Full wwPDB X-ray Structure Validation Report ⓘ

Nov 9, 2021 – 07:18 pm GMT

PDB ID	:	7Q52
Title	:	Crystal structure of S/T protein kinase PknG from Mycobacterium tuberculosis in complex with inhibitor L2W
Deposited on	:	2021-11-02
Resolution	:	2.35 Å(reported)

The following versions of software and data (see references ⓘ) were used in the production of this report:

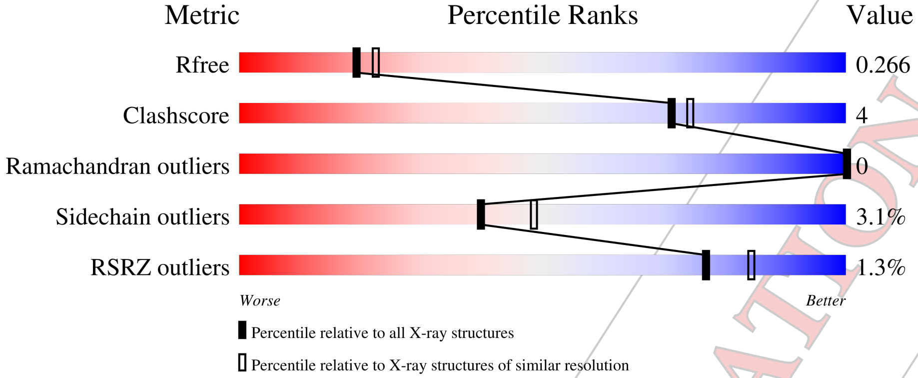
MolProbity	:	4.02b-467
Mogul	:	1.8.4 (270009), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.23.2
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0267
CCP4	:	7.1.010 (Gargrove)
Ideal geometry (proteins)	:	Engh &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

1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

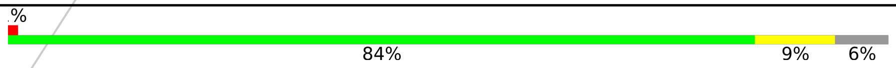
The reported resolution of this entry is 2.35 Å.

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	1164 (2.36-2.36)
Clashscore	141614	1232 (2.36-2.36)
Ramachandran outliers	138981	1211 (2.36-2.36)
Sidechain outliers	138945	1212 (2.36-2.36)
RSRZ outliers	127900	1150 (2.36-2.36)

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	AAA	333	

2 Entry composition ⓘ

There are 5 unique types of molecules in this entry. The entry contains 2464 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 Serine/threonine-protein kinase PknG.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	AAA	312	2410	1532	413	454	11	0	1	0

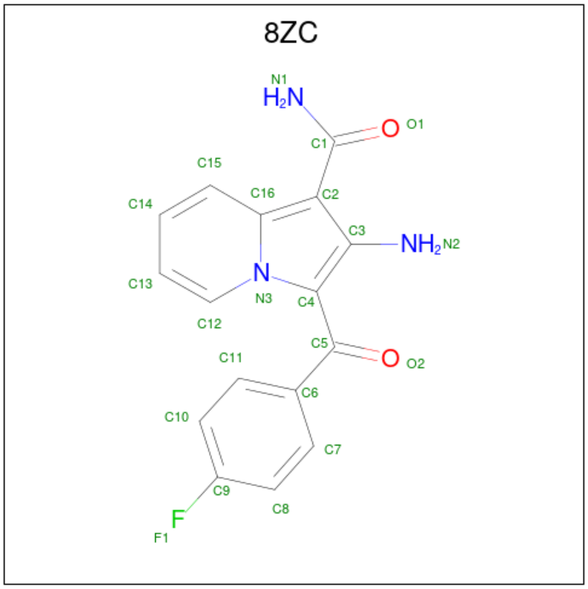
There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
AAA	1	GLY	-	expression tag	UNP P9WI73

- Molecule 2 is FE (III) ION (three-letter code: FE) (formula: Fe).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Fe		
2	AAA	1	1	1	0	0

- Molecule 3 is 2-azanyl-3-(4-fluorophenyl)carbonyl-indolizine-1-carboxamide (three-letter code: 8ZC) (formula: $C_{16}H_{12}FN_3O_2$) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	F	N	O		
3	AAA	1	22	16	1	3	2	0	0

- Molecule 4 is SODIUM ION (three-letter code: NA) (formula: Na).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Na		
4	AAA	2	2	2	0	0

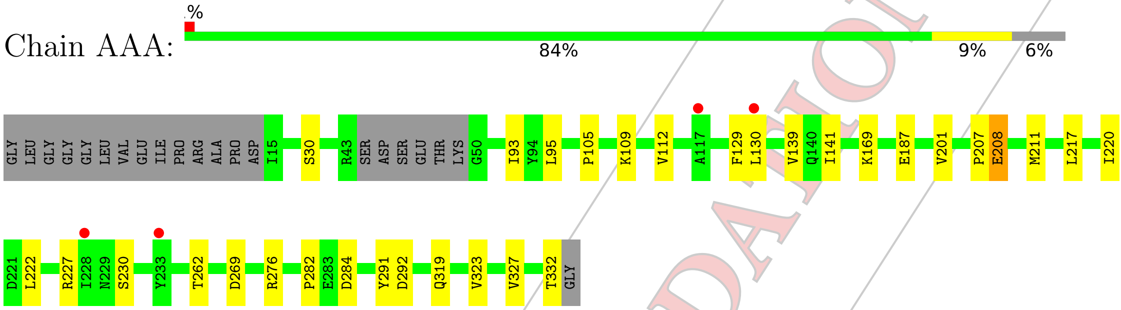
- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
5	AAA	29	29	29	0	0

3 Residue-property plots ⓘ

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: Serine/threonine-protein kinase PknG



4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	77.33Å 37.02Å 106.78Å 90.00° 105.50° 90.00°	Depositor
Resolution (Å)	51.50 – 2.35 51.45 – 2.35	Depositor EDS
% Data completeness (in resolution range)	98.4 (51.50-2.35) 98.4 (51.45-2.35)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.67 (at 2.34Å)	Xtriage
Refinement program	REFMAC 5.8.0267	Depositor
R, R_{free}	0.215 , 0.268 0.213 , 0.266	Depositor DCC
R_{free} test set	627 reflections (5.12%)	wwPDB-VP
Wilson B-factor (Å^2)	67.1	Xtriage
Anisotropy	0.033	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	(Not available) , (Not available)	EDS
L-test for twinning[2]	$< \|L\| > = 0.53$, $< L^2 > = 0.36$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	2464	wwPDB-VP
Average B, all atoms (Å^2)	74.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 10.38% 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: 8ZC, FE, NA

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	AAA	0.31	0/2465	0.67	0/3357

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	AAA	2410	0	2370	17	0
2	AAA	1	0	0	0	0
3	AAA	22	0	0	0	0
4	AAA	2	0	0	0	0
5	AAA	29	0	0	1	0
All	All	2464	0	2370	17	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 4.

All (17) 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:AAA:211:MET:HG3	1:AAA:220:ILE:HG21	1.67	0.76
1:AAA:169:LYS:HD3	1:AAA:208[B]:GLU:HG3	1.82	0.60
1:AAA:141:ILE:HD12	1:AAA:141:ILE:H	1.67	0.59
1:AAA:95:LEU:HD21	1:AAA:105:PRO:HB3	1.86	0.58
1:AAA:141:ILE:HD11	1:AAA:222:LEU:CD1	2.37	0.55
1:AAA:291:TYR:CE1	1:AAA:327:VAL:HG13	2.44	0.52
1:AAA:211:MET:HG3	1:AAA:220:ILE:CG2	2.37	0.51
1:AAA:130:LEU:O	1:AAA:141:ILE:HD13	2.10	0.51
1:AAA:169:LYS:HD3	1:AAA:208[B]:GLU:CG	2.45	0.47
1:AAA:187:GLU:HB3	1:AAA:217:LEU:HD12	1.97	0.46
1:AAA:93:ILE:CD1	1:AAA:109:LYS:HD3	2.45	0.46
1:AAA:292:ASP:OD1	5:AAA:501:HOH:O	2.21	0.44
1:AAA:319:GLN:O	1:AAA:323:VAL:HG23	2.18	0.43
1:AAA:227:ARG:O	1:AAA:230:SER:HB2	2.18	0.43
1:AAA:129:PHE:CD1	1:AAA:129:PHE:C	2.92	0.42
1:AAA:269:ASP:O	1:AAA:282:PRO:HG3	2.21	0.40
1:AAA:207:PRO:HD3	1:AAA:262:THR:HG23	2.03	0.40

There are no symmetry-related clashes.

5.3 Torsion angles

5.3.1 Protein backbone

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	AAA	309/333 (93%)	298 (96%)	11 (4%)	0	100	100

There are no Ramachandran outliers to report.

5.3.2 Protein sidechains

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	AAA	256/273 (94%)	247 (96%)	9 (4%)	36 / 44

All (9) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	AAA	30	SER
1	AAA	112	VAL
1	AAA	139	VAL
1	AAA	201	VAL
1	AAA	208[A]	GLU
1	AAA	208[B]	GLU
1	AAA	276	ARG
1	AAA	284	ASP
1	AAA	332	THR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates ⓘ

There are no monosaccharides in this entry.

5.6 Ligand geometry ⓘ

Of 4 ligands modelled in this entry, 3 are monoatomic - leaving 1 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

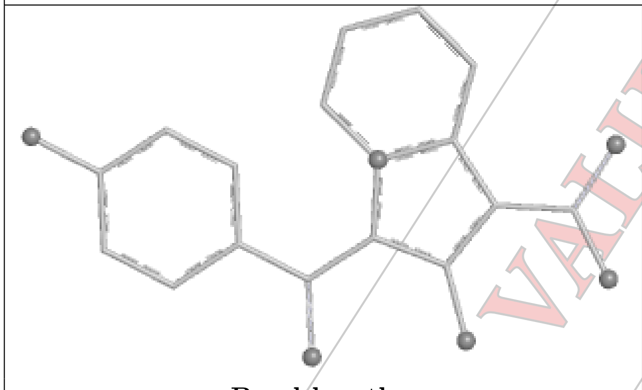
There are no torsion outliers.

There are no ring outliers.

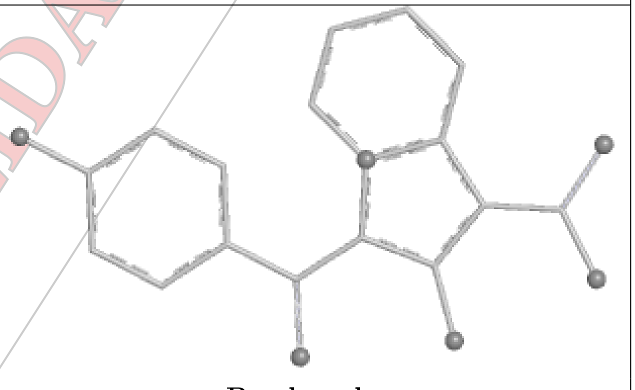
No monomer is involved in short contacts.

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.

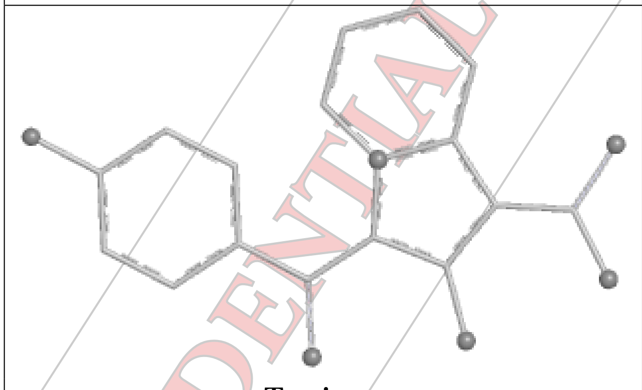
Ligand 8ZC AAA 402



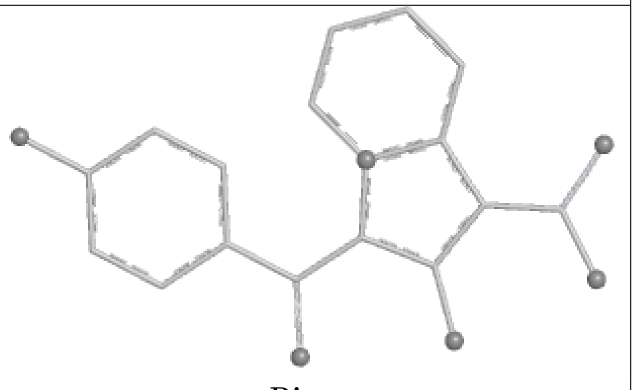
Bond lengths



Bond angles



Torsions



Rings

5.7 Other polymers ⓘ

There are no such residues in this entry.

5.8 Polymer linkage issues ⓘ

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($\AA^2$)	Q<0.9
1	AAA	312/333 (93%)	0.05	4 (1%)	77	84	49, 71, 104, 128	0

All (4) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	AAA	233	TYR	2.9
1	AAA	130	LEU	2.2
1	AAA	117	ALA	2.2
1	AAA	228	ILE	2.0

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

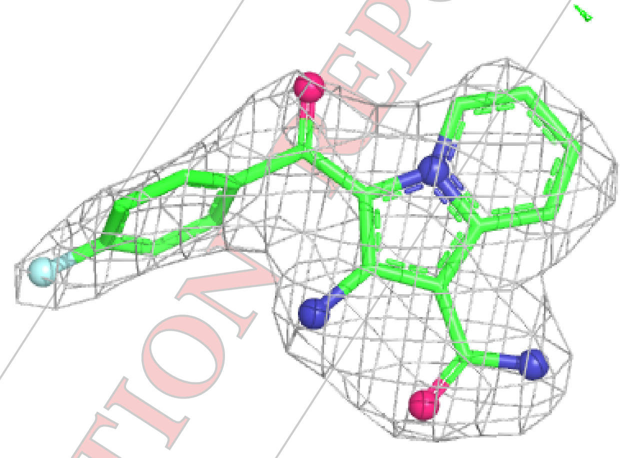
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists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

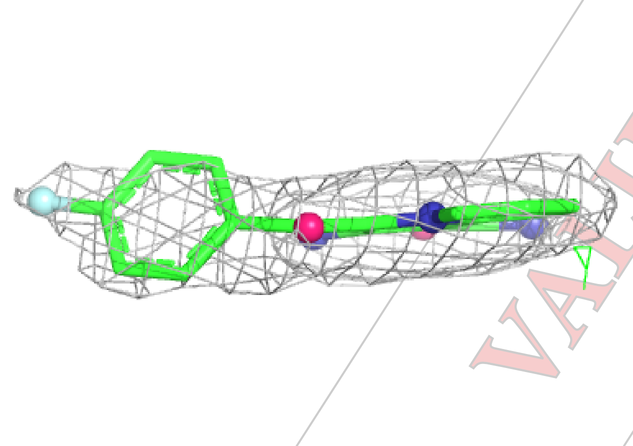
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors($\AA^2$)	Q<0.9
2	FE	AAA	401	1/1	0.91	0.11	75,75,75,75	0
3	8ZC	AAA	402	22/22	0.93	0.17	50,57,77,79	22
4	NA	AAA	403	1/1	0.93	0.13	78,78,78,78	0
4	NA	AAA	404	1/1	0.94	0.41	68,68,68,68	0

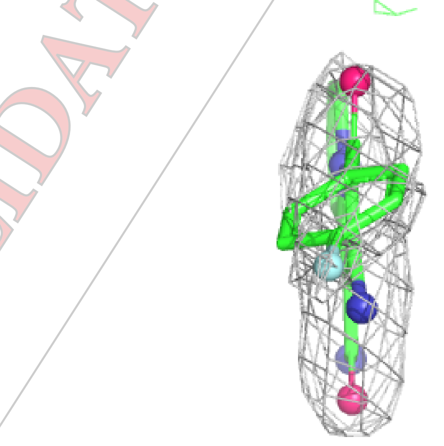
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation to approximate a three-dimensional view.

Electron density around 8ZC AAA 402:

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 ⓘ

There are no such residues in this entry.