wwPDB X-ray Structure Validation Summary Report (i)

Dec 2, 2023 – 03:05 pm GMT

PDB ID	:	2J0U
Title	:	The crystal structure of eIF4AIII-Barentsz complex at 3.0 A resolution
Authors	:	Bono, F.; Ebert, J.; Lorentzen, E.; Conti, E.
Deposited on	:	2006-08-04
Resolution	:	3.00 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 (i) 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 (i)) were used in the production of this report:

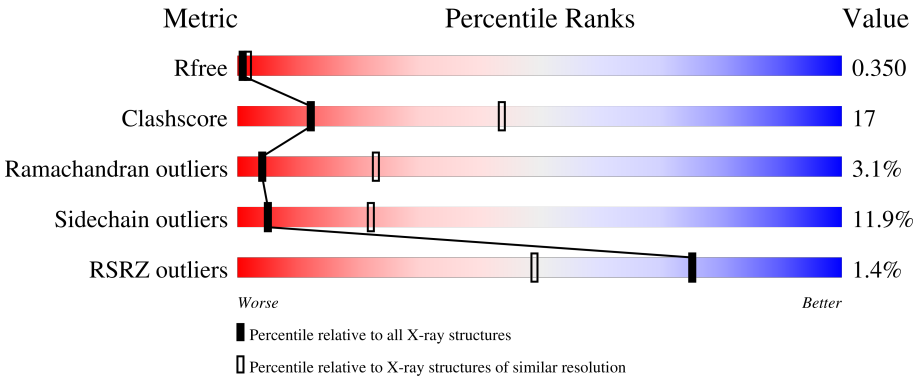
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

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 3.00 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	2092 (3.00-3.00)
Clashscore	141614	2416 (3.00-3.00)
Ramachandran outliers	138981	2333 (3.00-3.00)
Sidechain outliers	138945	2336 (3.00-3.00)
RSRZ outliers	127900	1990 (3.00-3.00)

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	374	61% 29% 7%
2	B	374	2% 64% 24% 7%
3	T	114	5% 92%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 5139 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 ATP-DEPENDENT RNA HELICASE DDX48.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	366	Total	C	N	O	S	0	0	0
			2654	1682	460	497	15			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	94	VAL	ILE	conflict	UNP P38919
A	122	VAL	ILE	conflict	UNP P38919
A	378	VAL	ILE	conflict	UNP P38919
A	388	VAL	ILE	conflict	UNP P38919

- Molecule 2 is a protein called ATP-DEPENDENT RNA HELICASE DDX48.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	347	Total	C	N	O	S	0	0	0
			2413	1543	412	446	12			

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	94	VAL	ILE	conflict	UNP P38919
B	122	VAL	ILE	conflict	UNP P38919
B	137	SER	CYS	conflict	UNP P38919
B	225	VAL	ILE	conflict	UNP P38919
B	378	VAL	ILE	conflict	UNP P38919
B	388	VAL	ILE	conflict	UNP P38919

- Molecule 3 is a protein called PROTEIN CASC3.

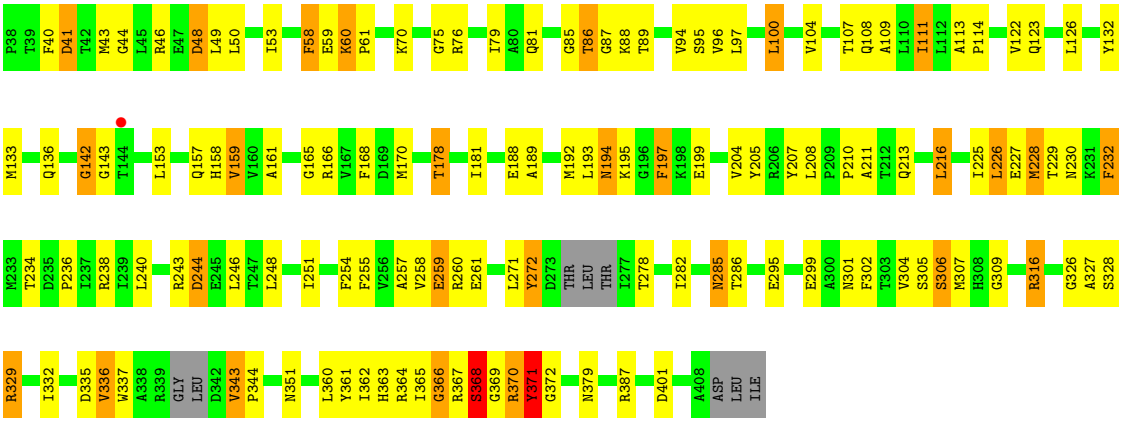
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
3	T	9	Total	C	N	O	0	0	0
			72	47	12	13			

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.

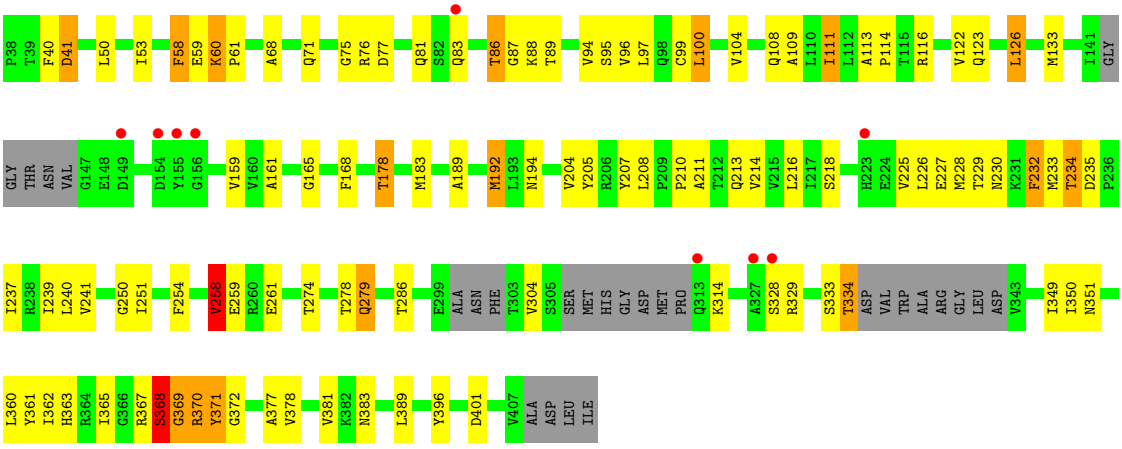
• Molecule 1: ATP-DEPENDENT RNA HELICASE DDX48

Chain A:



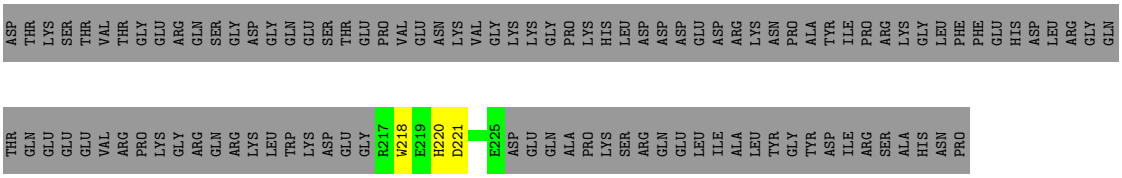
• Molecule 2: ATP-DEPENDENT RNA HELICASE DDX48

Chain B:



• Molecule 3: PROTEIN CASC3

Chain T:


ASP THR LYS SER THR VAL THR GLY GLU ARG GLN SER GLY ASP GLY GLN GLU SER THR GLU PRO VAL GLU ASN LYS VAL GLY LYS LYS GLY PRO LYS HIS LEU ASP ASP ASP GLU ASP ARG LYS ASN PRO ALA TYR ILE PRO ARG LYS GLY LEU PHE GLU HIS ASP LEU ARG GLY GLN
THR GLN GLU GLU GLU VAL ARG PRO LYS GLY ARG GLN ARG LYS LEU TRP LYS ASP GLU GLY R217 W218 E219 H220 D221 E225 ASP GLU GLN ALA PRO LYS SER ARG GLN GLU LEU ILE ALA LEU TYR GLY TYR ASP ILE ARG SER ALA HIS ASN PRO

4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, α, β, γ	71.76Å 107.19Å 243.95Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	15.00 – 3.00 14.96 – 3.00	Depositor EDS
% Data completeness (in resolution range)	100.0 (15.00-3.00) 94.3 (14.96-3.00)	Depositor EDS
R_{merge}	0.22	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.63 (at 3.01Å)	Xtriage
Refinement program	REFMAC 5.2.0005	Depositor
R, R_{free}	0.273 , 0.328 0.298 , 0.350	Depositor DCC
R_{free} test set	902 reflections (4.99%)	wwPDB-VP
Wilson B-factor (Å^2)	70.0	Xtriage
Anisotropy	0.177	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.31 , 43.6	EDS
L-test for twinning[2]	$< \|L\| > = 0.45$, $< L^2 > = 0.27$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.84	EDS
Total number of atoms	5139	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 3.99% 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

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	0.56	0/2694	0.74	1/3671 (0.0%)
2	B	0.53	0/2450	0.70	2/3355 (0.1%)
3	T	0.59	0/75	0.56	0/102
All	All	0.55	0/5219	0.72	3/7128 (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	A	0	5
2	B	1	3
All	All	1	8

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	368	SER	N-CA-C	5.79	126.64	111.00
1	A	159	VAL	CG1-CB-CG2	5.57	119.82	110.90
2	B	258	VAL	CG1-CB-CG2	5.07	119.01	110.90

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
2	B	368	SER	CA

5 of 8 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	210	PRO	Peptide

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Group
1	A	261	GLU	Peptide
1	A	343	VAL	Peptide
1	A	366	GLY	Peptide
1	A	370	ARG	Peptide

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	2654	0	2472	97	0
2	B	2413	0	2137	70	0
3	T	72	0	43	4	0
All	All	5139	0	4652	167	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 17.

The worst 5 of 167 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)
2:B:370:ARG:O	2:B:372:GLY:N	2.02	0.93
1:A:75:GLY:C	1:A:213:GLN:HE22	1.77	0.87
2:B:225:VAL:O	2:B:228:MET:HB3	1.87	0.75
2:B:367:ARG:HA	2:B:368:SER:CB	2.19	0.73
1:A:75:GLY:C	1:A:213:GLN:NE2	2.41	0.72

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	360/374 (96%)	314 (87%)	34 (9%)	12 (3%)	4 21
2	B	337/374 (90%)	294 (87%)	34 (10%)	9 (3%)	5 26
3	T	7/114 (6%)	5 (71%)	1 (14%)	1 (14%)	0 1
All	All	704/862 (82%)	613 (87%)	69 (10%)	22 (3%)	4 23

5 of 22 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	86	THR
1	A	336	VAL
1	A	370	ARG
2	B	86	THR
2	B	211	ALA

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	251/329 (76%)	217 (86%)	34 (14%)	4 17
2	B	208/329 (63%)	187 (90%)	21 (10%)	7 29
3	T	5/100 (5%)	5 (100%)	0	100 100
All	All	464/758 (61%)	409 (88%)	55 (12%)	5 22

5 of 55 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	316	ARG
2	B	58	PHE
2	B	334	THR
2	B	234	THR
1	A	328	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 7 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	253	GLN
1	A	285	ASN
2	B	213	GLN
1	A	351	ASN
1	A	213	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 monosaccharides 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

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, 95^{th} 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	366/374 (97%)	-0.19	1 (0%) 94 84	40, 53, 62, 69	0
2	B	347/374 (92%)	-0.01	9 (2%) 56 27	33, 52, 61, 68	0
3	T	9/114 (7%)	-0.34	0 100 100	60, 66, 67, 67	0
All	All	722/862 (83%)	-0.11	10 (1%) 75 49	33, 53, 62, 69	0

The worst 5 of 10 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	155	TYR	4.8
2	B	223	HIS	2.7
2	B	156	GLY	2.7
2	B	149	ASP	2.6
2	B	83	GLN	2.6

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 monosaccharides in this entry.

6.4 Ligands

There are no ligands in this entry.

6.5 Other polymers (i)

There are no such residues in this entry.