



# wwPDB X-ray Structure Validation Summary Report i

Feb 26, 2014 – 04:53 PM GMT

PDB ID : 4G3E  
Title : Crystal structure of murine NF-kappaB inducing kinase (NIK) bound to a 6-alkynylindoline (cmp1)  
Authors : Hymowitz, S.G.; de Leon-Boenig G.  
Deposited on : 2012-07-13  
Resolution : 2.50 Å (reported)

This is a wwPDB 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/ValidationPDFNotes.html>

---

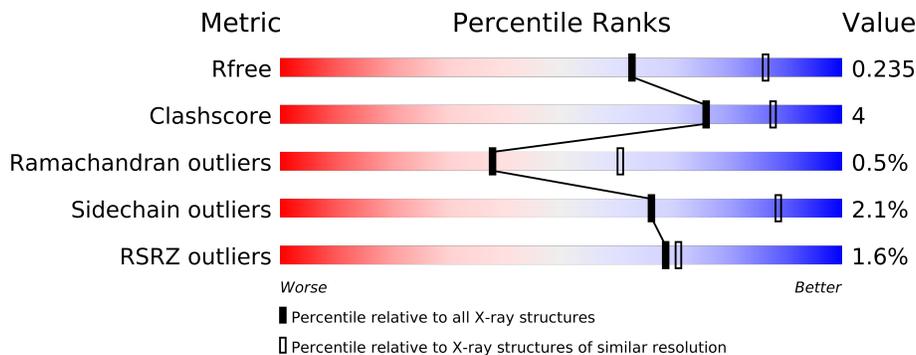
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.15 2013  
Xtriage (Phenix) : dev-1323  
EDS : stable22639  
Percentile statistics : 21963  
Refmac : 5.8.0049  
CCP4 : 6.3.0 (Settle)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)  
Validation Pipeline (wwPDB-VP) : stable22683

# 1 Overall quality at a glance

The reported resolution of this entry is 2.50 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))
$R_{free}$	66092	2784 (2.50-2.50)
Clashscore	79885	3562 (2.50-2.50)
Ramachandran outliers	78287	3480 (2.50-2.50)
Sidechain outliers	78261	3482 (2.50-2.50)
RSRZ outliers	66119	2785 (2.50-2.50)

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. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	352	
1	B	352	

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 5094 atoms, of which 0 are hydrogen and 0 are deuterium.

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.

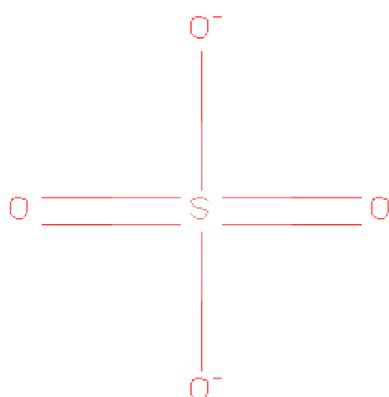
- Molecule 1 is a protein called NF-kappa-beta-inducingkinase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	314	2451	1552	436	442	21	0	1	0
1	B	325	2513	1588	445	459	21	0	0	0

There are 10 discrepancies between the modelled and reference sequences:

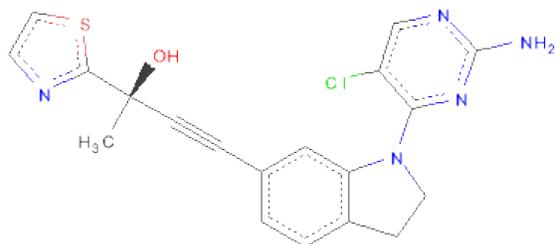
Chain	Residue	Modelled	Actual	Comment	Reference
A	327	GLY	-	EXPRESSION TAG	UNP Q9WUL6
A	328	SER	-	EXPRESSION TAG	UNP Q9WUL6
A	676	GLY	-	EXPRESSION TAG	UNP Q9WUL6
A	677	ASN	-	EXPRESSION TAG	UNP Q9WUL6
A	678	SER	-	EXPRESSION TAG	UNP Q9WUL6
B	327	GLY	-	EXPRESSION TAG	UNP Q9WUL6
B	328	SER	-	EXPRESSION TAG	UNP Q9WUL6
B	676	GLY	-	EXPRESSION TAG	UNP Q9WUL6
B	677	ASN	-	EXPRESSION TAG	UNP Q9WUL6
B	678	SER	-	EXPRESSION TAG	UNP Q9WUL6

- Molecule 2 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



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

- Molecule 3 is (2R)-4-[1-(2-AMINO-5-CHLOROPYRIMIDIN-4-YL)-2,3-DIHYDRO-1H-INDOL-6-YL]-2-(1,3-THIAZOL-2-YL)BUT-3-YN-2-OL (three-letter code: 0WC) (formula: C<sub>19</sub>H<sub>16</sub>ClN<sub>5</sub>OS).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	
3	A	1	Total	C	Cl	N	O	S	0	0
			27	19	1	5	1	1		
3	B	1	Total	C	Cl	N	O	S	0	0
			27	19	1	5	1	1		

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	34	Total	O	0	0
			34	34		
4	B	27	Total	O	0	0
			27	27		



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 43	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	143.15Å 143.15Å 45.15Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	47.73 – 2.50 47.72 – 2.50	Depositor EDS
% Data completeness (in resolution range)	99.5 (47.73-2.50) 99.6 (47.72-2.50)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.07	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.19 (at 2.51Å)	Xtriage
Refinement program	REFMAC 5.6.0117	Depositor
R, $R_{free}$	0.189 , 0.240 0.183 , 0.235	Depositor DCC
$R_{free}$ test set	3259 reflections (11.30%)	DCC
Wilson B-factor (Å <sup>2</sup> )	43.9	Xtriage
Anisotropy	0.040	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.36 , 26.9	EDS
Estimated twinning fraction	0.033 for h,-k,-l	Xtriage
L-test for twinning	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtriage
Outliers	0 of 32110 reflections	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	5094	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	45.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.19% of the height of the origin peak. No significant pseudotranslation is detected.*

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

## 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: SO4, OWC

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	0.56	3/2510 (0.1%)	0.62	2/3390 (0.1%)
1	B	0.55	4/2570 (0.2%)	0.61	0/3474
All	All	0.55	7/5080 (0.1%)	0.62	2/6864 (0.0%)

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	B	0	1

The worst 5 of 7 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	667	TRP	CD2-CE2	5.44	1.47	1.41
1	A	401	TRP	CD2-CE2	5.34	1.47	1.41
1	B	598	TRP	CD2-CE2	5.18	1.47	1.41
1	A	362	TRP	CD2-CE2	5.16	1.47	1.41
1	B	466	TRP	CD2-CE2	5.13	1.47	1.41

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	430	VAL	CB-CA-C	-5.88	100.24	111.40
1	A	514	LEU	CA-CB-CG	5.01	126.82	115.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	405	GLN	Peptide

## 5.2 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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	2451	0	38	4	0
1	B	2513	0	12	9	0
2	A	10	0	0	1	0
2	B	5	0	0	1	0
3	A	27	0	16	4	0
3	B	27	0	16	4	0
4	A	34	0	0	0	0
4	B	27	0	0	0	0
All	All	5094	0	82	20	0

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 4.

The worst 5 of 20 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
3:B:702:0WC:CL	3:B:702:0WC:H12	2.27	0.71
1:A:397:GLU:OE1	1:A:432:LYS:NZ	2.34	0.59
3:B:702:0WC:CL	3:B:702:0WC:C11	2.90	0.56
3:A:703:0WC:CL	3:A:703:0WC:H12	2.42	0.56
3:A:703:0WC:H13	3:A:703:0WC:N16	2.25	0.52

There are no symmetry-related clashes.

## 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	307/352 (87%)	292 (95%)	13 (4%)	2 (1%)	30	50
1	B	319/352 (91%)	304 (95%)	14 (4%)	1 (0%)	50	73
All	All	626/704 (89%)	596 (95%)	27 (4%)	3 (0%)	38	60

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	405	GLN
1	A	446	CYS
1	A	405	GLN

### 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	264/292 (90%)	257 (97%)	7 (3%)	57	83
1	B	271/292 (93%)	266 (98%)	5 (2%)	71	91
All	All	535/584 (92%)	523 (98%)	12 (2%)	66	88

5 of 12 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	475	GLU
1	A	667	TRP
1	B	543	LEU
1	A	437	VAL
1	B	437	VAL

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

### 5.3.3 RNA ⓘ

There are no RNA chains 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 ⓘ

5 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	SO4	A	701	-	4,4,4	0.45	0	6,6,6	0.08	0
2	SO4	A	702	-	4,4,4	0.65	0	6,6,6	0.37	0
3	0WC	A	703	-	30,30,30	1.01	2 (6%)	42,44,44	1.99	12 (28%)
2	SO4	B	701	-	4,4,4	0.46	0	6,6,6	0.27	0
3	0WC	B	702	-	30,30,30	1.09	0	42,44,44	2.15	15 (35%)

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	SO4	A	701	-	-	0/0/0/0	0/0/0/0
2	SO4	A	702	-	-	0/0/0/0	0/0/0/0
3	0WC	A	703	-	-	0/8/25/25	0/2/4/4
2	SO4	B	701	-	-	0/0/0/0	0/0/0/0
3	0WC	B	702	-	-	0/8/25/25	0/2/4/4

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	703	0WC	C21-C15	2.25	1.42	1.40
3	A	703	0WC	C17-N19	-2.04	1.32	1.35

The worst 5 of 27 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	703	0WC	C25-N24-C23	5.68	117.77	104.24
3	B	702	0WC	C25-N24-C23	5.39	117.07	104.24
3	B	702	0WC	C14-C13-C9	-5.16	117.40	122.61
3	A	703	0WC	C15-C21-CL	4.49	125.02	121.48
3	A	703	0WC	C14-C13-C9	-4.38	118.19	122.61

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

## 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, 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	314/352 (89%)	-0.06	7 (2%) 59 61	25, 39, 79, 115	0
1	B	325/352 (92%)	-0.09	3 (0%) 81 82	22, 41, 83, 118	0
All	All	639/704 (90%)	-0.07	10 (1%) 68 71	22, 40, 83, 118	0

The worst 5 of 10 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	405	GLN	3.9
1	A	606	LEU	3.6
1	A	608	LEU	2.9
1	B	675	ARG	2.7
1	A	407	ARG	2.6

### 6.2 Non-standard residues in protein, DNA, RNA chains

There are no non-standard protein/DNA/RNA residues in this entry.

### 6.3 Carbohydrates

There are no carbohydrates in this entry.

### 6.4 Ligands

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	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
3	OWC	A	703	27/27	0.14	0.48	25,36,49,54	0
3	OWC	B	702	27/27	0.12	-0.48	26,32,40,47	0
2	SO4	A	701	5/5	0.16	-1.13	54,54,63,69	0
2	SO4	A	702	5/5	0.12	-1.69	40,58,67,80	0
2	SO4	B	701	5/5	0.11	-3.52	53,54,56,57	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.