wwPDB X-ray Structure Validation Summary Report (i)

Nov 22, 2023 – 07:47 PM JST

PDB ID	:	8GQE
Title	:	Crystal structure of the W285A mutant of UVR8 in complex with RUP2
Authors	:	Wang, Y.D.; Wang, L.X.; Guan, Z.Y.; chang, H.F.; Yin, P.
Deposited on	:	2022-08-30
Resolution	:	2.00 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 (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 (i)) were used in the production of this report:

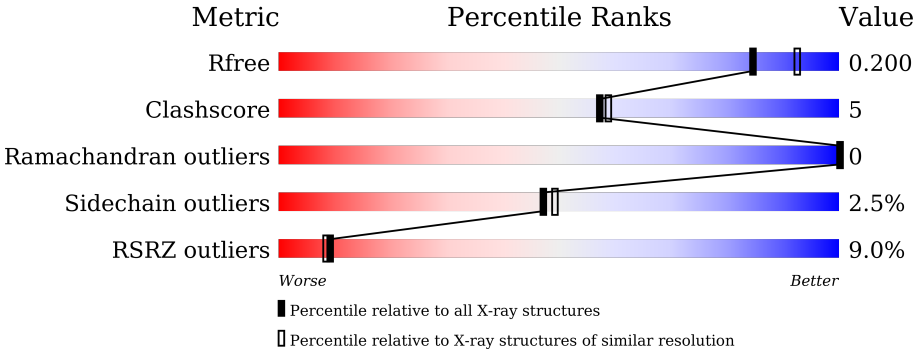
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (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 2.00 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	8085 (2.00-2.00)
Clashscore	141614	9178 (2.00-2.00)
Ramachandran outliers	138981	9054 (2.00-2.00)
Sidechain outliers	138945	9053 (2.00-2.00)
RSRZ outliers	127900	7900 (2.00-2.00)

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	369	6% 85% 9% 6%
2	B	443	10% 78% 9% • 13%

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 6064 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 WD repeat-containing protein RUP2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	348	2642	1646	475	507	14	0	0	0

There are 22 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	19	MET	-	initiating methionine	UNP Q9FFA7
A	367	LEU	-	expression tag	UNP Q9FFA7
A	368	GLU	-	expression tag	UNP Q9FFA7
A	369	ASP	-	expression tag	UNP Q9FFA7
A	370	TYR	-	expression tag	UNP Q9FFA7
A	371	LYS	-	expression tag	UNP Q9FFA7
A	372	ASP	-	expression tag	UNP Q9FFA7
A	373	HIS	-	expression tag	UNP Q9FFA7
A	374	ASP	-	expression tag	UNP Q9FFA7
A	375	GLY	-	expression tag	UNP Q9FFA7
A	376	ASP	-	expression tag	UNP Q9FFA7
A	377	TYR	-	expression tag	UNP Q9FFA7
A	378	LYS	-	expression tag	UNP Q9FFA7
A	379	ASP	-	expression tag	UNP Q9FFA7
A	380	HIS	-	expression tag	UNP Q9FFA7
A	381	ASP	-	expression tag	UNP Q9FFA7
A	382	ILE	-	expression tag	UNP Q9FFA7
A	383	ASP	-	expression tag	UNP Q9FFA7
A	384	TYR	-	expression tag	UNP Q9FFA7
A	385	LYS	-	expression tag	UNP Q9FFA7
A	386	ASP	-	expression tag	UNP Q9FFA7
A	387	ASP	-	expression tag	UNP Q9FFA7

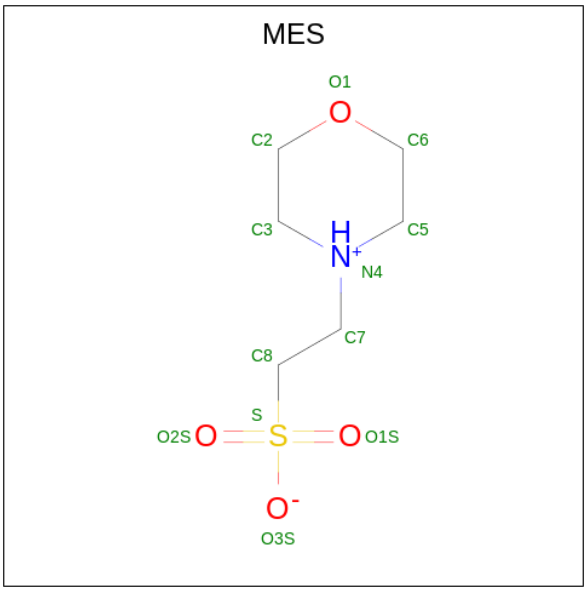
- Molecule 2 is a protein called Ultraviolet-B receptor UVR8.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	386	2960	1843	534	569	14	0	5	0

There are 34 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	11	MET	-	initiating methionine	UNP Q9FN03
B	285	ALA	TRP	engineered mutation	UNP Q9FN03
B	422	LEU	-	expression tag	UNP Q9FN03
B	423	GLU	-	expression tag	UNP Q9FN03
B	424	SER	-	expression tag	UNP Q9FN03
B	425	ALA	-	expression tag	UNP Q9FN03
B	426	TRP	-	expression tag	UNP Q9FN03
B	427	SER	-	expression tag	UNP Q9FN03
B	428	HIS	-	expression tag	UNP Q9FN03
B	429	PRO	-	expression tag	UNP Q9FN03
B	430	GLN	-	expression tag	UNP Q9FN03
B	431	PHE	-	expression tag	UNP Q9FN03
B	432	GLU	-	expression tag	UNP Q9FN03
B	433	LYS	-	expression tag	UNP Q9FN03
B	434	GLY	-	expression tag	UNP Q9FN03
B	435	GLY	-	expression tag	UNP Q9FN03
B	436	GLY	-	expression tag	UNP Q9FN03
B	437	SER	-	expression tag	UNP Q9FN03
B	438	GLY	-	expression tag	UNP Q9FN03
B	439	GLY	-	expression tag	UNP Q9FN03
B	440	GLY	-	expression tag	UNP Q9FN03
B	441	SER	-	expression tag	UNP Q9FN03
B	442	GLY	-	expression tag	UNP Q9FN03
B	443	GLY	-	expression tag	UNP Q9FN03
B	444	SER	-	expression tag	UNP Q9FN03
B	445	ALA	-	expression tag	UNP Q9FN03
B	446	TRP	-	expression tag	UNP Q9FN03
B	447	SER	-	expression tag	UNP Q9FN03
B	448	HIS	-	expression tag	UNP Q9FN03
B	449	PRO	-	expression tag	UNP Q9FN03
B	450	GLN	-	expression tag	UNP Q9FN03
B	451	PHE	-	expression tag	UNP Q9FN03
B	452	GLU	-	expression tag	UNP Q9FN03
B	453	LYS	-	expression tag	UNP Q9FN03

- Molecule 3 is 2-(N-MORPHOLINO)-ETHANESULFONIC ACID (three-letter code: MES) (formula: $C_6H_{13}NO_4S$) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	N	O	S		
3	B	1	12	6	1	4	1	0	0

- Molecule 4 is water.

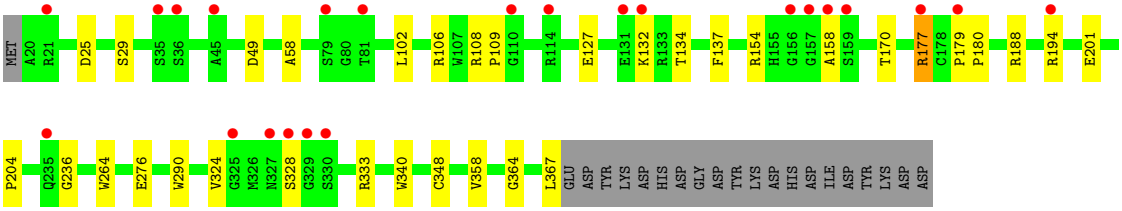
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
4	A	248	248	248	0	0
			Total	O		
4	B	202	202	202	0	0

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.

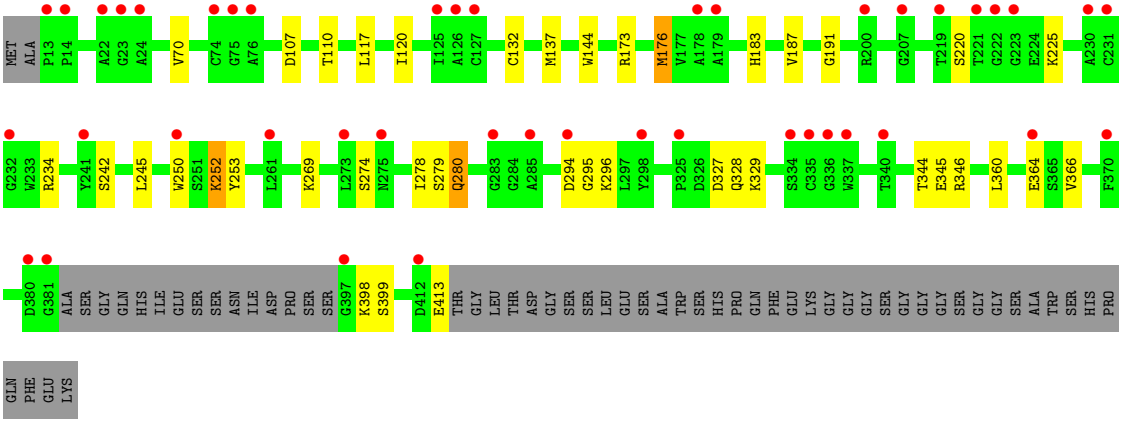
- Molecule 1: WD repeat-containing protein RUP2

Chain A:



- Molecule 2: Ultraviolet-B receptor UVR8

Chain B:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	I 4	Depositor
Cell constants a, b, c, α, β, γ	173.39Å 173.39Å 76.74Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	34.01 – 2.00 34.01 – 2.00	Depositor EDS
% Data completeness (in resolution range)	99.6 (34.01-2.00) 99.6 (34.01-2.00)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.36 (at 2.00Å)	Xtriage
Refinement program	PHENIX 1.19.1_4122	Depositor
R, R_{free}	0.173 , 0.200 0.173 , 0.200	Depositor DCC
R_{free} test set	3905 reflections (5.10%)	wwPDB-VP
Wilson B-factor (Å^2)	31.1	Xtriage
Anisotropy	0.340	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.36 , 51.0	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	0.012 for -k,-h,-l	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	6064	wwPDB-VP
Average B, all atoms (Å^2)	40.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.94% 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 (i)

5.1 Standard geometry (i)

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

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	0/2701	0.68	2/3668 (0.1%)
2	B	0.36	0/3030	0.56	0/4109
All	All	0.39	0/5731	0.62	2/7777 (0.0%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	177	ARG	NE-CZ-NH2	-17.00	111.80	120.30
1	A	177	ARG	NH1-CZ-NH2	5.57	125.53	119.40

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	2642	0	2551	19	0
2	B	2960	0	2824	32	0
3	B	12	0	13	3	0
4	A	248	0	0	3	1
4	B	202	0	0	7	0
All	All	6064	0	5388	54	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 5.

The worst 5 of 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)
3:B:501:MES:H52	4:B:722:HOH:O	1.37	1.20
1:A:276:GLU:OE1	4:A:401:HOH:O	1.84	0.94
1:A:127:GLU:HB2	1:A:137:PHE:HB3	1.69	0.74
3:B:501:MES:H32	4:B:638:HOH:O	1.89	0.73
1:A:132:LYS:HB2	1:A:134:THR:HG22	1.70	0.72

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)
4:A:570:HOH:O	4:A:577:HOH:O[6_554]	1.84	0.36

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	346/369 (94%)	334 (96%)	12 (4%)	0	100	100
2	B	387/443 (87%)	385 (100%)	2 (0%)	0	100	100
All	All	733/812 (90%)	719 (98%)	14 (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	284/304 (93%)	279 (98%)	5 (2%)	59	63
2	B	310/347 (89%)	299 (96%)	11 (4%)	36	35
All	All	594/651 (91%)	578 (97%)	16 (3%)	47	46

5 of 16 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	327	ASP
2	B	280	GLN
2	B	137	MET
2	B	252	LYS
2	B	107	ASP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
2	B	183	HIS
2	B	203	ASN

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.

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	MES	B	501	-	12,12,12	0.78	0	14,16,16	1.03	1 (7%)

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	MES	B	501	-	-	3/6/14/14	0/1/1/1

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	B	501	MES	C7-N4-C3	2.31	117.15	111.23

There are no chirality outliers.

All (3) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	B	501	MES	C8-C7-N4-C3
3	B	501	MES	C8-C7-N4-C5
3	B	501	MES	N4-C7-C8-S

There are no ring outliers.

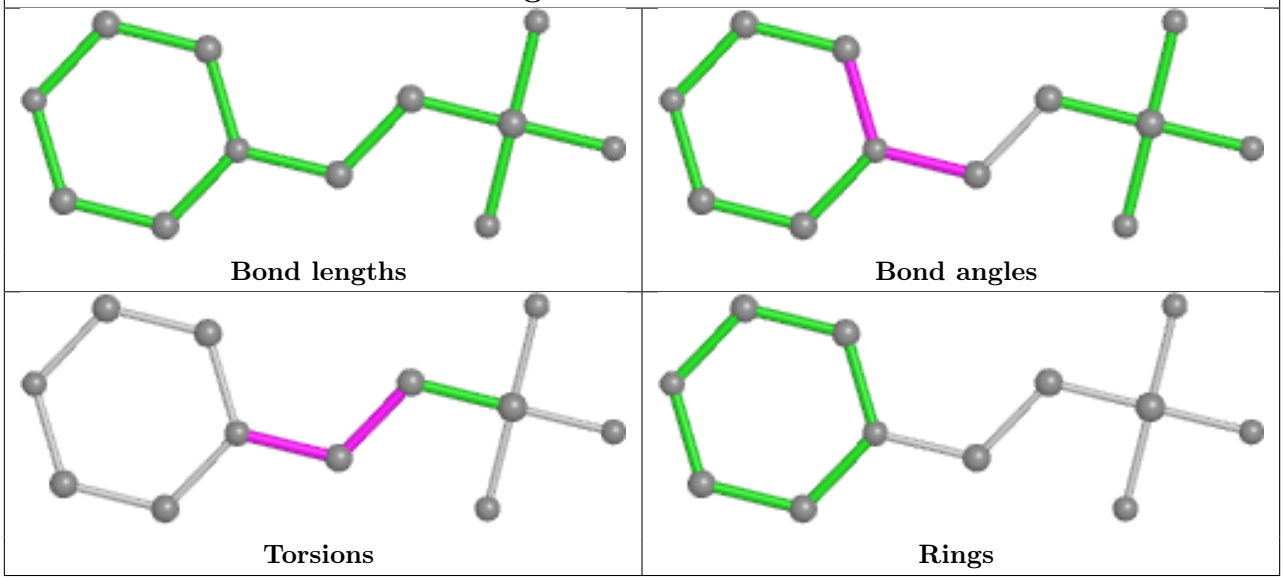
1 monomer is involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	B	501	MES	3	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 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 MES B 501



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	348/369 (94%)	0.17	23 (6%) 18 17	21, 31, 63, 117	0
2	B	386/443 (87%)	0.38	43 (11%) 5 4	23, 41, 68, 98	0
All	All	734/812 (90%)	0.28	66 (8%) 9 8	21, 36, 66, 117	0

The worst 5 of 66 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	381	GLY	7.7
1	A	158	ALA	5.8
1	A	328	SER	5.8
1	A	327	ASN	5.7
1	A	36	SER	5.2

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 monosaccharides in this entry.

6.4 Ligands (i)

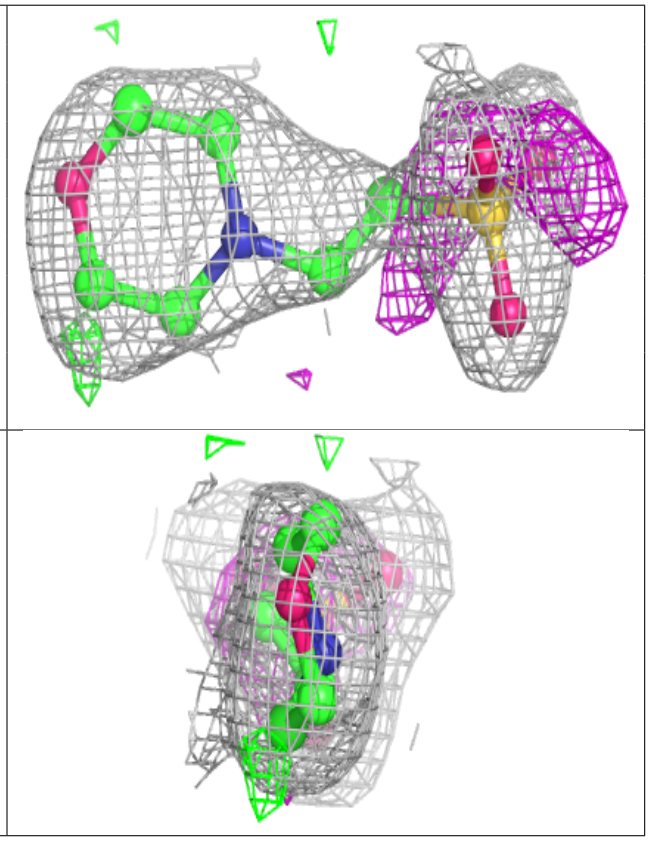
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists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

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
3	MES	B	501	12/12	0.89	0.26	40,44,52,69	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 MES B 501:

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.