



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 9, 2017 – 10:59 am GMT

PDB ID : 5TQ7
Title : Design and Synthesis of a pan-JAK Kinase Inhibitor Clinical Candidate (PF-06263276) Suitable for Inhaled and Topical Delivery for the Treatment of Inflammatory Diseases of the Lungs and Skin
Authors : Chrencik, J.; Jones, P.
Deposited on : 2016-10-23
Resolution : 2.10 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 : 1.7.2 (RC1), CSD as538be (2017)
Xtriage (Phenix) : 1.9-1692
EDS : recal28906
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)
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) : recal28906

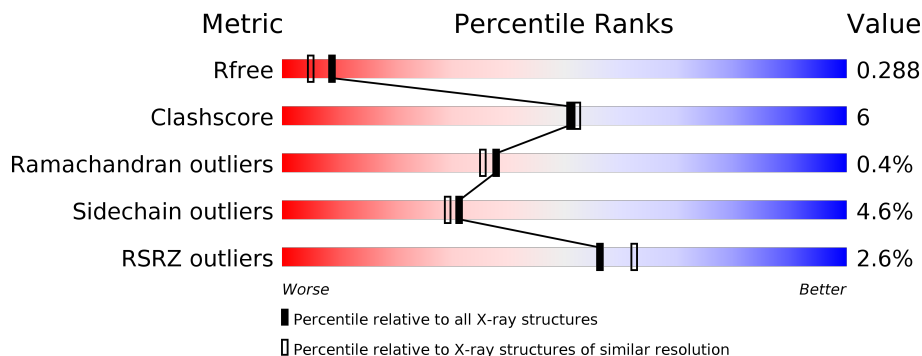
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.10 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}	100719	4243 (2.10-2.10)
Clashscore	112137	4788 (2.10-2.10)
Ramachandran outliers	110173	4740 (2.10-2.10)
Sidechain outliers	110143	4741 (2.10-2.10)
RSRZ outliers	101464	4275 (2.10-2.10)

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	304	<div> <div>2%</div> <div> <div></div> <div>79%</div> <div>11%</div> <div>•</div> <div>9%</div> </div> </div>
1	B	304	<div> <div>3%</div> <div> <div></div> <div>69%</div> <div>17%</div> <div>•</div> <div>11%</div> </div> </div>

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 4761 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 JAK2.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	277	Total	C	N	O	P	S	0	0	0
			2257	1441	379	422	2	13			
1	B	271	Total	C	N	O	P	S	0	0	0
			2218	1419	376	408	2	13			

There are 22 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	829	MET	-	initiating methionine	UNP O60674
A	830	GLY	-	expression tag	UNP O60674
A	831	HIS	-	expression tag	UNP O60674
A	832	HIS	-	expression tag	UNP O60674
A	833	HIS	-	expression tag	UNP O60674
A	834	HIS	-	expression tag	UNP O60674
A	835	HIS	-	expression tag	UNP O60674
A	836	HIS	-	expression tag	UNP O60674
A	1073	SER	MET	engineered mutation	UNP O60674
A	1076	THR	PHE	engineered mutation	UNP O60674
A	1126	VAL	ILE	engineered mutation	UNP O60674
B	829	MET	-	initiating methionine	UNP O60674
B	830	GLY	-	expression tag	UNP O60674
B	831	HIS	-	expression tag	UNP O60674
B	832	HIS	-	expression tag	UNP O60674
B	833	HIS	-	expression tag	UNP O60674
B	834	HIS	-	expression tag	UNP O60674
B	835	HIS	-	expression tag	UNP O60674
B	836	HIS	-	expression tag	UNP O60674
B	1073	SER	MET	engineered mutation	UNP O60674
B	1076	THR	PHE	engineered mutation	UNP O60674
B	1126	VAL	ILE	engineered mutation	UNP O60674

- Molecule 2 is {(3R,4R)-4-methyl-3-[methyl(7H-pyrrolo[2,3-d]pyrimidin-4-yl)amino]piperidin-1-yl}[(3R)-3-(phenylsulfonyl)pyrrolidin-1-yl]methanone (three-letter code: 7GT) (formula:

7GT

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	1	Total 34	C 24	N 6	O 3	S 1	0	0
2	B	1	Total 34	C 24	N 6	O 3	S 1	0	0

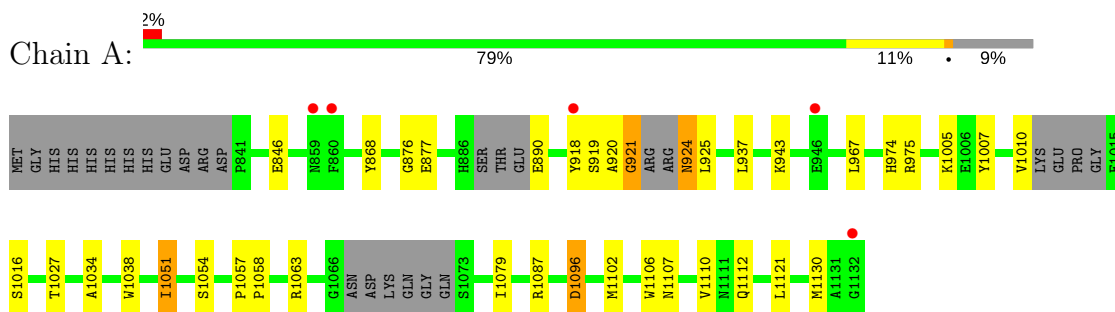
- Molecule 3 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	103	Total O 103 103	0	0
3	B	115	Total O 115 115	0	0

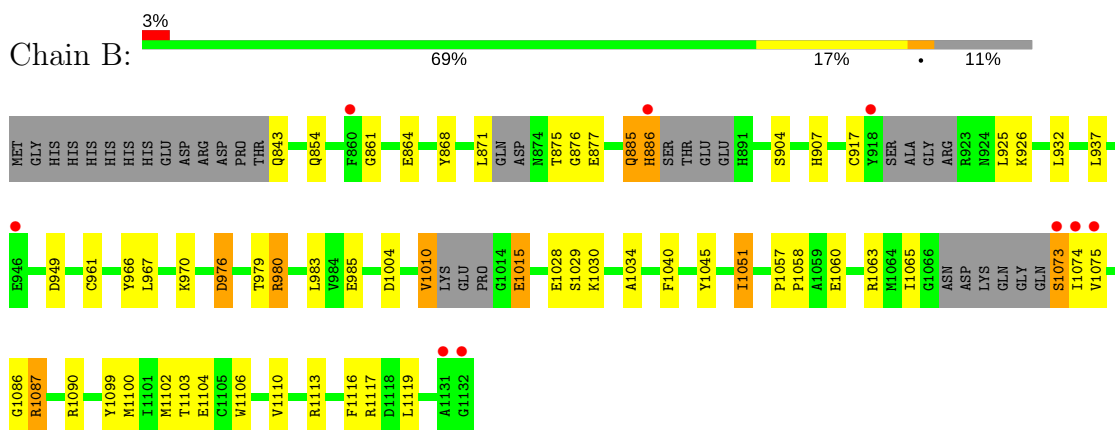
3 Residue-property plots [i](#)

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: Tyrosine-protein kinase JAK2



• Molecule 1: Tyrosine-protein kinase JAK2



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	37.06Å 185.11Å 52.49Å 90.00° 110.64° 90.00°	Depositor
Resolution (Å)	33.68 – 2.10 33.68 – 2.10	Depositor EDS
% Data completeness (in resolution range)	75.5 (33.68-2.10) 75.4 (33.68-2.10)	Depositor EDS
R_{merge}	0.20	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.92 (at 2.10Å)	Xtriage
Refinement program	REFMAC 5.8.0135	Depositor
R, R_{free}	0.226 , 0.286 0.231 , 0.288	Depositor DCC
R_{free} test set	1468 reflections (5.33%)	DCC
Wilson B-factor (Å ²)	23.1	Xtriage
Anisotropy	0.108	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.35 , 17.1	EDS
L-test for twinning ²	$\langle L \rangle = 0.45$, $\langle L^2 \rangle = 0.27$	Xtriage
Estimated twinning fraction	0.170 for h,-k,-h-l	Xtriage
F_o, F_c correlation	0.92	EDS
Total number of atoms	4761	wwPDB-VP
Average B, all atoms (Å ²)	26.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.20% 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: 7GT, PTR

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	0.65	0/2269	0.84	4/3054 (0.1%)
1	B	0.67	0/2228	0.87	4/2995 (0.1%)
All	All	0.66	0/4497	0.86	8/6049 (0.1%)

There are no bond length outliers.

The worst 5 of 8 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	976	ASP	CB-CG-OD2	-6.14	112.77	118.30
1	B	1113	ARG	NE-CZ-NH2	-6.14	117.23	120.30
1	A	918	TYR	CB-CA-C	-5.90	98.61	110.40
1	A	1096	ASP	CB-CG-OD1	5.86	123.57	118.30
1	A	921	GLY	N-CA-C	-5.54	99.26	113.10

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	2257	0	2205	25	0
1	B	2218	0	2181	32	0
2	A	34	0	0	1	0
2	B	34	0	0	1	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	A	103	0	0	1	0
3	B	115	0	0	4	0
All	All	4761	0	4386	58	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 6.

The worst 5 of 58 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:919:SER:CB	1:A:924:ASN:HD21	1.82	0.91
1:A:919:SER:HB2	1:A:924:ASN:ND2	1.91	0.85
1:A:919:SER:HB2	1:A:924:ASN:HD21	1.44	0.79
1:A:1034:ALA:CB	1:A:1110:VAL:HG13	2.15	0.76
1:B:1051:ILE:O	1:B:1051:ILE:HG22	1.93	0.69

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	265/304 (87%)	255 (96%)	9 (3%)	1 (0%)	38	35
1	B	257/304 (84%)	248 (96%)	8 (3%)	1 (0%)	38	35
All	All	522/608 (86%)	503 (96%)	17 (3%)	2 (0%)	38	35

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	1051	ILE
1	B	1051	ILE

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	243/275 (88%)	235 (97%)	8 (3%)	43	45
1	B	238/275 (86%)	224 (94%)	14 (6%)	23	19
All	All	481/550 (88%)	459 (95%)	22 (5%)	31	29

5 of 22 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	875	THR
1	B	886	HIS
1	B	1073	SER
1	B	877	GLU
1	B	885	GLN

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	854	GLN
1	A	924	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](#)

4 non-standard protein/DNA/RNA residues 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	$\# Z > 2$
1	PTR	A	1007	1	15,16,17	0.82	0	19,22,24	1.05	1 (5%)
1	PTR	A	1008	1	15,16,17	0.54	0	19,22,24	1.24	2 (10%)
1	PTR	B	1007	1	15,16,17	0.86	0	19,22,24	1.71	6 (31%)
1	PTR	B	1008	1	15,16,17	0.66	0	19,22,24	1.18	2 (10%)

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
1	PTR	A	1007	1	-	0/9/11/13	0/1/1/1
1	PTR	A	1008	1	-	0/9/11/13	0/1/1/1
1	PTR	B	1007	1	-	0/9/11/13	0/1/1/1
1	PTR	B	1008	1	-	0/9/11/13	0/1/1/1

There are no bond length outliers.

The worst 5 of 11 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	1007	PTR	CG-CB-CA	-4.16	105.90	114.29
1	A	1008	PTR	CG-CB-CA	-2.69	108.87	114.29
1	A	1008	PTR	CB-CA-C	-2.65	106.31	111.41
1	B	1007	PTR	O2P-P-OH	-2.54	96.57	105.63
1	A	1007	PTR	O3P-P-OH	-2.12	98.08	105.63

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
1	A	1007	PTR	1	0

5.5 Carbohydrates

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	$\# Z > 2$
2	7GT	A	4000	-	34,38,38	0.54	0	38,56,56	0.75	1 (2%)
2	7GT	B	4000	-	34,38,38	0.59	0	38,56,56	1.14	2 (5%)

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	7GT	A	4000	-	-	0/28/50/50	0/5/5/5
2	7GT	B	4000	-	-	0/28/50/50	0/5/5/5

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	4000	7GT	N18-C10-N8	-3.16	113.50	116.99
2	A	4000	7GT	C6-N5-C19	-2.13	112.08	120.99
2	B	4000	7GT	C10-N8-C7	5.22	126.74	121.17

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	4000	7GT	1	0
2	B	4000	7GT	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	275/304 (90%)	-0.17	5 (1%) 69 73	12, 24, 46, 60	0
1	B	269/304 (88%)	-0.14	9 (3%) 47 54	12, 23, 44, 56	0
All	All	544/608 (89%)	-0.16	14 (2%) 56 62	12, 23, 46, 60	0

The worst 5 of 14 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	1073	SER	5.7
1	A	1132	GLY	5.1
1	B	918	TYR	4.6
1	A	918	TYR	4.4
1	A	859	ASN	3.8

6.2 Non-standard residues in protein, DNA, RNA chains [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(Å ²)	Q<0.9
1	PTR	A	1007	16/17	0.91	0.12	-	25,37,51,54	0
1	PTR	B	1007	16/17	0.94	0.10	-	23,29,50,52	0
1	PTR	B	1008	16/17	0.95	0.10	-	22,34,45,47	0
1	PTR	A	1008	16/17	0.91	0.11	-	26,39,57,60	0

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	7GT	A	4000	34/34	0.90	0.17	0.93	19,27,65,72	0
2	7GT	B	4000	34/34	0.93	0.13	0.49	21,26,57,58	0

6.5 Other polymers [i](#)

There are no such residues in this entry.