



Full wwPDB X-ray Structure Validation Report ⓘ

May 16, 2020 – 11:02 am BST

PDB ID : 5XGJ
Title : Crystal structure of PI3K complex with an inhibitor
Authors : Song, K.; Yang, X.; Zhao, Y.; Jian, Z.
Deposited on : 2017-04-13
Resolution : 2.97 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 following versions of software and data (see [references ⓘ](#)) 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 : 2.11
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) : 2.11

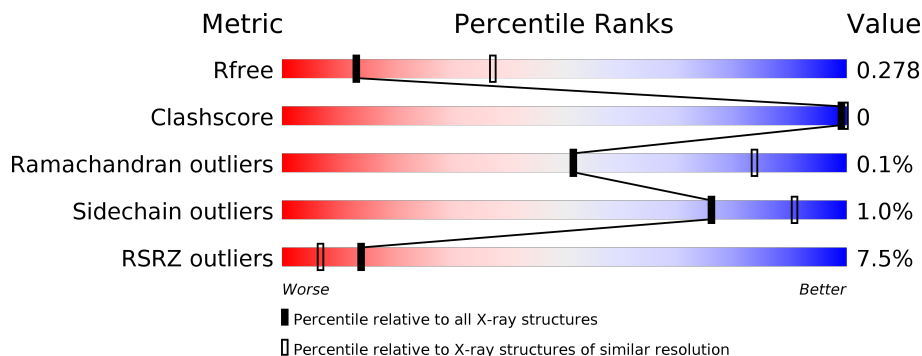
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.97 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	2754 (3.00-2.96)
Clashscore	141614	3103 (3.00-2.96)
Ramachandran outliers	138981	2993 (3.00-2.96)
Sidechain outliers	138945	2996 (3.00-2.96)
RSRZ outliers	127900	2644 (3.00-2.96)

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 on 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 $\leq 5\%$. 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	1048	
2	B	278	

2 Entry composition [i](#)

There are 5 unique types of molecules in this entry. The entry contains 10430 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.

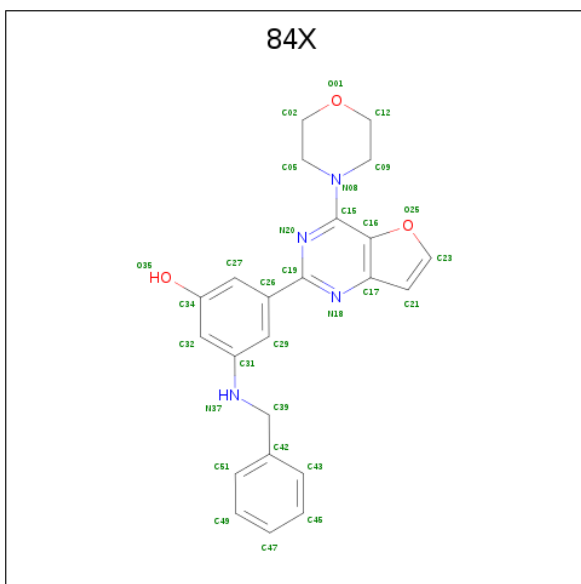
- Molecule 1 is a protein called Phosphatidylinositol 4,5-bisphosphate 3-kinase catalytic subunit alpha isoform.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	980	8013	5127	1371	1451	64	0	0	0

- Molecule 2 is a protein called Phosphatidylinositol 3-kinase regulatory subunit alpha.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	278	2358	1476	421	453	8	0	0	0

- Molecule 3 is 3-(4-morpholin-4-ylfuro[3,2-d]pyrimidin-2-yl)-5-[(phenylmethyl)amino]phenol (three-letter code: 84X) (formula: C₂₃H₂₂N₄O₃).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
3	A	1	30	23	4	3	0	0

- Molecule 4 is SULFATE ION (three-letter code: SO4) (formula: O₄S).



Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	B	1	Total	O S	0	0
			5	4 1		

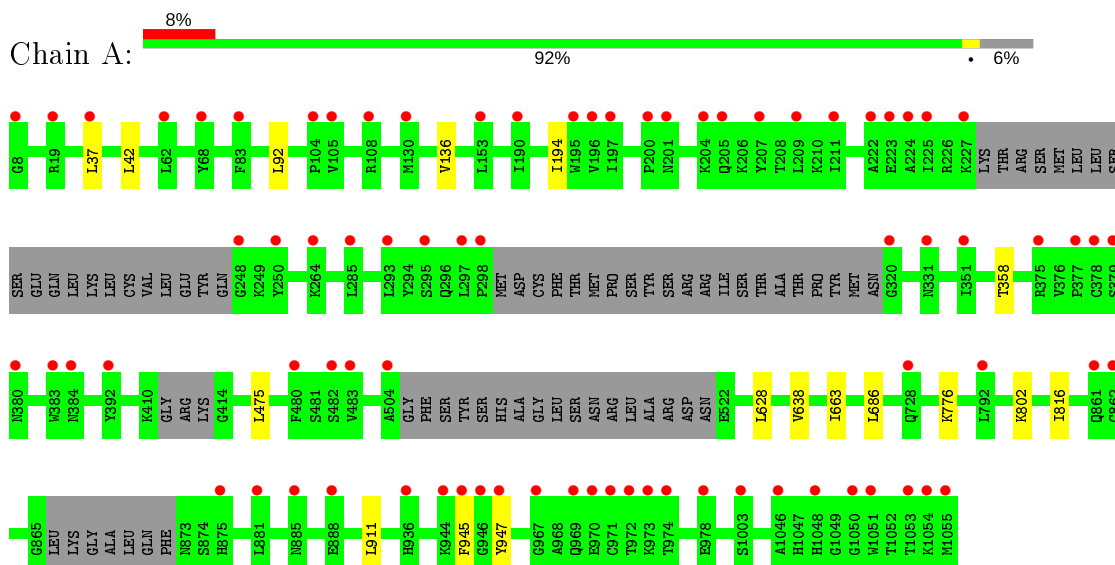
- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	17	Total	O	0	0
			17	17		
5	B	7	Total	O	0	0
			7	7		

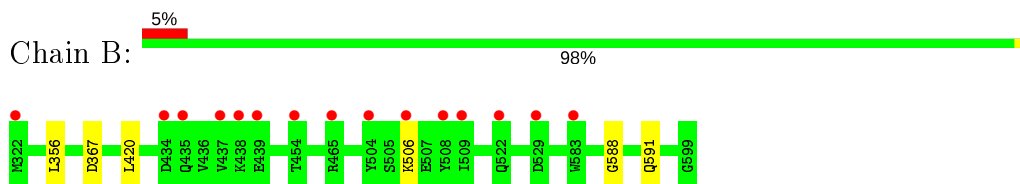
3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA 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: Phosphatidylinositol 4,5-bisphosphate 3-kinase catalytic subunit alpha isoform



- Molecule 2: Phosphatidylinositol 3-kinase regulatory subunit alpha



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	69.88Å 136.37Å 149.03Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	50.00 – 2.97 48.80 – 2.97	Depositor EDS
% Data completeness (in resolution range)	98.6 (50.00-2.97) 98.6 (48.80-2.97)	Depositor EDS
R_{merge}	0.16	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.06 (at 2.96Å)	Xtrriage
Refinement program	REFMAC 5.8.0135	Depositor
R, R_{free}	0.231 , 0.279 0.231 , 0.278	Depositor DCC
R_{free} test set	1400 reflections (4.72%)	wwPDB-VP
Wilson B-factor (Å ²)	84.9	Xtrriage
Anisotropy	0.015	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.31 , 42.8	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.93	EDS
Total number of atoms	10430	wwPDB-VP
Average B, all atoms (Å ²)	99.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.56% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ 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: SO4, 84X

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	0.36	0/8191	0.52	0/11066
2	B	0.36	0/2398	0.53	0/3212
All	All	0.36	0/10589	0.52	0/14278

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 [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	8013	0	8025	6	0
2	B	2358	0	2333	0	0
3	A	30	0	0	0	0
4	B	5	0	0	0	0
5	A	17	0	0	0	0
5	B	7	0	0	0	0
All	All	10430	0	10358	6	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 0.

All (6) 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:628:LEU:HD23	1:A:663:ILE:HD13	1.95	0.48
1:A:136:VAL:HG13	1:A:686:LEU:HD11	1.98	0.46
1:A:42:LEU:HD11	1:A:92:LEU:HD11	2.00	0.42
1:A:776:LYS:HB2	1:A:802:LYS:HE3	2.01	0.42
1:A:194:ILE:HD12	1:A:194:ILE:N	2.36	0.41
1:A:816:ILE:HG21	1:A:911:LEU:HD21	2.03	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	968/1048 (92%)	936 (97%)	32 (3%)	0	100	100
2	B	276/278 (99%)	269 (98%)	6 (2%)	1 (0%)	34	70
All	All	1244/1326 (94%)	1205 (97%)	38 (3%)	1 (0%)	51	83

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	588	GLY

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	895/955 (94%)	889 (99%)	6 (1%)	84	93
2	B	258/258 (100%)	253 (98%)	5 (2%)	57	82
All	All	1153/1213 (95%)	1142 (99%)	11 (1%)	76	91

All (11) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	37	LEU
1	A	358	THR
1	A	475	LEU
1	A	638	VAL
1	A	945	PHE
1	A	947	TYR
2	B	356	LEU
2	B	367	ASP
2	B	420	LEU
2	B	506	LYS
2	B	591	GLN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	201	ASN
1	A	205	GLN
1	A	827	GLN
2	B	453	ASN
2	B	457	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 carbohydrates in this entry.

5.6 Ligand geometry

2 ligands are 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
4	SO4	B	601	-	4,4,4	0.30	0	6,6,6	0.08	0
3	84X	A	1101	-	30,34,34	1.98	9 (30%)	37,47,47	2.22	5 (13%)

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	84X	A	1101	-	-	3/13/21/21	0/5/5/5

All (9) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	1101	84X	C15-N08	5.63	1.53	1.37
3	A	1101	84X	C31-N37	4.28	1.51	1.38
3	A	1101	84X	C51-C42	3.04	1.45	1.38
3	A	1101	84X	C15-N20	2.75	1.36	1.32
3	A	1101	84X	C39-C42	2.35	1.56	1.51
3	A	1101	84X	C39-N37	2.29	1.52	1.45
3	A	1101	84X	C19-N20	2.16	1.39	1.34
3	A	1101	84X	C17-N18	2.16	1.41	1.37
3	A	1101	84X	C16-C17	-2.09	1.35	1.42

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	1101	84X	N18-C19-N20	-10.22	118.08	126.11
3	A	1101	84X	C19-N18-C17	6.05	120.61	116.50
3	A	1101	84X	C15-N20-C19	2.74	122.60	116.17

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	1101	84X	C26-C19-N18	2.36	122.49	116.88
3	A	1101	84X	C09-N08-C15	2.32	125.35	118.73

There are no chirality outliers.

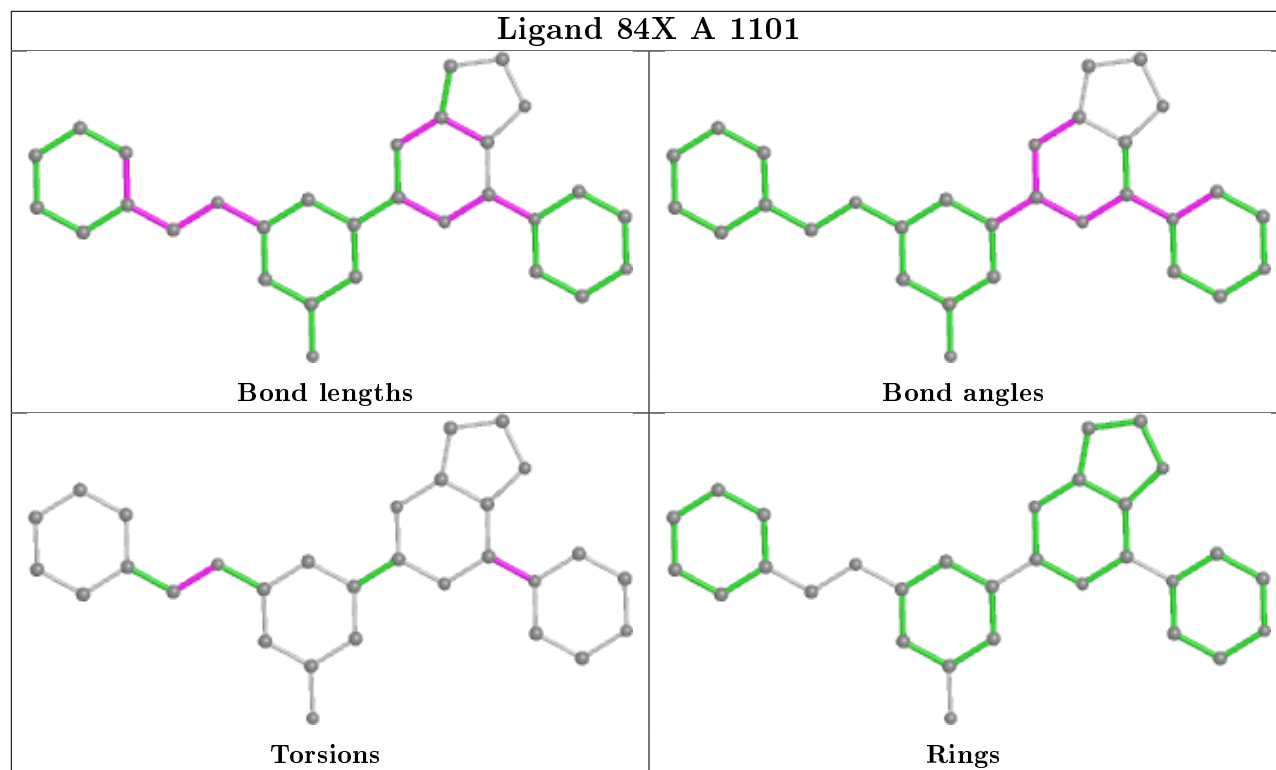
All (3) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	1101	84X	C16-C15-N08-C09
3	A	1101	84X	N20-C15-N08-C09
3	A	1101	84X	C42-C39-N37-C31

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 than 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.



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, 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	980/1048 (93%)	0.67	79 (8%) 12 6	61, 92, 136, 173	0
2	B	278/278 (100%)	0.41	15 (5%) 25 15	76, 100, 147, 170	0
All	All	1258/1326 (94%)	0.61	94 (7%) 14 7	61, 95, 141, 173	0

All (94) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	482	SER	8.3
1	A	946	GLY	7.5
2	B	437	VAL	6.4
1	A	383	TRP	5.9
1	A	1055	MET	5.7
1	A	201	ASN	5.4
1	A	378	CYS	5.0
2	B	439	GLU	5.0
1	A	320	GLY	4.8
1	A	204	LYS	4.4
1	A	225	ILE	4.4
1	A	377	PRO	4.3
1	A	205	GLN	4.3
1	A	375	ARG	4.3
2	B	435	GLN	4.2
1	A	227	LYS	4.2
1	A	888	GLU	4.2
1	A	947	TYR	4.0
1	A	504	ALA	3.8
1	A	380	ASN	3.8
2	B	438	LYS	3.7
1	A	62	LEU	3.6
1	A	1051	TRP	3.5
1	A	483	VAL	3.5

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	207	TYR	3.4
1	A	881	LEU	3.4
1	A	967	GLY	3.2
1	A	211	ILE	3.2
2	B	508	TYR	3.1
1	A	264	LYS	3.1
1	A	37	LEU	3.1
1	A	875	HIS	3.1
1	A	196	VAL	3.0
1	A	1046	ALA	3.0
1	A	1050	GLY	2.9
1	A	200	PRO	2.9
1	A	108	ARG	2.9
1	A	105	VAL	2.9
2	B	583	TRP	2.8
2	B	465	ARG	2.8
2	B	322	MET	2.8
1	A	862	CYS	2.8
1	A	1048	HIS	2.7
1	A	222	ALA	2.7
2	B	434	ASP	2.7
1	A	973	LYS	2.7
1	A	297	LEU	2.7
1	A	970	GLU	2.7
1	A	130	MET	2.6
2	B	509	ILE	2.6
1	A	195	TRP	2.6
1	A	209	LEU	2.6
1	A	285	LEU	2.6
2	B	504	TYR	2.6
1	A	971	CYS	2.6
1	A	197	ILE	2.6
1	A	969	GLN	2.5
2	B	522	GLN	2.5
1	A	248	GLY	2.5
1	A	944	LYS	2.5
1	A	331	ASN	2.4
1	A	936	HIS	2.4
1	A	68	TYR	2.4
1	A	104	PRO	2.4
1	A	1054	LYS	2.4
1	A	885	ASN	2.4

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	293	LEU	2.4
1	A	480	PHE	2.4
1	A	19	ARG	2.3
1	A	223	GLU	2.3
1	A	153	LEU	2.3
1	A	351	ILE	2.3
2	B	454	THR	2.3
1	A	728	GLN	2.3
1	A	861	GLN	2.3
1	A	384	ASN	2.3
1	A	1003	SER	2.3
1	A	190	ILE	2.3
1	A	250	TYR	2.3
1	A	298	PRO	2.2
1	A	295	SER	2.2
1	A	974	THR	2.2
1	A	978	GLU	2.2
1	A	392	TYR	2.1
1	A	792	LEU	2.1
2	B	506	LYS	2.1
1	A	945	PHE	2.1
1	A	379	SER	2.1
1	A	1053	THR	2.1
2	B	529	ASP	2.1
1	A	224	ALA	2.0
1	A	972	THR	2.0
1	A	83	PHE	2.0
1	A	8	GLY	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 carbohydrates in this entry.

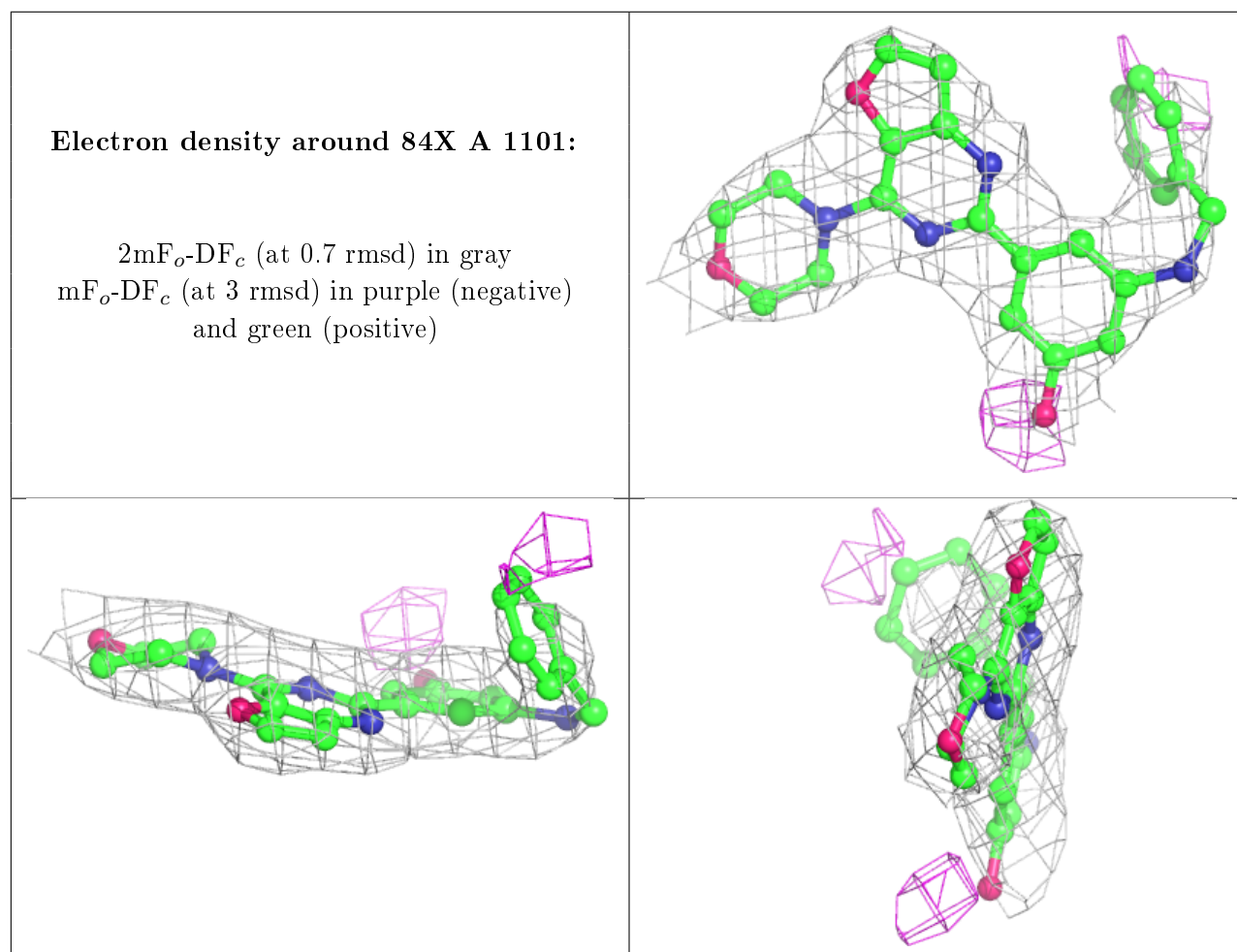
6.4 Ligands [i](#)

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	84X	A	1101	30/30	0.93	0.28	73,76,93,94	0
4	SO4	B	601	5/5	0.95	0.13	97,97,99,101	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.



6.5 Other polymers [i](#)

There are no such residues in this entry.