wwPDB X-ray Structure Validation Summary Report (i)

Nov 25, 2021 – 12:12 pm GMT

PDB ID	:	6ZG9
Title	:	Structure of M1-StaR-T4L in complex with GSK1034702 at 2.5A
Authors	:	Rucktooa, P.; Cooke, R.M.
Deposited on	:	2020-06-18
Resolution	:	2.50 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 following versions of software and data (see references (i)) were used in the production of this report:

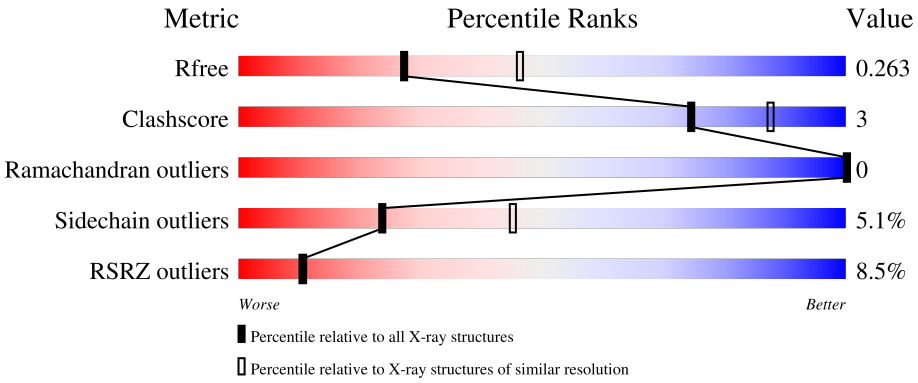
MolProbity	:	4.02b-467
Mogul	:	1.8.4 (270009), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.23.2
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0267
CCP4	:	7.1.010 (Gargrove)
Ideal geometry (proteins)	:	Engh & 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 2.50 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	4661 (2.50-2.50)
Clashscore	141614	5346 (2.50-2.50)
Ramachandran outliers	138981	5231 (2.50-2.50)
Sidechain outliers	138945	5233 (2.50-2.50)
RSRZ outliers	127900	4559 (2.50-2.50)

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	455	8% 87% 11% .

2 Entry composition (i)

There are 6 unique types of molecules in this entry. The entry contains 3741 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 Muscarinic acetylcholine receptor M1,Endolysin,Muscarinic acetylcholine receptor M1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	446	3559	2313	600	622	24	0	1	0

There are 49 discrepancies between the modelled and reference sequences:

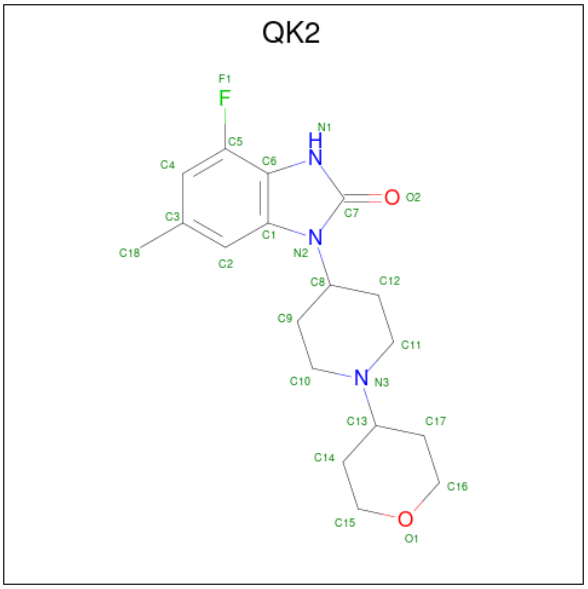
Chain	Residue	Modelled	Actual	Comment	Reference
A	20	MET	-	initiating methionine	UNP P11229
A	21	GLU	-	expression tag	UNP P11229
A	22	THR	-	expression tag	UNP P11229
A	23	VAL	-	expression tag	UNP P11229
A	24	GLU	-	expression tag	UNP P11229
A	25	MET	-	expression tag	UNP P11229
A	26	VAL	-	expression tag	UNP P11229
A	27	ALA	PHE	engineered mutation	UNP P11229
A	29	ALA	GLY	conflict	UNP P11229
A	30	THR	ILE	conflict	UNP P11229
A	31	VAL	THR	conflict	UNP P11229
A	32	ALA	THR	engineered mutation	UNP P11229
A	44	ILE	LEU	engineered mutation	UNP P11229
A	46	LEU	VAL	engineered mutation	UNP P11229
A	47	MET	LEU	conflict	UNP P11229
A	48	LEU	ILE	conflict	UNP P11229
A	50	ILE	PHE	conflict	UNP P11229
A	54	ARG	THR	conflict	UNP P11229
A	55	GLN	GLU	conflict	UNP P11229
A	57	GLN	LYS	conflict	UNP P11229
A	64	ALA	LEU	engineered mutation	UNP P11229
A	65	PHE	LEU	conflict	UNP P11229
A	76	ALA	THR	engineered mutation	UNP P11229
A	84	VAL	THR	engineered mutation	UNP P11229
A	86	ILE	LEU	conflict	UNP P11229
A	87	ILE	LEU	conflict	UNP P11229

Continued on next page...

Continued from previous page...

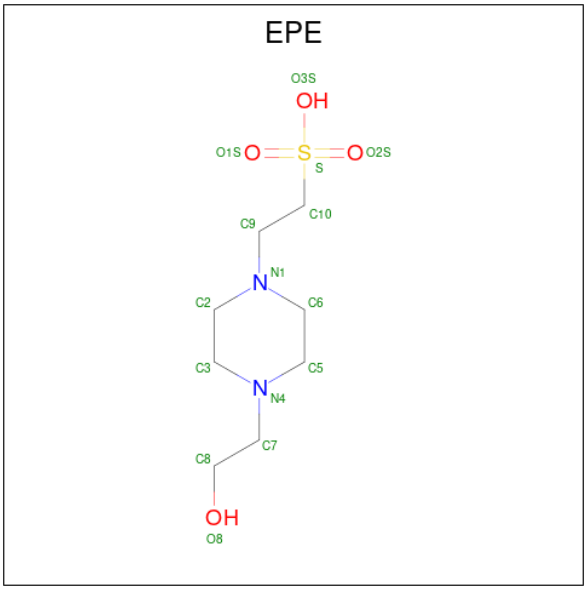
Chain	Residue	Modelled	Actual	Comment	Reference
A	95	ALA	THR	engineered mutation	UNP P11229
A	101	ALA	TRP	engineered mutation	UNP P11229
A	112	ALA	SER	engineered mutation	UNP P11229
A	143	LEU	ALA	engineered mutation	UNP P11229
A	196	THR	ALA	engineered mutation	UNP P11229
A	1012	GLY	ARG	conflict	UNP P00720
A	1054	THR	CYS	conflict	UNP P00720
A	1097	ALA	CYS	conflict	UNP P00720
A	1137	ARG	ILE	conflict	UNP P00720
A	362	ALA	LYS	engineered mutation	UNP P11229
A	364	LEU	ALA	engineered mutation	UNP P11229
A	411	ALA	SER	engineered mutation	UNP P11229
A	435	ALA	CYS	conflict	UNP P11229
A	439	HIS	-	expression tag	UNP P11229
A	440	HIS	-	expression tag	UNP P11229
A	441	HIS	-	expression tag	UNP P11229
A	442	HIS	-	expression tag	UNP P11229
A	443	HIS	-	expression tag	UNP P11229
A	444	HIS	-	expression tag	UNP P11229
A	445	HIS	-	expression tag	UNP P11229
A	446	HIS	-	expression tag	UNP P11229
A	447	HIS	-	expression tag	UNP P11229
A	448	HIS	-	expression tag	UNP P11229

- Molecule 2 is 7-fluoranyl-5-methyl-3-[1-(oxan-4-yl)piperidin-4-yl]-1 {H}-benzimidazol-2 -one (three-letter code: QK2) (formula: $C_{18}H_{24}FN_3O_2$) (labeled as "Ligand of Interest" by depositor).



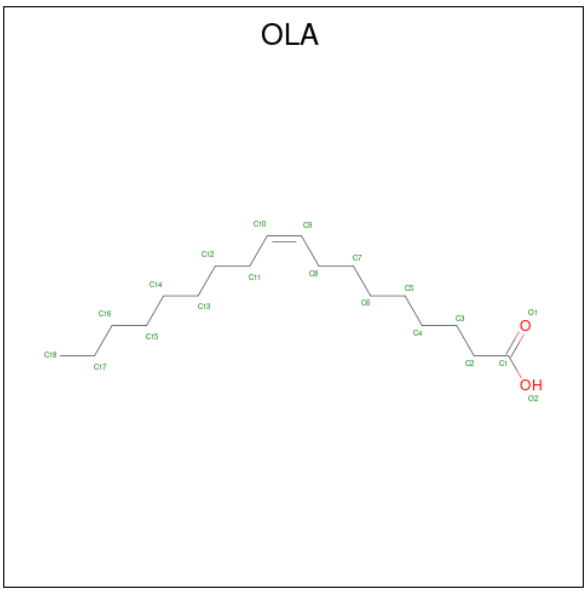
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	1	Total	C	F	N	O	0	0
			24	18	1	3	2		

- Molecule 3 is 4-(2-HYDROXYETHYL)-1-PIPERAZINE ETHANESULFONIC ACID (three-letter code: EPE) (formula: $C_8H_{18}N_2O_4S$).



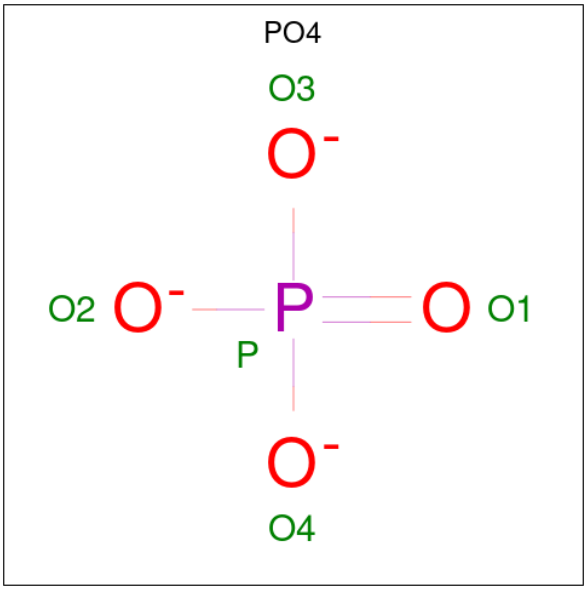
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	A	1	Total	C	N	O	S	0	0
			15	8	2	4	1		

- Molecule 4 is OLEIC ACID (three-letter code: OLA) (formula: $C_{18}H_{34}O_2$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total C O 20 18 2	0	0
4	A	1	Total C O 20 18 2	0	0
4	A	1	Total C O 15 13 2	0	0
4	A	1	Total C O 20 18 2	0	0
4	A	1	Total C O 20 18 2	0	0
4	A	1	Total C O 20 18 2	0	0

- Molecule 5 is PHOSPHATE ION (three-letter code: PO4) (formula: O_4P).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	1	Total O P 5 4 1	0	0

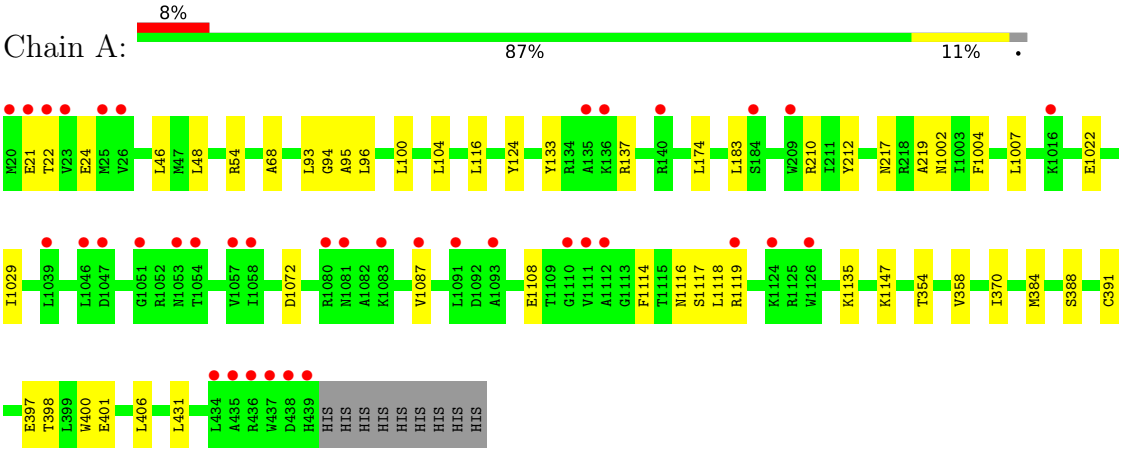
- Molecule 6 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	23	Total O 23 23	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: Muscarinic acetylcholine receptor M1,Endolysin,Muscarinic acetylcholine receptor M1



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	62.36Å 66.57Å 153.10Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	33.97 – 2.50 33.97 – 2.50	Depositor EDS
% Data completeness (in resolution range)	61.7 (33.97-2.50) 61.7 (33.97-2.50)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.21 (at 2.51Å)	Xtriage
Refinement program	BUSTER 2.11.7	Depositor
R, R_{free}	0.203 , 0.242 0.224 , 0.263	Depositor DCC
R_{free} test set	732 reflections (5.21%)	wwPDB-VP
Wilson B-factor (Å^2)	36.1	Xtriage
Anisotropy	0.242	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	(Not available) , (Not available)	EDS
L-test for twinning[2]	$< \|L\| > = 0.46$, $< L^2 > = 0.29$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.90	EDS
Total number of atoms	3741	wwPDB-VP
Average B, all atoms (Å^2)	52.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.27% 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: OLA, QK2, PO4, EPE

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	0/3639	0.60	0/4947

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	3559	0	3640	23	0
2	A	24	0	0	0	0
3	A	15	0	18	0	0
4	A	115	0	185	10	0
5	A	5	0	0	0	0
6	A	23	0	0	0	0
All	All	3741	0	3843	24	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 3.

The worst 5 of 24 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:212:TYR:OH	1:A:358:VAL:HG11	1.80	0.80
1:A:384:MET:HE1	1:A:400:TRP:HA	1.71	0.71
1:A:94:GLY:HA2	4:A:1206:OLA:H21	1.77	0.66
1:A:94:GLY:HA3	4:A:1206:OLA:H71	1.77	0.64
1:A:95:ALA:H	4:A:1206:OLA:H42	1.65	0.61

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	443/455 (97%)	437 (99%)	6 (1%)	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	377/385 (98%)	358 (95%)	19 (5%)	24 46

5 of 19 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	354	THR
1	A	406	LEU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	431	LEU
1	A	391	CYS
1	A	1072	ASP

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

Mol	Chain	Res	Type
1	A	1140	ASN
1	A	422	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)

9 ligands are 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.

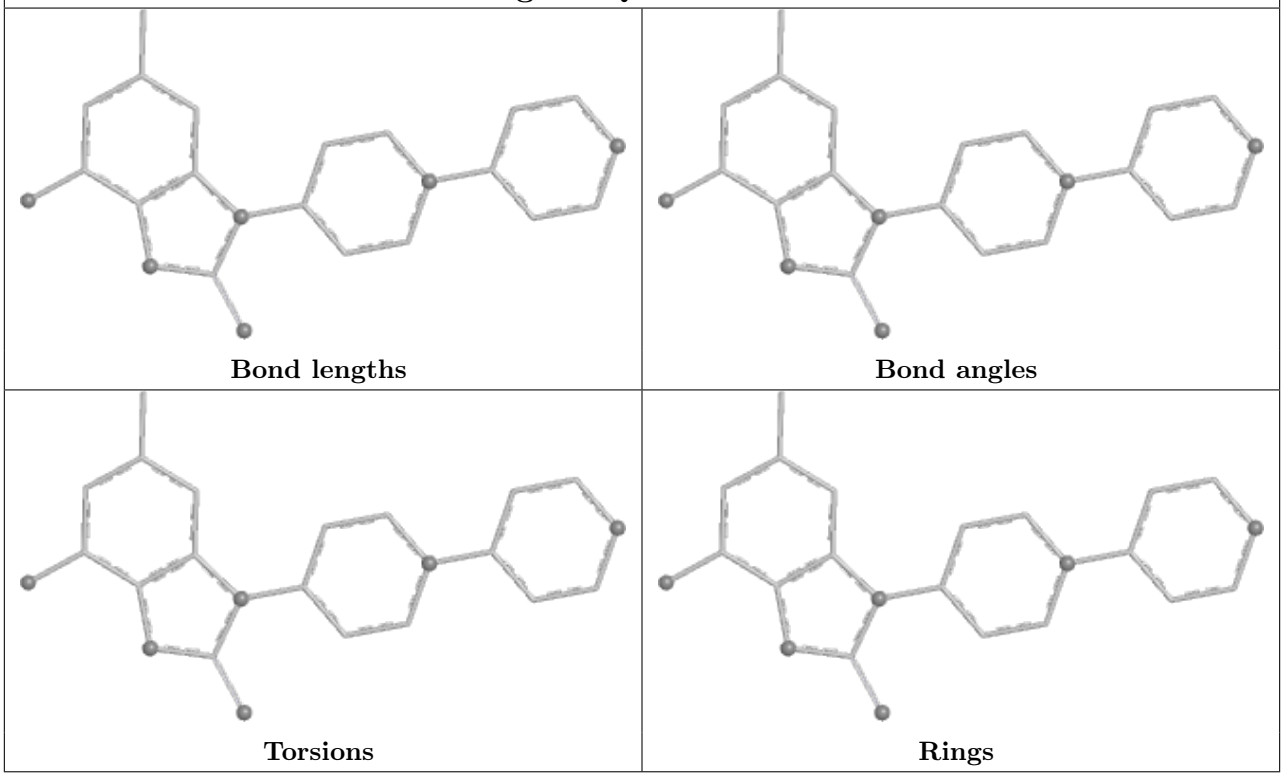
There are no ring outliers.

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 QK2 A 1201



5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
1	A	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	A	219:ALA	C	1002:ASN	N	2.66

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	446/455 (98%)	0.42	38 (8%)	10	10	18, 48, 89, 135	0

The worst 5 of 38 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	1057	VAL	9.0
1	A	437	TRP	6.8
1	A	1054	THR	5.4
1	A	439	HIS	4.8
1	A	1016	LYS	4.6

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
3	EPE	A	1202	15/15	0.76	0.25	102,106,118,119	0
4	OLA	A	1205	15/20	0.78	0.23	39,44,54,55	0

Continued on next page...

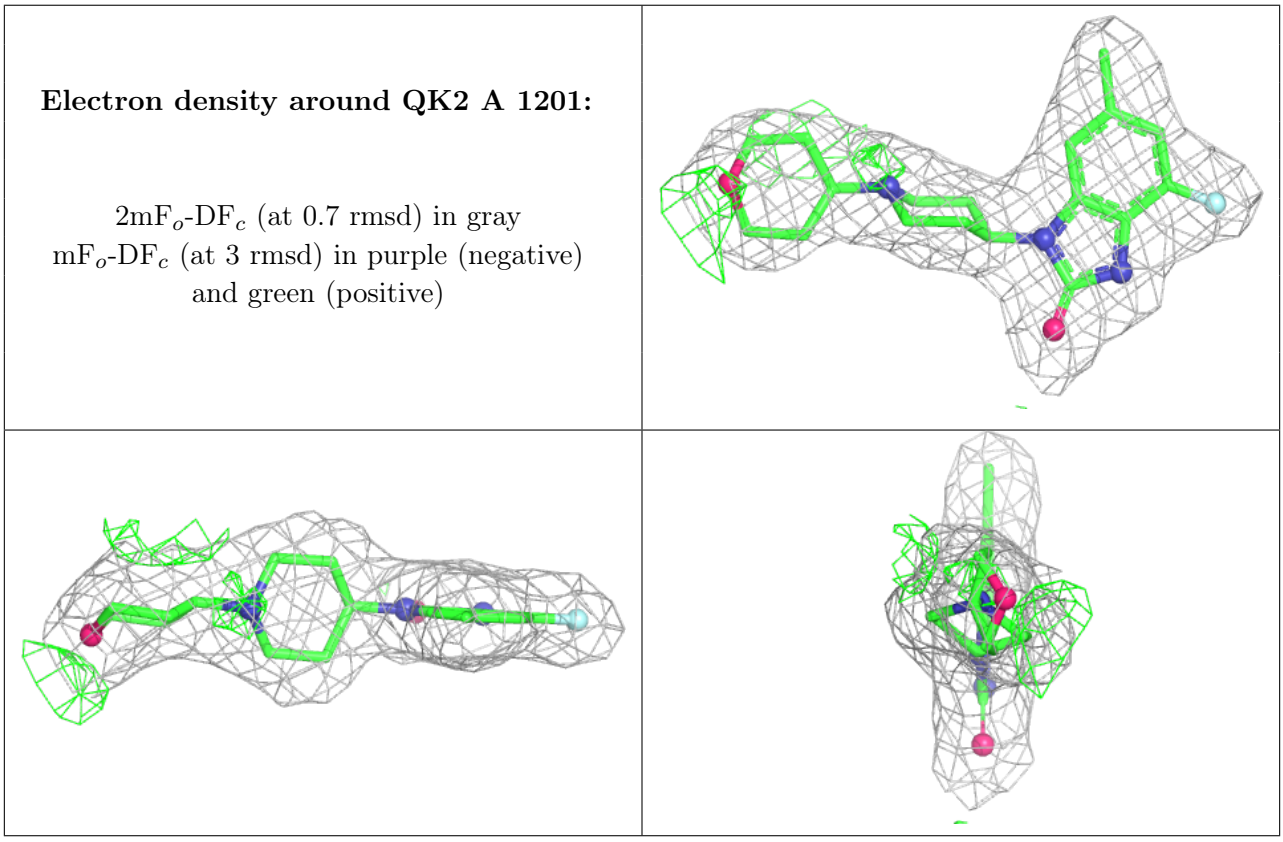
Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
4	OLA	A	1207	20/20	0.81	0.25	40,48,56,57	0
4	OLA	A	1206	20/20	0.82	0.25	38,41,56,57	0
4	OLA	A	1208	20/20	0.84	0.21	38,56,67,68	0
5	PO4	A	1209	5/5	0.88	0.36	100,100,101,101	0
4	OLA	A	1203	20/20	0.89	0.22	34,42,55,56	0
4	OLA	A	1204	20/20	0.92	0.19	16,26,39,41	0
2	QK2	A	1201	24/24	0.95	0.19	9,25,41,41	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 QK2 A 1201:

$2mF_o$-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.