



wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 10, 2017 – 05:46 AM EDT

PDB ID : 2O61
Title : Crystal Structure of NFkB, IRF7, IRF3 bound to the interferon-b enhancer
Authors : Panne, D.
Deposited on : unknown
Resolution : 2.80 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
Xtriage (Phenix) : 1.9-1692
EDS : rb-20030345
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) : rb-20030345

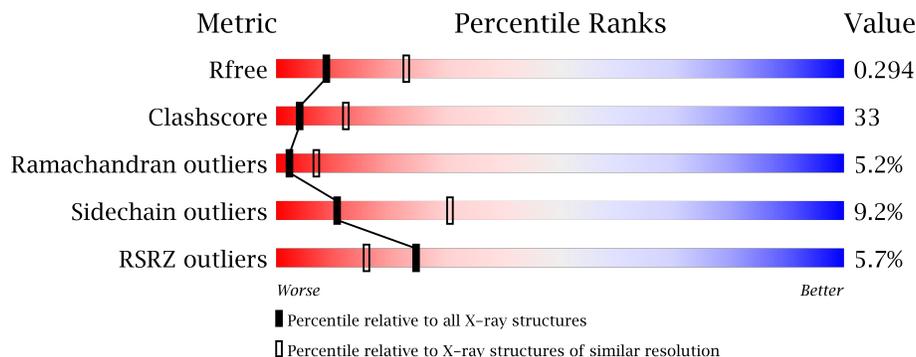
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.80 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	2583 (2.80-2.80)
Clashscore	112137	3033 (2.80-2.80)
Ramachandran outliers	110173	2983 (2.80-2.80)
Sidechain outliers	110143	2985 (2.80-2.80)
RSRZ outliers	101464	2610 (2.80-2.80)

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	E	36	
2	F	34	
3	A	540	
4	B	314	

2 Entry composition [i](#)

There are 5 unique types of molecules in this entry. The entry contains 7912 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 DNA chain called 36-MER.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
1	E	36	750	357	147	211	35	0	0	0

- Molecule 2 is a DNA chain called 34-MER.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
2	F	34	678	328	110	207	33	0	0	0

- Molecule 3 is a protein called Transcription factor p65/Interferon regulatory factor 7/Interferon regulatory factor 3 fusion protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	A	498	3988	2505	745	722	16	0	0	0

There are 48 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	18	MET	-	INITIATING METHIONINE	UNP Q04206
A	19	GLY	-	CLONING ARTIFACT	UNP Q04206
A	292	GLY	-	LINKER	UNP Q04206
A	293	SER	-	LINKER	UNP Q04206
A	294	LEU	-	LINKER	UNP Q04206
A	295	SER	-	LINKER	UNP Q04206
A	296	SER	-	LINKER	UNP Q04206
A	297	GLY	-	LINKER	UNP Q04206
A	298	SER	-	LINKER	UNP Q04206
A	299	SER	-	LINKER	UNP Q04206
A	300	LEU	-	LINKER	UNP Q04206
A	301	SER	-	LINKER	UNP Q04206
A	302	SER	-	LINKER	UNP Q04206

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	303	PRO	-	LINKER	UNP Q04206
A	304	SER	-	LINKER	UNP Q04206
A	305	ALA	-	LINKER	UNP Q04206
A	306	GLY	-	LINKER	UNP Q04206
A	1118	GLN	PRO	ENGINEERED	UNP Q92985
A	1126	GLY	-	LINKER	UNP Q92985
A	1127	SER	-	LINKER	UNP Q92985
A	1128	LEU	-	LINKER	UNP Q92985
A	1129	SER	-	LINKER	UNP Q92985
A	1130	SER	-	LINKER	UNP Q92985
A	1131	ASP	-	LINKER	UNP Q92985
A	1132	SER	-	LINKER	UNP Q92985
A	1133	SER	-	LINKER	UNP Q92985
A	1134	LEU	-	LINKER	UNP Q92985
A	1135	SER	-	LINKER	UNP Q92985
A	1136	SER	-	LINKER	UNP Q92985
A	1137	PRO	-	LINKER	UNP Q92985
A	1138	SER	-	LINKER	UNP Q92985
A	1139	ALA	-	LINKER	UNP Q92985
A	1140	LEU	-	LINKER	UNP Q92985
A	1141	SER	-	LINKER	UNP Q92985
A	1142	PRO	-	LINKER	UNP Q92985
A	1143	LYS	-	LINKER	UNP Q92985
A	1144	PRO	-	LINKER	UNP Q92985
A	1145	ARG	-	LINKER	UNP Q92985
A	1146	ILE	-	LINKER	UNP Q92985
A	2112	GLY	-	EXPRESSION TAG	UNP Q14653
A	2113	LEU	-	EXPRESSION TAG	UNP Q14653
A	2114	GLU	-	EXPRESSION TAG	UNP Q14653
A	2115	HIS	-	EXPRESSION TAG	UNP Q14653
A	2116	HIS	-	EXPRESSION TAG	UNP Q14653
A	2117	HIS	-	EXPRESSION TAG	UNP Q14653
A	2118	HIS	-	EXPRESSION TAG	UNP Q14653
A	2119	HIS	-	EXPRESSION TAG	UNP Q14653
A	2120	HIS	-	EXPRESSION TAG	UNP Q14653

- Molecule 4 is a protein called Nuclear factor NF-kappa-B p105 subunit.

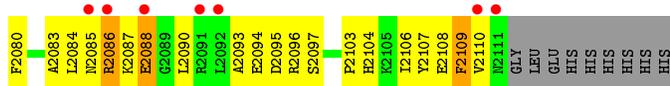
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
4	B	314	2471	1566	430	462	13	0	0	0

There is a discrepancy between the modelled and reference sequences:

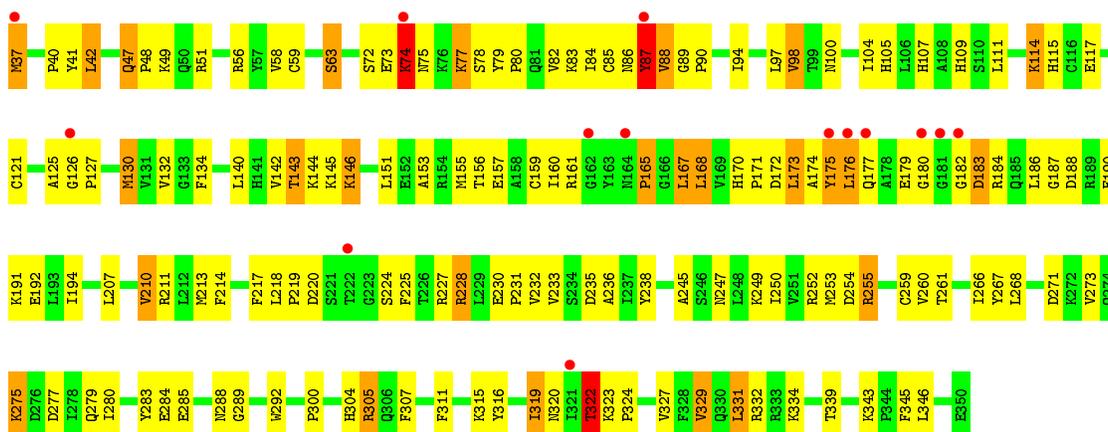
Chain	Residue	Modelled	Actual	Comment	Reference
B	37	MET	-	INITIATING METHIONINE	UNP P19838

- Molecule 5 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	E	6	Total O 6 6	0	0
5	F	4	Total O 4 4	0	0
5	A	10	Total O 10 10	0	0
5	B	5	Total O 5 5	0	0



● Molecule 4: Nuclear factor NF-kappa-B p105 subunit



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	95.22Å 116.37Å 134.24Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	40.65 – 2.80 40.50 – 2.77	Depositor EDS
% Data completeness (in resolution range)	99.8 (40.65-2.80) 99.0 (40.50-2.77)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.43 (at 2.77Å)	Xtrriage
Refinement program	CNS 1.1	Depositor
R, R_{free}	0.245 , 0.278 0.253 , 0.294	Depositor DCC
R_{free} test set	1863 reflections (4.99%)	DCC
Wilson B-factor (Å ²)	62.5	Xtrriage
Anisotropy	0.449	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.30 , 48.1	EDS
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.92	EDS
Total number of atoms	7912	wwPDB-VP
Average B, all atoms (Å ²)	64.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.47% of the height of the origin peak. No significant pseudotranslation is detected.*

¹ Intensities estimated from amplitudes.

² Theoretical 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 [i](#)

5.1 Standard geometry [i](#)

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	E	0.47	0/845	0.82	2/1306 (0.2%)
2	F	0.42	0/755	0.79	0/1160
3	A	0.41	0/4094	0.67	2/5547 (0.0%)
4	B	0.41	0/2524	0.68	0/3407
All	All	0.42	0/8218	0.71	4/11420 (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	E	1	0

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed($^{\circ}$)	Ideal($^{\circ}$)
1	E	35	DG	N9-C1'-C2'	6.43	124.82	112.60
1	E	35	DG	O4'-C1'-N9	5.13	111.59	108.00
3	A	1099	LEU	CA-CB-CG	5.10	127.03	115.30
3	A	175	LEU	CA-CB-CG	5.02	126.84	115.30

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
1	E	35	DG	C1'

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	E	750	0	407	48	0
2	F	678	0	388	60	0
3	A	3988	0	3878	224	0
4	B	2471	0	2468	177	0
5	A	10	0	0	0	0
5	B	5	0	0	0	0
5	E	6	0	0	0	0
5	F	4	0	0	0	0
All	All	7912	0	7141	490	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 33.

The worst 5 of 490 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)
4:B:47:GLN:HG3	4:B:48:PRO:HD2	1.29	1.10
3:A:2009:LEU:HG	3:A:2087:LYS:HE3	1.32	1.09
2:F:30:DC:H2''	2:F:31:DT:H5''	1.33	1.08
3:A:2086:ARG:CG	3:A:2086:ARG:HH11	1.67	1.07
3:A:2029:LYS:HD3	3:A:2029:LYS:H	1.20	1.06

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	
3	A	490/540 (91%)	399 (81%)	60 (12%)	31 (6%)	1	4
4	B	312/314 (99%)	264 (85%)	37 (12%)	11 (4%)	4	14
All	All	802/854 (94%)	663 (83%)	97 (12%)	42 (5%)	2	7

5 of 42 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	A	81	PRO
3	A	169	SER
3	A	191	THR
3	A	1074	SER
3	A	1128	LEU

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	
3	A	423/463 (91%)	384 (91%)	39 (9%)	11	30
4	B	269/269 (100%)	244 (91%)	25 (9%)	10	30
All	All	692/732 (94%)	628 (91%)	64 (9%)	11	30

5 of 64 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	A	1131	ASP
3	A	2073	LEU
4	B	285	GLU
3	A	2012	LEU
3	A	2031	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 27 such sidechains are listed below:

Mol	Chain	Res	Type
3	A	263	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
3	A	1111	ASN
4	B	304	HIS
3	A	271	GLN
3	A	138	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](#)

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](#)

There are no ligands in this entry.

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	E	36/36 (100%)	0.07	2 (5%) 25 16	30, 51, 124, 142	0
2	F	34/34 (100%)	-0.05	1 (2%) 52 41	27, 48, 71, 123	0
3	A	498/540 (92%)	0.21	33 (6%) 19 11	27, 60, 122, 139	0
4	B	314/314 (100%)	0.16	14 (4%) 34 24	28, 63, 116, 140	0
All	All	882/924 (95%)	0.18	50 (5%) 24 16	27, 60, 121, 142	0

The worst 5 of 50 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	A	1132	SER	8.3
4	B	181	GLY	6.8
3	A	2016	LEU	6.3
4	B	176	LEU	6.0
4	B	180	GLY	5.2

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](#)

There are no ligands in this entry.

6.5 Other polymers

There are no such residues in this entry.