WORLDWIDE PDB PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report ⓘ

Jun 18, 2024 – 01:21 AM EDT

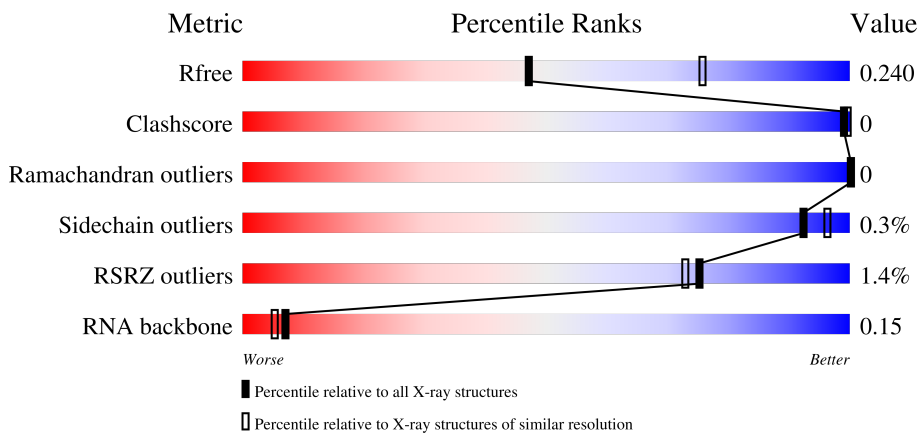|		
PDB ID	:	5W5I
Title	:	Human IFIT1 dimer with PPP-AAAA
Authors	:	Abbas, Y.M.; Martinez-Montero, S.; Damha, M.J.; Nagar, B.
Deposited on	:	2017-06-15
Resolution	:	2.65 Å(reported)

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 2.65 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	Percentile Ranks	Value
Rfree		0.240
Clashscore		0
Ramachandran outliers		0
Sidechain outliers		0.3%
RSRZ outliers		1.4%
RNA backbone		0.15

Worse — *Better*

■ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	1332 (2.68-2.64)
Clashscore	141614	1374 (2.68-2.64)
Ramachandran outliers	138981	1349 (2.68-2.64)
Sidechain outliers	138945	1349 (2.68-2.64)
RSRZ outliers	127900	1318 (2.68-2.64)
RNA backbone	3102	1010 (2.96-2.36)

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% 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	479	89% • 10%
1	C	479	90% • 9%
2	B	4	100%
2	D	4	100%

WORLDWIDE PDB
PROTEIN DATA BANK

2   Entry composition ⓘ

There are 3 unique types of molecules in this entry. The entry contains 14520 atoms, of which 7219 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-induced protein with tetratricopeptide repeats 1.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	430	Total 7048	C 2242	H 3537	N 606	O 644	S 19	0	0	0
1	C	436	Total 7153	C 2272	H 3592	N 617	O 653	S 19	0	0	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	0	SER	-	expression tag	UNP P09914
C	0	SER	-	expression tag	UNP P09914

- Molecule 2 is a RNA chain called RNA (5'-D(*(ATP))-R(P*AP*AP*A)-3').

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
2	B	4	Total 142	C 40	H 45	N 20	O 31	P 6	0	0	0
2	D	4	Total 142	C 40	H 45	N 20	O 31	P 6	0	0	0

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	8	Total 8	O 8	0	0
3	B	1	Total 1	O 1	0	0
3	C	21	Total 21	O 21	0	0
3	D	5	Total 5	O 5	0	0

3 Residue-property plots ⓘ

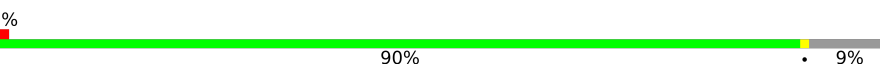These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. 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.

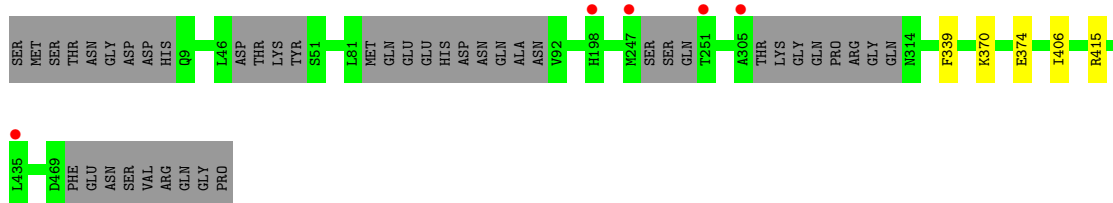• Molecule 1: Interferon-induced protein with tetratricopeptide repeats 1

Chain A:
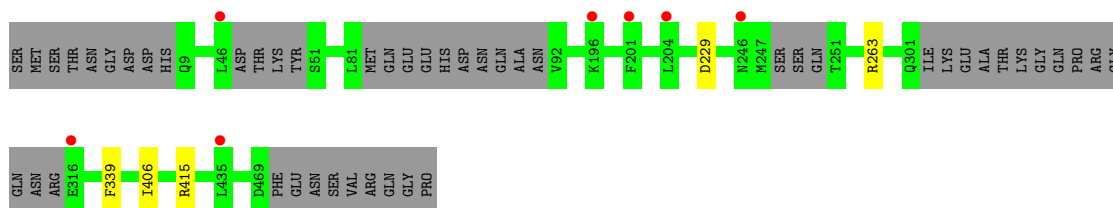
• Molecule 1: Interferon-induced protein with tetratricopeptide repeats 1

Chain C:

• Molecule 2: RNA (5'-D(*(ATP))-R(P*AP*AP*A)-3')

Chain B:

There are no outlier residues recorded for this chain.

• Molecule 2: RNA (5'-D(*(ATP))-R(P*AP*AP*A)-3')

Chain D:

There are no outlier residues recorded for this chain.

4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	I 4	Depositor
Cell constants a, b, c, α, β, γ	184.79Å 184.79Å 88.86Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	44.40 – 2.65 44.40 – 2.65	Depositor EDS
% Data completeness (in resolution range)	88.6 (44.40-2.65) 88.6 (44.40-2.65)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.06	Depositor
$< I/\sigma(I) >$ [1]	4.98 (at 2.65Å)	Xtriage
Refinement program	PHENIX (1.10_2142: ???)	Depositor
R, R_{free}	0.219 , 0.242 0.219 , 0.240	Depositor DCC
R_{free} test set	1922 reflections (4.97%)	wwPDB-VP
Wilson B-factor (Å^2)	40.5	Xtriage
Anisotropy	0.065	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.37 , 30.3	EDS
L-test for twinning[2]	$< |L| > = 0.48, < L^2 > = 0.30$	Xtriage
Estimated twinning fraction	0.057 for -k,-h,-l	Xtriage
F_o,F_c correlation	0.91	EDS
Total number of atoms	14520	wwPDB-VP
Average B, all atoms (Å^2)	57.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.68% 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: ATP

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	0.25	0/3575	0.37	0/4799
1	C	0.24	0/3625	0.36	0/4865
2	B	0.16	0/74	0.67	0/113
2	D	0.32	0/74	0.81	0/113
All	All	0.24	0/7348	0.38	0/9890

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	3511	3537	3535	2	0
1	C	3561	3592	3589	2	0
2	B	97	45	45	0	0
2	D	97	45	45	0	0
3	A	8	0	0	0	0
3	B	1	0	0	0	0
3	C	21	0	0	0	0
3	D	5	0	0	0	0
All	All	7301	7219	7214	4	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 0.

All (4) 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:406:ILE:O	1:A:415:ARG:NH1	2.32	0.62
1:A:229:ASP:OD1	1:A:263:ARG:NH1	2.40	0.54
1:C:406:ILE:O	1:C:415:ARG:NH1	2.41	0.54
1:C:370:LYS:HE2	1:C:374:GLU:OE2	2.11	0.51

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	420/479 (88%)	413 (98%)	7 (2%)	0	100	100
1	C	426/479 (89%)	418 (98%)	8 (2%)	0	100	100
All	All	846/958 (88%)	831 (98%)	15 (2%)	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	
1	A	375/418 (90%)	374 (100%)	1 (0%)	92	96

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	C	380/418 (91%)	379 (100%)	1 (0%)	92	96
All	All	755/836 (90%)	753 (100%)	2 (0%)	92	96

All (2) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	339	PHE
1	C	339	PHE

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

5.3.3 RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
2	B	2/4 (50%)	0	0
2	D	2/4 (50%)	0	0
All	All	4/8 (50%)	0	0

There are no RNA backbone outliers to report.

There are no RNA pucker outliers to report.

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 monosaccharides in this entry.

5.6 Ligand geometry ⓘ

There are no ligands in this entry.

5.7 Other polymers ⓘ

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(Å²)	Q<0.9
1	A	430/479 (89%)	0.22	7 (1%)	72	69	26, 56, 87, 111	0
1	C	436/479 (91%)	0.04	5 (1%)	80	79	19, 44, 78, 110	0
2	B	3/4 (75%)	-0.09	0	100	100	43, 43, 44, 52	0
2	D	3/4 (75%)	0.19	0	100	100	28, 28, 29, 38	0
All	All	872/966 (90%)	0.13	12 (1%)	75	73	19, 50, 84, 111	0

The worst 5 of 12 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	247	MET	3.7
1	C	435	LEU	3.3
1	C	305	ALA	3.1
1	C	251	THR	2.7
1	A	46	LEU	2.4

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 monosaccharides 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.