wwPDB X-ray Structure Validation Summary Report (i)

May 21, 2020 – 06:17 am BST

PDB ID	:	3WKM
Title	:	The periplasmic PDZ tandem fragment of the RseP homologue from Aquifex aeolicus in complex with the Fab fragment
Authors	:	Nogi, T.; Tabata, S.; Tamura-kawakami, K.; Takagi, J.
Deposited on	:	2013-10-28
Resolution	:	2.20 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 following versions of software and data (see references (1)) were used in the production of this report:

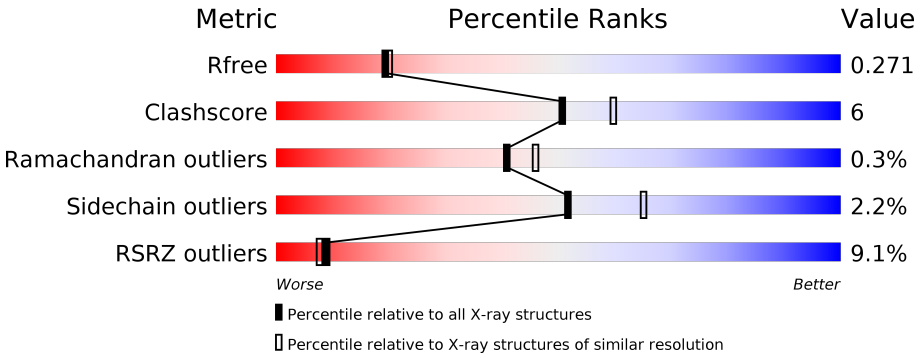
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

1 Overall quality at a glance (i)

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

The reported resolution of this entry is 2.20 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	4898 (2.20-2.20)
Clashscore	141614	5594 (2.20-2.20)
Ramachandran outliers	138981	5503 (2.20-2.20)
Sidechain outliers	138945	5504 (2.20-2.20)
RSRZ outliers	127900	4800 (2.20-2.20)

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 <=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	180	77% 20%
1	B	180	2% 84% 14%
2	H	225	5% 84% 12%
2	I	225	20% 78% 8% 14%
3	L	218	86% 12%
3	M	218	21% 84% 13%

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 9781 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 Putative zinc metalloprotease aq_1964.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	176	Total	C	N	O	S	0	6	0
			1433	928	249	255	1			
1	B	179	Total	C	N	O	S	0	3	0
			1427	922	245	259	1			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	113	GLY	-	EXPRESSION TAG	UNP O67776
A	114	SER	-	EXPRESSION TAG	UNP O67776
B	113	GLY	-	EXPRESSION TAG	UNP O67776
B	114	SER	-	EXPRESSION TAG	UNP O67776

- Molecule 2 is a protein called MOUSE IGG1-KAPPA FAB (HEAVY CHAIN).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	H	216	Total	C	N	O	S	0	1	0
			1647	1042	271	327	7			
2	I	194	Total	C	N	O	S	0	3	0
			1502	947	250	299	6			

- Molecule 3 is a protein called MOUSE IGG1-KAPPA FAB (LIGHT CHAIN).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	L	215	Total	C	N	O	S	0	3	0
			1692	1056	284	345	7			
3	M	214	Total	C	N	O	S	0	4	0
			1693	1057	283	346	7			

- Molecule 4 is water.

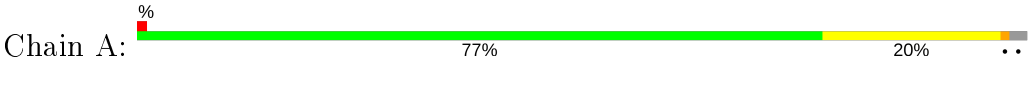
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	67	Total O 67 67	0	0
4	H	71	Total O 71 71	0	0
4	L	87	Total O 87 87	0	0
4	B	53	Total O 53 53	0	0
4	I	49	Total O 49 49	0	0
4	M	60	Total O 60 60	0	0

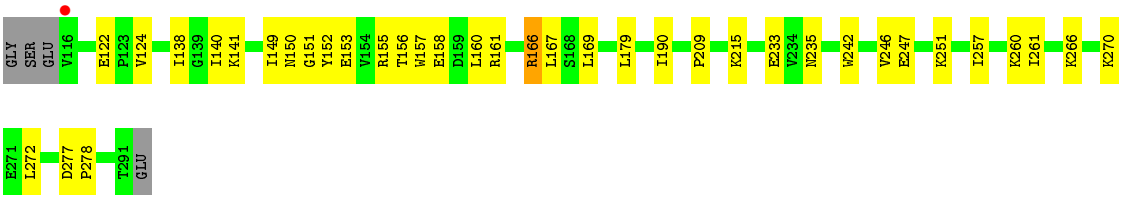
3 Residue-property plots

These plots are drawn for all protein, RNA and DNA 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: Putative zinc metalloprotease aq_1964

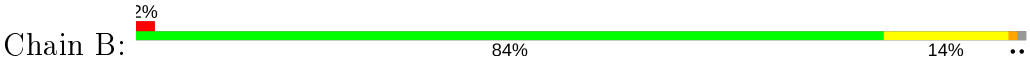
Chain A:

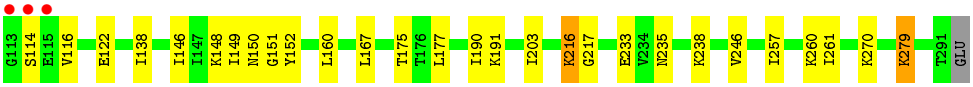




- Molecule 1: Putative zinc metalloprotease aq_1964

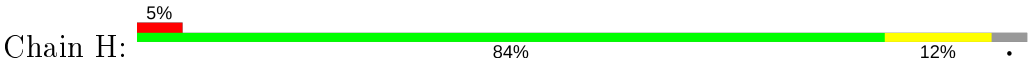
Chain B:

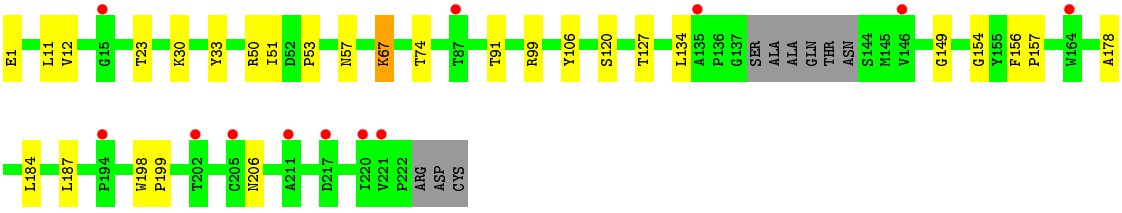




- Molecule 2: MOUSE IGG1-KAPPA FAB (HEAVY CHAIN)

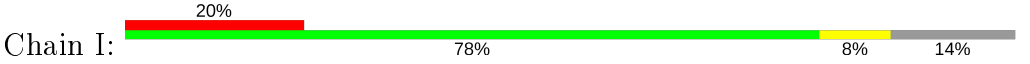
Chain H:

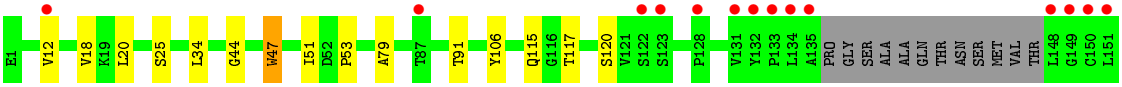


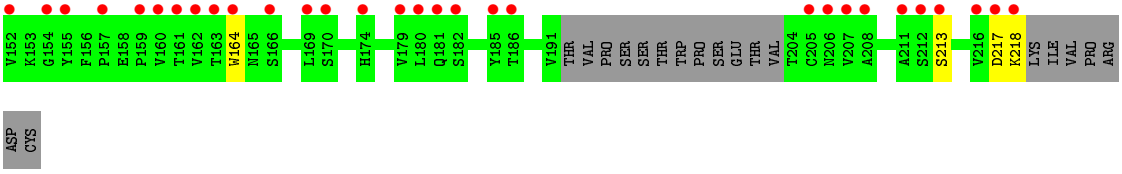


- Molecule 2: MOUSE IGG1-KAPPA FAB (HEAVY CHAIN)

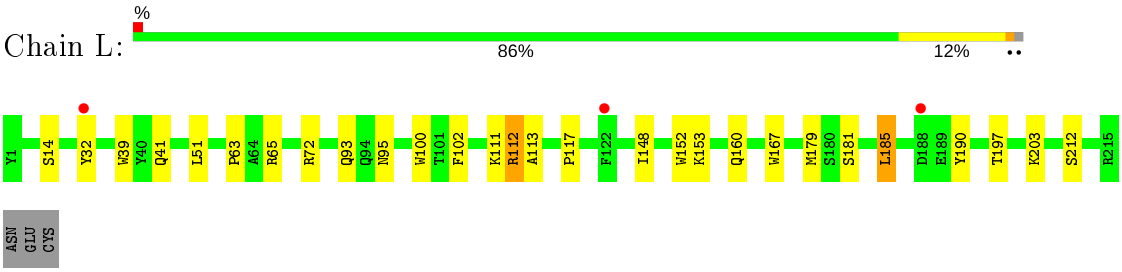
Chain I:



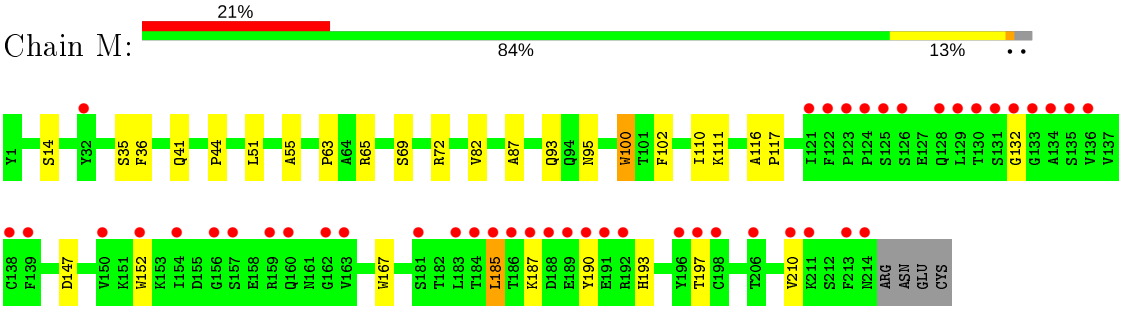




- Molecule 3: MOUSE IGG1-KAPPA FAB (LIGHT CHAIN)



- Molecule 3: MOUSE IGG1-KAPPA FAB (LIGHT CHAIN)



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α, β, γ	51.24Å 86.04Å 92.54Å 100.84° 96.92° 101.32°	Depositor
Resolution (Å)	47.44 – 2.20 47.44 – 2.20	Depositor EDS
% Data completeness (in resolution range)	98.2 (47.44-2.20) 98.2 (47.44-2.20)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.84 (at 2.20Å)	Xtriage
Refinement program	REFMAC 5.6.0117	Depositor
R, R_{free}	0.224 , 0.272 0.224 , 0.271	Depositor DCC
R_{free} test set	3771 reflections (5.03%)	wwPDB-VP
Wilson B-factor (Å^2)	32.0	Xtriage
Anisotropy	0.459	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 36.9	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	9781	wwPDB-VP
Average B, all atoms (Å^2)	39.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.00% 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	0.45	1/1456 (0.1%)	0.58	0/1960
1	B	0.46	0/1450	0.59	0/1954
2	H	0.53	0/1690	0.59	0/2312
2	I	0.52	2/1538 (0.1%)	0.56	0/2097
3	L	0.52	3/1732 (0.2%)	0.63	0/2357
3	M	0.51	3/1733 (0.2%)	0.60	0/2358
All	All	0.50	9/9599 (0.1%)	0.59	0/13038

The worst 5 of 9 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	L	167	TRP	CD2-CE2	5.48	1.48	1.41
3	M	167	TRP	CD2-CE2	5.30	1.47	1.41
3	L	39	TRP	CD2-CE2	5.16	1.47	1.41
3	M	152	TRP	CD2-CE2	5.16	1.47	1.41
2	I	164	TRP	CD2-CE2	5.09	1.47	1.41

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	1433	0	1547	36	0
1	B	1427	0	1528	31	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	H	1647	0	1600	14	0
2	I	1502	0	1447	10	0
3	L	1692	0	1610	11	0
3	M	1693	0	1609	17	0
4	A	67	0	0	2	0
4	B	53	0	0	1	0
4	H	71	0	0	1	0
4	I	49	0	0	0	0
4	L	87	0	0	0	0
4	M	60	0	0	4	0
All	All	9781	0	9341	107	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 6.

The worst 5 of 107 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:138[B]:ILE:CD1	1:B:138[B]:ILE:HD12	1.76	1.15
1:A:138[B]:ILE:HD12	1:B:138[B]:ILE:CD1	1.83	1.08
2:H:23:THR:HG22	4:H:316:HOH:O	1.54	1.03
1:A:138[B]:ILE:HD12	1:B:138[B]:ILE:HD12	1.37	1.03
1:A:150:ASN:CG	1:A:151:GLY:H	1.65	0.96

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	180/180 (100%)	177 (98%)	3 (2%)	0	100	100
1	B	180/180 (100%)	174 (97%)	6 (3%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	H	213/225 (95%)	202 (95%)	11 (5%)	0	100	100
2	I	191/225 (85%)	181 (95%)	9 (5%)	1 (0%)	29	31
3	L	216/218 (99%)	209 (97%)	5 (2%)	2 (1%)	17	16
3	M	216/218 (99%)	202 (94%)	12 (6%)	2 (1%)	17	16
All	All	1196/1246 (96%)	1145 (96%)	46 (4%)	5 (0%)	41	37

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	L	203	LYS
3	L	72	ARG
2	I	213	SER
3	M	72[A]	ARG
3	M	72[B]	ARG

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	158/155 (102%)	154 (98%)	4 (2%)	47	60
1	B	157/155 (101%)	154 (98%)	3 (2%)	57	71
2	H	187/193 (97%)	181 (97%)	6 (3%)	39	50
2	I	168/193 (87%)	165 (98%)	3 (2%)	59	72
3	L	192/192 (100%)	188 (98%)	4 (2%)	53	67
3	M	192/192 (100%)	187 (97%)	5 (3%)	46	58
All	All	1054/1080 (98%)	1029 (98%)	25 (2%)	52	62

5 of 25 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	L	160	GLN
3	L	212	SER

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
3	M	185	LEU
3	L	185	LEU
1	B	216	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
3	L	96	ASN

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 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

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	176/180 (97%)	-0.11	1 (0%) 89 88	20, 30, 44, 59	0
1	B	179/180 (99%)	-0.04	3 (1%) 70 68	21, 31, 53, 78	0
2	H	216/225 (96%)	0.34	12 (5%) 24 23	18, 43, 67, 75	0
2	I	194/225 (86%)	0.92	44 (22%) 0 0	20, 47, 90, 102	0
3	L	215/218 (98%)	0.14	3 (1%) 75 73	16, 32, 56, 66	0
3	M	214/218 (98%)	0.93	46 (21%) 0 0	17, 35, 100, 110	0
All	All	1194/1246 (95%)	0.38	109 (9%) 9 8	16, 34, 85, 110	0

The worst 5 of 109 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	113	GLY	7.3
3	M	124	PRO	6.2
3	M	121	ILE	6.2
2	I	182	SER	6.1
3	M	123	PRO	5.9

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 carbohydrates 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.