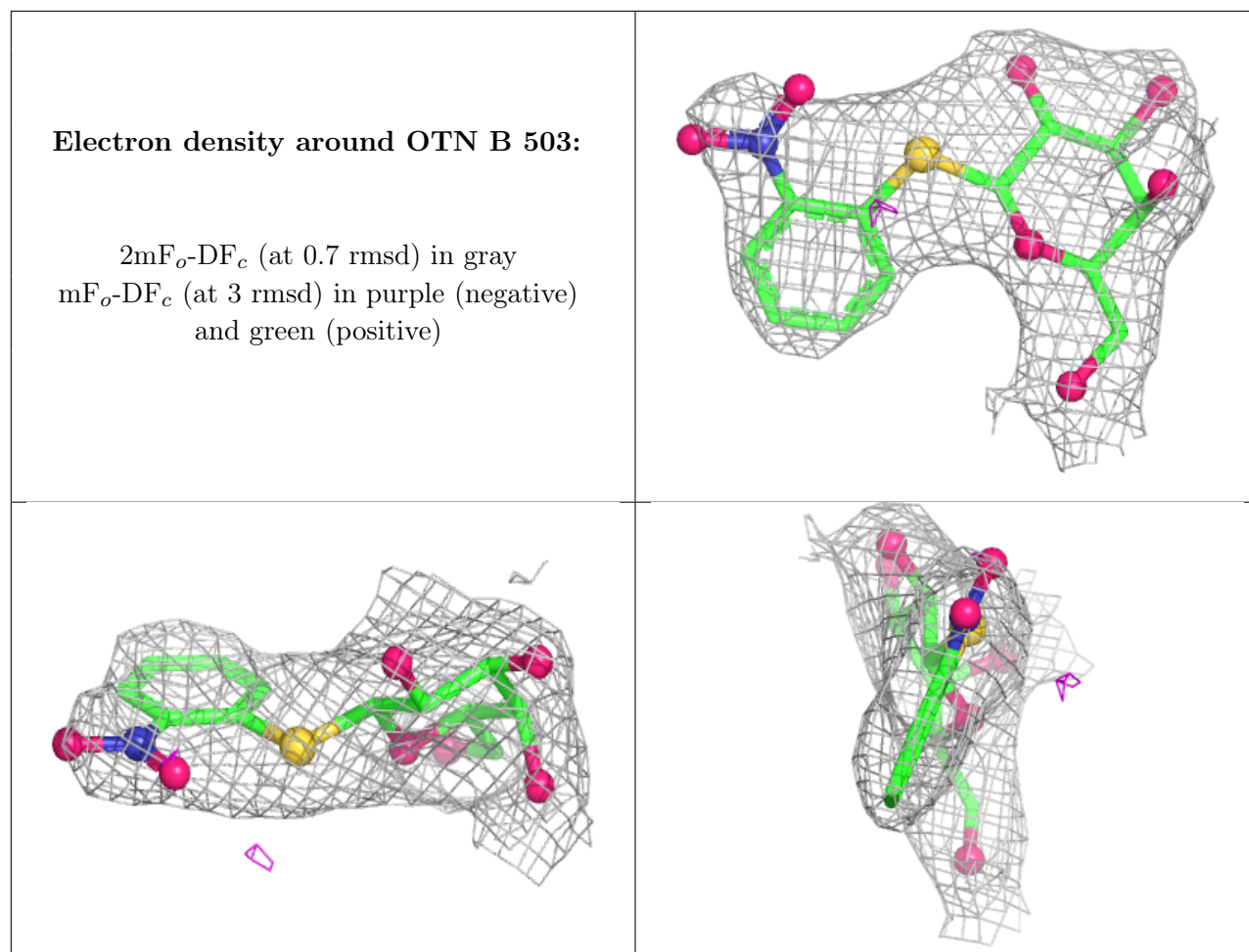
Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(\AA^2)	Q<0.9
2	GOL	B	502	6/6	0.95	0.08	10,13,17,17	0
2	GOL	A	501	6/6	0.97	0.07	15,16,25,29	0
3	OTN	B	503	21/21	0.97	0.06	9,17,36,48	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.