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

Jan 7, 2019 – 10:37 AM EST

PDB ID	:	5H3D
EMDB ID:	:	EMD-2546
Title	:	Helical structure of membrane tubules decorated by ACAP1 (BARPH doamin) protein by cryo-electron microscopy and MD simulation
Authors	:	Chan, C.; Pang, X.Y.; Zhang, Y.; Sun, F.; Fan, J.
Deposited on	:	2016-10-22
Resolution	:	14.00 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.

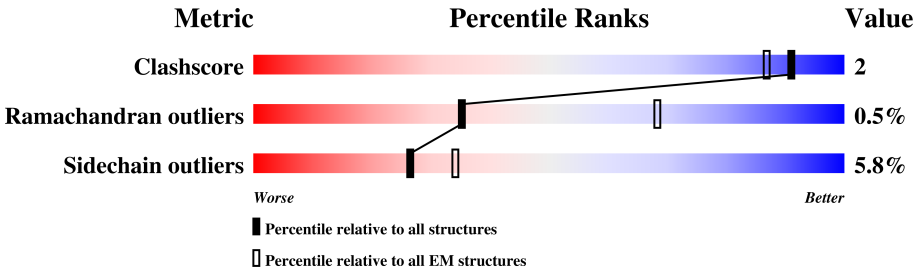
MolProbity	:	4.02b-467
Percentile statistics	:	20171227.v01 (using entries in the PDB archive December 27th 2017)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP)	:	rb-20031633

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 14.00 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)	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	382	66% 24% • • 5%
1	B	382	70% 19% • 5%
1	C	382	70% 20% 6% 5%
1	D	382	67% 23% • • 5%

2 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 23314 atoms, of which 11698 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 Arf-GAP with coiled-coil, ANK repeat and PH domain-containing protein 1.

Mol	Chain	Residues	Atoms						AltConf	Trace
1	A	364	Total	C	H	N	O	S	0	0
			5849	1825	2934	534	543	13		
1	B	361	Total	C	H	N	O	S	0	0
			5808	1810	2915	531	539	13		
1	C	364	Total	C	H	N	O	S	0	0
			5849	1825	2934	534	543	13		
1	D	361	Total	C	H	N	O	S	0	0
			5808	1810	2915	531	539	13		

There are 20 discrepancies between the modelled and reference sequences:

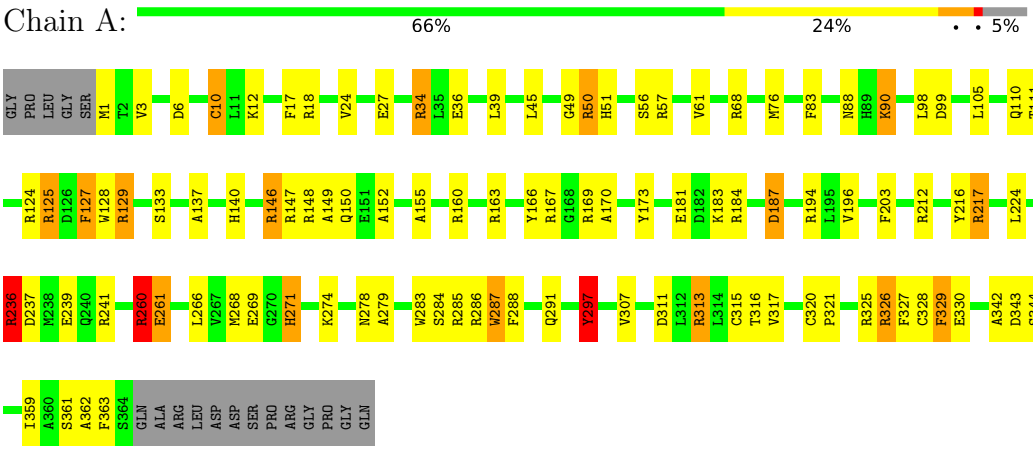
Chain	Residue	Modelled	Actual	Comment	Reference
A	-4	GLY	-	expression tag	UNP Q15027
A	-3	PRO	-	expression tag	UNP Q15027
A	-2	LEU	-	expression tag	UNP Q15027
A	-1	GLY	-	expression tag	UNP Q15027
A	0	SER	-	expression tag	UNP Q15027
B	-4	GLY	-	expression tag	UNP Q15027
B	-3	PRO	-	expression tag	UNP Q15027
B	-2	LEU	-	expression tag	UNP Q15027
B	-1	GLY	-	expression tag	UNP Q15027
B	0	SER	-	expression tag	UNP Q15027
C	-4	GLY	-	expression tag	UNP Q15027
C	-3	PRO	-	expression tag	UNP Q15027
C	-2	LEU	-	expression tag	UNP Q15027
C	-1	GLY	-	expression tag	UNP Q15027
C	0	SER	-	expression tag	UNP Q15027
D	-4	GLY	-	expression tag	UNP Q15027
D	-3	PRO	-	expression tag	UNP Q15027
D	-2	LEU	-	expression tag	UNP Q15027
D	-1	GLY	-	expression tag	UNP Q15027
D	0	SER	-	expression tag	UNP Q15027

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.

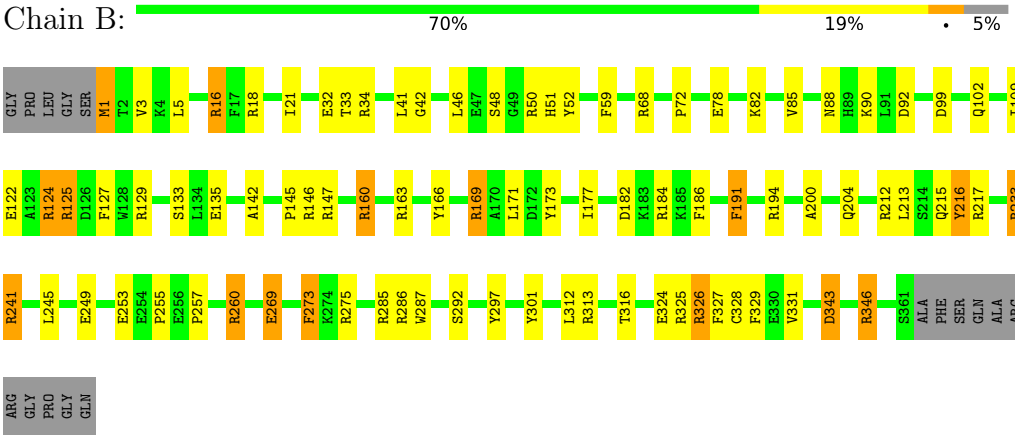
• Molecule 1: Arf-GAP with coiled-coil, ANK repeat and PH domain-containing protein 1

Chain A:



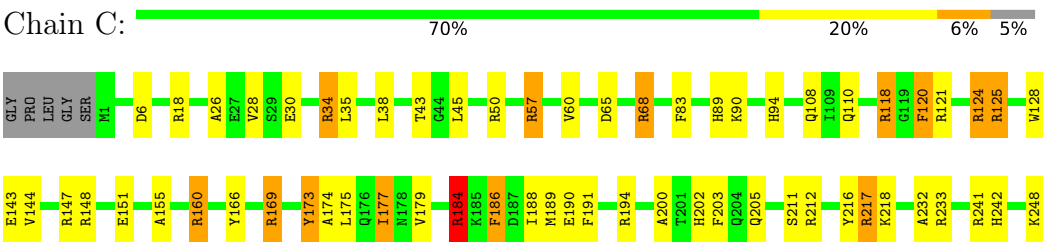
• Molecule 1: Arf-GAP with coiled-coil, ANK repeat and PH domain-containing protein 1

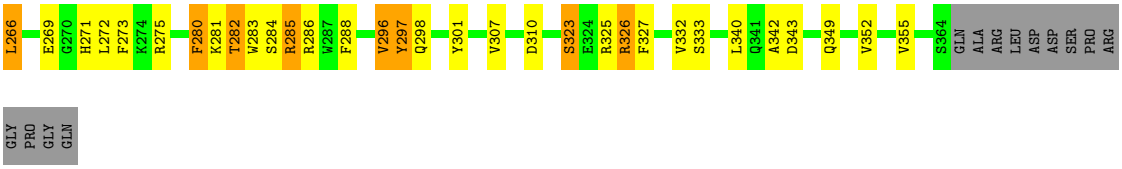
Chain B:



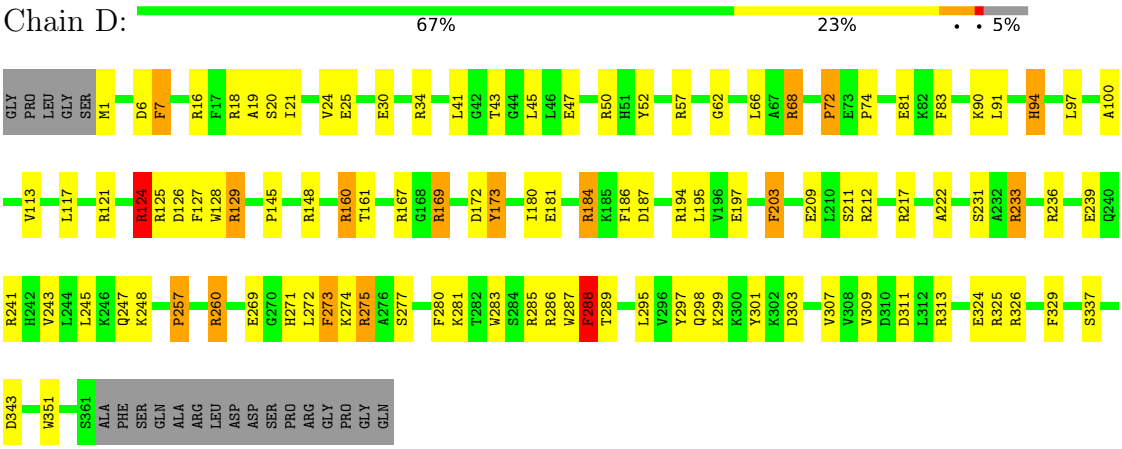
• Molecule 1: Arf-GAP with coiled-coil, ANK repeat and PH domain-containing protein 1

Chain C:





• Molecule 1: Arf-GAP with coiled-coil, ANK repeat and PH domain-containing protein 1



4 Experimental information (i)

Property	Value	Source
Reconstruction method	HELICAL	Depositor
Imposed symmetry	HELICAL, twist=42.72°, rise=23.20 Å, axial sym=C3	Depositor
Number of segments used	304	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	PHASE FLIPPING ONLY	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ($e^-/Å^2$)	20	Depositor
Minimum defocus (nm)	2500	Depositor
Maximum defocus (nm)	3500	Depositor
Magnification	75000	Depositor
Image detector	GATAN ULTRASCAN 4000 (4k x 4k)	Depositor

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	#$\|Z\|$ >2	RMSZ	#$\|Z\|$ >2
1	A	1.62	16/2963 (0.5%)	2.28	114/3987 (2.9%)
1	B	1.64	18/2940 (0.6%)	2.18	92/3956 (2.3%)
1	C	1.60	15/2963 (0.5%)	2.21	94/3987 (2.4%)
1	D	1.61	16/2940 (0.5%)	2.16	92/3956 (2.3%)
All	All	1.62	65/11806 (0.6%)	2.21	392/15886 (2.5%)

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	12
1	B	0	10
1	C	0	12
1	D	0	15
All	All	0	49

The worst 5 of 65 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	287	TRP	CD2-CE2	8.13	1.51	1.41
1	B	78	GLU	CB-CG	7.08	1.65	1.52
1	A	239	GLU	CD-OE1	6.70	1.33	1.25
1	C	151	GLU	CB-CG	6.64	1.64	1.52
1	B	48	SER	CA-CB	6.54	1.62	1.52

The worst 5 of 392 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	241	ARG	NE-CZ-NH1	24.61	132.60	120.30
1	A	148	ARG	NE-CZ-NH2	21.06	130.83	120.30
1	D	124	ARG	NE-CZ-NH1	20.73	130.66	120.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	D	16	ARG	NE-CZ-NH2	20.09	130.34	120.30
1	C	325	ARG	NE-CZ-NH1	19.80	130.20	120.30

There are no chirality outliers.

5 of 49 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	115	GLU	Peptide
1	A	125	ARG	Sidechain
1	A	127	PHE	Sidechain
1	A	17	PHE	Sidechain
1	A	34	ARG	Sidechain

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	2915	2934	2936	10	0
1	B	2893	2915	2917	8	0
1	C	2915	2934	2936	13	0
1	D	2893	2915	2917	14	0
All	All	11616	11698	11706	39	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 39 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:D:94:HIS:CD2	1:D:97:LEU:HD23	2.31	0.65
1:C:202:HIS:CE1	1:D:203:PHE:CE2	2.91	0.57
1:C:202:HIS:CE1	1:D:203:PHE:CZ	2.97	0.52
1:A:271:HIS:CE1	1:A:287:TRP:CE3	3.00	0.49
1:C:296:VAL:HG12	1:C:307:VAL:HA	1.95	0.49

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	362/382 (95%)	352 (97%)	8 (2%)	2 (1%)	27	70
1	B	359/382 (94%)	348 (97%)	9 (2%)	2 (1%)	27	70
1	C	362/382 (95%)	345 (95%)	15 (4%)	2 (1%)	27	70
1	D	359/382 (94%)	340 (95%)	18 (5%)	1 (0%)	43	81
All	All	1442/1528 (94%)	1385 (96%)	50 (4%)	7 (0%)	35	74

5 of 7 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	266	LEU
1	C	323	SER
1	B	257	PRO
1	C	266	LEU
1	D	257	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	312/325 (96%)	293 (94%)	19 (6%)	20	52
1	B	310/325 (95%)	291 (94%)	19 (6%)	20	52
1	C	312/325 (96%)	294 (94%)	18 (6%)	22	52
1	D	310/325 (95%)	294 (95%)	16 (5%)	25	56
All	All	1244/1300 (96%)	1172 (94%)	72 (6%)	26	52

5 of 72 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	260	ARG
1	C	124	ARG
1	D	260	ARG
1	B	269	GLU
1	B	331	VAL

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 7 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	207	HIS
1	D	207	HIS
1	C	225	HIS
1	B	51	HIS
1	D	51	HIS

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.