wwPDB X-ray Structure Validation Summary Report (i)

Mar 5, 2024 – 03:49 AM EST

PDB ID	:	2EFD
Title	:	Ara7/AtVps9a
Authors	:	Uejima, T.; Ihara, K.; Wakatsuki, S.
Deposited on	:	2007-02-22
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 (1)) were used in the production of this report:

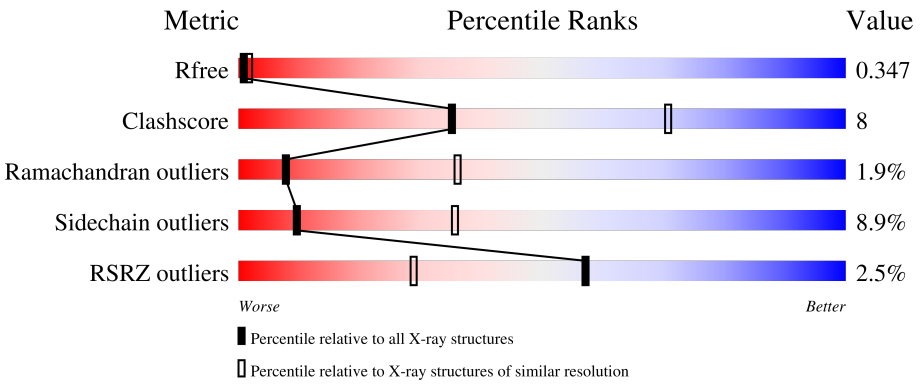
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation 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	267	66% 24% • 8%
1	C	267	2% 74% 16% • 8%
2	B	181	2% 59% 19% 6% 16%
2	D	181	6% 58% 26% 16%

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 6266 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 Similarity to vacuolar protein sorting-associated protein VPS9.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	246	1953	1242	324	375	12	0	0	0
1	C	246	1953	1242	324	375	12	0	0	0

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-1	GLY	-	expression tag	UNP Q9LT31
A	0	SER	-	expression tag	UNP Q9LT31
C	-1	GLY	-	expression tag	UNP Q9LT31
C	0	SER	-	expression tag	UNP Q9LT31

- Molecule 2 is a protein called Small GTP-binding protein-like.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	152	1180	752	202	222	4	0	0	0
2	D	152	1180	752	202	222	4	0	0	0

There are 4 discrepancies between the modelled and reference sequences:

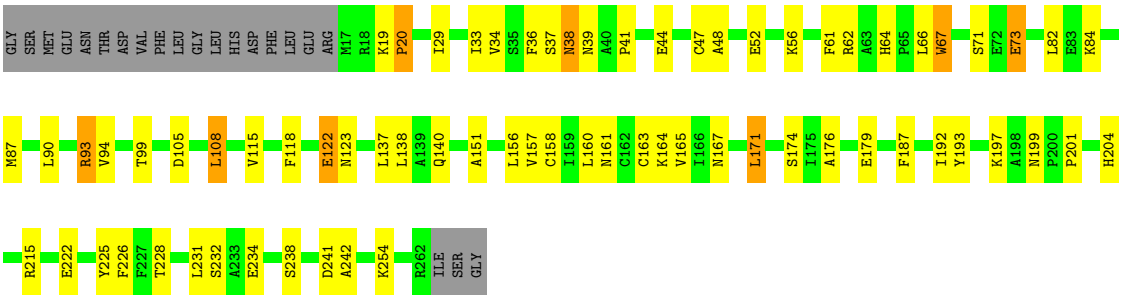
Chain	Residue	Modelled	Actual	Comment	Reference
B	-1	GLY	-	expression tag	UNP Q9LT31
B	0	SER	-	expression tag	UNP Q9LT31
D	-1	GLY	-	expression tag	UNP Q9LT31
D	0	SER	-	expression tag	UNP Q9LT31

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.

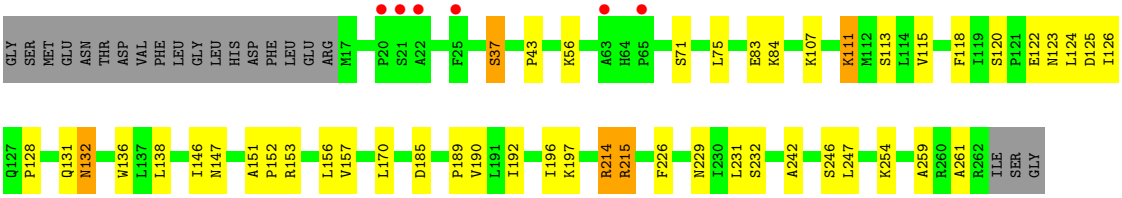
• Molecule 1: Similarity to vacuolar protein sorting-associated protein VPS9

Chain A:



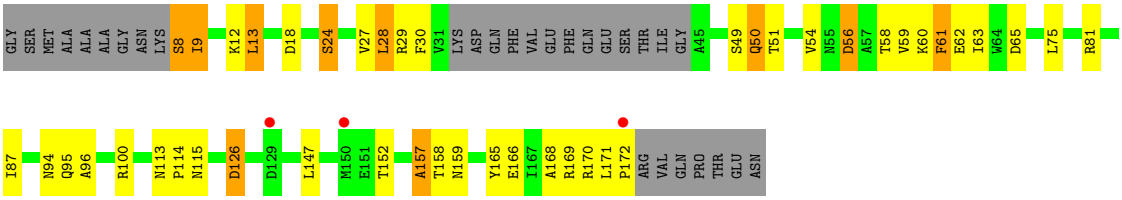
• Molecule 1: Similarity to vacuolar protein sorting-associated protein VPS9

Chain C:



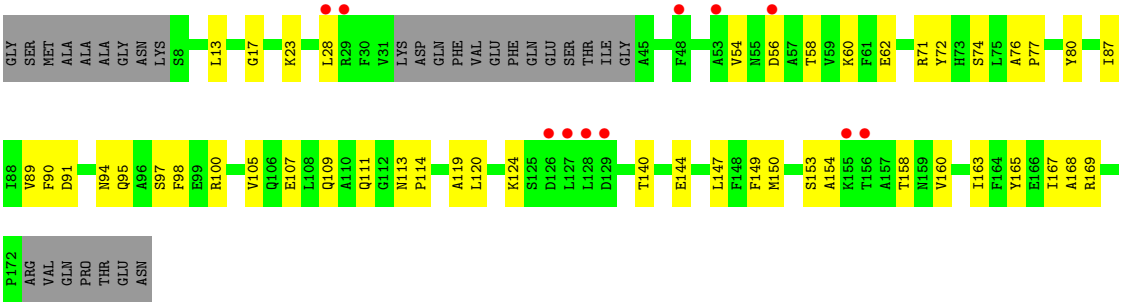
• Molecule 2: Small GTP-binding protein-like

Chain B:



• Molecule 2: Small GTP-binding protein-like

Chain D:


GLY SER MET ALA ALA ALA GLY ASN LYS S8 L13 G17 K23 L28 R29 F30 V31 LYS ASP GLN PHE VAL GLU PHE GLN GLU SER THR ILE GLY A45 F48 A53 V54 N55 D56 A57 T58 V59 K60 F61 E62 R71 Y72 H73 S74 L75 A76 P77 Y80 I87
I88 V89 F90 D91 N94 Q95 A96 S97 F98 E99 R100 V105 Q106 E107 L108 Q109 A110 Q111 G112 N113 P114 A119 L120 K124 S125 D126 L127 L128 D129 T140 E144 L147 F148 F149 M150 S153 A154 K155 T156 A157 T158 N159 V160 I163 F164 Y165 E166 I167 A168 R169
P172 ARG VAL GLN PRO THR GLU ASN

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 61	Depositor
Cell constants a, b, c, α, β, γ	189.54Å 189.54Å 75.17Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	38.95 – 3.00 38.94 – 3.00	Depositor EDS
% Data completeness (in resolution range)	95.9 (38.95-3.00) 95.9 (38.94-3.00)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	0.09	Depositor
$< I/\sigma(I) >$ [1]	3.23 (at 3.01Å)	Xtriage
Refinement program	REFMAC 5.2.0005	Depositor
R, R_{free}	0.306 , 0.351 0.304 , 0.347	Depositor DCC
R_{free} test set	1494 reflections (5.02%)	wwPDB-VP
Wilson B-factor (Å^2)	74.0	Xtriage
Anisotropy	0.085	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.29 , 25.1	EDS
L-test for twinning[2]	$< \|L\| > = 0.43$, $< L^2 > = 0.25$	Xtriage
Estimated twinning fraction	0.076 for h,-h-k,-l	Xtriage
F_o,F_c correlation	0.85	EDS
Total number of atoms	6266	wwPDB-VP
Average B, all atoms (Å^2)	74.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.17% 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 (i)

5.1 Standard geometry (i)

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	0.48	0/1991	0.61	0/2687
1	C	0.46	0/1991	0.57	0/2687
2	B	0.48	0/1200	0.62	0/1622
2	D	0.43	0/1200	0.59	0/1622
All	All	0.46	0/6382	0.60	0/8618

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 (i)

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	1953	0	1937	39	0
1	C	1953	0	1937	28	0
2	B	1180	0	1174	20	0
2	D	1180	0	1174	21	0
All	All	6266	0	6222	102	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 8.

The worst 5 of 102 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:D:105:VAL:HG12	2:D:109:GLN:HE21	1.42	0.83
1:C:128:PRO:HA	1:C:131:GLN:HG3	1.63	0.79
1:A:48:ALA:O	1:A:52:GLU:HB2	1.87	0.73
1:C:111:LYS:O	1:C:115:VAL:HG23	1.89	0.71
1:A:41:PRO:HA	1:A:93:ARG:HH21	1.55	0.71

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	244/267 (91%)	209 (86%)	31 (13%)	4 (2%)	9 40
1	C	244/267 (91%)	218 (89%)	23 (9%)	3 (1%)	13 48
2	B	148/181 (82%)	132 (89%)	13 (9%)	3 (2%)	7 34
2	D	148/181 (82%)	123 (83%)	20 (14%)	5 (3%)	3 20
All	All	784/896 (88%)	682 (87%)	87 (11%)	15 (2%)	8 36

5 of 15 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	67	TRP
2	B	95	GLN
2	D	56	ASP
1	A	215	ARG
2	D	124	LYS

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	217/235 (92%)	198 (91%)	19 (9%)	10 36
1	C	217/235 (92%)	206 (95%)	11 (5%)	24 60
2	B	120/143 (84%)	102 (85%)	18 (15%)	3 14
2	D	120/143 (84%)	108 (90%)	12 (10%)	7 29
All	All	674/756 (89%)	614 (91%)	60 (9%)	9 35

5 of 60 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	56	ASP
2	D	147	LEU
2	B	158	THR
2	D	100	ARG
2	D	169	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 8 such sidechains are listed below:

Mol	Chain	Res	Type
2	D	143	GLN
2	D	139	GLN
2	D	94	ASN
2	D	10	ASN
2	D	109	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 (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	246/267 (92%)	-0.12	0 100 100	43, 56, 94, 101	0
1	C	246/267 (92%)	0.03	6 (2%) 59 30	49, 65, 114, 127	0
2	B	152/181 (83%)	0.22	3 (1%) 65 36	46, 86, 106, 109	0
2	D	152/181 (83%)	0.40	11 (7%) 15 4	65, 92, 116, 118	0
All	All	796/896 (88%)	0.09	20 (2%) 57 29	43, 69, 113, 127	0

The worst 5 of 20 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	D	29	ARG	7.3
2	D	53	ALA	4.9
2	D	156	THR	4.7
2	D	126	ASP	3.7
2	D	127	LEU	3.6

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.