



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 1, 2016 – 11:03 AM GMT

PDB ID : 3NZ0  
Title : Non-phosphorylated TYK2 kinase with CMP6  
Authors : Eigenbrot, C.; Ultsch, M.  
Deposited on : 2010-07-15  
Resolution : 2.00 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](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

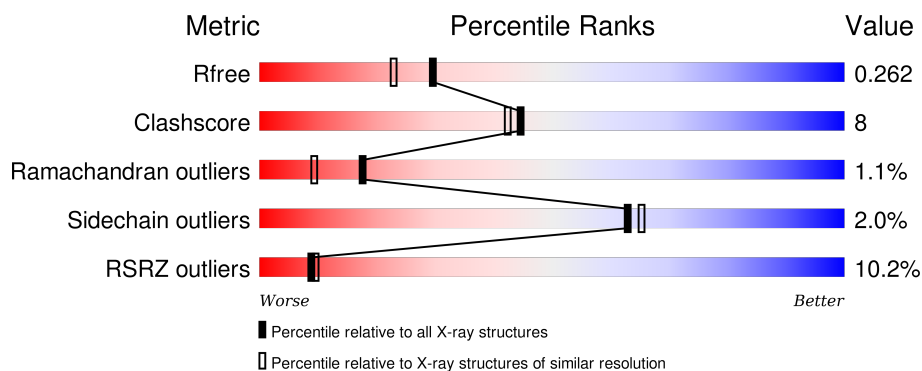
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*


The reported resolution of this entry is 2.00 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	6249 (2.00-2.00)
Clashscore	102246	7340 (2.00-2.00)
Ramachandran outliers	100387	7248 (2.00-2.00)
Sidechain outliers	100360	7247 (2.00-2.00)
RSRZ outliers	91569	6262 (2.00-2.00)

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	302	

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 2550 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.

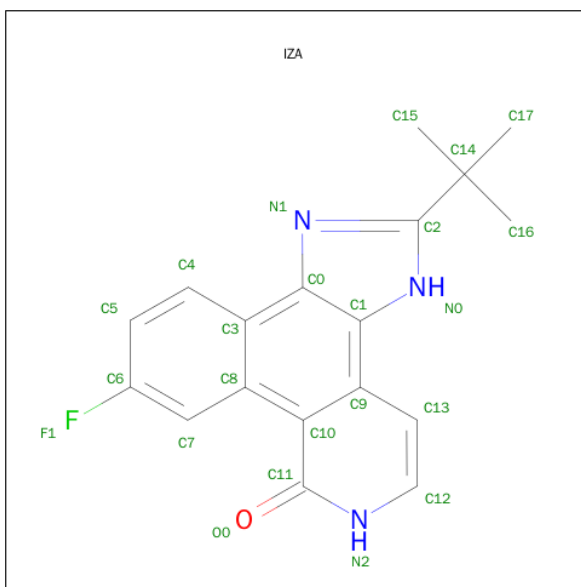
- Molecule 1 is a protein called Non-receptor tyrosine-protein kinase TYK2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	285	2354	1510	407	423	14	0	2	0

There are 11 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	882	MET	-	EXPRESSION TAG	UNP P29597
A	883	GLY	-	EXPRESSION TAG	UNP P29597
A	884	SER	-	EXPRESSION TAG	UNP P29597
A	1023	ASN	ASP	ENGINEERED MUTATION	UNP P29597
A	1177	ARG	-	EXPRESSION TAG	UNP P29597
A	1178	HIS	-	EXPRESSION TAG	UNP P29597
A	1179	HIS	-	EXPRESSION TAG	UNP P29597
A	1180	HIS	-	EXPRESSION TAG	UNP P29597
A	1181	HIS	-	EXPRESSION TAG	UNP P29597
A	1182	HIS	-	EXPRESSION TAG	UNP P29597
A	1183	HIS	-	EXPRESSION TAG	UNP P29597

- Molecule 2 is 2-TERT-BUTYL-9-FLUORO-3,6-DIHYDRO-7H-BENZ[H]-IMIDAZ[4,5-F]IS OQUINOLINE-7-ONE (three-letter code: IZA) (formula: C<sub>18</sub>H<sub>16</sub>FN<sub>3</sub>O).



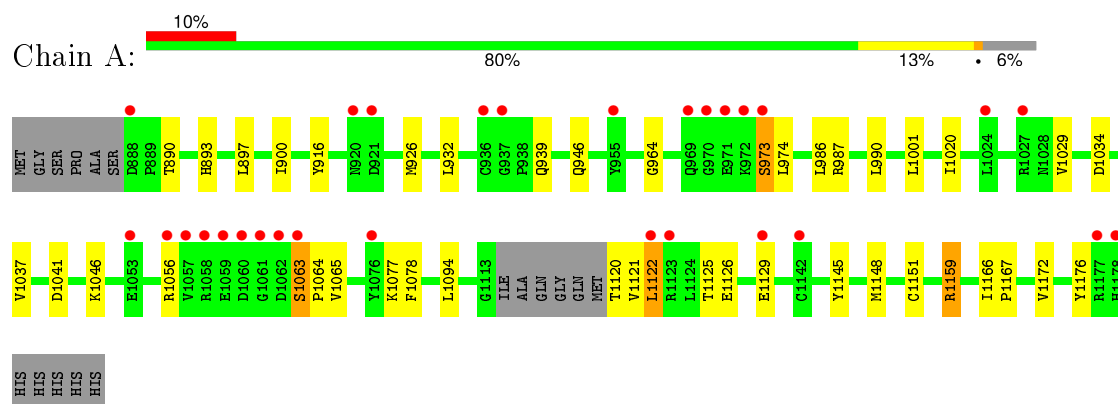
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	1	Total	C	F	N	O	0	0
			23	18	1	3	1		

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	173	Total	O	0	0
			173	173		



- Molecule 1: Non-receptor tyrosine-protein kinase TYK2



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	56.59Å 64.91Å 83.02Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	28.30 – 2.00 28.30 – 2.00	Depositor EDS
% Data completeness (in resolution range)	99.6 (28.30-2.00) 99.6 (28.30-2.00)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.09	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.61 (at 2.00Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, $R_{free}$	0.209 , 0.253 0.213 , 0.262	Depositor DCC
$R_{free}$ test set	1096 reflections (5.45%)	DCC
Wilson B-factor (Å <sup>2</sup> )	24.8	Xtriage
Anisotropy	0.045	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.37 , 50.7	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.51$ , $\langle L^2 \rangle = 0.34$	Xtriage
Outliers	0 of 21201 reflections	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	2550	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	26.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.65% 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: IZA

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	0/2418	0.60	1/3271 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed( $^{\circ}$ )	Ideal( $^{\circ}$ )
1	A	1159	ARG	NE-CZ-NH1	6.56	123.58	120.30

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	A	2354	0	2309	36	0
2	A	23	0	16	1	0
3	A	173	0	0	3	0
All	All	2550	0	2325	37	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 8.

All (37) 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:1056:ARG:NH2	1:A:1077:LYS:HE2	1.87	0.89
1:A:890:THR:HG21	3:A:99:HOH:O	1.72	0.88
1:A:900:ILE:CD1	1:A:926:MET:CE	2.55	0.85
1:A:900:ILE:CD1	1:A:926:MET:HE1	2.08	0.83
1:A:1056:ARG:HH21	1:A:1077:LYS:HE2	1.44	0.79
1:A:1001:LEU:HD13	1:A:1172:VAL:HG12	1.66	0.77
1:A:890:THR:HG23	1:A:964:GLY:HA2	1.73	0.71
1:A:900:ILE:CD1	1:A:926:MET:HE3	2.20	0.70
1:A:900:ILE:HD13	1:A:926:MET:HE1	1.75	0.69
1:A:900:ILE:HD12	1:A:926:MET:HE3	1.76	0.67
1:A:1001:LEU:CD1	1:A:1172:VAL:HG12	2.28	0.64
1:A:900:ILE:HD12	1:A:926:MET:CE	2.28	0.63
1:A:1063:SER:OG	1:A:1064:PRO:HA	1.99	0.62
1:A:1125:THR:O	1:A:1129:GLU:HG2	2.00	0.62
1:A:1056:ARG:NH2	1:A:1077:LYS:CE	2.64	0.59
1:A:986:LEU:HD13	1:A:1094:LEU:HD21	1.85	0.58
1:A:946:GLN:HE22	1:A:1046:LYS:NZ	2.00	0.58
1:A:890:THR:HG22	1:A:890:THR:O	2.06	0.54
1:A:1063:SER:CB	1:A:1065:VAL:HG13	2.41	0.50
1:A:990:LEU:HD11	1:A:1094:LEU:HD23	1.93	0.50
1:A:1145:TYR:HA	1:A:1148:MET:HE3	1.93	0.49
1:A:1151:CYS:O	1:A:1159:ARG:HD2	2.14	0.48
1:A:893:HIS:O	1:A:897:LEU:HD13	2.14	0.47
1:A:1166:ILE:HB	1:A:1167:PRO:HD3	1.97	0.47
1:A:900:ILE:HD11	1:A:926:MET:CE	2.44	0.46
1:A:973:SER:OG	1:A:974:LEU:N	2.48	0.46
1:A:1120:THR:N	3:A:65:HOH:O	2.51	0.43
1:A:1063:SER:HB2	1:A:1065:VAL:N	2.34	0.43
1:A:1121:VAL:HG12	3:A:139:HOH:O	2.18	0.42
1:A:1122:LEU:HD22	1:A:1126:GLU:OE1	2.18	0.42
1:A:897:LEU:HD12	1:A:916:TYR:CD1	2.55	0.42
1:A:1172:VAL:HG13	1:A:1176:TYR:CE2	2.56	0.41
1:A:946:GLN:HE22	1:A:1046:LYS:HZ2	1.68	0.41
2:A:1:IZA:H7	2:A:1:IZA:O0	2.20	0.41
1:A:890:THR:HG23	1:A:964:GLY:CA	2.45	0.40
1:A:1020:ILE:HD12	1:A:1078:PHE:CD2	2.57	0.40
1:A:1029:VAL:HG13	1:A:1037:VAL:HG13	2.04	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	283/302 (94%)	275 (97%)	5 (2%)	3 (1%)	17 9

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	973	SER
1	A	1063	SER
1	A	1041	ASP

### 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	254/265 (96%)	249 (98%)	5 (2%)	63 65

All (5) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	932	LEU
1	A	939	GLN
1	A	987	ARG
1	A	1034	ASP
1	A	1122	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	939	GLN
1	A	946	GLN
1	A	1102	GLN
1	A	1164	ASN

### 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 ⓘ

1 ligand is 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	IZA	A	1	-	22,26,26	2.73	9 (40%)	25,41,41	1.85	5 (20%)

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	IZA	A	1	-	-	0/6/6/6	0/4/4/4

All (9) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	1	IZA	C4-C3	-5.49	1.29	1.41
2	A	1	IZA	C5-C6	-4.58	1.27	1.37
2	A	1	IZA	C3-C0	-2.85	1.39	1.45
2	A	1	IZA	C9-C1	-2.54	1.40	1.45
2	A	1	IZA	F1-C6	-2.20	1.31	1.36
2	A	1	IZA	C13-C12	3.28	1.40	1.36
2	A	1	IZA	C7-C8	3.34	1.47	1.41
2	A	1	IZA	C11-C10	4.89	1.49	1.41
2	A	1	IZA	C8-C10	6.07	1.50	1.41

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	1	IZA	C8-C3-C0	-3.57	116.85	119.77
2	A	1	IZA	C13-C12-N2	-3.28	121.01	123.91
2	A	1	IZA	C7-C8-C3	-2.47	116.14	118.89
2	A	1	IZA	C12-N2-C11	2.21	122.20	116.57
2	A	1	IZA	C4-C5-C6	5.79	123.99	118.77

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
2	A	1	IZA	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, 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	285/302 (94%)	0.66	29 (10%) 9 9	13, 22, 50, 82	0

All (29) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	1061	GLY	15.3
1	A	1060	ASP	11.7
1	A	1062	ASP	9.6
1	A	1178	HIS	8.0
1	A	1059	GLU	6.4
1	A	1063	SER	6.0
1	A	971	GLU	5.9
1	A	972	LYS	5.8
1	A	888	ASP	5.7
1	A	1058	ARG	5.6
1	A	1122	LEU	3.8
1	A	1056	ARG	3.7
1	A	1076	TYR	3.7
1	A	1057	VAL	3.5
1	A	969	GLN	3.4
1	A	1177	ARG	3.4
1	A	973	SER	3.1
1	A	937	GLY	3.1
1	A	936	CYS	3.1
1	A	920	ASN	3.0
1	A	1053	GLU	2.6
1	A	955[A]	TYR	2.6
1	A	1142	CYS	2.5
1	A	970	GLY	2.5
1	A	1123	ARG	2.4
1	A	1027	ARG	2.3
1	A	921	ASP	2.1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	1024	LEU	2.1
1	A	1129	GLU	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, 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	IZA	A	1	23/23	0.94	0.15	0.05	15,17,19,19	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.