



wwPDB X-ray Structure Validation Summary Report i

Feb 26, 2014 – 09:28 PM GMT

PDB ID : 2JLE
Title : NOVEL INDAZOLE NNRTIS CREATED USING MOLECULAR TEMPLATE HYBRIDIZATION BASED ON CRYSTALLOGRAPHIC OVERLAYS
Authors : Jones, L.H.; Allan, G.; Barba, O.; Burt, C.; Corbau, R.; Dupont, T.; Irving, S.; Mowbray, C.E.; Phillips, C.; Swain, N.A.; Webster, R.; Westby, M.
Deposited on : 2008-09-08
Resolution : 2.90 Å(reported)

This is a wwPDB 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/ValidationPDFNotes.html>

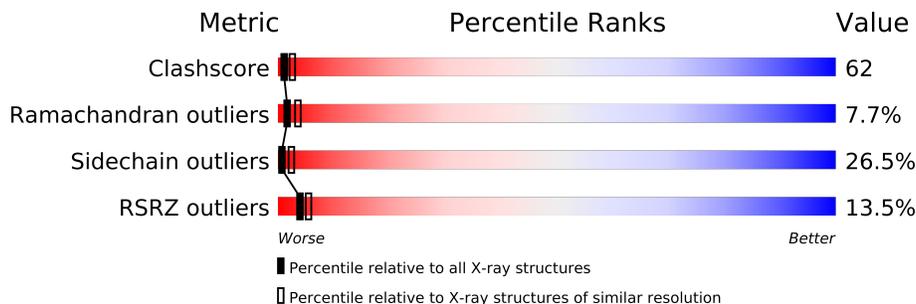
The following versions of software and data (see [references](#)) were used in the production of this report:

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 2.90 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))
Clashscore	79885	1326 (2.90-2.90)
Ramachandran outliers	78287	1290 (2.90-2.90)
Sidechain outliers	78261	1292 (2.90-2.90)
RSRZ outliers	66119	1054 (2.90-2.90)

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. 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	566	
1	B	566	

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 8120 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.

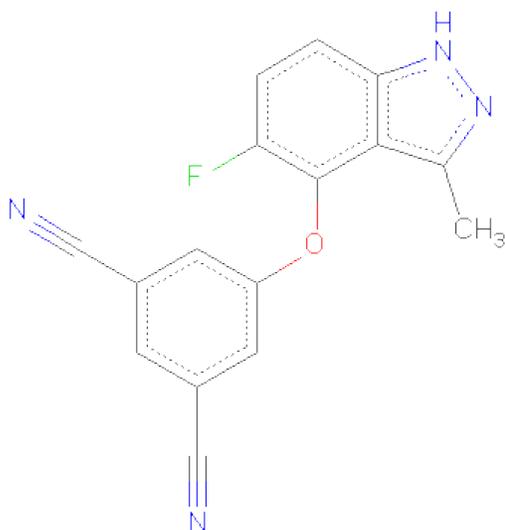
- Molecule 1 is a protein called REVERSE TRANSCRIPTASE/RNASEH.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	545	Total	C	N	O	S	0	0	0
			4445	2875	742	820	8			
1	B	416	Total	C	N	O	S	0	0	1
			3414	2218	569	620	7			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	103	LYS	ARG	CONFLICT	UNP Q72547
A	350	LYS	ARG	CONFLICT	UNP Q72547
B	103	LYS	ARG	CONFLICT	UNP Q72547
B	350	LYS	ARG	CONFLICT	UNP Q72547

- Molecule 2 is 5-[(5-FLUORO-3-METHYL-1H-INDAZOL-4-YL)OXY]BENZENE-1,3-DICARBONITRILE (three-letter code: I15) (formula: C₁₆H₉FN₄O).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	F	N	O		
2	A	1	22	16	1	4	1	0	0

- Molecule 3 is water.

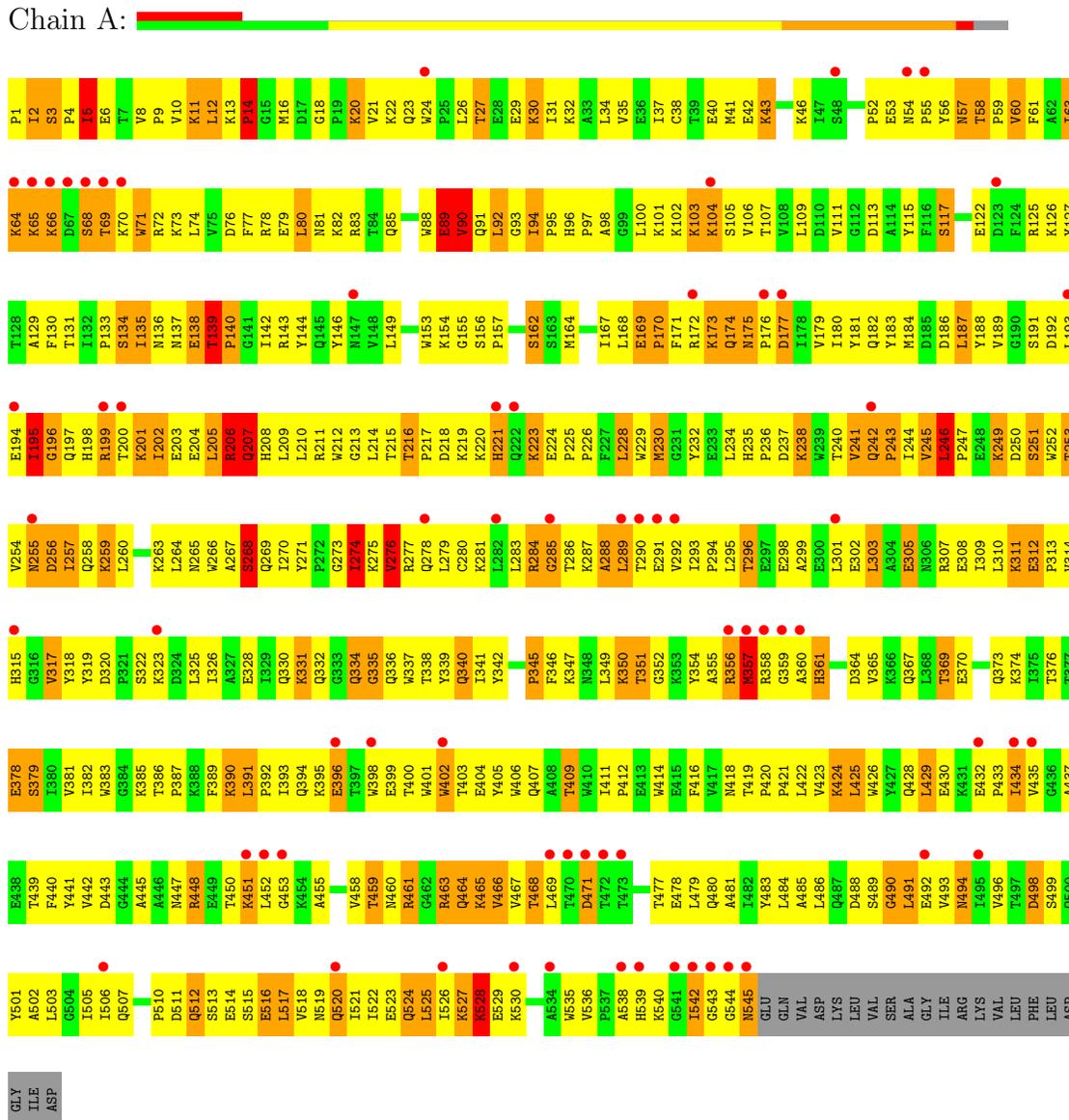
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	152	Total	O	0	0
			152	152		
3	B	87	Total	O	0	0
			87	87		

3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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: REVERSE TRANSCRIPTASE/RNASEH

Chain A:



- Molecule 1: REVERSE TRANSCRIPTASE/RNASEH

Chain B:

4 Data and refinement statistics

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, α , β , γ	117.20Å 154.60Å 155.00Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	14.13 – 2.90 29.94 – 2.90	Depositor EDS
% Data completeness (in resolution range)	96.7 (14.13-2.90) 90.5 (29.94-2.90)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.24 (at 2.90Å)	Xtrriage
Refinement program	BUSTER/TNT	Depositor
R, R_{free}	0.261 , 0.351 0.276 , (Not available)	Depositor DCC
R_{free} test set	No test flags present.	DCC
Wilson B-factor (Å ²)	66.3	Xtrriage
Anisotropy	0.213	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.38 , 79.6	EDS
Estimated twinning fraction	No twinning to report.	Xtrriage
L-test for twinning	$\langle L \rangle = 0.58$, $\langle L^2 \rangle = 0.44$	Xtrriage
Outliers	3 of 30525 reflections (0.010%)	Xtrriage
F_o, F_c correlation	0.90	EDS
Total number of atoms	8120	wwPDB-VP
Average B, all atoms (Å ²)	60.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.04% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities 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: I15

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	0.90	3/4562 (0.1%)	1.08	8/6199 (0.1%)
1	B	0.90	0/3510	1.07	8/4772 (0.2%)
All	All	0.90	3/8072 (0.0%)	1.08	16/10971 (0.1%)

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	516	GLU	CG-CD	6.18	1.61	1.51
1	A	516	GLU	CB-CG	5.39	1.62	1.52
1	A	432	GLU	CG-CD	5.20	1.59	1.51

The worst 5 of 16 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	420	PRO	C-N-CD	-9.69	99.27	120.60
1	A	139	THR	C-N-CD	-9.58	99.53	120.60
1	B	312	GLU	C-N-CD	-9.18	100.40	120.60
1	B	344	GLU	C-N-CD	-7.69	103.68	120.60
1	B	132	ILE	C-N-CD	-7.13	104.90	120.60

There are no chirality outliers.

There are no planarity outliers.

5.2 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 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	4445	0	4493	620	0
1	B	3414	0	3443	399	0
2	A	22	0	9	1	0
3	A	152	0	0	25	0
3	B	87	0	0	10	0
All	All	8120	0	7945	989	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 62.

The worst 5 of 989 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:104:LYS:HB3	1:A:192:ASP:HA	1.20	1.18
1:B:103:LYS:HE3	1:B:179:VAL:HG21	1.24	1.15
1:A:64:LYS:HE3	1:A:69:THR:HA	1.22	1.10
1:A:174:GLN:HA	1:A:174:GLN:HE21	1.14	1.07
1:A:288:ALA:HB3	1:A:291:GLU:HB2	1.33	1.06

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	543/566 (96%)	415 (76%)	86 (16%)	42 (8%)	1	3
1	B	412/566 (73%)	338 (82%)	42 (10%)	32 (8%)	1	3
All	All	955/1132 (84%)	753 (79%)	128 (13%)	74 (8%)	1	3

5 of 74 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	2	ILE

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	14	PRO
1	A	90	VAL
1	A	135	ILE
1	A	139	THR

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	487/505 (96%)	356 (73%)	131 (27%)	1	2
1	B	375/505 (74%)	278 (74%)	97 (26%)	1	2
All	All	862/1010 (85%)	634 (74%)	228 (26%)	1	2

5 of 228 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	428	GLN
1	A	520	GLN
1	B	353	LYS
1	A	439	THR
1	A	468	THR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 32 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	367	GLN
1	A	494	ASN
1	B	255	ASN
1	A	480	GLN
1	A	507	GLN

5.3.3 RNA [i](#)

There are no RNA chains in this entry.

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

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

5.5 Carbohydrates (i)

There are no carbohydrates in this entry.

5.6 Ligand geometry (i)

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	I15	A	1546	-	24,24,24	2.34	6 (25%)	33,34,34	1.68	9 (27%)

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	I15	A	1546	-	-	0/8/8/8	0/1/3/3

The worst 5 of 6 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	1546	I15	C10-C4	-7.24	1.45	1.50
2	A	1546	I15	C4-C1	4.15	1.47	1.39
2	A	1546	I15	C8-C12	3.33	1.43	1.36
2	A	1546	I15	F1-C6	3.29	1.43	1.35
2	A	1546	I15	C12-C6	2.78	1.44	1.38

The worst 5 of 9 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	1546	I15	C10-C4-N9	3.22	126.67	119.77
2	A	1546	I15	C12-C6-C2	-3.19	115.23	122.52
2	A	1546	I15	C16-C14-C11	-2.99	115.21	119.24
2	A	1546	I15	C12-C8-C3	2.57	122.79	119.38
2	A	1546	I15	C17-C16-C14	2.40	124.24	119.73

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, 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	545/566 (96%)	0.61	68 (12%) 5 6	34, 59, 86, 107	0
1	B	416/566 (73%)	0.80	62 (14%) 3 4	36, 58, 95, 119	0
All	All	961/1132 (84%)	0.69	130 (13%) 4 5	34, 59, 90, 119	0

The worst 5 of 130 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	357	MET	12.0
1	B	284	ARG	8.3
1	A	359	GLY	6.2
1	A	453	GLY	6.1
1	B	278	GLN	5.6

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, 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	RSR	LLDF	B-factors(\AA^2)	Q<0.9
2	I15	A	1546	22/22	0.20	-0.62	42,51,55,57	0

6.5 Other polymers

There are no such residues in this entry.