



Full wwPDB X-ray Structure Validation Report ⓘ

May 24, 2020 – 04:08 am BST

PDB ID : 5YVI
Title : Crystal structure of Karyopherin beta2 in complex with FUS(456-526)
Authors : Yoshizawa, T.; Fung, H.Y.J.; Chook, Y.M.
Deposited on : 2017-11-25
Resolution : 2.90 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
Xtriage (Phenix)	:	1.13
EDS	:	2.11
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)	:	2.11

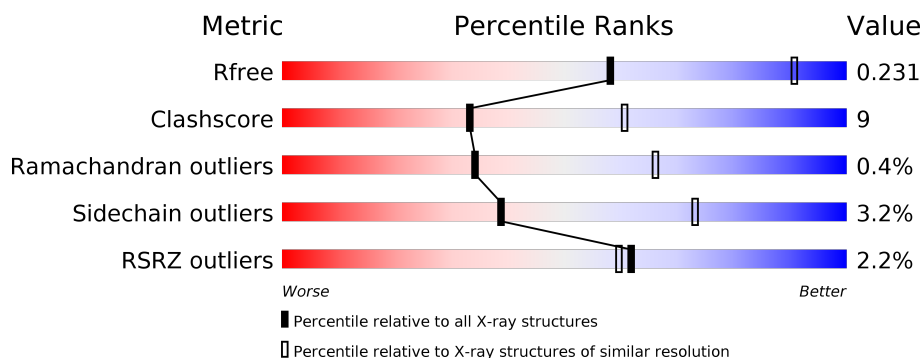
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.90 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	1957 (2.90-2.90)
Clashscore	141614	2172 (2.90-2.90)
Ramachandran outliers	138981	2115 (2.90-2.90)
Sidechain outliers	138945	2117 (2.90-2.90)
RSRZ outliers	127900	1906 (2.90-2.90)

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 on 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	868	<div> <div>2%</div> <div> <div></div> <div>76%</div> <div>18%</div> <div>• 5%</div> </div> </div>
2	B	73	<div> <div>16%</div> <div>10%</div> <div>74%</div> </div>

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 6760 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 Transportin-1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	828	6594	4229	1100	1214	51	0	0	0

There are 9 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-1	GLY	-	expression tag	UNP Q92973
A	0	SER	-	expression tag	UNP Q92973
A	361	GLY	-	linker	UNP Q92973
A	362	GLY	-	linker	UNP Q92973
A	363	SER	-	linker	UNP Q92973
A	364	GLY	-	linker	UNP Q92973
A	365	GLY	-	linker	UNP Q92973
A	366	SER	-	linker	UNP Q92973
A	367	GLY	-	linker	UNP Q92973

- Molecule 2 is a protein called RNA-binding protein FUS.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	19	166	96	38	31	1	0	0	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	454	GLY	-	expression tag	UNP P35637
B	455	SER	-	expression tag	UNP P35637

4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, α , β , γ	129.04Å 156.10Å 68.25Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	44.90 – 2.90 44.90 – 2.73	Depositor EDS
% Data completeness (in resolution range)	93.6 (44.90-2.90) 85.6 (44.90-2.73)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.86 (at 2.73Å)	Xtriage
Refinement program	PHENIX (1.11.1 _2575: ???)	Depositor
R, R_{free}	0.179 , 0.231 0.179 , 0.231	Depositor DCC
R_{free} test set	1602 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å ²)	47.3	Xtriage
Anisotropy	0.161	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.36 , 53.6	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	6760	wwPDB-VP
Average B, all atoms (Å ²)	51.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.60% 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

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	0.42	0/6734	0.56	0/9142
2	B	0.55	0/169	0.66	0/220
All	All	0.43	0/6903	0.56	0/9362

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	6594	0	6636	111	0
2	B	166	0	156	13	0
All	All	6760	0	6792	116	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 9.

All (116) 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:514:ARG:HD2	2:B:518:ARG:NH2	1.84	0.91
2:B:514:ARG:CD	2:B:518:ARG:NH2	2.37	0.86

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:B:514:ARG:HD2	2:B:518:ARG:HH21	1.40	0.82
1:A:5:TRP:O	1:A:6:LYS:NZ	2.16	0.73
1:A:163:LEU:HD13	1:A:206:THR:HG23	1.69	0.73
1:A:668:ILE:HD11	1:A:690:LEU:HD21	1.71	0.71
1:A:254:MET:CE	1:A:292:VAL:HG11	2.21	0.70
1:A:163:LEU:HB2	1:A:172:LEU:HD12	1.74	0.68
1:A:770:ASN:HD21	2:B:509:GLY:HA2	1.58	0.67
1:A:546:GLY:HA3	2:B:518:ARG:NH1	2.09	0.67
1:A:214:ASP:O	1:A:218:GLU:HG2	1.95	0.67
1:A:759:ASN:ND2	1:A:759:ASN:O	2.28	0.66
1:A:254:MET:HE3	1:A:292:VAL:HG11	1.77	0.66
1:A:642:ILE:HD11	1:A:680:MET:HE3	1.77	0.66
1:A:725:CYS:O	1:A:729:THR:HG23	1.96	0.65
1:A:739:MET:HG3	1:A:743:MET:HE2	1.79	0.65
1:A:820:MET:O	1:A:822:SER:N	2.31	0.64
1:A:642:ILE:HD11	1:A:680:MET:CE	2.27	0.64
1:A:736:SER:HA	1:A:743:MET:HE3	1.79	0.63
1:A:828:VAL:HG12	1:A:832:PHE:HD1	1.64	0.63
1:A:388:ASN:HD21	1:A:426:GLU:HG2	1.64	0.62
1:A:743:MET:HG3	1:A:743:MET:O	1.99	0.61
1:A:418:ILE:O	1:A:461:THR:HG21	2.01	0.61
1:A:479:LYS:HE3	1:A:518:TYR:HE1	1.64	0.61
1:A:765:LYS:O	1:A:769:GLU:HG3	2.00	0.61
1:A:127:PRO:O	1:A:168:LEU:HD11	2.01	0.60
1:A:450:LYS:HD3	1:A:450:LYS:H	1.67	0.59
1:A:502:SER:HB3	2:B:522:ARG:HG2	1.83	0.59
1:A:122:GLU:HB3	1:A:124:GLN:HE22	1.67	0.58
1:A:877:PRO:HD2	1:A:880:LEU:HD23	1.86	0.58
1:A:592:SER:OG	2:B:518:ARG:NH2	2.30	0.58
1:A:779:GLY:HA3	1:A:820:MET:SD	2.43	0.57
1:A:381:ALA:O	1:A:385:VAL:HG23	2.05	0.56
1:A:32:THR:O	1:A:36:LYS:HG3	2.06	0.56
2:B:514:ARG:HD3	2:B:518:ARG:NH2	2.21	0.55
2:B:509:GLY:O	2:B:510:LYS:HB3	2.07	0.55
1:A:109:THR:O	1:A:112:ILE:HG22	2.08	0.54
1:A:88:THR:CG2	1:A:92:LYS:HE3	2.38	0.54
1:A:493:ASN:HD22	1:A:496:VAL:H	1.55	0.54
1:A:119:SER:OG	1:A:120:LYS:N	2.41	0.53
1:A:121:GLY:O	1:A:125:ASN:ND2	2.31	0.53
1:A:288:ILE:HG12	1:A:292:VAL:HG23	1.91	0.53
1:A:124:GLN:NE2	1:A:124:GLN:H	2.08	0.52

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:254:MET:HE1	1:A:292:VAL:HG11	1.90	0.52
1:A:749:MET:O	1:A:753:GLN:NE2	2.43	0.52
1:A:546:GLY:HA3	2:B:518:ARG:HH12	1.74	0.52
1:A:736:SER:HA	1:A:743:MET:CE	2.40	0.51
1:A:804:ILE:HD11	1:A:810:LYS:HG3	1.93	0.51
1:A:294:VAL:HG13	1:A:295:ARG:HG2	1.93	0.51
1:A:773:ILE:HD13	1:A:809:GLU:HB3	1.93	0.50
1:A:699:PHE:CZ	1:A:739:MET:HE3	2.46	0.50
1:A:699:PHE:HZ	1:A:739:MET:HE3	1.76	0.50
1:A:660:GLU:HB2	1:A:698:CYS:HB3	1.93	0.50
1:A:814:PHE:O	1:A:817:ILE:HG22	2.12	0.49
1:A:770:ASN:ND2	2:B:509:GLY:HA2	2.26	0.49
1:A:493:ASN:HD22	1:A:496:VAL:HG23	1.76	0.49
1:A:584:PHE:HB2	1:A:585:PRO:HD3	1.94	0.49
1:A:457:ILE:O	1:A:461:THR:HG23	2.12	0.49
1:A:153:GLN:O	1:A:157:GLU:HG3	2.13	0.49
1:A:227:GLU:O	1:A:232:ARG:NH2	2.39	0.48
1:A:783:PRO:HB3	1:A:819:THR:HG22	1.95	0.48
1:A:398:ILE:HG21	1:A:424:ILE:HD13	1.96	0.48
1:A:852:MET:O	1:A:856:ILE:HD12	2.13	0.48
1:A:587:LEU:HB2	1:A:643:VAL:CG1	2.44	0.48
1:A:23:GLN:OE1	1:A:64:THR:HG22	2.14	0.47
1:A:88:THR:HG22	1:A:92:LYS:HE3	1.96	0.47
1:A:92:LYS:HD3	1:A:126:TRP:CE2	2.49	0.47
1:A:703:LYS:HB3	1:A:704:PRO:HD3	1.96	0.47
1:A:482:MET:HG2	1:A:504:PHE:HE1	1.80	0.47
1:A:163:LEU:O	1:A:164:ASP:HB2	2.15	0.46
1:A:22:SER:HA	1:A:33:VAL:HG11	1.98	0.46
1:A:625:LEU:HD13	1:A:633:TYR:CG	2.51	0.46
1:A:793:PHE:HZ	1:A:820:MET:HE1	1.80	0.46
1:A:782:CYS:O	1:A:786:VAL:HG23	2.16	0.46
1:A:243:LEU:HD23	1:A:282:THR:HG22	1.96	0.45
1:A:545:ILE:HD13	1:A:565:LEU:HD11	1.99	0.45
1:A:48:ASN:OD1	1:A:87:VAL:HG13	2.16	0.45
1:A:536:LYS:HE3	1:A:537:ASN:HD21	1.81	0.45
1:A:498:GLU:HB2	1:A:537:ASN:ND2	2.32	0.45
1:A:261:MET:HG3	1:A:279:PHE:CD2	2.52	0.44
1:A:224:ALA:HB1	1:A:260:TYR:CE1	2.52	0.44
1:A:312:ASP:HA	1:A:315:ILE:HD12	1.99	0.44
1:A:625:LEU:HA	1:A:625:LEU:HD23	1.84	0.44
1:A:254:MET:HE3	1:A:292:VAL:HG21	1.99	0.44

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:443:LEU:HD22	1:A:458:THR:HG23	2.00	0.43
1:A:7:PRO:HG3	1:A:52:PHE:CZ	2.52	0.43
1:A:508:GLU:HA	1:A:515:LEU:HD11	1.99	0.43
1:A:807:ASN:HB2	1:A:808:GLU:OE1	2.18	0.43
1:A:214:ASP:CG	1:A:249:ARG:HH11	2.21	0.43
1:A:242:LEU:HA	1:A:242:LEU:HD23	1.78	0.43
1:A:602:LEU:HA	1:A:602:LEU:HD23	1.75	0.43
1:A:229:PRO:HB3	1:A:269:ASP:OD2	2.19	0.43
1:A:122:GLU:HB3	1:A:124:GLN:NE2	2.31	0.42
1:A:168:LEU:HB3	1:A:169:ASP:H	1.67	0.42
1:A:828:VAL:HG12	1:A:832:PHE:CD1	2.49	0.42
1:A:254:MET:HE3	1:A:292:VAL:CG1	2.48	0.42
1:A:502:SER:HB3	2:B:522:ARG:CG	2.48	0.42
1:A:11:GLY:O	1:A:15:ILE:HG12	2.19	0.42
1:A:260:TYR:CZ	1:A:264:ARG:HD2	2.55	0.42
1:A:770:ASN:HD21	2:B:510:LYS:H	1.66	0.42
1:A:398:ILE:HA	1:A:398:ILE:HD13	1.81	0.42
1:A:479:LYS:HE3	1:A:518:TYR:CE1	2.50	0.42
1:A:310:TYR:HB3	1:A:315:ILE:HG13	2.02	0.41
1:A:539:LEU:HD22	1:A:585:PRO:HG2	2.02	0.41
1:A:395:LEU:HA	1:A:395:LEU:HD23	1.86	0.41
1:A:26:ASP:HB3	1:A:29:ILE:HG22	2.03	0.41
1:A:113:LEU:HA	1:A:113:LEU:HD12	1.77	0.41
1:A:251:LEU:HA	1:A:251:LEU:HD23	1.79	0.41
1:A:130:LEU:HB2	1:A:131:PRO:HD3	2.03	0.41
1:A:501:CYS:HB3	1:A:540:ILE:HG21	2.03	0.41
1:A:614:ASN:O	1:A:618:LYS:HG3	2.21	0.41
1:A:868:ASN:N	1:A:868:ASN:OD1	2.54	0.41
1:A:191:ARG:HD2	1:A:223:LEU:HD21	2.03	0.41
1:A:233:LYS:HG3	1:A:275:GLU:HG3	2.02	0.41
1:A:58:LYS:H	1:A:58:LYS:HG2	1.75	0.40
1:A:808:GLU:H	1:A:808:GLU:CD	2.16	0.40

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	818/868 (94%)	803 (98%)	12 (2%)	3 (0%)	34	66
2	B	17/73 (23%)	15 (88%)	2 (12%)	0	100	100
All	All	835/941 (89%)	818 (98%)	14 (2%)	3 (0%)	34	66

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	535	HIS
1	A	598	GLN
1	A	862	ASN

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	741/776 (96%)	719 (97%)	22 (3%)	41	75
2	B	17/43 (40%)	15 (88%)	2 (12%)	5	16
All	All	758/819 (93%)	734 (97%)	24 (3%)	39	73

All (24) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	59	SER
1	A	81	GLN
1	A	119	SER
1	A	149	PHE
1	A	163	LEU
1	A	164	ASP
1	A	186	SER
1	A	204	SER
1	A	249	ARG

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	270	GLU
1	A	450	LYS
1	A	488	ARG
1	A	501	CYS
1	A	625	LEU
1	A	649	SER
1	A	759	ASN
1	A	812	SER
1	A	822	SER
1	A	847	ASP
1	A	863	GLN
1	A	869	TRP
1	A	887	PHE
2	B	512	ASP
2	B	513	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (16) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	81	GLN
1	A	124	GLN
1	A	153	GLN
1	A	388	ASN
1	A	493	ASN
1	A	497	GLN
1	A	535	HIS
1	A	537	ASN
1	A	554	HIS
1	A	563	GLN
1	A	629	GLN
1	A	632	GLN
1	A	685	GLN
1	A	726	ASN
1	A	770	ASN
1	A	791	GLN

5.3.3 RNA ⓘ

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	A	828/868 (95%)	-0.22	19 (2%) 60 58	16, 47, 95, 126	0
2	B	19/73 (26%)	-0.22	0 100 100	18, 39, 86, 92	0
All	All	847/941 (90%)	-0.22	19 (2%) 62 59	16, 47, 93, 126	0

All (19) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	870	ARG	3.3
1	A	792	GLN	3.0
1	A	829	ILE	3.0
1	A	789	MET	2.9
1	A	762	ASN	2.9
1	A	791	GLN	2.8
1	A	869	TRP	2.7
1	A	872	PHE	2.7
1	A	535	HIS	2.7
1	A	860	PHE	2.6
1	A	27	THR	2.6
1	A	833	ILE	2.5
1	A	828	VAL	2.5
1	A	857	LEU	2.3
1	A	871	ARG	2.3
1	A	876	PHE	2.2
1	A	751	LEU	2.1
1	A	28	THR	2.1
1	A	31	ARG	2.0

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 [i](#)

There are no such residues in this entry.