wwPDB X-ray Structure Validation Summary Report (i)

Jan 2, 2024 – 07:19 pm GMT

PDB ID	:	5AGC
Title	:	Crystallographic forms of the Vps75 tetramer
Authors	:	Hammond, C.M.; Sundaramoorthy, R.; Owen-Hughes, T.
Deposited on	:	2015-01-29
Resolution	:	4.00 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:

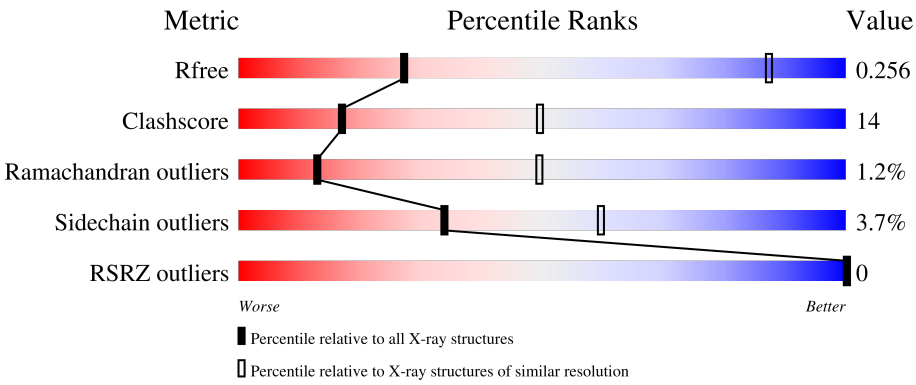
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation 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 4.00 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	1087 (4.30-3.70)
Clashscore	141614	1148 (4.30-3.70)
Ramachandran outliers	138981	1108 (4.30-3.70)
Sidechain outliers	138945	1099 (4.30-3.70)
RSRZ outliers	127900	1028 (4.34-3.66)

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	264	59% 21% • 18%
1	B	264	56% 25% • 17%
1	C	264	58% 22% • 18%
1	D	264	58% 22% • 18%

2 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 7250 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 VACUOLAR PROTEIN SORTING-ASSOCIATED PROTEIN 75.

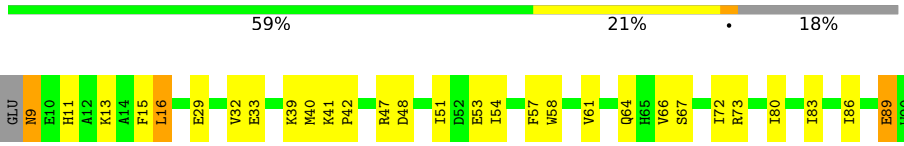
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	217	1810	1168	294	343	5	0	0	0
			Total	C	N	O	S			
1	B	220	1836	1182	298	351	5	0	0	0
			Total	C	N	O	S			
1	C	216	1802	1164	292	341	5	0	0	0
			Total	C	N	O	S			
1	D	216	1802	1164	292	341	5	0	0	0

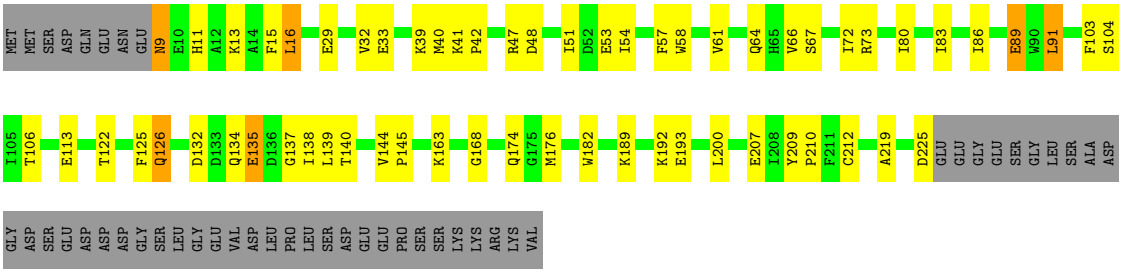
3 Residue-property plots (i)

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: VACUOLAR PROTEIN SORTING-ASSOCIATED PROTEIN 75

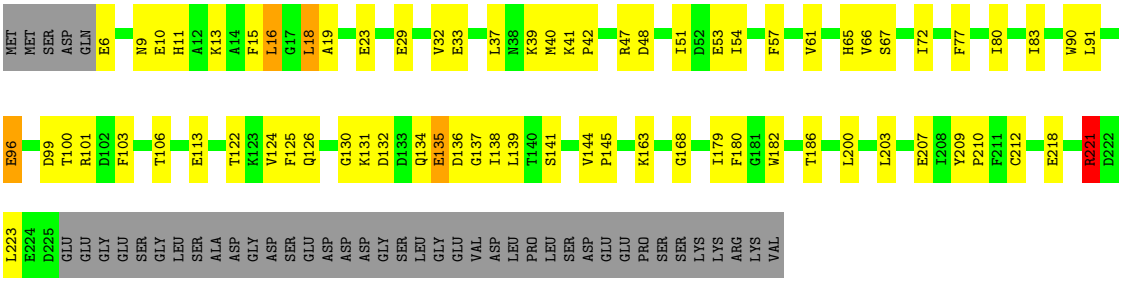
Chain A:





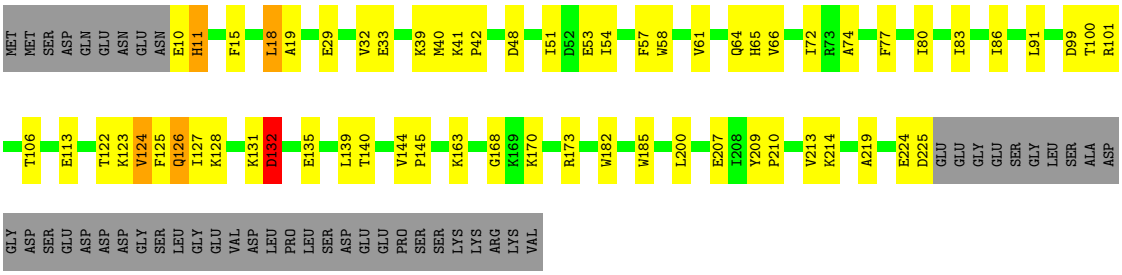
• Molecule 1: VACUOLAR PROTEIN SORTING-ASSOCIATED PROTEIN 75

Chain B:



• Molecule 1: VACUOLAR PROTEIN SORTING-ASSOCIATED PROTEIN 75

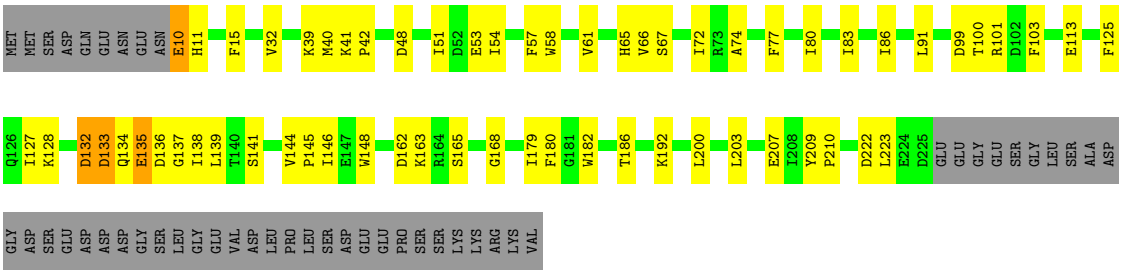
Chain C:



• Molecule 1: VACUOLAR PROTEIN SORTING-ASSOCIATED PROTEIN 75

Chain D:

58%
22%
18%


MET MET SER ASP GLN GLU ASN GLU ASN E10 H11 F15 V32 K39 M40 K41 P42 D48 I51 D52 E53 I54 F57 W58 V61 H65 V66 S67 I72 R73 A74 F77 I80 I83 I86 L91 D99 T100 R101 D102 F103 E113 F125
Q126 I127 K128 D132 D133 Q134 E135 D136 G137 I138 L139 T140 S141 V144 P145 I146 E147 W148 D162 K163 R164 S165 G168 I179 F180 G181 W182 T186 K192 L200 L203 E207 I208 Y209 P210 D222 L223 E224 D225 GLU GLU GLY GLU SER GLY LEU SER ALA ASP
GLY ASP SER GLU ASP ASP ASP GLY SER LEU GLY GLU VAL ASP LEU PRO LEU SER ASP GLU GLU PRO SER SER LYS LYS ARG LYS VAL

4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	82.75Å 92.43Å 160.75Å 90.00° 93.85° 90.00°	Depositor
Resolution (Å)	50.06 – 4.00 50.06 – 4.00	Depositor EDS
% Data completeness (in resolution range)	98.9 (50.06-4.00) 98.9 (50.06-4.00)	Depositor EDS
R_{merge}	0.14	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	3.30 (at 4.00Å)	Xtriage
Refinement program	REFMAC 5.6.0117	Depositor
R, R_{free}	0.229 , 0.265 0.247 , 0.256	Depositor DCC
R_{free} test set	495 reflections (4.82%)	wwPDB-VP
Wilson B-factor (Å^2)	124.4	Xtriage
Anisotropy	0.121	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.27 , 98.0	EDS
L-test for twinning[2]	$< \|L\| > = 0.44$, $< L^2 > = 0.26$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.90	EDS
Total number of atoms	7250	wwPDB-VP
Average B, all atoms (Å^2)	139.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.93% 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	0/1858	0.74	0/2503
1	B	0.49	0/1884	0.76	2/2538 (0.1%)
1	C	0.49	0/1850	0.73	0/2492
1	D	0.52	1/1850 (0.1%)	0.71	1/2492 (0.0%)
All	All	0.50	1/7442 (0.0%)	0.73	3/10025 (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	1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	D	132	ASP	C-N	-5.11	1.22	1.34

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	221	ARG	CA-C-N	-7.34	101.05	117.20
1	B	221	ARG	CA-C-O	6.51	133.77	120.10
1	D	132	ASP	C-N-CA	-6.01	106.68	121.70

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	9	ASN	Peptide

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	1810	0	1751	42	0
1	B	1836	0	1769	54	0
1	C	1802	0	1745	55	0
1	D	1802	0	1745	60	0
All	All	7250	0	7010	199	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 14.

The worst 5 of 199 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:C:74:ALA:CB	1:D:74:ALA:HB2	1.52	1.35
1:C:74:ALA:HB3	1:D:74:ALA:CB	1.53	1.34
1:D:132:ASP:O	1:D:133:ASP:CB	1.71	1.34
1:B:47:ARG:NH2	1:B:212:CYS:SG	2.24	1.11
1:A:47:ARG:NH2	1:A:212:CYS:SG	2.24	1.11

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	215/264 (81%)	203 (94%)	10 (5%)	2 (1%)	17 55
1	B	218/264 (83%)	205 (94%)	9 (4%)	4 (2%)	8 41

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	C	214/264 (81%)	203 (95%)	9 (4%)	2 (1%)	17	55
1	D	214/264 (81%)	202 (94%)	10 (5%)	2 (1%)	17	55
All	All	861/1056 (82%)	813 (94%)	38 (4%)	10 (1%)	13	49

5 of 10 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	223	LEU
1	D	133	ASP
1	A	135	GLU
1	B	135	GLU
1	B	221	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	194/235 (83%)	187 (96%)	7 (4%)	35	61
1	B	197/235 (84%)	188 (95%)	9 (5%)	27	54
1	C	193/235 (82%)	185 (96%)	8 (4%)	30	57
1	D	193/235 (82%)	188 (97%)	5 (3%)	46	67
All	All	777/940 (83%)	748 (96%)	29 (4%)	34	60

5 of 29 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	136	ASP
1	D	113	GLU
1	C	11	HIS
1	D	10	GLU
1	C	10	GLU

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

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	217/264 (82%)	-0.53	0	100	100	84, 132, 189, 244	0
1	B	220/264 (83%)	-0.43	0	100	100	88, 135, 207, 243	0
1	C	216/264 (81%)	-0.46	0	100	100	89, 131, 181, 244	0
1	D	216/264 (81%)	-0.55	0	100	100	96, 139, 180, 239	0
All	All	869/1056 (82%)	-0.49	0	100	100	84, 135, 192, 244	0

There are no RSRZ outliers to report.

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.