wwPDB EM Map/Model Validation Report (i)

Apr 10, 2016 – 01:34 PM BST

PDB ID	:	2HI5
EMDB ID:	:	EMD-1240
Title	:	Model for bacteriophage fd from cryo-EM
Authors	:	Wang, Y.A.; Yu, X.; Overman, S.; Tsuboi, M.; Thomas, G.J.; Egelman, E.H.
Deposited on	:	2006-06-29
Resolution	:	8.00 Å(reported)

This is a wwPDB EM Map/Model Validation Report for a publicly released PDB/EMDB entry.
For rigid body fitted models, validation errors reported here could
stem from errors in the original structure(s) used in the fitting.
We welcome your comments at validation@mail.wwpdb.org
A user guide is available at
http://wwpdb.org/validation/2016/EMValidationReportHelp

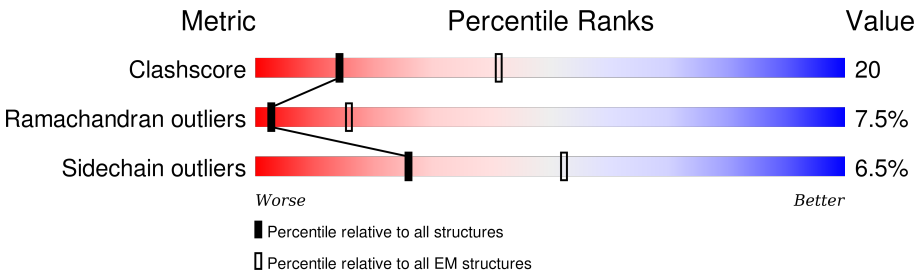
MolProbity	:	4.02b-467
Mogul	:	unknown
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP)	:	trunk27241

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 8.00 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	114402	924
Ramachandran outliers	111179	726
Sidechain outliers	111093	686

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	50	32% 36% 14% • 16%

2 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 315 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 Coat protein B.

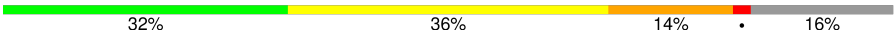
Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
1	A	42	315	211	48	55	1	0	0

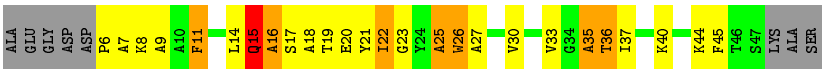
3 Residue-property plots (i)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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: Coat protein B

Chain A:





4 Experimental information (i)

Property	Value	Source
Reconstruction method	HELICAL	Depositor
Imposed symmetry	POINT, Not provided	Depositor
Number of images	Not provided	Depositor
Resolution determination method	0.5 FSC cut-off	Depositor
CTF correction method	phase flipping of entire images	Depositor
Microscope	Tecnai F20	Depositor
Voltage (kV)	200	Depositor
Electron dose ($e^-/\text{Å}^2$)	Not provided	Depositor
Minimum defocus (nm)	1400	Depositor
Maximum defocus (nm)	3800	Depositor
Magnification	50000	Depositor
Image detector	KODAK SO163 FILM	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	4.84	4/322 (1.2%)	3.46	8/435 (1.8%)

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	4

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	15	GLN	C-N	-54.20	0.09	1.34
1	A	35	ALA	C-N	-52.68	0.12	1.34
1	A	25	ALA	C-N	-38.95	0.44	1.34
1	A	26	TRP	NE1-CE2	-7.18	1.28	1.37

The worst 5 of 8 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	35	ALA	O-C-N	-38.74	60.72	122.70
1	A	15	GLN	O-C-N	-35.76	65.48	122.70
1	A	35	ALA	C-N-CA	-20.74	69.86	121.70
1	A	15	GLN	CA-C-N	19.83	160.82	117.20
1	A	25	ALA	C-N-CA	-18.20	76.21	121.70

There are no chirality outliers.

All (4) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	15	GLN	Mainchain,Peptide

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Group
1	A	25	ALA	Mainchain
1	A	35	ALA	Mainchain

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	315	0	324	13	81
All	All	315	0	324	13	81

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 20.

The worst 5 of 13 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:14:LEU:HA	1:A:17:SER:OG	1.58	1.01
1:A:14:LEU:CA	1:A:17:SER:OG	2.38	0.68
1:A:22:ILE:HG23	1:A:26:TRP:CH2	2.29	0.67
1:A:33:VAL:O	1:A:37:ILE:HG13	2.07	0.55
1:A:23:GLY:HA2	1:A:26:TRP:CE3	2.44	0.53

The worst 5 of 81 symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:11:PHE:CD2	1:A:22:ILE:CG2[1_556]	0.24	1.96
1:A:7:ALA:CB	1:A:19:THR:N[1_556]	0.25	1.95
1:A:7:ALA:C	1:A:19:THR:CB[1_556]	0.67	1.53
1:A:8:LYS:N	1:A:19:THR:OG1[1_556]	0.87	1.33
1:A:11:PHE:CE1	1:A:22:ILE:O[1_556]	0.87	1.33

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	40/50 (80%)	35 (88%)	2 (5%)	3 (8%)	1 21

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	16	ALA
1	A	36	THR
1	A	15	GLN

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	31/36 (86%)	29 (94%)	2 (6%)	21 58

All (2) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	11	PHE
1	A	22	ILE

Some 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 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.