



wwPDB X-ray Structure Validation Summary Report ⓘ

Jul 14, 2016 – 07:40 PM EDT

PDB ID : 5FI4
Title : Discovery of imidazo[1,2-a]-pyridine inhibitors of pan-PI3 kinases that are efficacious in a mouse xenograft model
Authors : Elling, R.A.; Knapp, M.S.; Han, W.; Daniel, L.M.; Xy, Y.; Burger, M.T.; Ni, Z.; Smith, A.; Lan, J.; Williams, T.; Verhagen, J.; Huh, K.; Merritt, H.; Chan, J.; Kaufman, S.; Voliva, C.F.; Pecchi, S.
Deposited on : 2015-12-22
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

<http://wwpdb.org/validation/2016/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	:	1.7.1 (RC1), CSD as537be (2016)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20027790
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	rb-20027790

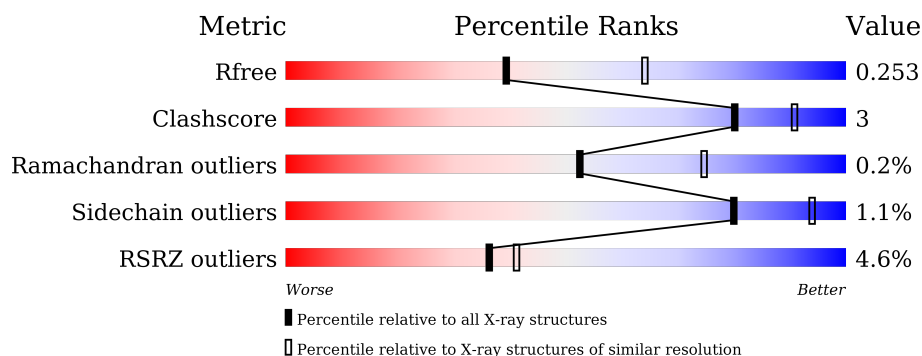
1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.50 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}	91344	3553 (2.50-2.50)
Clashscore	102246	4242 (2.50-2.50)
Ramachandran outliers	100387	4156 (2.50-2.50)
Sidechain outliers	100360	4158 (2.50-2.50)
RSRZ outliers	91569	3562 (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 on the lower bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria. 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	1074	<div> <div>3%</div> <div>84%</div> <div>10%</div> <div>6%</div> </div>
2	B	323	<div> <div>7%</div> <div>70%</div> <div>7%</div> <div>23%</div> </div>

2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 10499 atoms, of which 23 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
1	A	1009	Total	C	N	O	S	0	5	0
			8126	5203	1381	1477	65			

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	232	LYS	MET	engineered mutation	UNP P42336
A	233	LYS	LEU	engineered mutation	UNP P42336
A	1069	HIS	-	expression tag	UNP P42336
A	1070	HIS	-	expression tag	UNP P42336
A	1071	HIS	-	expression tag	UNP P42336
A	1072	HIS	-	expression tag	UNP P42336
A	1073	HIS	-	expression tag	UNP P42336
A	1074	HIS	-	expression tag	UNP P42336

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	249	Total	C	N	O	S	0	0	0
			2054	1288	361	400	5			

There are 8 discrepancies between the modelled and reference sequences:

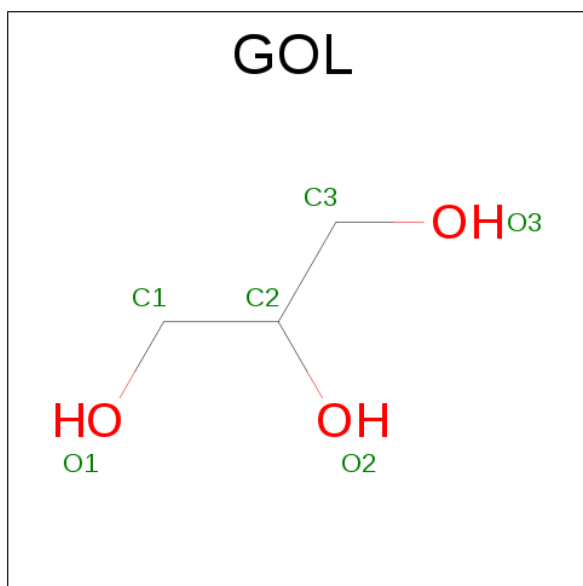
Chain	Residue	Modelled	Actual	Comment	Reference
B	295	MET	-	initiating methionine	UNP P27986
B	296	GLU	-	expression tag	UNP P27986
B	297	TYR	-	expression tag	UNP P27986
B	298	MET	-	expression tag	UNP P27986
B	299	PRO	-	expression tag	UNP P27986
B	300	MET	-	expression tag	UNP P27986
B	301	GLU	-	expression tag	UNP P27986

Continued on next page...

Continued from previous page...

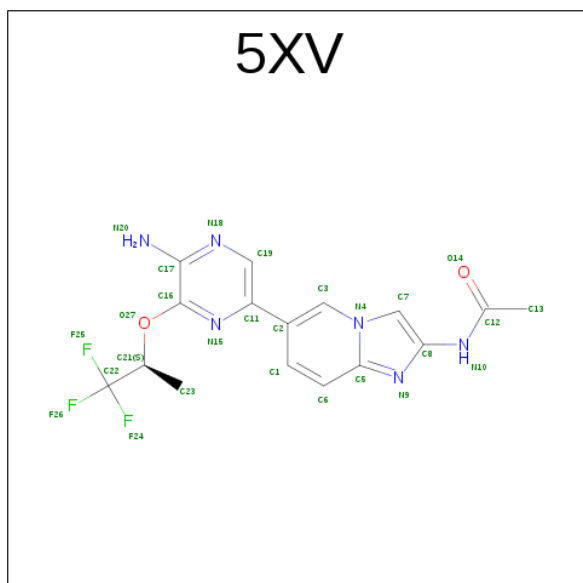
Chain	Residue	Modelled	Actual	Comment	Reference
B	306	TYR	THR	engineered mutation	UNP P27986

- Molecule 3 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	1	Total	C	H	O	0	0
			14	3	8	3		

- Molecule 4 is {N}-[6-[5-azanyl-6-[(2 {S})-1,1,1-tris(fluoranyl)propan-2-yl]oxy-pyrazin-2-yl]imidazo[1,2-a]pyridin-2-yl]ethanamide (three-letter code: 5XV) (formula: $C_{16}H_{15}F_3N_6O_2$).



Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
4	A	1	Total	C	F	H	N	O	0	0
			42	16	3	15	6	2		

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	237	Total	O	0	0
			237	237		
5	B	26	Total	O	0	0
			26	26		

4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	105.20Å 105.32Å 134.39Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	38.56 – 2.50 38.38 – 2.50	Depositor EDS
% Data completeness (in resolution range)	100.0 (38.56-2.50) 100.0 (38.38-2.50)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.65 (at 2.51Å)	Xtriage
Refinement program	BUSTER 2.11.5	Depositor
R, R_{free}	0.194 , 0.245 0.199 , 0.253	Depositor DCC
R_{free} test set	2669 reflections (5.38%)	DCC
Wilson B-factor (Å ²)	44.0	Xtriage
Anisotropy	0.270	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.31 , 57.3	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	0.015 for k,h,-l	Xtriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	10499	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 3.76% 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: 5XV, GOL

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	0.51	0/8323	0.69	0/11279
2	B	0.47	0/2085	0.67	0/2799
All	All	0.50	0/10408	0.68	0/14078

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	8126	0	7969	58	0
2	B	2054	0	1957	13	0
3	A	6	8	8	1	0
4	A	27	15	15	0	0
5	A	237	0	0	0	0
5	B	26	0	0	1	0
All	All	10476	23	9949	68	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 68 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:181:ILE:HD12	1:A:278:MET:HG2	1.71	0.72
1:A:1053:THR:HG23	1:A:1054:LYS:H	1.54	0.71
1:A:873:ASN:HD22	1:A:875:HIS:HB2	1.56	0.70
1:A:25:LEU:HD13	2:B:497:GLN:HG3	1.77	0.67
1:A:492:ILE:HG21	1:A:584:TYR:HD2	1.61	0.66

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	1004/1074 (94%)	959 (96%)	43 (4%)	2 (0%)	52	75
2	B	237/323 (73%)	232 (98%)	5 (2%)	0	100	100
All	All	1241/1397 (89%)	1191 (96%)	48 (4%)	2 (0%)	52	75

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	946	GLY
1	A	412	ARG

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	888/980 (91%)	880 (99%)	8 (1%)	84	95
2	B	215/301 (71%)	211 (98%)	4 (2%)	65	87
All	All	1103/1281 (86%)	1091 (99%)	12 (1%)	80	94

5 of 12 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	901	CYS
1	A	937	PHE
2	B	416	TYR
1	A	526	ASN
2	B	387	ASP

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

Mol	Chain	Res	Type
1	A	873	ASN
2	B	415	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 [i](#)

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	$\# Z > 2$
3	GOL	A	1101	-	5,5,5	0.10	0	5,5,5	0.40	0
4	5XV	A	1102	-	25,29,29	2.85	5 (20%)	25,43,43	1.90	7 (28%)

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	GOL	A	1101	-	-	0/4/4/4	0/0/0/0
4	5XV	A	1102	-	-	0/16/18/18	0/3/3/3

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	A	1102	5XV	C5-N9	-10.86	1.23	1.33
4	A	1102	5XV	C3-N4	-3.43	1.31	1.37
4	A	1102	5XV	C8-N10	-2.72	1.35	1.40
4	A	1102	5XV	C6-C5	4.33	1.47	1.40
4	A	1102	5XV	C1-C2	4.59	1.49	1.39

The worst 5 of 7 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	A	1102	5XV	C1-C6-C5	-3.86	114.98	119.70
4	A	1102	5XV	F26-C22-C21	-3.47	105.64	112.19
4	A	1102	5XV	O14-C12-C13	-2.95	116.63	122.07
4	A	1102	5XV	C19-C11-N15	-2.43	117.31	120.02
4	A	1102	5XV	O14-C12-N10	2.48	126.11	123.08

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	1101	GOL	1	0

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	1009/1074 (93%)	-0.06	37 (3%) 45 50	25, 47, 82, 116	0
2	B	249/323 (77%)	0.44	21 (8%) 14 14	33, 62, 93, 107	0
All	All	1258/1397 (90%)	0.04	58 (4%) 36 41	25, 50, 88, 116	0

The worst 5 of 58 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	242	CYS	6.1
2	B	400	SER	5.5
1	A	231	SER	5.0
1	A	1057	TRP	4.6
1	A	872	PHE	4.4

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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	LLDF	B-factors(\AA^2)	Q<0.9
3	GOL	A	1101	6/6	0.88	0.18	0.68	65,68,70,70	0
4	5XV	A	1102	27/27	0.97	0.14	-0.31	24,33,53,56	0

6.5 Other polymers [i](#)

There are no such residues in this entry.