



Full wwPDB X-ray Structure Validation Report ⓘ

Mar 8, 2018 – 09:39 pm GMT

PDB ID : 5KBR
Title : Pak1 in complex with 7-azaindole inhibitor
Authors : Ferguson, A.
Deposited on : 2016-06-03
Resolution : 2.36 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 : 1.7.3 (157068), CSD as539be (2018)
Xtriage (Phenix) : 1.13
EDS : trunk30967
Percentile statistics : 20171227.v01 (using entries in the PDB archive December 27th 2017)
Refmac : 5.8.0158
CCP4 : 7.0 (Gargrove)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : trunk30967

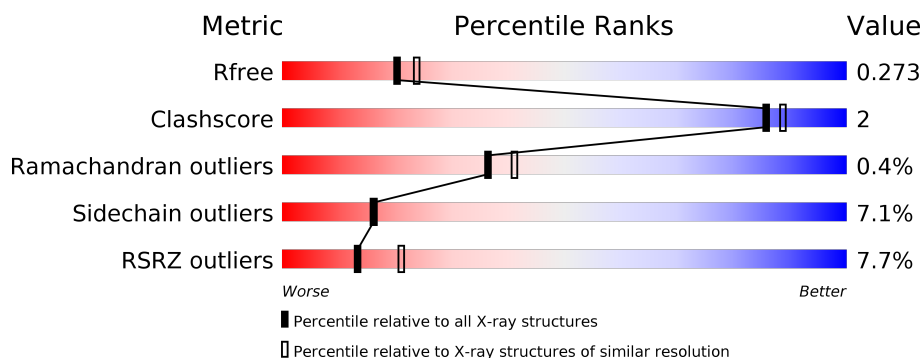
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.36 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}	111664	1015 (2.36-2.36)
Clashscore	122126	1081 (2.36-2.36)
Ramachandran outliers	120053	1066 (2.36-2.36)
Sidechain outliers	120020	1067 (2.36-2.36)
RSRZ outliers	108989	1002 (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 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	289	<div> <div>10%</div> <div> <div></div> <div>75%</div> <div>15%</div> <div>10%</div> </div> </div>
1	B	289	<div> <div>4%</div> <div> <div></div> <div>82%</div> <div>8%</div> <div>10%</div> </div> </div>

2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 4198 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.

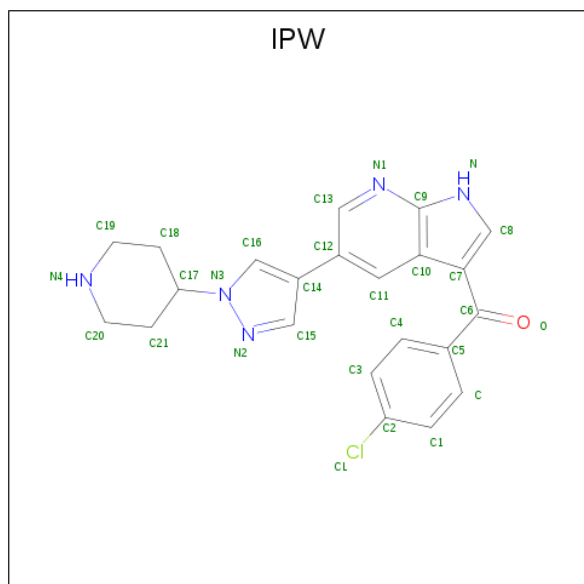
- Molecule 1 is a protein called Serine/threonine-protein kinase PAK 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	259	Total	C	N	O	S	0	0	0
			2030	1293	341	382	14			
1	B	260	Total	C	N	O	S	0	0	0
			2038	1293	342	388	15			

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	389	ASN	ASP	engineered mutation	UNP Q13153
A	423	GLU	THR	engineered mutation	UNP Q13153
A	503	ASP	GLU	engineered mutation	UNP Q13153
B	389	ASN	ASP	engineered mutation	UNP Q13153
B	423	GLU	THR	engineered mutation	UNP Q13153
B	503	ASP	GLU	engineered mutation	UNP Q13153

- Molecule 2 is (4-chlorophenyl)-[5-(1-piperidin-4-ylpyrazol-4-yl)-1 {H}-pyrrolo[2,3-b]pyridin-3-yl]methanone (three-letter code: IPW) (formula: C₂₂H₂₀ClN₅O).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	1	Total	C	Cl	N	O	0	0
			29	22	1	5	1		
2	B	1	Total	C	Cl	N	O	0	0
			29	22	1	5	1		

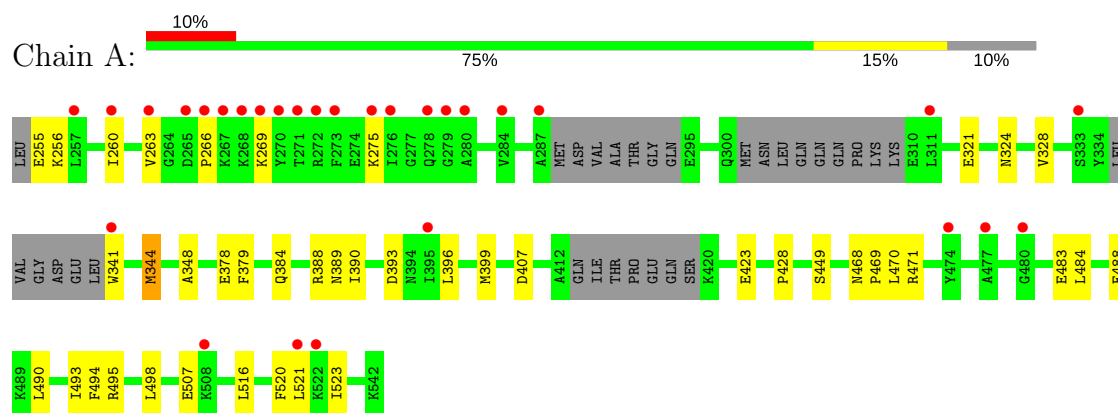
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	28	Total	O	0	0
			28	28		
3	B	44	Total	O	0	0
			44	44		

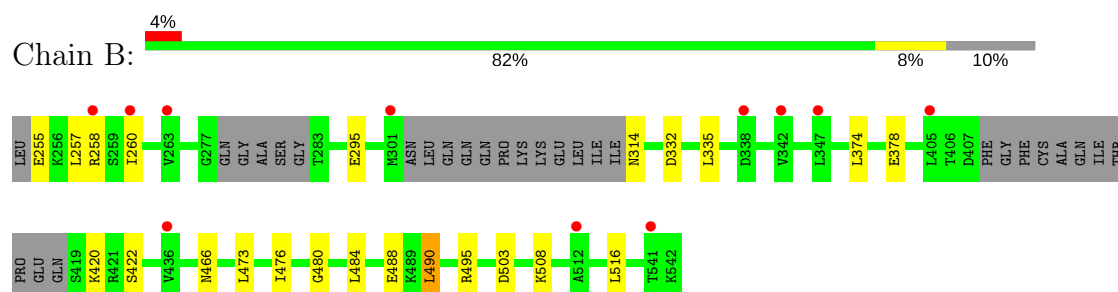
3 Residue-property plots

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: Serine/threonine-protein kinase PAK 1



• Molecule 1: Serine/threonine-protein kinase PAK 1



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	61.82Å 81.83Å 65.91Å 90.00° 107.05° 90.00°	Depositor
Resolution (Å)	63.02 – 2.36 63.02 – 2.36	Depositor EDS
% Data completeness (in resolution range)	99.4 (63.02-2.36) 99.5 (63.02-2.36)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.87 (at 2.37Å)	Xtriage
Refinement program	BUSTER 2.11.7	Depositor
R, R_{free}	0.195 , 0.255 0.211 , 0.273	Depositor DCC
R_{free} test set	1309 reflections (5.08%)	wwPDB-VP
Wilson B-factor (Å ²)	57.1	Xtriage
Anisotropy	0.358	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.33 , 67.8	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	4198	wwPDB-VP
Average B, all atoms (Å ²)	91.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.45% 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

5.1 Standard geometry

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

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	0.53	0/2062	0.70	0/2782
1	B	0.53	0/2069	0.72	0/2794
All	All	0.53	0/4131	0.71	0/5576

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	2030	0	2065	15	0
1	B	2038	0	2075	5	0
2	A	29	0	0	0	0
2	B	29	0	0	0	0
3	A	28	0	0	1	0
3	B	44	0	0	1	0
All	All	4198	0	4140	20	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 2.

All (20) 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:328:VAL:HG21	1:A:396:LEU:HD12	1.77	0.66
1:A:520:PHE:O	1:A:523:ILE:HG13	2.08	0.54
1:A:328:VAL:HG12	1:A:344:MET:HG2	1.90	0.53
1:A:266:PRO:HB2	1:A:341:TRP:CE3	2.47	0.50
1:B:476:ILE:O	1:B:480:GLY:HA2	2.13	0.47
1:A:428:PRO:HG2	1:A:469:PRO:HB3	1.95	0.47
1:A:321:GLU:HG3	3:A:727:HOH:O	2.14	0.47
1:A:493:ILE:HD13	1:A:520:PHE:HA	1.98	0.46
1:B:490:LEU:HD12	1:B:495:ARG:HD3	1.98	0.46
1:A:256:LYS:O	1:A:260:ILE:HG12	2.16	0.45
1:A:516:LEU:HA	1:A:521:LEU:HD12	1.98	0.45
1:A:324:ASN:HB2	1:A:379:PHE:CD2	2.52	0.44
1:B:374:LEU:HB3	1:B:516:LEU:HG	2.00	0.43
1:B:495:ARG:HG3	3:B:740:HOH:O	2.19	0.42
1:A:348:ALA:HB3	1:A:399:MET:HG2	2.02	0.42
1:A:470:LEU:HD23	1:A:470:LEU:HA	1.91	0.42
1:A:390:ILE:HB	1:A:449:SER:HB2	2.02	0.42
1:A:388:ARG:HH12	1:A:423:GLU:HG2	1.84	0.42
1:A:490:LEU:HD22	1:A:494:PHE:CD2	2.56	0.40
1:B:255:GLU:HB2	1:B:257:LEU:HD23	2.02	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	249/289 (86%)	239 (96%)	8 (3%)	2 (1%)	21	22
1	B	252/289 (87%)	244 (97%)	8 (3%)	0	100	100
All	All	501/578 (87%)	483 (96%)	16 (3%)	2 (0%)	36	41

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	275	LYS
1	A	407	ASP

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	A	223/250 (89%)	207 (93%)	16 (7%)	16	15
1	B	226/250 (90%)	210 (93%)	16 (7%)	16	16
All	All	449/500 (90%)	417 (93%)	32 (7%)	16	16

All (32) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	255	GLU
1	A	263	VAL
1	A	269	LYS
1	A	344	MET
1	A	378	GLU
1	A	384	GLN
1	A	389	ASN
1	A	393	ASP
1	A	468	ASN
1	A	471	ARG
1	A	483	GLU
1	A	484	LEU
1	A	488	GLU
1	A	495	ARG
1	A	498	LEU
1	A	507	GLU
1	B	258	ARG
1	B	260	ILE
1	B	295	GLU
1	B	314	ASN
1	B	332	ASP

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	335	LEU
1	B	378	GLU
1	B	420	LYS
1	B	422	SER
1	B	466	ASN
1	B	473	LEU
1	B	484	LEU
1	B	488	GLU
1	B	490	LEU
1	B	503	ASP
1	B	508	LYS

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	383	ASN
1	A	389	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 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	$\# Z > 2$
2	IPW	A	601	-	30,33,33	0.90	2 (6%)	30,47,47	0.88	0
2	IPW	B	601	-	30,33,33	1.00	2 (6%)	30,47,47	1.42	2 (6%)

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
2	IPW	A	601	-	-	0/8/24/24	0/5/5/5
2	IPW	B	601	-	-	0/8/24/24	0/5/5/5

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	601	IPW	C7-C6	-2.83	1.46	1.50
2	A	601	IPW	C7-C6	-2.13	1.47	1.50
2	A	601	IPW	C7-C10	2.83	1.45	1.42
2	B	601	IPW	C7-C10	3.17	1.45	1.42

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	601	IPW	C7-C6-C5	-5.16	114.27	119.33
2	B	601	IPW	C13-C12-C14	-2.55	117.19	121.80

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

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(Å ²)	Q<0.9
1	A	259/289 (89%)	0.93	29 (11%) 5 8	53, 93, 156, 184	0
1	B	260/289 (89%)	0.47	11 (4%) 36 49	52, 79, 121, 159	0
All	All	519/578 (89%)	0.70	40 (7%) 13 21	52, 84, 142, 184	0

All (40) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	266	PRO	9.9
1	A	273	PHE	7.0
1	A	276	ILE	7.0
1	A	272	ARG	6.5
1	A	265	ASP	6.2
1	A	271	THR	5.7
1	A	270	TYR	5.7
1	A	474	TYR	4.6
1	A	268	LYS	4.1
1	A	477	ALA	4.1
1	B	258	ARG	4.0
1	A	263	VAL	3.9
1	A	341	TRP	3.8
1	A	333	SER	3.6
1	A	278	GLN	3.3
1	A	284	VAL	3.2
1	B	342	VAL	3.1
1	B	338	ASP	2.9
1	A	267	LYS	2.9
1	A	311	LEU	2.9
1	A	275	LYS	2.9
1	A	287	ALA	2.8
1	A	257	LEU	2.8
1	A	260	ILE	2.7

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	260	ILE	2.7
1	A	508	LYS	2.7
1	A	522	LYS	2.6
1	A	280	ALA	2.6
1	A	480	GLY	2.6
1	A	269	LYS	2.5
1	B	263	VAL	2.5
1	B	405	LEU	2.5
1	B	301	MET	2.4
1	B	541	THR	2.3
1	A	395	ILE	2.3
1	A	279	GLY	2.2
1	B	512	ALA	2.2
1	B	347	LEU	2.1
1	A	521	LEU	2.1
1	B	436	VAL	2.1

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
2	IPW	A	601	29/29	0.85	0.27	97,106,112,113	0
2	IPW	B	601	29/29	0.86	0.31	102,106,127,130	0

6.5 Other polymers [i](#)

There are no such residues in this entry.