Full wwPDB X-ray Structure Validation Report ⓘ

May 31, 2022 – 01:09 pm BST

PDB ID	:	7OOT
Title	:	X-ray Structure of Interferon Regulatory Factor 4 DNA binding domain bound to an interferon-stimulated response element
Authors	:	Agnarelli, A.; El Omari, K.; Alt, A.O.; Mancini, E.J.
Deposited on	:	2021-05-28
Resolution	:	2.25 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*
A user guide is available at
https://www.wwpdb.org/validation/2017/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:

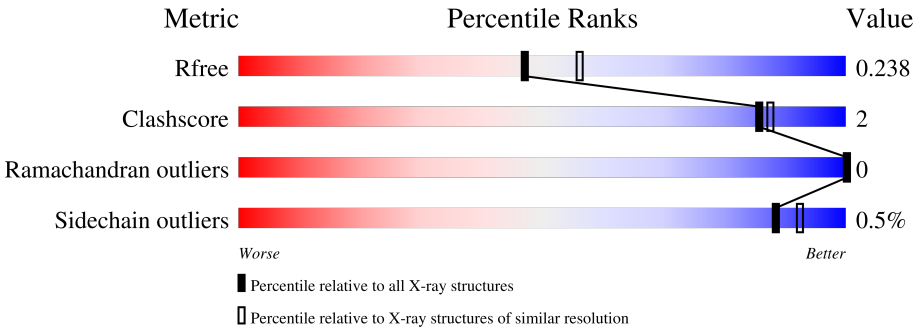
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.28.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0267
CCP4	:	7.1.010 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.28.1

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.25 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}	130704	1377 (2.26-2.26)
Clashscore	141614	1487 (2.26-2.26)
Ramachandran outliers	138981	1449 (2.26-2.26)
Sidechain outliers	138945	1450 (2.26-2.26)

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 <=5%

Mol	Chain	Length	Quality of chain
1	A	141	74% · 21%
1	B	141	76% · 21%
2	C	20	95% 5%
3	E	20	80% 20%

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 2755 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 Interferon regulatory factor 4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	111	Total	C	N	O	S	0	0	0
			927	596	162	168	1			
1	B	111	Total	C	N	O	S	0	0	0
			927	596	162	168	1			

There are 42 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-1	MET	-	initiating methionine	UNP Q15306
A	0	GLY	-	expression tag	UNP Q15306
A	1	SER	-	expression tag	UNP Q15306
A	2	HIS	-	expression tag	UNP Q15306
A	3	HIS	-	expression tag	UNP Q15306
A	4	HIS	-	expression tag	UNP Q15306
A	5	HIS	-	expression tag	UNP Q15306
A	6	HIS	-	expression tag	UNP Q15306
A	7	HIS	-	expression tag	UNP Q15306
A	8	SER	-	expression tag	UNP Q15306
A	9	ALA	-	expression tag	UNP Q15306
A	10	ALA	-	expression tag	UNP Q15306
A	11	LEU	-	expression tag	UNP Q15306
A	12	GLU	-	expression tag	UNP Q15306
A	13	VAL	-	expression tag	UNP Q15306
A	14	LEU	-	expression tag	UNP Q15306
A	15	PHE	-	expression tag	UNP Q15306
A	16	GLN	-	expression tag	UNP Q15306
A	17	GLY	-	expression tag	UNP Q15306
A	18	PRO	-	expression tag	UNP Q15306
A	19	GLY	-	expression tag	UNP Q15306
B	-1	MET	-	initiating methionine	UNP Q15306
B	0	GLY	-	expression tag	UNP Q15306
B	1	SER	-	expression tag	UNP Q15306
B	2	HIS	-	expression tag	UNP Q15306

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	3	HIS	-	expression tag	UNP Q15306
B	4	HIS	-	expression tag	UNP Q15306
B	5	HIS	-	expression tag	UNP Q15306
B	6	HIS	-	expression tag	UNP Q15306
B	7	HIS	-	expression tag	UNP Q15306
B	8	SER	-	expression tag	UNP Q15306
B	9	ALA	-	expression tag	UNP Q15306
B	10	ALA	-	expression tag	UNP Q15306
B	11	LEU	-	expression tag	UNP Q15306
B	12	GLU	-	expression tag	UNP Q15306
B	13	VAL	-	expression tag	UNP Q15306
B	14	LEU	-	expression tag	UNP Q15306
B	15	PHE	-	expression tag	UNP Q15306
B	16	GLN	-	expression tag	UNP Q15306
B	17	GLY	-	expression tag	UNP Q15306
B	18	PRO	-	expression tag	UNP Q15306
B	19	GLY	-	expression tag	UNP Q15306

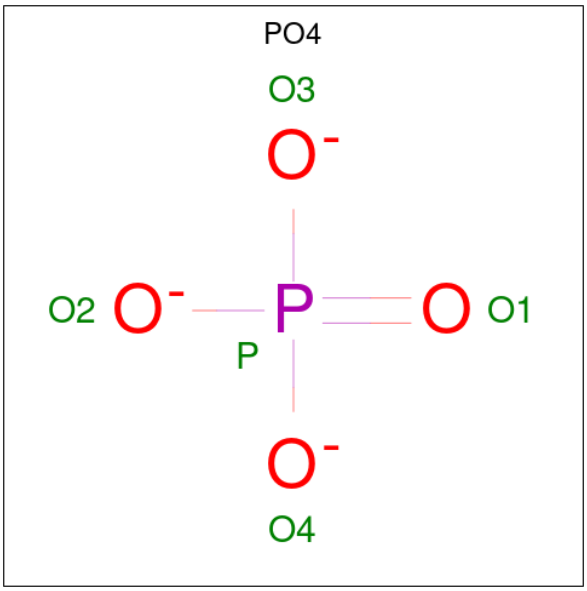
- Molecule 2 is a DNA chain called DNA (5'-D(P*TP*CP*AP*AP*CP*TP*GP*AP*AP*AP *CP*CP*GP*AP*GP*AP*AP*AP*GP*C)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	20	Total	C	N	O	P	0	0	0
			412	195	84	113	20			

- Molecule 3 is a DNA chain called DNA (5'-D(P*AP*GP*CP*TP*TP*TP*CP*TP*CP*GP *GP*TP*TP*TP*CP*AP*GP*TP*TP*G)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	E	20	Total	C	N	O	P	0	0	0
			408	196	65	127	20			

- Molecule 4 is PHOSPHATE ION (three-letter code: PO4) (formula: O_4P).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	B	1	Total O P 5 4 1	0	0

- Molecule 5 is water.

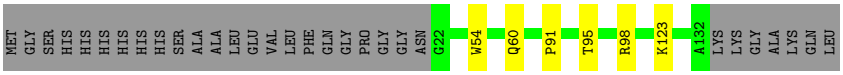
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	25	Total O 25 25	0	0
5	B	28	Total O 28 28	0	0
5	C	10	Total O 10 10	0	0
5	E	13	Total O 13 13	0	0

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. 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. 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.

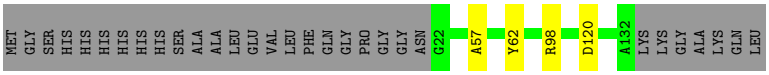
• Molecule 1: Interferon regulatory factor 4

Chain A:



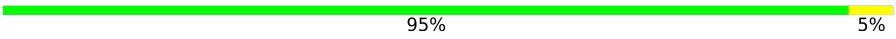
• Molecule 1: Interferon regulatory factor 4

Chain B:



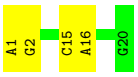
• Molecule 2: DNA (5'-D(P*TP*CP*AP*AP*CP*TP*GP*AP*AP*AP*CP*CP*GP*AP*GP*AP*AP*AP*GP*C)-3')

Chain C:



• Molecule 3: DNA (5'-D(P*AP*GP*CP*TP*TP*TP*CP*TP*CP*GP*GP*TP*TP*TP*CP*AP*GP*TP*TP*G)-3')

Chain E:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	64.78Å 83.25Å 88.53Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	44.27 – 2.25 44.26 – 2.25	Depositor EDS
% Data completeness (in resolution range)	99.1 (44.27-2.25) 99.1 (44.26-2.25)	Depositor EDS
R_{merge}	0.11	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.42 (at 2.24Å)	Xtriage
Refinement program	PHENIX 1.18.2_3874	Depositor
R, R_{free}	0.202 , 0.238 0.202 , 0.238	Depositor DCC
R_{free} test set	1116 reflections (4.82%)	wwPDB-VP
Wilson B-factor (Å^2)	45.6	Xtriage
Anisotropy	0.320	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	(Not available) , (Not available)	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	2755	wwPDB-VP
Average B, all atoms (Å^2)	54.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.74% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: PO4

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	0.46	0/953	0.56	0/1286
1	B	0.43	0/953	0.53	0/1286
2	C	0.89	0/464	0.87	0/713
3	E	0.93	0/454	1.13	0/699
All	All	0.63	0/2824	0.74	0/3984

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-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 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	927	0	910	5	0
1	B	927	0	910	3	0
2	C	412	0	223	1	0
3	E	408	0	230	3	0
4	B	5	0	0	0	0
5	A	25	0	0	3	0
5	B	28	0	0	1	0
5	C	10	0	0	0	0
5	E	13	0	0	0	0
All	All	2755	0	2273	11	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 2.

All (11) 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:98:ARG:NH1	5:A:202:HOH:O	2.34	0.59
1:A:98:ARG:NH2	5:A:203:HOH:O	2.36	0.58
3:E:15:DC:H2'	3:E:16:DA:C8	2.37	0.58
1:B:98:ARG:NH2	5:B:301:HOH:O	2.39	0.54
1:A:60:GLN:HA	5:A:213:HOH:O	2.10	0.50
1:A:54:TRP:CD2	1:A:123:LYS:HD2	2.48	0.48
1:A:91:PRO:O	1:A:95:THR:HG23	2.14	0.47
3:E:1:DA:H2"	3:E:2:DG:C8	2.52	0.43
1:B:98:ARG:HH22	2:C:6:DT:H72	1.84	0.43
1:B:57:ALA:HA	1:B:62:TYR:CD1	2.56	0.41
3:E:1:DA:H5'	3:E:1:DA:N3	2.35	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	109/141 (77%)	105 (96%)	4 (4%)	0	100	100
1	B	109/141 (77%)	106 (97%)	3 (3%)	0	100	100
All	All	218/282 (77%)	211 (97%)	7 (3%)	0	100	100

There are no Ramachandran outliers to report.

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	97/119 (82%)	97 (100%)	0	100 100
1	B	97/119 (82%)	96 (99%)	1 (1%)	76 84
All	All	194/238 (82%)	193 (100%)	1 (0%)	88 92

All (1) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	120	ASP

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

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 monosaccharides 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	$\#\|Z\| > 2$
4	PO4	B	200	-	4,4,4	0.86	0	6,6,6	1.08	0

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

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

Unable to reproduce the depositors R factor - this section is therefore empty.

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

Unable to reproduce the depositors R factor - this section is therefore empty.

6.3 Carbohydrates

Unable to reproduce the depositors R factor - this section is therefore empty.

6.4 Ligands

Unable to reproduce the depositors R factor - this section is therefore empty.

6.5 Other polymers

Unable to reproduce the depositors R factor - this section is therefore empty.