Full wwPDB X-ray Structure Validation Report ⓘ

Nov 8, 2022 – 05:06 pm GMT

PDB ID	:	7O5V
Title	:	Crystal structure of holo-H44A mutant of Hydroxy ketone aldolase (SwHKA) from Sphingomonas wittichii RW1, in complex with Hydroxypyruvate
Authors	:	Laustsen, J.; Justo, I.; Marsden, S.R.; Hanefeld, U.; Bento, I.
Deposited on	:	2021-04-09
Resolution	:	1.95 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:

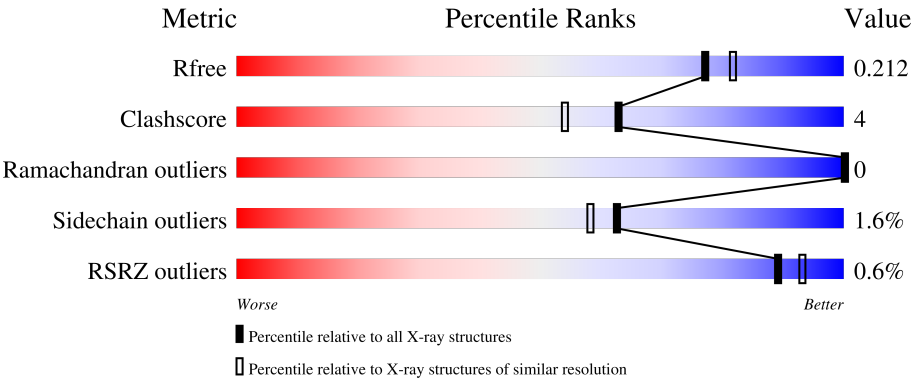
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.31.2
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0267
CCP4	:	7.1.010 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.31.2

1 Overall quality at a glance (i)

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

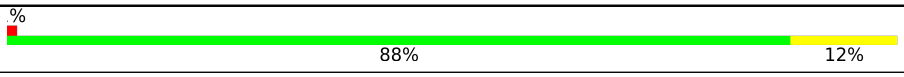
The reported resolution of this entry is 1.95 Å.

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	2580 (1.96-1.96)
Clashscore	141614	2705 (1.96-1.96)
Ramachandran outliers	138981	2678 (1.96-1.96)
Sidechain outliers	138945	2678 (1.96-1.96)
RSRZ outliers	127900	2539 (1.96-1.96)

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	AAA	251	
1	BBB	251	

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
5	3PY	AAA	304	-	X	-	-

2 Entry composition (i)

There are 6 unique types of molecules in this entry. The entry contains 7727 atoms, of which 3761 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 HpcH/HpaI aldolase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	AAA	250	Total	C	H	N	O	S	41	0	0
			3723	1171	1864	331	349	8			
1	BBB	251	Total	C	H	N	O	S	41	3	0
			3782	1188	1894	337	355	8			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
AAA	1	HIS	-	expression tag	UNP A5VH82
AAA	44	ALA	HIS	engineered mutation	UNP A5VH82
BBB	1	HIS	-	expression tag	UNP A5VH82
BBB	44	ALA	HIS	engineered mutation	UNP A5VH82

- 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	AAA	1	Total	Mg	0	0
			1	1		
2	BBB	1	Total	Mg	0	0
			1	1		

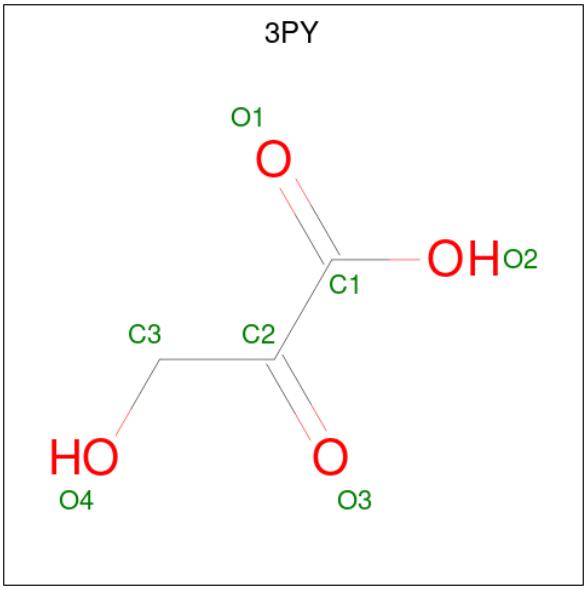
- Molecule 3 is BROMIDE ION (three-letter code: BR) (formula: Br).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	AAA	2	Total	Br	0	0
			2	2		

- Molecule 4 is POTASSIUM ION (three-letter code: K) (formula: K).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	AAA	1	Total K 1 1	0	0
4	BBB	1	Total K 1 1	0	0

- Molecule 5 is 3-HYDROXYPYRUVIC ACID (three-letter code: 3PY) (formula: $C_3H_4O_4$) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	AAA	1	Total C H O 10 3 3 4	1	0

- Molecule 6 is water.

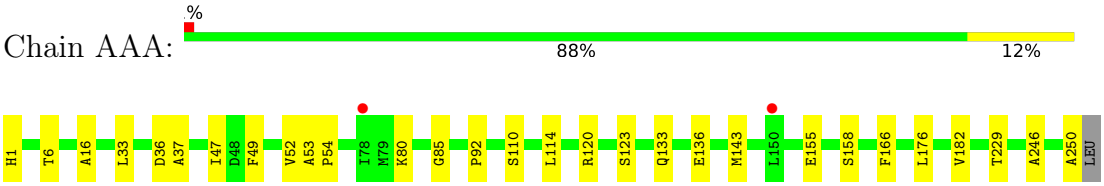
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	AAA	112	Total O 112 112	0	0
6	BBB	94	Total O 94 94	0	3

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.

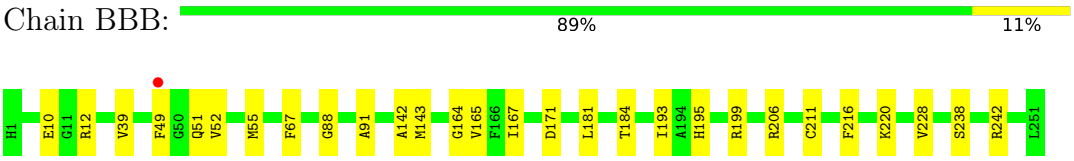
• Molecule 1: HpcH/HpaI aldolase

Chain AAA:



• Molecule 1: HpcH/HpaI aldolase

Chain BBB:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	H 3	Depositor
Cell constants a, b, c, α, β, γ	70.91Å 70.91Å 222.57Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	53.77 – 1.95 53.77 – 1.95	Depositor EDS
% Data completeness (in resolution range)	96.3 (53.77-1.95) 96.3 (53.77-1.95)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.60 (at 1.95Å)	Xtriage
Refinement program	REFMAC 5.8.0267	Depositor
R, R_{free}	0.162 , 0.212 0.166 , 0.212	Depositor DCC
R_{free} test set	1484 reflections (5.06%)	wwPDB-VP
Wilson B-factor (Å^2)	31.2	Xtriage
Anisotropy	0.202	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	(Not available) , (Not available)	EDS
L-test for twinning[2]	$< \|L\| > = 0.42$, $< L^2 > = 0.24$	Xtriage
Estimated twinning fraction	0.289 for -h-k,k,-l	Xtriage
Reported twinning fraction	0.712 for H, K, L 0.288 for K, H, -L	Depositor
Outliers	0 of 29317 reflections	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	7727	wwPDB-VP
Average B, all atoms (Å^2)	35.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 9.15% 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: MG, K, 3PY, BR

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	AAA	0.80	0/1894	0.91	0/2576
1	BBB	0.77	0/1926	0.88	0/2617
All	All	0.79	0/3820	0.90	0/5193

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	AAA	0	1

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	AAA	85	GLY	Peptide

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	AAA	1859	1864	1858	15	1
1	BBB	1888	1894	1888	18	0
2	AAA	1	0	0	0	0
2	BBB	1	0	0	0	0
3	AAA	2	0	0	0	0
4	AAA	1	0	0	0	0
4	BBB	1	0	0	0	0
5	AAA	7	3	3	0	0
6	AAA	112	0	0	1	0
6	BBB	94	0	0	1	0
All	All	3966	3761	3749	32	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 4.

All (32) 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:BBB:181:LEU:HD11	1:BBB:184:THR:HA	1.66	0.75
1:AAA:155:GLU:O	1:AAA:158:SER:OG	2.13	0.63
1:BBB:49:PHE:O	1:BBB:52:VAL:HG12	2.02	0.59
1:BBB:181:LEU:CD1	1:BBB:184:THR:HA	2.35	0.56
1:AAA:16:ALA:O	1:AAA:229:THR:HA	2.05	0.56
1:AAA:53:ALA:N	1:AAA:54:PRO:HD2	2.21	0.55
1:BBB:67:PHE:CD1	1:BBB:88:GLY:HA3	2.44	0.53
1:AAA:16:ALA:HA	1:AAA:37:ALA:O	2.10	0.52
1:AAA:49:PHE:O	1:AAA:52:VAL:HG12	2.10	0.52
1:AAA:143:MET:HA	1:AAA:166:PHE:O	2.10	0.51
1:BBB:164:GLY:HA3	1:BBB:206:ARG:O	2.12	0.50
1:AAA:246:ALA:O	1:AAA:250:ALA:HA	2.13	0.49
1:BBB:39:VAL:HG22	1:BBB:67:PHE:HB3	1.96	0.47
1:AAA:33:LEU:HD22	1:BBB:242:ARG:HG2	1.97	0.47
1:BBB:167:ILE:HD13	1:BBB:193:ILE:HG23	1.97	0.47
1:AAA:36:ASP:OD1	6:AAA:401:HOH:O	2.19	0.46
1:BBB:10:GLU:OE2	1:BBB:12:ARG:NE	2.41	0.45
1:AAA:1:HIS:CG	1:AAA:6:THR:HG21	2.52	0.45
1:AAA:1:HIS:ND1	1:AAA:6:THR:HG21	2.31	0.45
1:BBB:67:PHE:CG	1:BBB:88:GLY:HA3	2.52	0.45
1:BBB:181:LEU:HD13	6:BBB:461:HOH:O	2.17	0.44
1:BBB:211:CYS:SG	1:BBB:228:VAL:HG13	2.59	0.43
1:BBB:195:HIS:O	1:BBB:199:ARG:HG2	2.18	0.43
1:BBB:91:ALA:O	1:BBB:143:MET:HB2	2.19	0.42

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:AAA:92:PRO:HA	1:AAA:143:MET:HB2	2.02	0.42
1:BBB:51:GLN:O	1:BBB:55:MET:HG3	2.20	0.42
1:BBB:181:LEU:C	1:BBB:181:LEU:HD12	2.41	0.41
1:AAA:120:ARG:NE	1:AAA:120:ARG:HA	2.35	0.41
1:BBB:142:ALA:O	1:BBB:165:VAL:HA	2.21	0.41
1:BBB:216:PHE:O	1:BBB:220:LYS:HG2	2.20	0.41
1:AAA:133:GLN:O	1:AAA:136:GLU:HG2	2.22	0.40
1:AAA:176:LEU:HD23	1:AAA:176:LEU:HA	1.91	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)
1:AAA:47:ILE:O	1:AAA:80:LYS:HZ2[2_555]	1.54	0.06

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	AAA	248/251 (99%)	234 (94%)	14 (6%)	0	100	100
1	BBB	252/251 (100%)	244 (97%)	8 (3%)	0	100	100
All	All	500/502 (100%)	478 (96%)	22 (4%)	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	AAA	186/187 (100%)	182 (98%)	4 (2%)	52 44
1	BBB	189/187 (101%)	187 (99%)	2 (1%)	73 71
All	All	375/374 (100%)	369 (98%)	6 (2%)	62 58

All (6) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	AAA	110	SER
1	AAA	114	LEU
1	AAA	123	SER
1	AAA	182	VAL
1	BBB	171	ASP
1	BBB	238	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

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 monosaccharides in this entry.

5.6 Ligand geometry (i)

Of 7 ligands modelled in this entry, 6 are monoatomic - leaving 1 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	$\#\|Z\| > 2$
5	3PY	AAA	304	2	5,6,6	7.29	2 (40%)	6,7,7	2.72	4 (66%)

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
5	3PY	AAA	304	2	-	2/5/6/6	-

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	AAA	304	3PY	C2-C1	-15.99	1.32	1.53
5	AAA	304	3PY	O2-C1	-2.60	1.23	1.30

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
5	AAA	304	3PY	O1-C1-C2	-4.28	116.00	121.72
5	AAA	304	3PY	O3-C2-C3	-3.52	115.06	120.50
5	AAA	304	3PY	O4-C3-C2	-2.44	106.11	112.48
5	AAA	304	3PY	O2-C1-C2	2.36	120.43	113.97

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	AAA	304	3PY	O3-C2-C3-O4
5	AAA	304	3PY	C1-C2-C3-O4

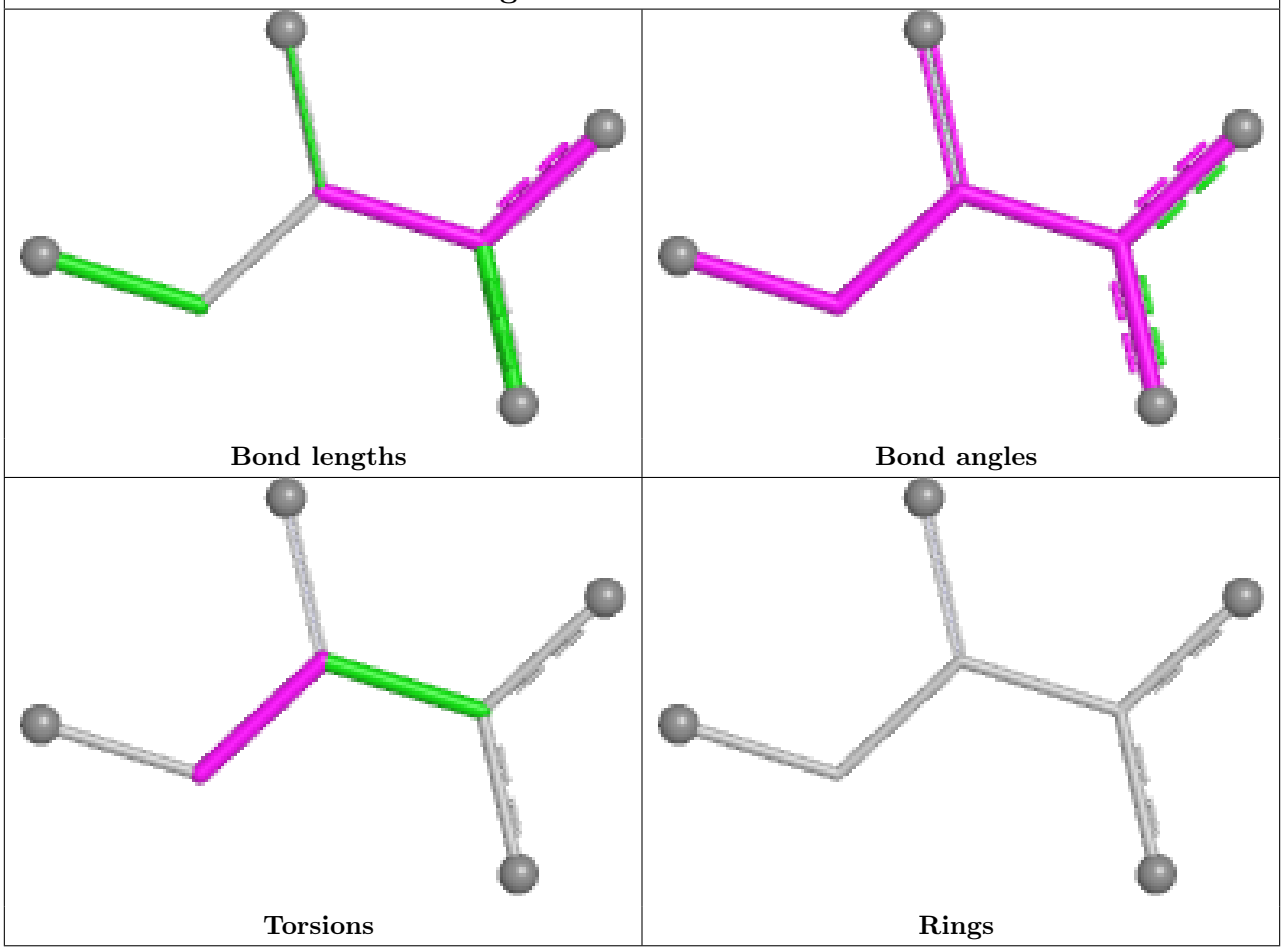
There are no ring outliers.

No monomer is involved in short contacts.

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 then 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.

Ligand 3PY AAA 304



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

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	AAA	250/251 (99%)	0.13	2 (0%)	86	90	25, 33, 43, 57	0
1	BBB	251/251 (100%)	0.12	1 (0%)	92	95	26, 35, 44, 53	0
All	All	501/502 (99%)	0.13	3 (0%)	89	93	25, 34, 44, 57	0

All (3) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	AAA	78	ILE	2.3
1	BBB	49	PHE	2.1
1	AAA	150	LEU	2.1

6.2 Non-standard residues in protein, DNA, RNA chains

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates

There are no monosaccharides in this entry.

6.4 Ligands

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	BBB	301	1/1	0.92	0.04	33,33,33,33	1
5	3PY	AAA	304	7/7	0.93	0.12	26,36,41,45	1

Continued on next page...

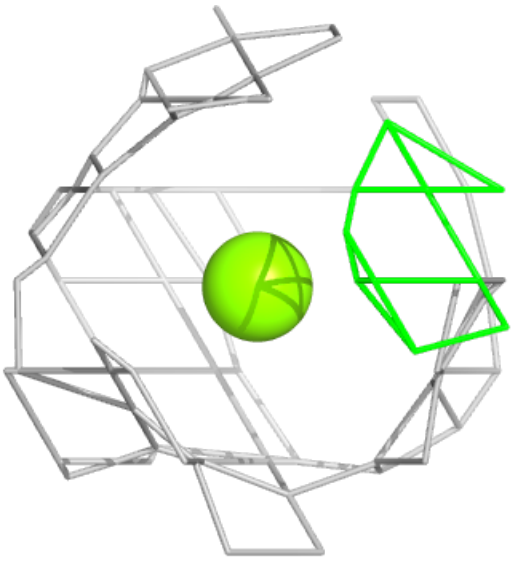
Continued from previous page...

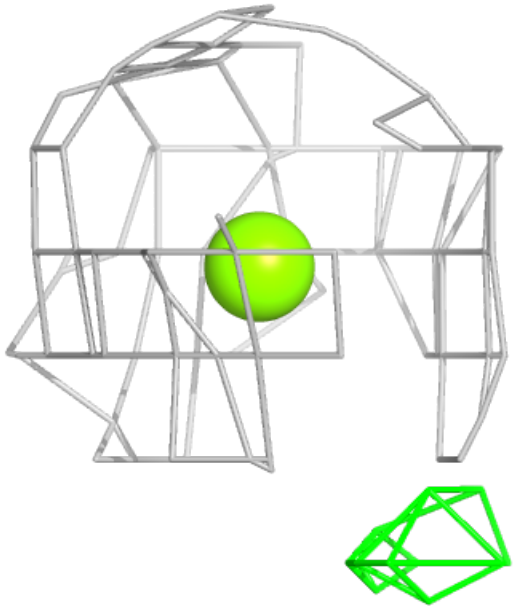
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
4	K	BBB	302	1/1	0.98	0.07	33,33,33,33	0
2	MG	AAA	301	1/1	0.98	0.07	32,32,32,32	0
3	BR	AAA	305	1/1	0.99	0.18	37,37,37,37	0
4	K	AAA	303	1/1	0.99	0.06	31,31,31,31	0
3	BR	AAA	302	1/1	1.00	0.04	38,38,38,38	0

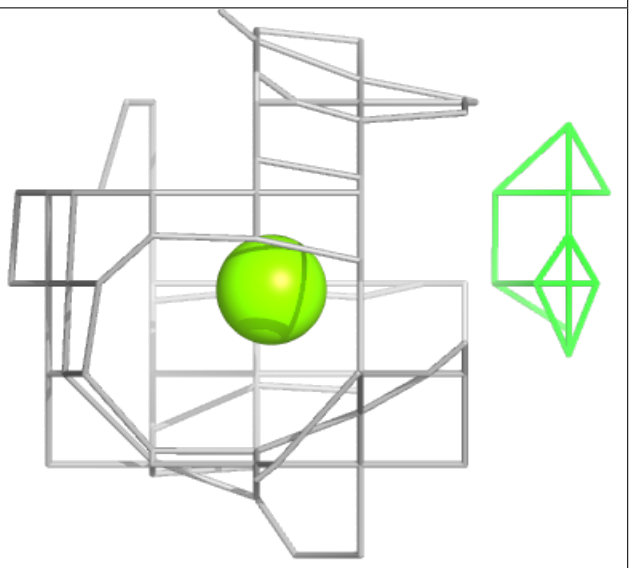
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation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

Electron density around MG BBB 301:

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

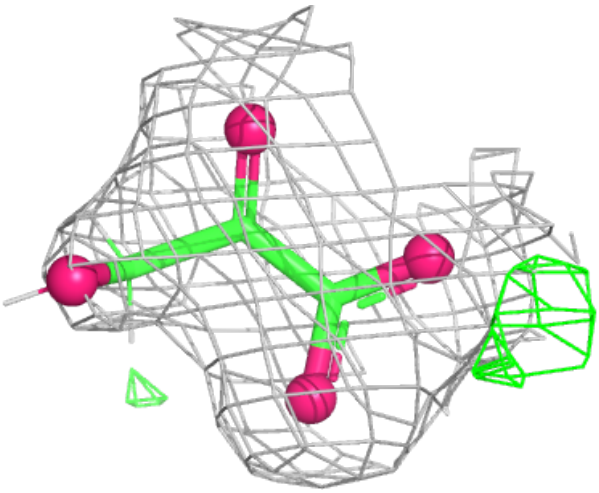


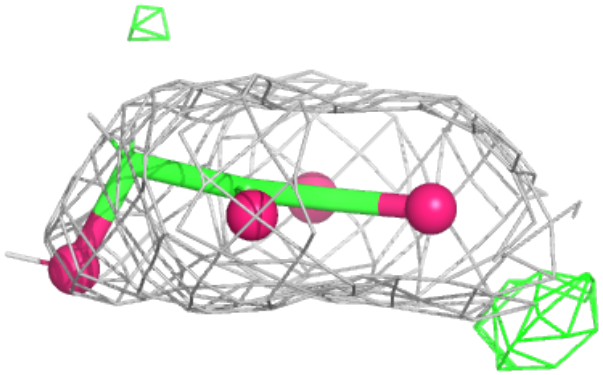


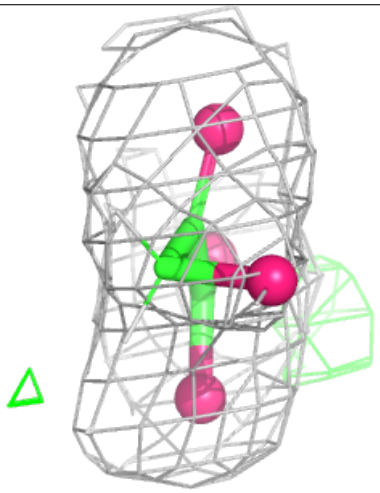


Electron density around 3PY AAA 304:

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

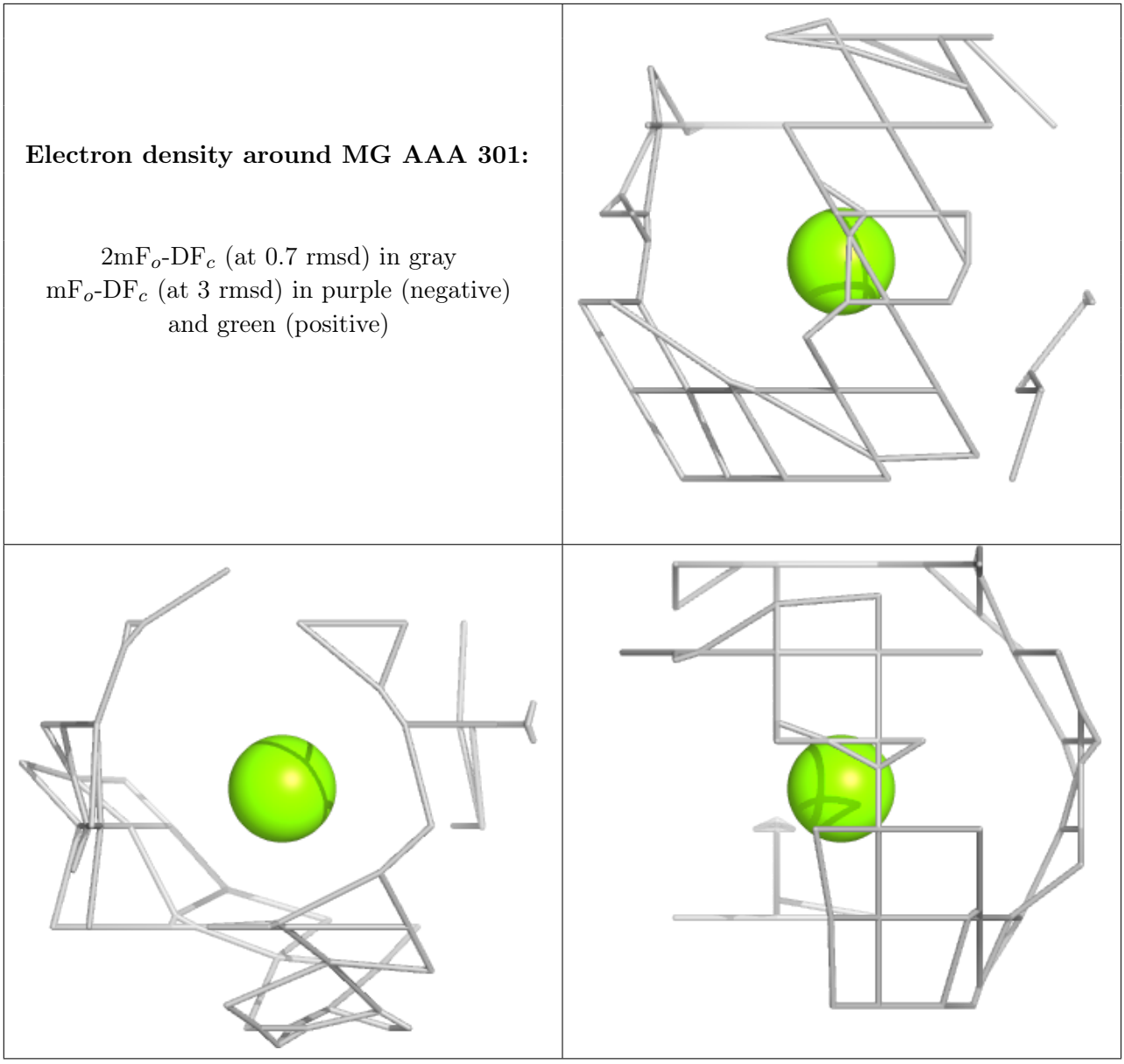






Electron density around MG AAA 301:

$2mF_o$-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.