



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 5, 2026 – 08:10 PM UTC

PDB ID : 8JKN / pdb_00008jkn
Title : T95R mutant IRF4 DNA-binding domain bound to an DNA containing GAAA motif
Authors : Wang, G.; Feng, X.; Ding, J.
Deposited on : 2023-06-01
Resolution : 2.92 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

<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

The types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

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

MolProbity	:	4-5-2 with Phenix2.0
Xtriage (Phenix)	:	2.0
EDS	:	3.0
Percentile statistics	:	20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4	:	9.0.010 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.49

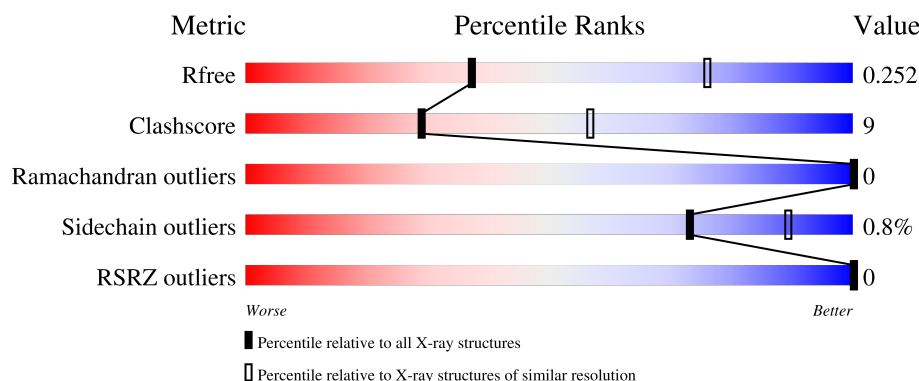
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.92 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}	180053	2995 (2.94-2.90)
Clashscore	190562	3213 (2.94-2.90)
Ramachandran outliers	187476	3128 (2.94-2.90)
Sidechain outliers	187428	3130 (2.94-2.90)
RSRZ outliers	180081	2995 (2.94-2.90)

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 of the lower bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria respectively. 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	19	
1	E	19	
2	B	19	
2	F	19	
3	C	116	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	D	116	 72% 21% 8%
3	G	116	 65% 29% 6%
3	H	116	 80% 10% 6%

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 5220 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 GAAA-Forward.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	19	Total	C	N	O	P	0	0	0
			389	184	80	106	19			
1	E	19	Total	C	N	O	P	0	0	0
			389	184	80	106	19			

- Molecule 2 is a DNA chain called GAAA-Reverse.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	19	Total	C	N	O	P	0	0	0
			390	187	62	122	19			
2	F	19	Total	C	N	O	P	0	0	0
			390	187	62	122	19			

- Molecule 3 is a protein called Interferon regulatory factor 4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	109	Total	C	N	O	S	0	0	0
			921	592	164	164	1			
3	D	107	Total	C	N	O	S	0	0	0
			898	578	161	158	1			
3	G	109	Total	C	N	O	S	0	0	0
			921	592	164	164	1			
3	H	109	Total	C	N	O	S	0	0	0
			918	591	163	163	1			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	95	ARG	THR	engineered mutation	UNP F2Z3D5
D	95	ARG	THR	engineered mutation	UNP F2Z3D5
G	95	ARG	THR	engineered mutation	UNP F2Z3D5
H	95	ARG	THR	engineered mutation	UNP F2Z3D5

- Molecule 4 is water.

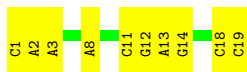
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	1	Total 1	O 1	0	0
4	B	1	Total 1	O 1	0	0
4	F	2	Total 2	O 2	0	0

3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes 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: GAAA-Forward

Chain A:  47% 53%



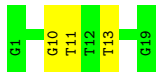
- Molecule 1: GAAA-Forward

Chain E:  21% 74% 5%



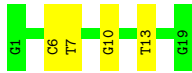
- Molecule 2: GAAA-Reverse

Chain B:  84% 16%



- Molecule 2: GAAA-Reverse

Chain F:  79% 21%



- Molecule 3: Interferon regulatory factor 4

Chain C:  78% 16% 6%

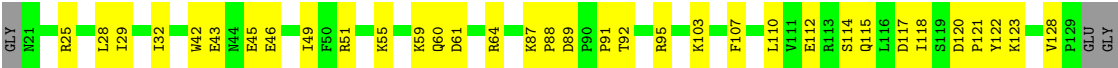


- Molecule 3: Interferon regulatory factor 4

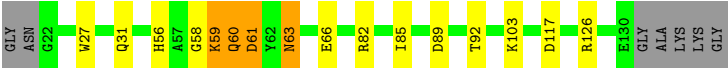
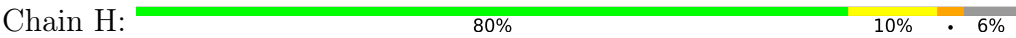
Chain D:  72% 21% 8%



• Molecule 3: Interferon regulatory factor 4



• Molecule 3: Interferon regulatory factor 4



4 Data and refinement statistics

Property	Value	Source
Space group	P 31 2 1	Depositor
Cell constants a, b, c, α , β , γ	118.64Å 118.64Å 154.65Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	48.75 – 2.92 48.75 – 2.92	Depositor EDS
% Data completeness (in resolution range)	99.8 (48.75-2.92) 99.8 (48.75-2.92)	Depositor EDS
R_{merge}	0.26	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.05 (at 2.91Å)	Xtriage
Refinement program	PHENIX (1.20.1_4487: ???)	Depositor
R, R_{free}	0.220 , 0.244 0.230 , 0.252	Depositor DCC
R_{free} test set	1413 reflections (5.05%)	wwPDB-VP
Wilson B-factor (Å ²)	52.0	Xtriage
Anisotropy	0.644	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.27 , 19.6	EDS
L-test for twinning ²	$\langle L \rangle = 0.43$, $\langle L^2 \rangle = 0.25$	Xtriage
Estimated twinning fraction	0.179 for -h,-k,l	Xtriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	5220	wwPDB-VP
Average B, all atoms (Å ²)	53.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.34% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical 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	A	0.40	0/438	0.81	3/672 (0.4%)
1	E	0.39	0/438	0.79	3/672 (0.4%)
2	B	0.36	0/434	0.62	0/669
2	F	0.34	0/434	0.55	0/669
3	C	0.25	0/947	0.40	0/1277
3	D	0.23	0/922	0.39	0/1241
3	G	0.22	0/947	0.38	0/1277
3	H	0.28	0/944	0.78	5/1273 (0.4%)
All	All	0.29	0/5504	0.59	11/7750 (0.1%)

There are no bond length outliers.

The worst 5 of 11 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	H	60	GLN	CB-CA-C	-16.12	88.58	112.11
1	A	1	DC	OP2-P-O3'	-11.33	74.02	108.00
1	E	1	DC	OP1-P-O3'	-10.44	76.67	108.00
3	H	60	GLN	N-CA-C	-8.77	99.45	111.28
3	H	59	LYS	N-CA-C	-8.71	96.98	109.96

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	389	0	211	7	0
1	E	389	0	211	10	0
2	B	390	0	219	4	0
2	F	390	0	219	3	0
3	C	921	0	908	13	1
3	D	898	0	890	17	0
3	G	921	0	908	29	0
3	H	918	0	904	14	1
4	A	1	0	0	0	0
4	B	1	0	0	0	0
4	F	2	0	0	0	0
All	All	5220	0	4470	83	1

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 9.

The worst 5 of 83 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)
3:G:46:GLU:N	3:G:46:GLU:OE2	2.01	0.92
1:E:14:DG:O6	3:H:103:LYS:NZ	2.11	0.82
3:G:59:LYS:HD2	3:G:61:ASP:H	1.51	0.75
3:G:112:GLU:OE2	3:G:112:GLU:N	2.22	0.72
3:G:89:ASP:OD2	3:G:92:THR:OG1	2.07	0.70

All (1) 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)
3:C:45:GLU:OE1	3:H:126:ARG:NH1[6_555]	2.18	0.02

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	C	107/116 (92%)	100 (94%)	7 (6%)	0	100	100
3	D	103/116 (89%)	98 (95%)	5 (5%)	0	100	100
3	G	107/116 (92%)	105 (98%)	2 (2%)	0	100	100
3	H	107/116 (92%)	103 (96%)	4 (4%)	0	100	100
All	All	424/464 (91%)	406 (96%)	18 (4%)	0	100	100

There are no Ramachandran outliers to report.

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	
3	C	97/100 (97%)	97 (100%)	0	100	100
3	D	94/100 (94%)	92 (98%)	2 (2%)	47	76
3	G	97/100 (97%)	97 (100%)	0	100	100
3	H	96/100 (96%)	95 (99%)	1 (1%)	68	87
All	All	384/400 (96%)	381 (99%)	3 (1%)	73	89

All (3) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	D	65	GLU
3	D	85	ILE
3	H	63	ASN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
3	C	26	GLN
3	H	115	GLN

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 oligosaccharides 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	A	19/19 (100%)	-1.90	0 100 100	32, 39, 57, 74	0
1	E	19/19 (100%)	-1.77	0 100 100	35, 44, 60, 67	0
2	B	19/19 (100%)	-1.86	0 100 100	31, 42, 63, 81	0
2	F	19/19 (100%)	-1.71	0 100 100	33, 42, 76, 78	0
3	C	109/116 (93%)	-1.39	0 100 100	33, 47, 81, 96	0
3	D	107/116 (92%)	-1.22	0 100 100	41, 59, 80, 98	0
3	G	109/116 (93%)	-1.26	0 100 100	38, 56, 81, 97	0
3	H	109/116 (93%)	-1.21	0 100 100	38, 55, 80, 96	0
All	All	510/540 (94%)	-1.35	0 100 100	31, 54, 81, 98	0

There are no RSRZ outliers to report.

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 oligosaccharides in this entry.

6.4 Ligands [i](#)

There are no ligands in this entry.

6.5 Other polymers [i](#)

There are no such residues in this entry.