Full wwPDB X-ray Structure Validation Report ⓘ

Nov 13, 2023 – 06:00 PM EST

PDB ID	:	8UWV
Title	:	Crystal structure of BT3984 SusD-like from Bacteroides thetaiotaomicron VPI-5482 at 1.1 A resolution (Space group P21)
Authors	:	Sastre, D.E.; Navarro, M.V.A.S.; Sundberg, E.J.
Deposited on	:	2023-11-08
Resolution	:	1.08 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:

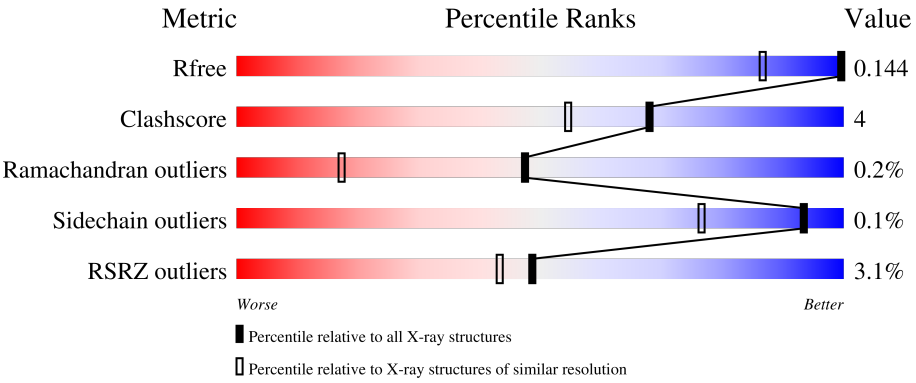
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.36
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

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 1.08 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	1386 (1.12-1.04)
Clashscore	141614	1021 (1.10-1.06)
Ramachandran outliers	138981	1381 (1.12-1.04)
Sidechain outliers	138945	1379 (1.12-1.04)
RSRZ outliers	127900	1359 (1.12-1.04)

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	513	3% 89% 7% .
1	B	513	3% 90% 6% .

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 16781 atoms, of which 7416 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 BT3984 SusD-like.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	495	Total	C	H	N	O	S	1	4	0
			7561	2453	3681	659	755	13			
1	B	496	Total	C	H	N	O	S	36	11	0
			7647	2478	3735	660	761	13			

- Molecule 2 is MAGNESIUM ION (three-letter code: MG) (formula: Mg) (labeled as "Ligand of Interest" by depositor).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	1	Total	Mg	0	0
			1	1		

- Molecule 3 is water.

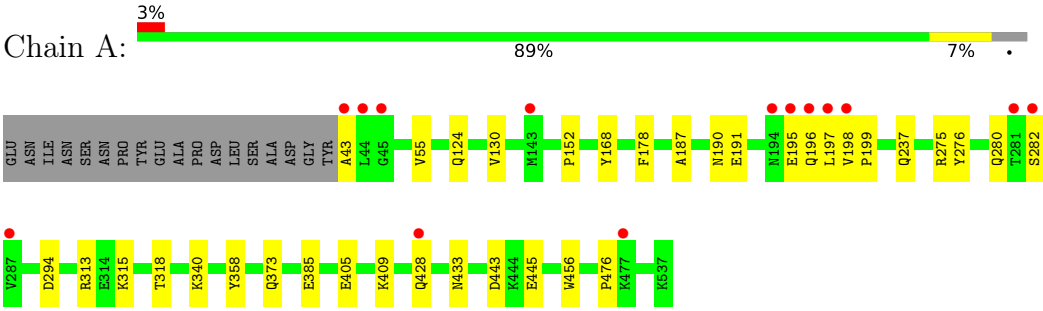
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	787	Total	O	0	0
			787	787		
3	B	785	Total	O	0	0
			785	785		

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.

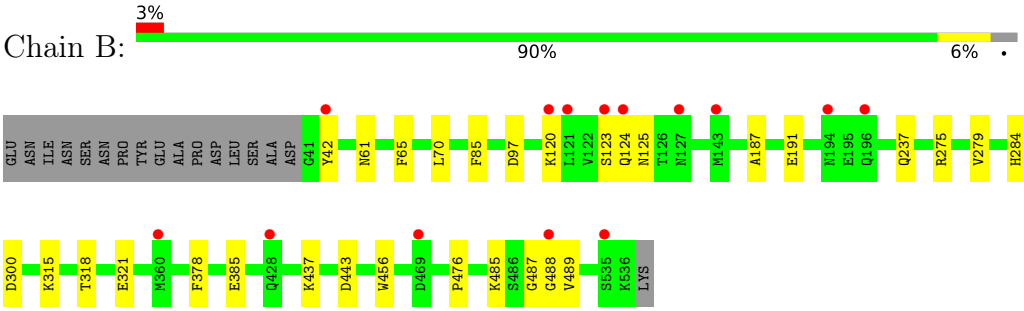
- Molecule 1: BT3984 SusD-like

Chain A:



- Molecule 1: BT3984 SusD-like

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	56.37Å 82.59Å 100.27Å 90.00° 98.88° 90.00°	Depositor
Resolution (Å)	46.18 – 1.08 46.18 – 1.08	Depositor EDS
% Data completeness (in resolution range)	91.5 (46.18-1.08) 91.5 (46.18-1.08)	Depositor EDS
R_{merge}	0.04	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.48 (at 1.08Å)	Xtriage
Refinement program	PHENIX (1.18.2_3874: ???)	Depositor
R, R_{free}	0.121 , 0.139 0.128 , 0.144	Depositor DCC
R_{free} test set	17779 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å^2)	9.0	Xtriage
Anisotropy	0.372	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.37 , 61.2	EDS
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.98	EDS
Total number of atoms	16781	wwPDB-VP
Average B, all atoms (Å^2)	14.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 27.72 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 2.0930e-03.*

[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: MG

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	0.61	1/3991 (0.0%)	0.81	4/5435 (0.1%)
1	B	0.60	0/4045	0.79	4/5512 (0.1%)
All	All	0.61	1/8036 (0.0%)	0.80	8/10947 (0.1%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	445	GLU	CD-OE2	-5.16	1.20	1.25

All (8) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	97	ASP	CB-CG-OD1	5.51	123.26	118.30
1	A	168	TYR	CB-CG-CD1	5.34	124.21	121.00
1	A	313	ARG	NE-CZ-NH2	-5.28	117.66	120.30
1	A	276	TYR	CA-CB-CG	5.13	123.14	113.40
1	B	97	ASP	CB-CG-OD2	-5.11	113.70	118.30
1	B	378	PHE	CB-CG-CD1	5.08	124.35	120.80
1	A	178	PHE	CB-CG-CD1	5.06	124.34	120.80
1	B	85	PHE	CB-CG-CD1	5.04	124.33	120.80

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	3880	3681	3712	28	0
1	B	3912	3735	3725	28	0
2	A	1	0	0	0	0
3	A	787	0	0	21	0
3	B	785	0	0	21	0
All	All	9365	7416	7437	57	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 (57) 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:196:GLN:HG2	3:A:709:HOH:O	1.64	0.98
1:A:405:GLU:O	1:A:409:LYS:HE2	1.64	0.97
1:B:488:GLY:O	3:B:601:HOH:O	1.87	0.92
1:B:61:ASN:O	3:B:603:HOH:O	1.90	0.88
1:A:443:ASP:OD2	3:A:701:HOH:O	1.92	0.88
1:A:433:ASN:HB2	3:A:1171:HOH:O	1.73	0.87
1:B:61:ASN:C	3:B:603:HOH:O	2.13	0.87
1:A:428:GLN:NE2	3:A:704:HOH:O	2.07	0.86
1:A:237:GLN:OE1	3:A:702:HOH:O	1.93	0.85
1:A:385:GLU:OE1	3:A:703:HOH:O	1.96	0.83
1:B:123:SER:HA	3:B:615:HOH:O	1.77	0.83
1:B:321:GLU:OE2	3:B:602:HOH:O	2.00	0.78
1:A:373:GLN:HG3	3:A:713:HOH:O	1.86	0.74
1:B:294:ASP:OD2	3:B:605:HOH:O	2.09	0.71
1:B:385:GLU:HG2	3:B:630:HOH:O	1.92	0.70
1:B:237:GLN:OE1	3:B:604:HOH:O	2.08	0.69
1:B:385:GLU:CG	3:B:630:HOH:O	2.40	0.69
1:A:294:ASP:OD2	3:A:705:HOH:O	2.12	0.67
1:B:275:ARG:NH2	3:B:612:HOH:O	2.28	0.66
1:A:43:ALA:N	3:A:708:HOH:O	2.29	0.66
1:A:340:LYS:NZ	3:A:707:HOH:O	2.26	0.65
1:A:195:GLU:OE1	3:A:706:HOH:O	2.15	0.64
1:B:485:LYS:CE	3:B:610:HOH:O	2.49	0.60
1:B:443:ASP:OD2	3:B:606:HOH:O	2.16	0.60
1:B:315:LYS:HE2	3:B:838:HOH:O	2.03	0.58
1:B:284:HIS:NE2	3:B:602:HOH:O	1.90	0.56
1:B:385:GLU:CD	3:B:630:HOH:O	2.44	0.56

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:198:VAL:HG22	1:A:199:PRO:HD2	1.89	0.55
1:B:279[A]:VAL:HG21	1:B:284:HIS:CD2	2.44	0.53
1:A:315:LYS:HE2	3:A:921:HOH:O	2.07	0.53
1:A:190:ASN:ND2	3:A:714:HOH:O	2.41	0.53
1:A:130:VAL:HG12	1:A:197:LEU:HD11	1.88	0.53
1:A:196:GLN:CG	3:A:709:HOH:O	2.38	0.52
1:A:187:ALA:O	1:A:191:GLU:HG3	2.10	0.51
1:A:340:LYS:HG3	3:A:1153:HOH:O	2.09	0.51
1:B:456:TRP:CD2	1:B:476:PRO:HG3	2.45	0.51
1:A:358:TYR:HD2	3:A:1095:HOH:O	1.94	0.50
1:B:120:LYS:HE2	1:B:124:GLN:HE21	1.77	0.50
1:A:318:THR:HG21	3:A:727:HOH:O	2.12	0.48
1:A:428:GLN:NE2	3:A:716:HOH:O	2.46	0.48
1:B:279[A]:VAL:CG2	1:B:289:CYS:CB	2.92	0.48
1:A:280:GLN:O	1:A:283:ASP:HB3	2.14	0.47
1:B:318:THR:HG21	3:B:612:HOH:O	2.14	0.46
1:B:279[A]:VAL:CG2	1:B:289:CYS:HB2	2.45	0.46
1:B:487:GLY:N	3:B:616:HOH:O	2.41	0.46
1:A:282:SER:O	1:A:283:ASP:HB2	2.16	0.45
1:A:456:TRP:CD2	1:A:476:PRO:HG3	2.51	0.45
1:A:152:PRO:HA	3:A:725:HOH:O	2.17	0.45
1:A:275:ARG:NH2	3:A:727:HOH:O	2.51	0.44
1:B:489:VAL:HA	3:B:601:HOH:O	2.18	0.43
1:A:124:GLN:NE2	3:A:717:HOH:O	2.46	0.43
1:B:187:ALA:O	1:B:191:GLU:HG3	2.18	0.43
1:B:437:LYS:NZ	3:B:620:HOH:O	2.51	0.43
1:B:70:LEU:C	1:B:70:LEU:HD23	2.39	0.42
1:B:42:TYR:CD2	1:B:125:ASN:HB3	2.55	0.41
1:B:65:PHE:HD2	3:B:603:HOH:O	2.04	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	497/513 (97%)	480 (97%)	15 (3%)	2 (0%)	34 10
1	B	505/513 (98%)	494 (98%)	11 (2%)	0	100 100
All	All	1002/1026 (98%)	974 (97%)	26 (3%)	2 (0%)	47 17

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	283	ASP
1	A	55	VAL

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	409/420 (97%)	409 (100%)	0	100 100
1	B	413/420 (98%)	412 (100%)	1 (0%)	93 76
All	All	822/840 (98%)	821 (100%)	1 (0%)	93 76

All (1) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	300	ASP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (9) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	190	ASN
1	A	192	ASN
1	A	247	ASN
1	A	373	GLN
1	A	413	ASN
1	A	428	GLN
1	B	124	GLN
1	B	192	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	247	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 monosaccharides in this entry.

5.6 Ligand geometry

Of 1 ligands modelled in this entry, 1 is 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	495/513 (96%)	-0.22	17 (3%)	45	39	6, 10, 21, 36	0
1	B	496/513 (96%)	-0.24	14 (2%)	53	47	6, 10, 21, 38	0
All	All	991/1026 (96%)	-0.23	31 (3%)	49	43	6, 10, 21, 38	0

All (31) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	281	THR	4.9
1	A	287	VAL	4.8
1	A	196	GLN	4.7
1	A	286	GLY	4.6
1	A	282	SER	4.1
1	B	124	GLN	4.1
1	A	143	MET	3.7
1	B	143	MET	3.7
1	B	121	LEU	3.6
1	A	194	ASN	3.6
1	B	196	GLN	3.4
1	B	42	TYR	3.4
1	B	535	SER	3.2
1	A	198	VAL	3.2
1	A	283	ASP	3.2
1	B	120	LYS	3.0
1	B	360	MET	2.8
1	A	428	GLN	2.8
1	B	123	SER	2.7
1	A	197	LEU	2.7
1	B	428	GLN	2.7
1	A	44	LEU	2.6
1	B	127	ASN	2.6
1	A	477	LYS	2.5

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	43	ALA	2.4
1	B	469	ASP	2.4
1	A	285	GLY	2.2
1	A	45	GLY	2.2
1	B	488	GLY	2.2
1	A	195	GLU	2.1
1	B	194	ASN	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)

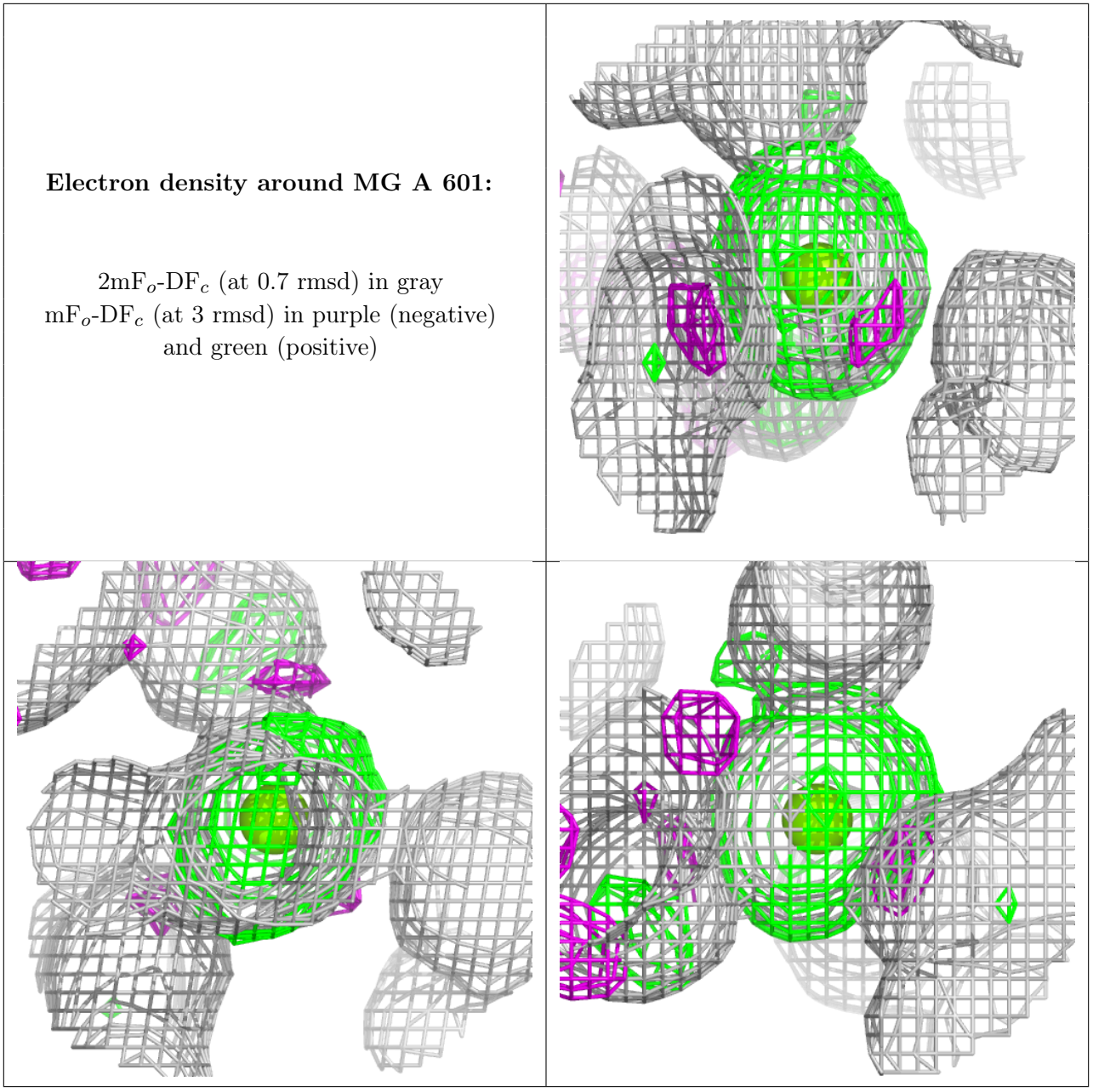
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9'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

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
2	MG	A	601	1/1	0.97	0.38	6,6,6,6	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 MG A 601:

$2mF_o$-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.