wwPDB/EMDataBank EM Map/Model Validation Summary Report (i)

Oct 16, 2018 – 10:14 AM EDT

PDB ID	:	5XMI
EMDB ID:	:	EMD-6733
Title	:	Cryo-EM Structure of the ATP-bound VPS4 mutant-E233Q hexamer (masked)
Authors	:	Sun, S.; Li, L.; Yang, F.; Wang, X.; Fan, F.; Li, X.; Wang, H.; Sui, S.
Deposited on	:	2017-05-15
Resolution	:	3.90 Å(reported)

This is a wwPDB/EMDataBank EM Map/Model Validation Summary Report for a publicly released PDB/EMDB entry.

We welcome your comments at *validation@mail.wwpdb.org*
A user guide is available at
https://www.wwpdb.org/validation/2017/EMValidationReportHelp
with specific help available everywhere you see the (i) symbol.

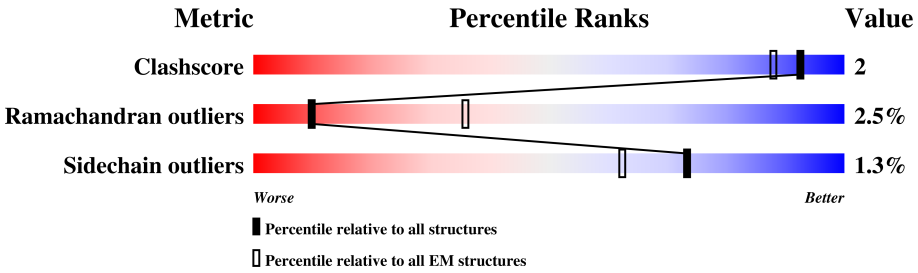
MolProbity	:	4.02b-467
Mogul	:	1.7.3 (157068), CSD as539be (2018)
Percentile statistics	:	20171227.v01 (using entries in the PDB archive December 27th 2017)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP)	:	rb-20031633

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
ELECTRON MICROSCOPY

The reported resolution of this entry is 3.90 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)	EM structures (#Entries)
Clashscore	136327	1886
Ramachandran outliers	132723	1663
Sidechain outliers	132532	1531

The table below summarises the geometric issues observed across the polymeric chains. The red, orange, yellow and green segments on the bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria. 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%

Mol	Chain	Length	Quality of chain
1	A	437	64% 8% • 27%
1	B	437	69% • 27%
1	C	437	67% 5% • 27%
1	D	437	68% 5% 27%
1	E	437	69% • • 27%
1	F	437	67% 6% 27%

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 14969 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, 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.

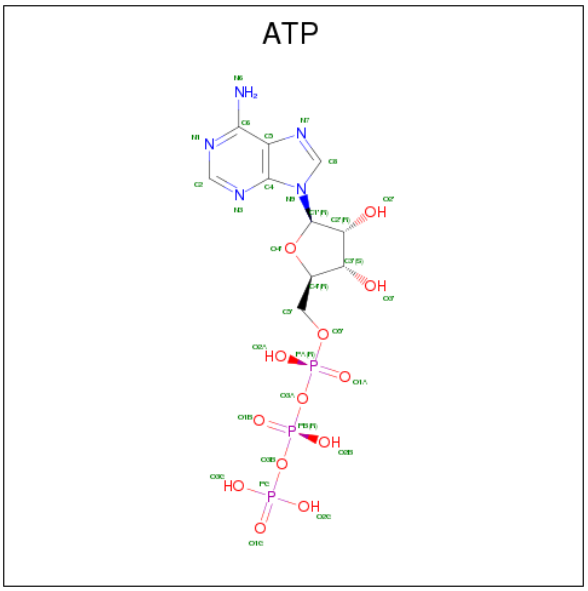
- Molecule 1 is a protein called Vacuolar protein sorting-associated protein 4.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	319	Total	C	N	O	S	0	0
			2469	1552	419	489	9		
1	B	319	Total	C	N	O	S	0	0
			2469	1552	419	489	9		
1	C	319	Total	C	N	O	S	0	0
			2469	1552	419	489	9		
1	D	319	Total	C	N	O	S	0	0
			2469	1552	419	489	9		
1	E	319	Total	C	N	O	S	0	0
			2469	1552	419	489	9		
1	F	319	Total	C	N	O	S	0	0
			2469	1552	419	489	9		

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	233	GLN	GLU	engineered mutation	UNP P52917
B	233	GLN	GLU	engineered mutation	UNP P52917
C	233	GLN	GLU	engineered mutation	UNP P52917
D	233	GLN	GLU	engineered mutation	UNP P52917
E	233	GLN	GLU	engineered mutation	UNP P52917
F	233	GLN	GLU	engineered mutation	UNP P52917

- Molecule 2 is ADENOSINE-5'-TRIPHOSPHATE (three-letter code: ATP) (formula: $C_{10}H_{16}N_5O_{13}P_3$).



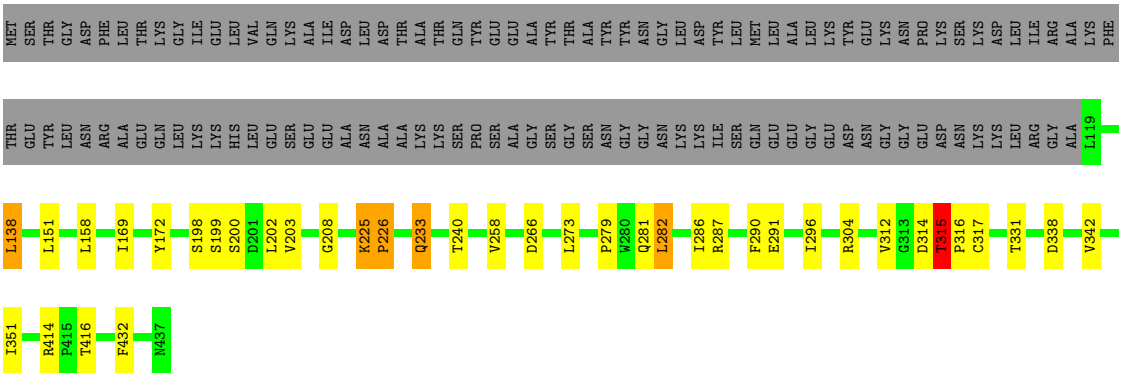
Mol	Chain	Residues	Atoms					AltConf
2	B	1	Total	C	N	O	P	0
			31	10	5	13	3	
2	C	1	Total	C	N	O	P	0
			31	10	5	13	3	
2	D	1	Total	C	N	O	P	0
			31	10	5	13	3	
2	E	1	Total	C	N	O	P	0
			31	10	5	13	3	
2	F	1	Total	C	N	O	P	0
			31	10	5	13	3	

3 Residue-property plots (i)

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. 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. 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.

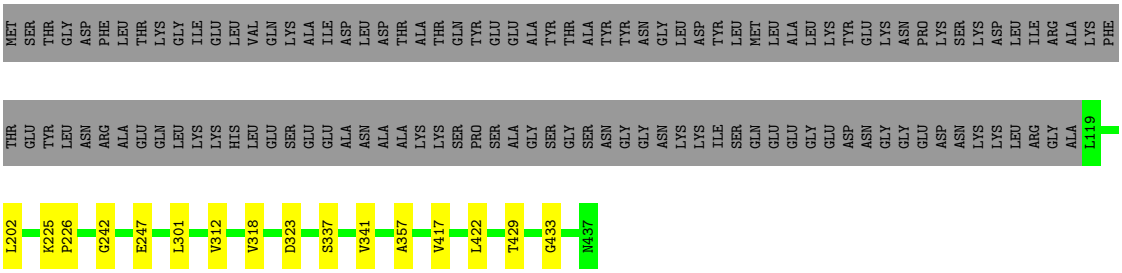
• Molecule 1: Vacuolar protein sorting-associated protein 4

Chain A:



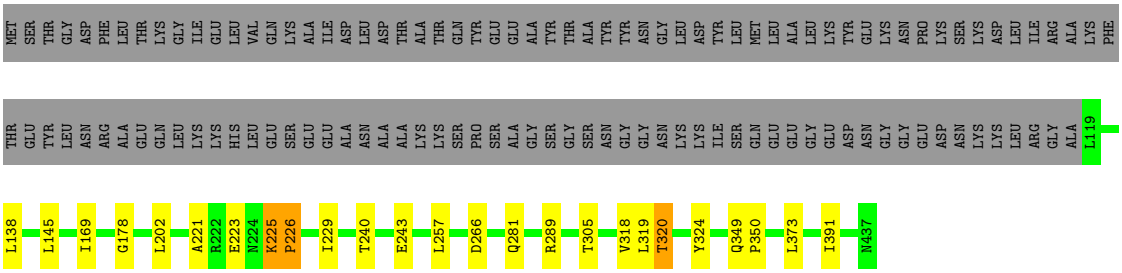
• Molecule 1: Vacuolar protein sorting-associated protein 4

Chain B:



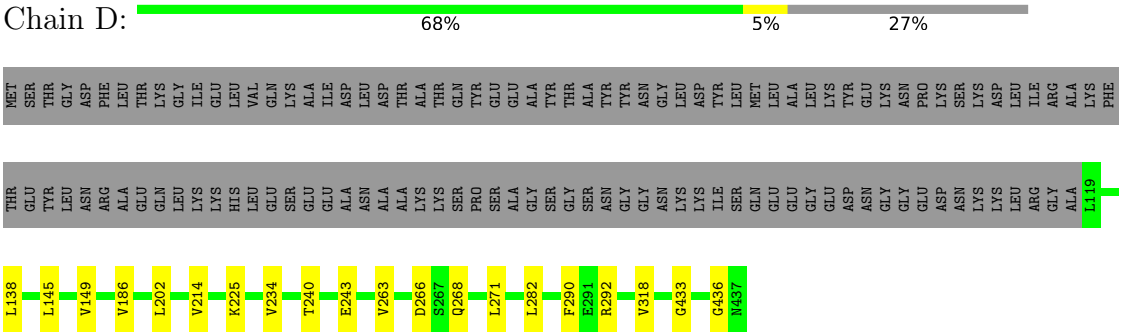
• Molecule 1: Vacuolar protein sorting-associated protein 4

Chain C:



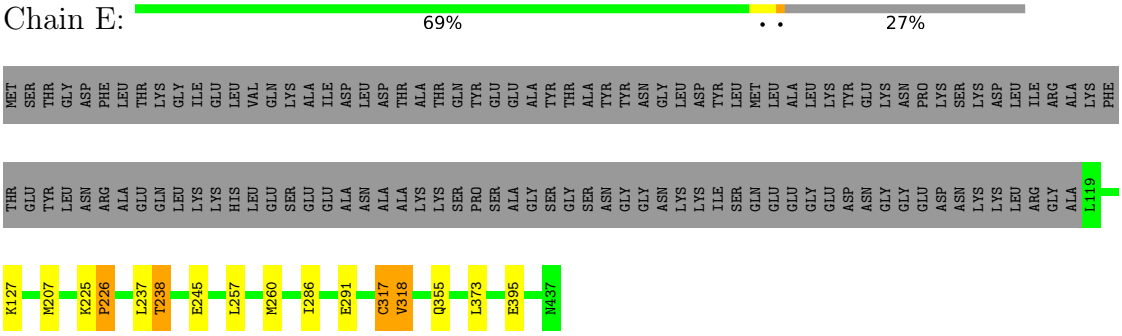
• Molecule 1: Vacuolar protein sorting-associated protein 4

Chain D:



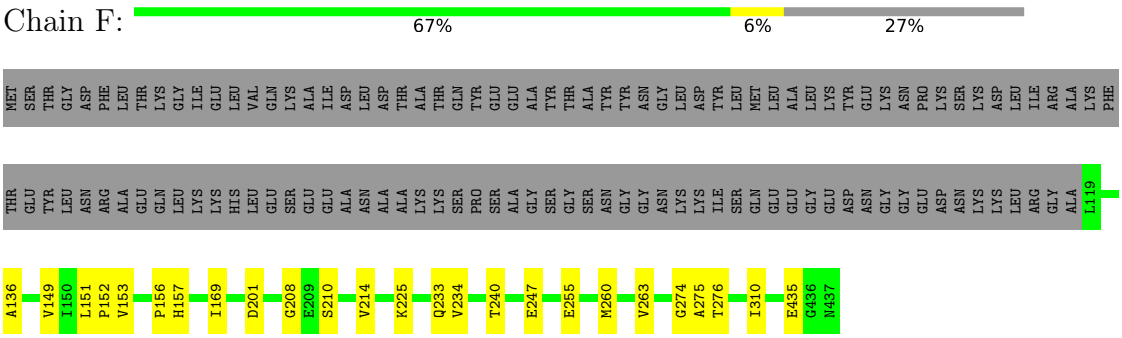
• Molecule 1: Vacuolar protein sorting-associated protein 4

Chain E:



• Molecule 1: Vacuolar protein sorting-associated protein 4

Chain F:



4 Experimental information (i)

Property	Value	Source
Reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	Depositor
Number of particles used	106918	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	PHASE FLIPPING AND AMPLITUDE CORRECTION	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ($e^-/Å^2$)	50	Depositor
Minimum defocus (nm)	Not provided	Depositor
Maximum defocus (nm)	Not provided	Depositor
Magnification	Not provided	Depositor
Image detector	GATAN K2 SUMMIT (4k x 4k)	Depositor

5 Model quality (i)

5.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: ATP

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	#$\|Z\|$ >2	RMSZ	#$\|Z\|$ >2
1	A	0.45	0/2511	0.68	0/3395
1	B	0.41	0/2511	0.57	0/3395
1	C	0.41	0/2511	0.54	0/3395
1	D	0.41	0/2511	0.55	0/3395
1	E	0.41	0/2511	0.57	0/3395
1	F	0.41	0/2511	0.56	0/3395
All	All	0.42	0/15066	0.58	0/20370

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	3
1	B	0	1
1	C	0	1
All	All	0	5

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (5) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	225	LYS	Peptide
1	A	315	THR	Peptide
1	A	416	THR	Peptide
1	B	225	LYS	Peptide

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Group
1	C	225	LYS	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	2469	0	2477	13	0
1	B	2469	0	2477	3	0
1	C	2469	0	2477	9	0
1	D	2469	0	2477	5	0
1	E	2469	0	2477	6	0
1	F	2469	0	2477	10	0
2	B	31	0	12	0	0
2	C	31	0	12	0	0
2	D	31	0	12	0	0
2	E	31	0	12	0	0
2	F	31	0	12	0	0
All	All	14969	0	14922	45	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 2.

The worst 5 of 45 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:172:TYR:HB2	1:A:279:PRO:HG3	1.79	0.65
1:B:417:VAL:HG13	1:B:422:LEU:HD22	1.82	0.61
1:E:237:LEU:HD12	1:E:238:THR:HG23	1.83	0.60
1:A:169:ILE:HB	1:A:273:LEU:HD23	1.83	0.60
1:A:315:THR:HG22	1:A:316:PRO:HD3	1.84	0.58

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 PDB entries followed by that with respect to all EM entries.

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	317/437 (72%)	276 (87%)	31 (10%)	10 (3%)	4 36
1	B	317/437 (72%)	282 (89%)	28 (9%)	7 (2%)	7 43
1	C	317/437 (72%)	283 (89%)	27 (8%)	7 (2%)	7 43
1	D	317/437 (72%)	277 (87%)	31 (10%)	9 (3%)	5 39
1	E	317/437 (72%)	284 (90%)	26 (8%)	7 (2%)	7 43
1	F	317/437 (72%)	283 (89%)	27 (8%)	7 (2%)	7 43
All	All	1902/2622 (72%)	1685 (89%)	170 (9%)	47 (2%)	10 41

5 of 47 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	226	PRO
1	A	290	PHE
1	A	312	VAL
1	A	314	ASP
1	B	226	PRO

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 PDB entries followed by that with respect to all EM entries.

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	272/367 (74%)	263 (97%)	9 (3%)	41 70
1	B	272/367 (74%)	270 (99%)	2 (1%)	85 93
1	C	272/367 (74%)	268 (98%)	4 (2%)	67 85

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
1	D	272/367 (74%)	271 (100%)	1 (0%)	92 96
1	E	272/367 (74%)	269 (99%)	3 (1%)	76 88
1	F	272/367 (74%)	270 (99%)	2 (1%)	85 93
All	All	1632/2202 (74%)	1611 (99%)	21 (1%)	73 86

5 of 21 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	202	LEU
1	C	223	GLU
1	E	318	VAL
1	A	414	ARG
1	E	355	GLN

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

Mol	Chain	Res	Type
1	C	281	GLN
1	F	277	ASN
1	E	216	GLN
1	A	281	GLN
1	C	437	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)

5 ligands are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. 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 $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
2	ATP	B	501	-	27,33,33	1.08	2 (7%)	27,52,52	1.93	4 (14%)
2	ATP	C	501	-	27,33,33	1.08	2 (7%)	27,52,52	1.99	5 (18%)
2	ATP	D	501	-	27,33,33	1.07	2 (7%)	27,52,52	1.96	4 (14%)
2	ATP	E	501	-	27,33,33	1.07	2 (7%)	27,52,52	1.94	4 (14%)
2	ATP	F	501	-	27,33,33	1.07	2 (7%)	27,52,52	2.01	5 (18%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	ATP	B	501	-	-	0/18/38/38	0/3/3/3
2	ATP	C	501	-	-	0/18/38/38	0/3/3/3
2	ATP	D	501	-	-	0/18/38/38	0/3/3/3
2	ATP	E	501	-	-	0/18/38/38	0/3/3/3
2	ATP	F	501	-	-	0/18/38/38	0/3/3/3

The worst 5 of 10 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	E	501	ATP	C2-N3	2.14	1.35	1.32
2	F	501	ATP	C2-N3	2.17	1.35	1.32
2	C	501	ATP	C2-N3	2.18	1.35	1.32
2	B	501	ATP	C2-N3	2.24	1.35	1.32
2	D	501	ATP	C2-N3	2.26	1.35	1.32

The worst 5 of 22 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	F	501	ATP	N3-C2-N1	-7.15	122.74	128.86
2	C	501	ATP	N3-C2-N1	-7.15	122.74	128.86
2	D	501	ATP	N3-C2-N1	-6.99	122.88	128.86
2	E	501	ATP	N3-C2-N1	-6.93	122.93	128.86
2	B	501	ATP	N3-C2-N1	-6.93	122.94	128.86

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

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.