



Full wwPDB X-ray Structure Validation Report ⓘ

Feb 1, 2016 – 01:51 PM GMT

PDB ID : 3V8W
Title : Crystal Structure of Interleukin-2 Inducible T-cell Kinase Itk Catalytic Domain with Thienopyrazolylindole Inhibitor 469
Authors : McLean, L.R.; Zhang, Y.
Deposited on : 2011-12-23
Resolution : 2.27 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
<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 (RC4), CSD as536be (2015)
Xtriage (Phenix) : 1.9-1692
EDS : rb-20026688
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) : trunk26865

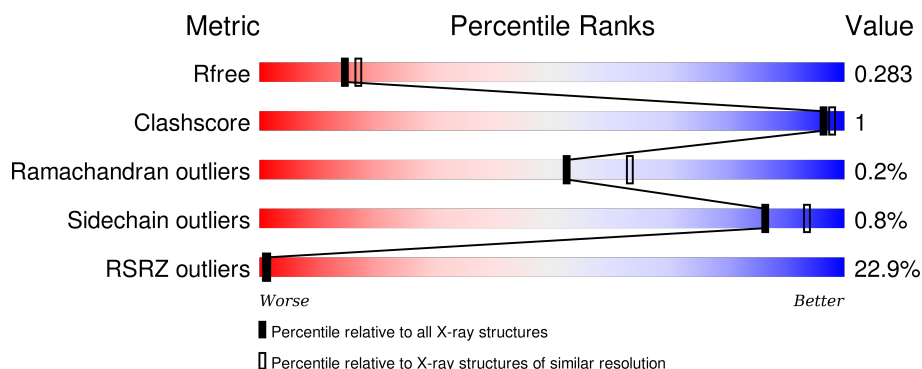
1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.27 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	5193 (2.30-2.26)
Clashscore	102246	5929 (2.30-2.26)
Ramachandran outliers	100387	5851 (2.30-2.26)
Sidechain outliers	100360	5850 (2.30-2.26)
RSRZ outliers	91569	5204 (2.30-2.26)

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	266	<div> <div>22%</div> <div>87%</div> <div>5%</div> <div>9%</div> </div>
1	B	266	<div> <div>20%</div> <div>87%</div> <div>•</div> <div>9%</div> </div>

2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 3792 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.

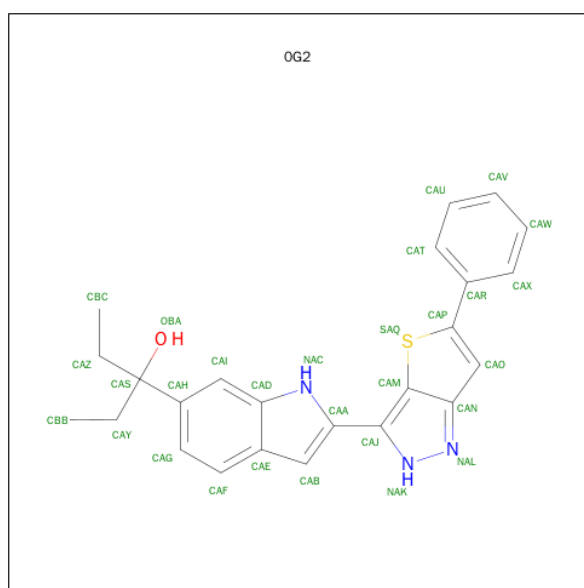
- Molecule 1 is a protein called Tyrosine-protein kinase ITK/TSK.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	243	Total	C	N	O	S	0	0	0
			1841	1179	311	336	15			
1	B	242	Total	C	N	O	S	0	0	0
			1839	1181	307	336	15			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	355	GLY	-	EXPRESSION TAG	UNP Q08881
A	356	SER	-	EXPRESSION TAG	UNP Q08881
B	355	GLY	-	EXPRESSION TAG	UNP Q08881
B	356	SER	-	EXPRESSION TAG	UNP Q08881

- Molecule 2 is 3-[2-(5-PHENYL-2H-THIENO[3,2-C]PYRAZOL-3-YL)-1H-INDOL-6-YL]PENTAN-3-OL (three-letter code: 0G2) (formula: C₂₄H₂₃N₃OS).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	1	Total	C	N	O	S	0	0
			29	24	3	1	1		
2	B	1	Total	C	N	O	S	0	0
			29	24	3	1	1		

- Molecule 3 is SULFATE ION (three-letter code: SO4) (formula: O₄S).



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

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	22	Total	O	0	0
			22	22		
4	B	27	Total	O	0	0
			27	27		

4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α , β , γ	146.08 Å 69.29 Å 68.58 Å 90.00° 108.04° 90.00°	Depositor
Resolution (Å)	38.50 – 2.27 38.50 – 2.27	Depositor EDS
% Data completeness (in resolution range)	99.7 (38.50-2.27) 99.7 (38.50-2.27)	Depositor EDS
R_{merge}	0.18	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.12 (at 2.27 Å)	Xtriage
Refinement program	BUSTER 2.9.1	Depositor
R, R_{free}	0.251 , 0.268 0.271 , 0.283	Depositor DCC
R_{free} test set	1521 reflections (5.31%)	DCC
Wilson B-factor (Å ²)	45.7	Xtriage
Anisotropy	0.297	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.36 , 38.6	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtriage
Outliers	1 of 30182 reflections (0.003%)	Xtriage
F_o, F_c correlation	0.92	EDS
Total number of atoms	3792	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 5.95% of the height of the origin peak. No significant pseudotranslation is detected.*

¹ Intensities estimated from amplitudes.

² Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ for acentric reflections are 0.5, 0.375 respectively for untwinned datasets, and 0.333, 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: 0G2, SO4

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	0.38	0/1879	0.56	0/2542
1	B	0.38	0/1879	0.55	0/2546
All	All	0.38	0/3758	0.56	0/5088

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	1841	0	1739	5	0
1	B	1839	0	1754	5	0
2	A	29	0	23	1	0
2	B	29	0	23	1	0
3	B	5	0	0	0	0
4	A	22	0	0	0	0
4	B	27	0	0	0	0
All	All	3792	0	3539	10	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 1.

All (10) 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:B:438:MET:H	2:B:702:OG2:H17	1.43	0.67
1:B:393:ILE:HG12	1:B:431:ILE:HG13	1.78	0.65
1:A:438:MET:H	2:A:701:OG2:H17	1.48	0.59
1:A:468:GLU:HG3	1:A:609:LEU:HD11	1.91	0.53
1:B:468:GLU:HG3	1:B:609:LEU:HD11	1.96	0.47
1:B:419:VAL:HG21	1:B:489:LEU:HD12	1.96	0.46
1:A:528:GLU:HB3	1:A:534:ARG:HB2	2.01	0.43
1:A:421:LEU:HD11	1:A:433:LEU:HB3	2.01	0.42
1:B:608:LEU:HA	1:B:611:GLN:HE21	1.84	0.41
1:A:608:LEU:HA	1:A:611:GLN:HE21	1.85	0.41

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	233/266 (88%)	226 (97%)	6 (3%)	1 (0%)	39	47
1	B	234/266 (88%)	228 (97%)	6 (3%)	0	100	100
All	All	467/532 (88%)	454 (97%)	12 (3%)	1 (0%)	52	63

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	398	MET

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	185/233 (79%)	183 (99%)	2 (1%)	80	89
1	B	190/233 (82%)	189 (100%)	1 (0%)	92	96
All	All	375/466 (80%)	372 (99%)	3 (1%)	86	93

All (3) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	500	ASP
1	A	590	ILE
1	B	421	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	592	ASN
1	A	611	GLN
1	B	611	GLN

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 carbohydrates in this entry.

5.6 Ligand geometry ⓘ

3 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	0G2	A	701	-	25,33,33	2.45	5 (20%)	23,49,49	1.04	2 (8%)
3	SO4	B	701	-	4,4,4	0.19	0	6,6,6	0.07	0
2	0G2	B	702	-	25,33,33	2.46	6 (24%)	23,49,49	1.09	3 (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
2	0G2	A	701	-	-	0/16/20/20	0/5/5/5
3	SO4	B	701	-	-	0/0/0/0	0/0/0/0
2	0G2	B	702	-	-	0/16/20/20	0/5/5/5

All (11) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	701	0G2	CAJ-CAA	-7.81	1.40	1.49
2	B	702	0G2	CAJ-CAA	-7.74	1.40	1.49
2	A	701	0G2	NAL-NAK	-6.84	1.24	1.37
2	B	702	0G2	NAL-NAK	-6.77	1.24	1.37
2	B	702	0G2	CAB-CAA	-3.32	1.33	1.40
2	A	701	0G2	CAB-CAA	-3.31	1.34	1.40
2	B	702	0G2	CAP-SAQ	-2.77	1.68	1.72
2	A	701	0G2	CAP-SAQ	-2.44	1.69	1.72
2	B	702	0G2	CAB-CAE	-2.18	1.33	1.41
2	A	701	0G2	CAB-CAE	-2.16	1.33	1.41
2	B	702	0G2	CAS-CAH	2.14	1.55	1.53

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	702	0G2	CAO-CAP-SAQ	-3.13	107.40	110.75
2	A	701	0G2	CAO-CAP-SAQ	-2.91	107.64	110.75

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	702	0G2	CAB-CAA-CAJ	-2.02	126.19	129.24
2	A	701	0G2	CAA-NAC-CAD	2.07	108.75	104.47
2	B	702	0G2	CAA-NAC-CAD	2.08	108.75	104.47

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	701	0G2	1	0
2	B	702	0G2	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

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	243/266 (91%)	1.33	59 (24%)  	36, 52, 77, 88	0
1	B	242/266 (90%)	1.26	52 (21%)  	35, 51, 74, 92	0
All	All	485/532 (91%)	1.30	111 (22%)  	35, 52, 76, 92	0

All (111) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	429	ALA	8.9
1	A	393	ILE	6.7
1	A	397	ALA	6.5
1	A	398	MET	5.9
1	B	426	LEU	5.3
1	B	572	SER	5.2
1	A	532	PHE	5.1
1	B	402	ASP	4.9
1	A	396	GLY	4.9
1	B	427	GLU	4.7
1	B	393	ILE	4.6
1	B	400	GLU	4.6
1	A	426	LEU	4.3
1	B	618	SER	4.2
1	B	575	PHE	4.1
1	B	573	THR	4.1
1	B	355	GLY	4.1
1	B	431	ILE	4.0
1	B	619	GLY	3.9
1	B	430	PRO	3.8
1	A	403	PHE	3.7
1	A	575	PHE	3.7
1	B	418	LEU	3.7
1	A	421	LEU	3.7

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	574	GLY	3.7
1	B	601	ASP	3.6
1	A	418	LEU	3.6
1	B	371	SER	3.6
1	A	416	PRO	3.6
1	A	502	GLY	3.5
1	B	398	MET	3.5
1	B	495	VAL	3.5
1	B	497	LYS	3.5
1	A	497	LYS	3.4
1	A	534	ARG	3.4
1	B	419	VAL	3.4
1	A	428	GLN	3.4
1	A	404	ILE	3.4
1	B	498	VAL	3.3
1	A	427	GLU	3.3
1	A	498	VAL	3.3
1	B	428	GLN	3.3
1	B	421	LEU	3.2
1	A	431	ILE	3.2
1	A	479	ILE	3.2
1	A	399	SER	3.2
1	A	618	SER	3.1
1	B	561	ASN	3.0
1	A	586	HIS	3.0
1	B	415	HIS	3.0
1	A	355	GLY	2.9
1	A	576	ARG	2.9
1	B	403	PHE	2.9
1	A	616	ALA	2.8
1	B	420	GLN	2.8
1	A	535	TYR	2.8
1	B	399	SER	2.8
1	A	429	ALA	2.8
1	A	577	LEU	2.8
1	A	430	PRO	2.8
1	B	574	GLY	2.7
1	B	496	ILE	2.7
1	B	530	PHE	2.7
1	B	532	PHE	2.6
1	A	360	PRO	2.6
1	A	578	TYR	2.6

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	551	VAL	2.6
1	B	404	ILE	2.6
1	A	466	VAL	2.5
1	B	394	ARG	2.5
1	A	601	ASP	2.5
1	A	409	VAL	2.5
1	B	564	ASN	2.5
1	A	402	ASP	2.5
1	B	425	CYS	2.4
1	B	616	ALA	2.4
1	A	417	LYS	2.4
1	A	419	VAL	2.4
1	B	417	LYS	2.4
1	B	372	GLY	2.4
1	B	567	VAL	2.4
1	B	521	PRO	2.4
1	A	420	GLN	2.3
1	A	598	ARG	2.3
1	A	562	ARG	2.3
1	B	422	TYR	2.3
1	A	579	LYS	2.3
1	A	375	GLY	2.2
1	B	360	PRO	2.2
1	A	496	ILE	2.2
1	B	477	CYS	2.2
1	A	619	GLY	2.2
1	A	617	GLU	2.2
1	A	394	ARG	2.2
1	A	536	SER	2.2
1	A	372	GLY	2.2
1	B	416	PRO	2.2
1	B	566	GLU	2.2
1	B	392	THR	2.2
1	B	442	CYS	2.2
1	A	529	VAL	2.1
1	A	489	LEU	2.1
1	A	490	VAL	2.1
1	A	600	GLU	2.1
1	A	483	LEU	2.1
1	B	406	GLU	2.1
1	A	495	VAL	2.1
1	B	522	VAL	2.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	548	MET	2.0
1	B	413	LEU	2.0
1	A	415	HIS	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. 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
2	0G2	B	702	29/29	0.97	0.16	-1.00	24,41,71,80	0
2	0G2	A	701	29/29	0.95	0.16	-1.15	25,39,75,121	0
3	SO4	B	701	5/5	0.96	0.13	-2.53	52,56,57,57	0

6.5 Other polymers [i](#)

There are no such residues in this entry.