Full wwPDB X-ray Structure Validation Report (i)

Nov 8, 2021 – 02:06 pm GMT

PDB ID	:	7PAV
Title	:	MALT1 in complex with compound 1
Authors	:	Kack, H.; Oster, L.
Deposited on	:	2021-07-30
Resolution	:	2.20 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 following versions of software and data (see references (i)) were used in the production of this report:

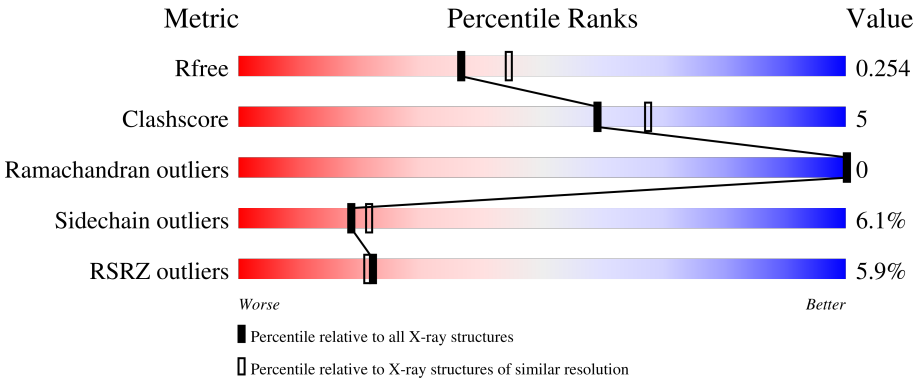
MolProbity	:	4.02b-467
Mogul	:	1.8.4 (270009), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.23.2
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0267
CCP4	:	7.1.010 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.23.2

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 2.20 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	4898 (2.20-2.20)
Clashscore	141614	5594 (2.20-2.20)
Ramachandran outliers	138981	5503 (2.20-2.20)
Sidechain outliers	138945	5504 (2.20-2.20)
RSRZ outliers	127900	4800 (2.20-2.20)

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	388	3% 79% 12% • 8%
1	B	388	8% 84% 10% • 5%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 5858 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 Mucosa-associated lymphoid tissue lymphoma translocation protein 1.

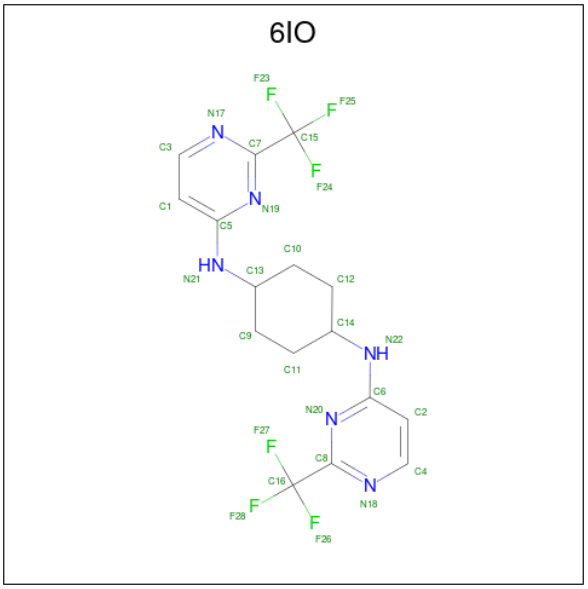
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	357	Total	C	N	O	S	0	1	0
			2834	1815	465	533	21			
1	B	368	Total	C	N	O	S	0	1	0
			2918	1876	471	548	23			

There are 22 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	338	MET	-	initiating methionine	UNP Q9UDY8
A	595	LYS	ASP	engineered mutation	UNP Q9UDY8
A	617	LYS	SER	engineered mutation	UNP Q9UDY8
A	666	ALA	HIS	engineered mutation	UNP Q9UDY8
A	681	GLU	HIS	engineered mutation	UNP Q9UDY8
A	720	HIS	-	expression tag	UNP Q9UDY8
A	721	HIS	-	expression tag	UNP Q9UDY8
A	722	HIS	-	expression tag	UNP Q9UDY8
A	723	HIS	-	expression tag	UNP Q9UDY8
A	724	HIS	-	expression tag	UNP Q9UDY8
A	725	HIS	-	expression tag	UNP Q9UDY8
B	338	MET	-	initiating methionine	UNP Q9UDY8
B	595	LYS	ASP	engineered mutation	UNP Q9UDY8
B	617	LYS	SER	engineered mutation	UNP Q9UDY8
B	666	ALA	HIS	engineered mutation	UNP Q9UDY8
B	681	GLU	HIS	engineered mutation	UNP Q9UDY8
B	720	HIS	-	expression tag	UNP Q9UDY8
B	721	HIS	-	expression tag	UNP Q9UDY8
B	722	HIS	-	expression tag	UNP Q9UDY8
B	723	HIS	-	expression tag	UNP Q9UDY8
B	724	HIS	-	expression tag	UNP Q9UDY8
B	725	HIS	-	expression tag	UNP Q9UDY8

- Molecule 2 is {N}1, {N}4-bis[2-(trifluoromethyl)pyrimidin-4-yl]cyclohexane-1,4-di

amine (three-letter code: 6IO) (formula: $C_{16}H_{16}F_6N_6$) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	B	1	Total C F N 28 16 6 6	0	0

- Molecule 3 is water.

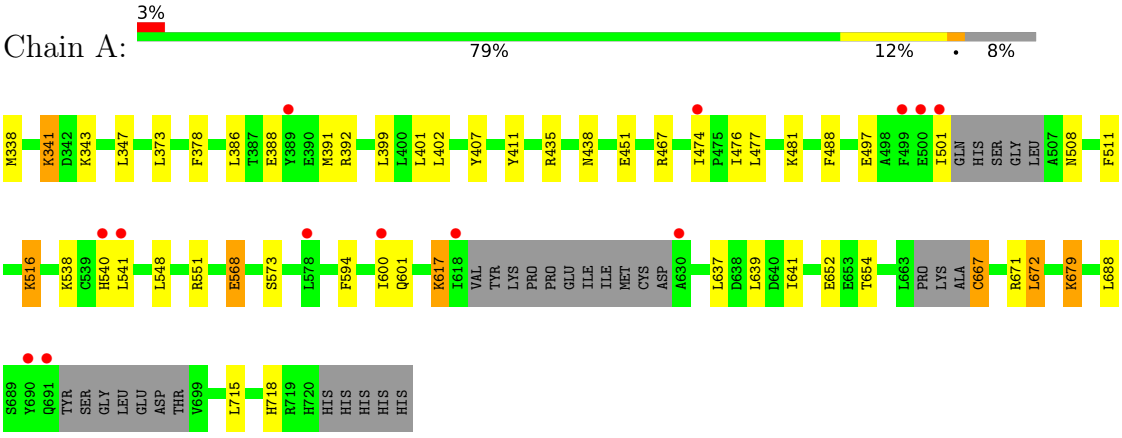
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	47	Total O 47 47	0	0
3	B	31	Total O 31 31	0	0

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.

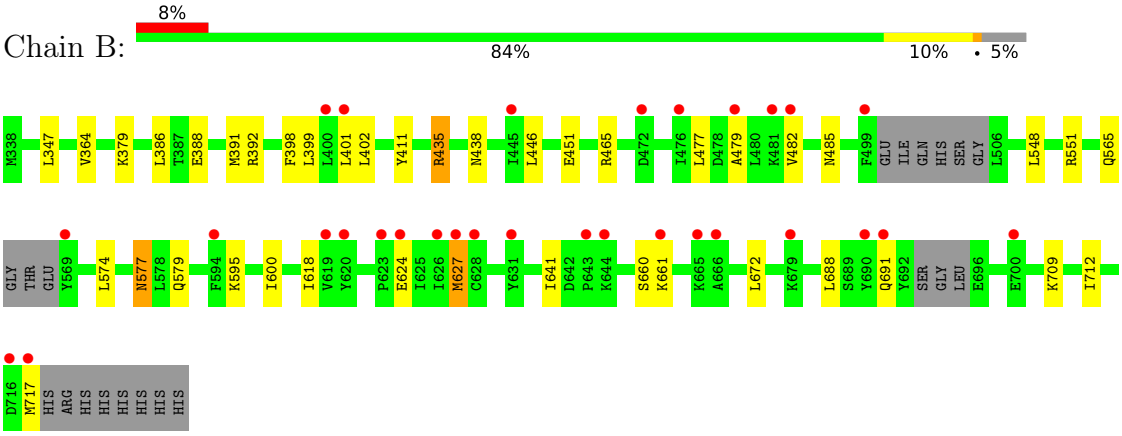
• Molecule 1: Mucosa-associated lymphoid tissue lymphoma translocation protein 1

Chain A:



• Molecule 1: Mucosa-associated lymphoid tissue lymphoma translocation protein 1

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	52.51Å 75.10Å 107.52Å 90.00° 94.16° 90.00°	Depositor
Resolution (Å)	53.62 – 2.20 53.62 – 2.20	Depositor EDS
% Data completeness (in resolution range)	69.4 (53.62-2.20) 69.4 (53.62-2.20)	Depositor EDS
R_{merge}	0.04	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.38 (at 2.20Å)	Xtriage
Refinement program	BUSTER 2.11.8	Depositor
R, R_{free}	0.208 , 0.256 0.207 , 0.254	Depositor DCC
R_{free} test set	1526 reflections (5.17%)	wwPDB-VP
Wilson B-factor (Å^2)	35.1	Xtriage
Anisotropy	0.105	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	(Not available) , (Not available)	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	5858	wwPDB-VP
Average B, all atoms (Å^2)	48.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.18% 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: 6IO

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	0.42	0/2882	0.60	0/3885
1	B	0.41	0/2970	0.62	0/4010
All	All	0.41	0/5852	0.61	0/7895

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	2834	0	2863	34	0
1	B	2918	0	2960	21	0
2	B	28	0	0	1	0
3	A	47	0	0	0	0
3	B	31	0	0	0	0
All	All	5858	0	5823	54	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.

All (54) 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:435:ARG:H	1:A:438:ASN:HD22	1.21	0.87
1:B:435:ARG:H	1:B:438:ASN:HD22	1.22	0.84
1:A:341:LYS:NZ	1:A:568:GLU:HG3	1.94	0.82
1:A:617:LYS:HZ2	1:A:667:CYS:N	1.82	0.78
1:A:341:LYS:HZ3	1:A:568:GLU:HG3	1.59	0.66
1:B:600:ILE:HD11	1:B:688:LEU:HD13	1.78	0.65
1:A:476:ILE:HB	1:A:477:LEU:HA	1.78	0.64
1:B:379:LYS:NZ	1:B:577:ASN:ND2	2.48	0.62
1:B:399:LEU:HA	1:B:402:LEU:HD12	1.83	0.60
1:A:341:LYS:HZ1	1:A:568:GLU:HG3	1.66	0.59
1:A:551:ARG:HG3	1:B:551:ARG:HG2	1.84	0.59
1:A:679:LYS:HZ2	1:A:718:HIS:HB3	1.68	0.58
1:A:617:LYS:NZ	1:A:667:CYS:N	2.50	0.58
1:B:379:LYS:HZ3	1:B:577:ASN:ND2	2.02	0.57
1:B:398:PHE:HA	1:B:401:LEU:HD12	1.85	0.57
1:A:540:HIS:CE1	1:A:541:LEU:HD23	2.40	0.56
1:A:617:LYS:HZ1	1:A:667:CYS:HA	1.71	0.56
1:A:679:LYS:HZ2	1:A:718:HIS:CD2	2.24	0.56
1:A:600:ILE:HD11	1:A:688:LEU:HD13	1.86	0.56
1:B:641:ILE:HD13	1:B:672:LEU:HD22	1.88	0.55
1:A:399:LEU:HD23	1:A:402:LEU:HD12	1.89	0.53
1:A:679:LYS:NZ	1:A:718:HIS:CD2	2.76	0.53
1:A:617:LYS:NZ	1:A:667:CYS:HA	2.24	0.53
1:A:679:LYS:HZ2	1:A:718:HIS:CG	2.26	0.53
1:B:347:LEU:HD22	1:B:411:TYR:HB3	1.91	0.53
1:B:386:LEU:HB2	1:B:391:MET:HG3	1.92	0.51
1:A:641:ILE:HD13	1:A:672:LEU:HD22	1.94	0.50
1:A:386:LEU:HB2	1:A:391:MET:HG3	1.93	0.50
1:A:508:ASN:ND2	1:A:516:LYS:NZ	2.61	0.49
1:A:347:LEU:HD22	1:A:411:TYR:HB3	1.93	0.49
1:B:600:ILE:HG12	1:B:618:ILE:HD13	1.95	0.48
1:A:679:LYS:HZ2	1:A:718:HIS:CB	2.26	0.48
1:A:654:THR:HB	1:A:671:ARG:HB2	1.96	0.48
1:B:627:MET:HB2	1:B:691:GLN:HB2	1.95	0.47
1:A:617:LYS:NZ	1:A:667:CYS:CA	2.78	0.47
1:A:617:LYS:HZ2	1:A:667:CYS:CA	2.29	0.45
1:B:641:ILE:CD1	1:B:672:LEU:HD22	2.47	0.44
1:A:508:ASN:ND2	1:A:516:LYS:HZ3	2.16	0.44
1:B:379:LYS:NZ	1:B:577:ASN:CG	2.71	0.44
1:A:540:HIS:ND1	1:A:541:LEU:HD23	2.32	0.43
1:A:488:PHE:HB2	1:A:551:ARG:HB3	2.00	0.43
1:A:511:PHE:CD1	1:A:548:LEU:HD11	2.54	0.42

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:388:GLU:O	1:A:392:ARG:HG3	2.18	0.42
1:A:373:LEU:O	1:A:378:PHE:HB2	2.20	0.42
1:B:379:LYS:HZ1	1:B:577:ASN:ND2	2.18	0.42
1:B:465:ARG:NH2	1:B:551:ARG:NH2	2.68	0.41
1:A:343:LYS:HG2	1:A:407:TYR:HB2	2.02	0.41
1:A:639:LEU:HD13	1:A:672:LEU:HD13	2.03	0.41
1:B:401:LEU:HD13	2:B:801:6IO:C8	2.50	0.41
1:B:388:GLU:O	1:B:392:ARG:HG3	2.21	0.41
1:A:594:PHE:CE1	1:A:600:ILE:HD12	2.57	0.40
1:B:477:LEU:C	1:B:479:ALA:H	2.25	0.40
1:B:435:ARG:H	1:B:438:ASN:ND2	2.03	0.40
1:B:595:LYS:HB3	1:B:595:LYS:HE3	1.98	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	348/388 (90%)	338 (97%)	10 (3%)	0	100	100
1	B	361/388 (93%)	337 (93%)	24 (7%)	0	100	100
All	All	709/776 (91%)	675 (95%)	34 (5%)	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	313/344 (91%)	292 (93%)	21 (7%)	16 18
1	B	324/344 (94%)	306 (94%)	18 (6%)	21 25
All	All	637/688 (93%)	598 (94%)	39 (6%)	18 21

All (39) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	338	MET
1	A	341	LYS
1	A	401	LEU
1	A	451	GLU
1	A	467	ARG
1	A	474	ILE
1	A	481	LYS
1	A	497	GLU
1	A	501	ILE
1	A	516	LYS
1	A	538	LYS
1	A	568	GLU
1	A	573	SER
1	A	601	GLN
1	A	617	LYS
1	A	637	LEU
1	A	652	GLU
1	A	667	CYS
1	A	672	LEU
1	A	679	LYS
1	A	715	LEU
1	B	364	VAL
1	B	435	ARG
1	B	446	LEU
1	B	451	GLU
1	B	482	VAL
1	B	485	ASN
1	B	548	LEU
1	B	565	GLN
1	B	574	LEU
1	B	577	ASN
1	B	579	GLN
1	B	624	GLU
1	B	627	MET

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	660	SER
1	B	661	LYS
1	B	709	LYS
1	B	712	ILE
1	B	717	MET

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (10) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	371	ASN
1	A	438	ASN
1	A	508	ASN
1	A	718	HIS
1	B	438	ASN
1	B	485	ASN
1	B	494	GLN
1	B	577	ASN
1	B	601	GLN
1	B	703	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 monosaccharides in this entry.

5.6 Ligand geometry (i)

1 ligand is modelled in this entry.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

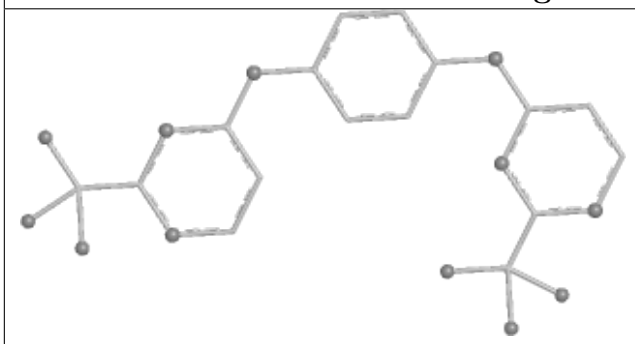
There are no torsion outliers.

There are no ring outliers.

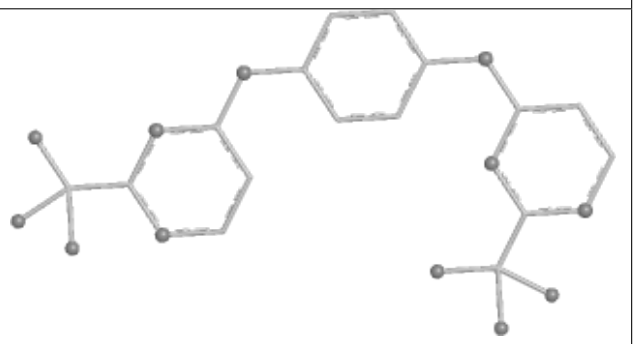
No monomer is involved in short contacts.

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.

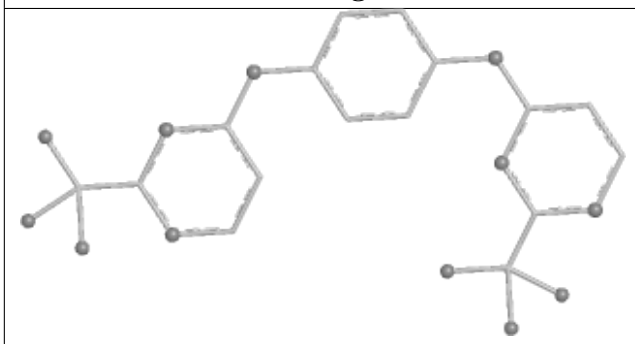
Ligand 6IO B 801



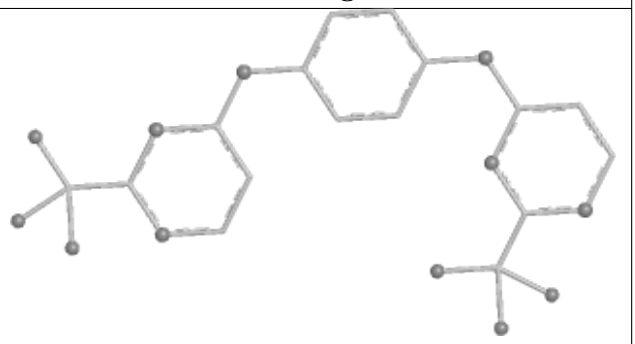
Bond lengths



Bond angles



Torsions



Rings

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	A	357/388 (92%)	0.18	13 (3%)	42	41	18, 43, 76, 91	0
1	B	368/388 (94%)	0.48	30 (8%)	11	10	19, 49, 87, 108	0
All	All	725/776 (93%)	0.33	43 (5%)	22	21	18, 45, 81, 108	0

All (43) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	627	MET	6.6
1	B	666	ALA	5.9
1	A	501	ILE	5.2
1	A	540	HIS	4.7
1	B	626	ILE	4.3
1	B	569	TYR	4.2
1	B	643	PRO	4.0
1	A	500	GLU	4.0
1	B	691	GLN	4.0
1	B	401	LEU	3.9
1	B	628	CYS	3.7
1	A	578	LEU	3.6
1	B	499	PHE	3.6
1	A	618	ILE	3.6
1	B	472	ASP	3.5
1	B	400	LEU	3.5
1	A	499	PHE	3.4
1	B	620	TYR	3.4
1	B	479	ALA	3.3
1	B	679	LYS	3.2
1	A	691	GLN	3.1
1	A	541	LEU	3.1
1	B	445	ILE	3.1
1	B	661	LYS	3.0

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	690	TYR	3.0
1	B	665	LYS	2.9
1	B	476	ILE	2.8
1	B	644	LYS	2.7
1	A	630	ALA	2.6
1	B	717	MET	2.4
1	B	631	TYR	2.4
1	B	481	LYS	2.4
1	B	482	VAL	2.4
1	B	623	PRO	2.2
1	A	600	ILE	2.2
1	B	624	GLU	2.2
1	B	619	VAL	2.1
1	B	700	GLU	2.1
1	B	690	TYR	2.1
1	B	594	PHE	2.1
1	A	474	ILE	2.1
1	B	716	ASP	2.1
1	A	389	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.

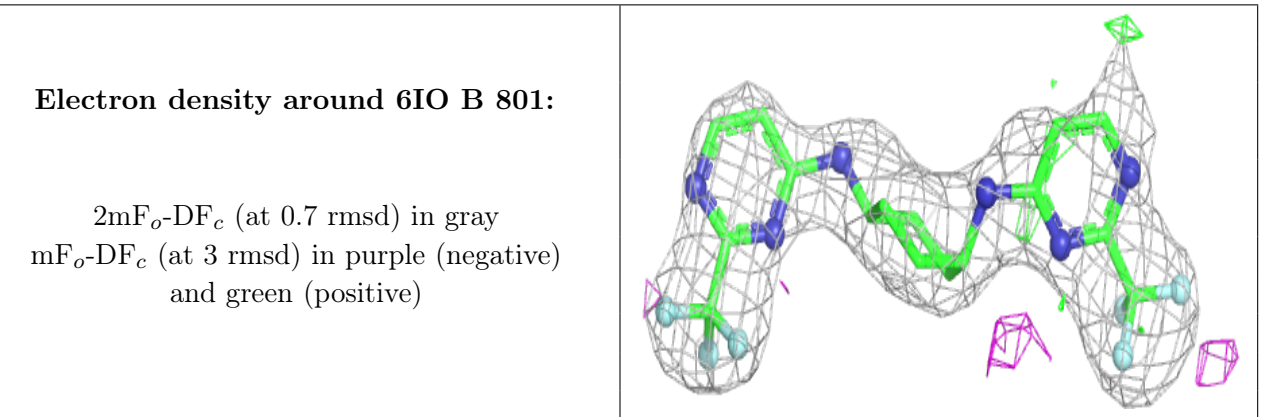
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
2	6IO	B	801	28/28	0.91	0.19	62,64,66,67	0

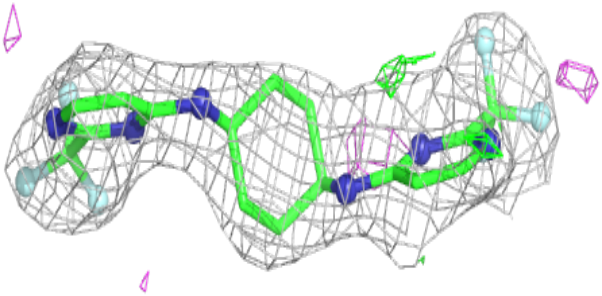
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers

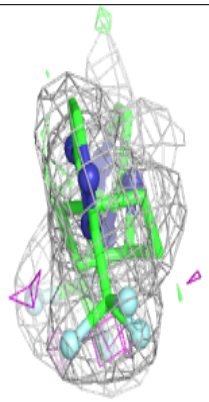
as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation to approximate a three-dimensional view.

Electron density around 6IO 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.