



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 08:44 AM GMT

PDB ID : 3FUP  
Title : Crystal structures of JAK1 and JAK2 inhibitor complexes  
Authors : Williams, N.K.; Bamert, R.S.; Patel, O.; Fantino, E.; Rossjohn, J.; Lucet, I.S.  
Deposited on : 2009-01-14  
Resolution : 2.40 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](mailto: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

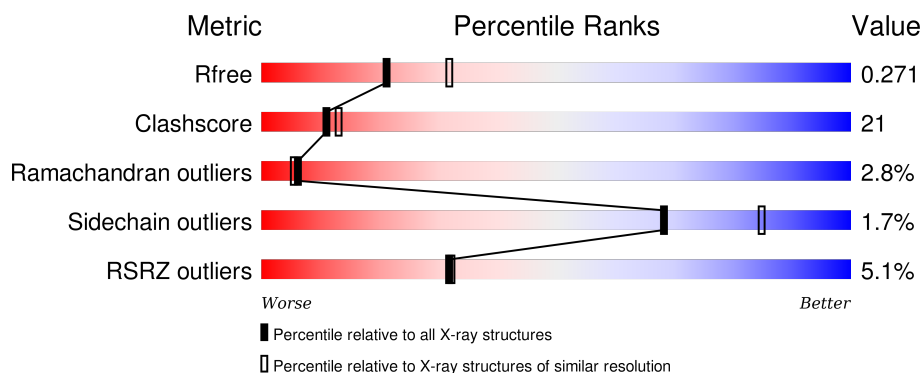
# 1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.40 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	2919 (2.40-2.40)
Clashscore	102246	3407 (2.40-2.40)
Ramachandran outliers	100387	3351 (2.40-2.40)
Sidechain outliers	100360	3352 (2.40-2.40)
RSRZ outliers	91569	2928 (2.40-2.40)

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  $\geq 3$ , 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	293	<div> <div>5%</div> <div>66%</div> <div>26%</div> <div>5% ..</div> </div>
1	B	293	<div> <div>4%</div> <div>59%</div> <div>32%</div> <div>5% ..</div> </div>

## 2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 4938 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.

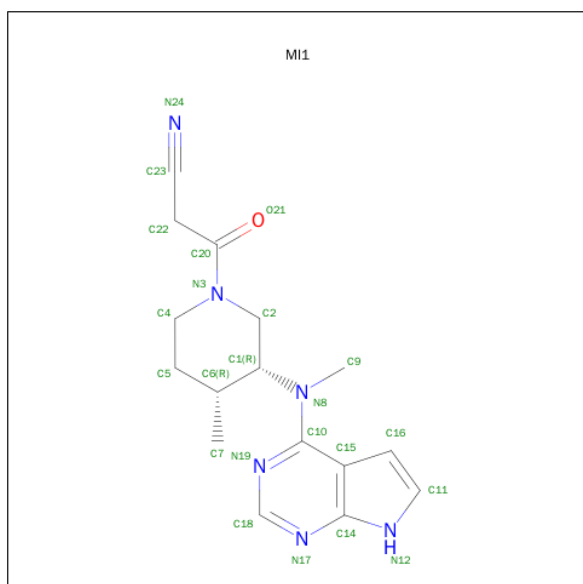
- Molecule 1 is a protein called Tyrosine-protein kinase JAK2.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	286	Total	C	N	O	P	S	0	0	0
			2370	1504	407	443	2	14			
1	B	289	Total	C	N	O	P	S	0	0	0
			2387	1515	410	446	2	14			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1129	GLN	ASN	ENGINEERED	UNP O60674
B	1129	GLN	ASN	ENGINEERED	UNP O60674

- Molecule 2 is 3-{(3R,4R)-4-METHYL-3-[METHYL(7H-PYRROLO[2,3-D]PYRIMIDIN-4-YL)AMINO]PIPERIDIN-1-YL}-3-OXOPROPANENITRILE (three-letter code: MI1) (formula: C<sub>16</sub>H<sub>20</sub>N<sub>6</sub>O).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total	C	N	O	0	0
			23	16	6	1		
2	B	1	Total	C	N	O	0	0
			23	16	6	1		

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	66	Total	O	0	0
			66	66		
3	B	69	Total	O	0	0
			69	69		



- Molecule 1: Tyrosine-protein kinase JAK2



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 41	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	110.49Å 110.49Å 70.32Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	110.43 – 2.40 78.13 – 2.40	Depositor EDS
% Data completeness (in resolution range)	94.3 (110.43-2.40) 94.4 (78.13-2.40)	Depositor EDS
$R_{merge}$	9.30	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.46 (at 2.40Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, $R_{free}$	0.178 , 0.250 0.215 , 0.271	Depositor DCC
$R_{free}$ test set	1601 reflections (5.37%)	DCC
Wilson B-factor (Å <sup>2</sup> )	38.4	Xtriage
Anisotropy	0.015	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.35 , 44.4	EDS
Estimated twinning fraction	0.037 for h,-k,-l	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.34$	Xtriage
Outliers	0 of 31434 reflections	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	4938	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	44.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.46% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>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 [i](#)

### 5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: MI1, 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	0.80	0/2386	0.84	10/3208 (0.3%)
1	B	0.66	1/2404 (0.0%)	0.72	9/3234 (0.3%)
All	All	0.73	1/4790 (0.0%)	0.78	19/6442 (0.3%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	6
1	B	0	16
All	All	0	22

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	1045	TYR	CD2-CE2	5.48	1.47	1.39

The worst 5 of 19 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	924	ASN	O-C-N	-8.81	108.60	122.70
1	A	1087	ARG	NE-CZ-NH1	5.73	123.17	120.30
1	A	924	ASN	CA-C-N	5.66	129.65	117.20
1	A	1128	ASP	CB-CG-OD2	5.45	123.20	118.30
1	B	949	ASP	CB-CG-OD2	5.28	123.06	118.30

There are no chirality outliers.

5 of 22 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	1000	VAL	Peptide
1	A	1001	LEU	Peptide
1	A	892	LEU	Peptide
1	A	918	TYR	Peptide
1	A	945	LYS	Peptide

## 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	2370	0	2334	69	2
1	B	2387	0	2345	123	1
2	A	23	0	18	7	0
2	B	23	0	20	5	0
3	A	66	0	0	3	0
3	B	69	0	0	8	0
All	All	4938	0	4717	199	2

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 21.

The worst 5 of 199 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:944:HIS:HA	1:A:946:GLU:OE1	1.39	1.18
1:B:867:ARG:NH1	1:B:876:GLY:O	1.86	1.09
1:B:969:THR:O	1:B:969:THR:HG22	1.59	1.01
1:B:1071:GLY:O	1:B:1072:GLN:HG3	1.58	1.01
1:A:919:SER:C	1:A:924:ASN:HB2	1.84	0.97

All (2) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:947:ARG:NH2	1:B:989:ARG:NH2[4_454]	1.66	0.54
1:A:872:GLN:O	1:A:1073:MET:SD[2_454]	2.15	0.05



## 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	A	280/293 (96%)	263 (94%)	12 (4%)	5 (2%)	11	13
1	B	283/293 (97%)	251 (89%)	21 (7%)	11 (4%)	4	3
All	All	563/586 (96%)	514 (91%)	33 (6%)	16 (3%)	6	5

5 of 16 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	1002	PRO
1	A	1069	LYS
1	B	901	ILE
1	B	943	LYS
1	B	970	LYS

### 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	259/265 (98%)	253 (98%)	6 (2%)	58	78
1	B	260/265 (98%)	257 (99%)	3 (1%)	78	90
All	All	519/530 (98%)	510 (98%)	9 (2%)	68	85

5 of 9 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	1000	VAL

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	B	1054	SER
1	B	875	THR
1	A	947	ARG
1	A	1010	VAL

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 13 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	1070	GLN
1	B	848	HIS
1	B	955	GLN
1	A	974	HIS
1	B	944	HIS

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

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	14,16,17	1.94	1 (7%)	18,22,24	0.69	0
1	PTR	A	1008	1	14,16,17	1.95	1 (7%)	18,22,24	0.69	0
1	PTR	B	1007	1	14,16,17	1.95	1 (7%)	18,22,24	0.69	0
1	PTR	B	1008	1	14,16,17	1.94	1 (7%)	18,22,24	0.70	0

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

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	1008	PTR	OH-CZ	-7.17	1.23	1.40
1	B	1007	PTR	OH-CZ	-7.16	1.23	1.40
1	B	1008	PTR	OH-CZ	-7.13	1.23	1.40
1	A	1007	PTR	OH-CZ	-7.12	1.23	1.40

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 6 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	A	1008	PTR	4	0
1	B	1008	PTR	2	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	MI1	A	1	-	22,25,25	2.39	3 (13%)	21,35,35	3.20	9 (42%)
2	MI1	B	1	-	22,25,25	2.11	2 (9%)	21,35,35	3.14	11 (52%)

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	MI1	A	1	-	-	0/14/28/28	0/3/3/3
2	MI1	B	1	-	-	0/14/28/28	0/3/3/3

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	1	MI1	C22-C23	-10.00	1.27	1.45
2	B	1	MI1	C22-C23	-8.41	1.30	1.45
2	A	1	MI1	C20-N3	-2.32	1.30	1.35
2	B	1	MI1	C14-N17	-2.28	1.33	1.37
2	A	1	MI1	C10-N19	2.03	1.37	1.34

The worst 5 of 20 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	1	MI1	N17-C18-N19	-10.61	120.77	128.89
2	B	1	MI1	N17-C18-N19	-7.77	122.95	128.89
2	B	1	MI1	C22-C23-N24	-6.97	153.04	177.66
2	A	1	MI1	C1-C2-N3	-3.63	106.11	111.11
2	A	1	MI1	C10-N8-C1	-2.93	117.92	121.16

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 12 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	1	MI1	7	0
2	B	1	MI1	5	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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	284/293 (96%)	0.16	16 (5%) 28 28	24, 39, 73, 106	0
1	B	287/293 (97%)	0.21	13 (4%) 37 38	30, 41, 71, 116	0
All	All	571/586 (97%)	0.18	29 (5%) 32 32	24, 40, 73, 116	0

The worst 5 of 29 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	1068	ASP	6.2
1	B	1013	PRO	5.1
1	A	919	SER	4.9
1	A	1070	GLN	4.5
1	A	860	PHE	4.5

### 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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	PTR	B	1007	16/17	0.91	0.13	-	43,58,77,82	0
1	PTR	A	1007	16/17	0.96	0.11	-	32,43,64,67	0
1	PTR	A	1008	16/17	0.93	0.14	-	28,56,67,67	0
1	PTR	B	1008	16/17	0.89	0.17	-	42,59,78,81	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, 95<sup>th</sup> 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( $\text{\AA}^2$ )	Q<0.9
2	MI1	B	1	23/23	0.94	0.17	0.49	28,35,48,55	0
2	MI1	A	1	23/23	0.94	0.16	-0.00	26,40,44,50	0

### 6.5 Other polymers [i](#)

There are no such residues in this entry.