wwPDB EM Map/Model Validation Report (i)

Apr 10, 2016 – 01:56 PM BST

PDB ID	:	4BTQ
EMDB ID:	:	EMD-1206
Title	:	Coordinates of the bacteriophage phi6 capsid subunits fitted into the cryoEM map EMD-1206
Authors	:	Nemecek, D.; Boura, E.; Wu, W.; Cheng, N.; Plevka, P.; Qiao, J.; Mindich, L.; Heymann, J.B.; Hurley, J.H.; Steven, A.C.
Deposited on	:	2013-06-18
Resolution	:	7.50 Å(reported)
Based on PDB ID	:	4K7H

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

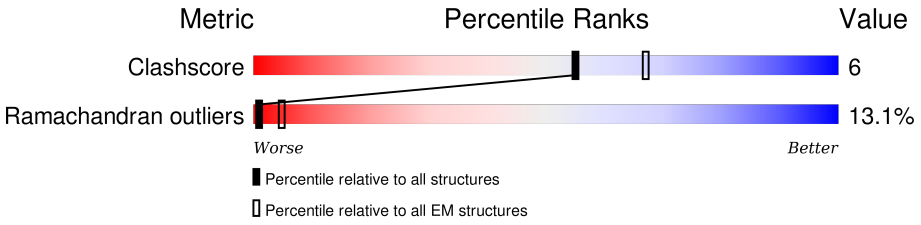
MolProbity	:	4.02b-467
Mogul	:	unknown
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP)	:	trunk27241

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 7.50 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

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	761	73% 23% .
1	B	761	66% 28% 5% .

2 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 6086 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 MAJOR INNER PROTEIN P1.

Mol	Chain	Residues	Atoms				AltConf	Trace
1	A	761	Total	C	N	O	0	0
			3043	1522	761	760		
1	B	761	Total	C	N	O	0	0
			3043	1522	761	760		

There are 2 discrepancies between the modelled and reference sequences:

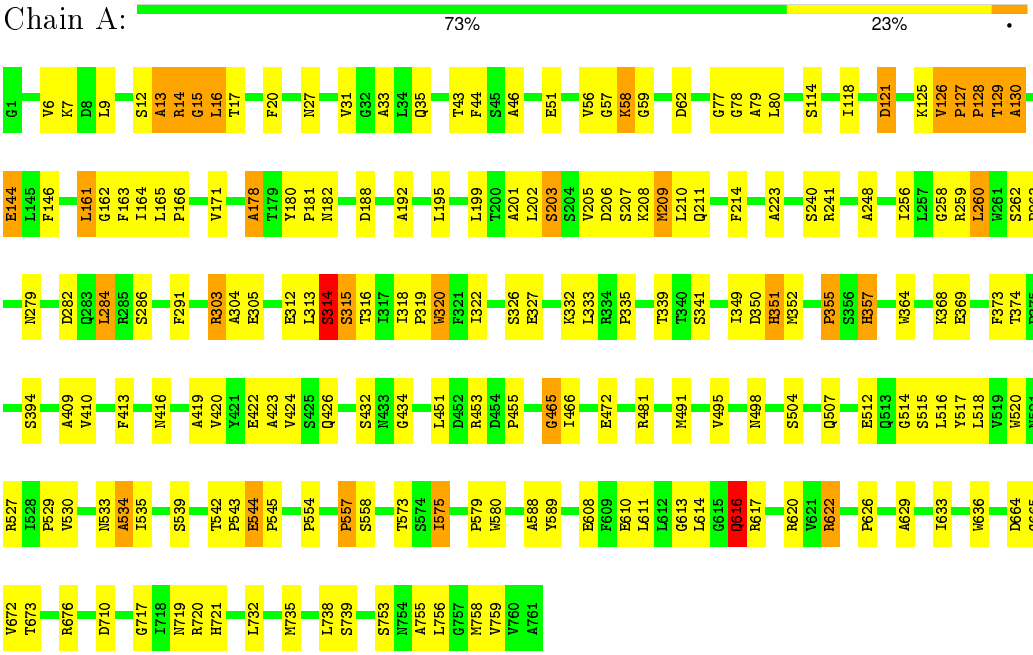
Chain	Residue	Modelled	Actual	Comment	Reference
A	1	GLY	-	EXPRESSION TAG	UNP P11126
B	1	GLY	-	EXPRESSION TAG	UNP P11126

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.

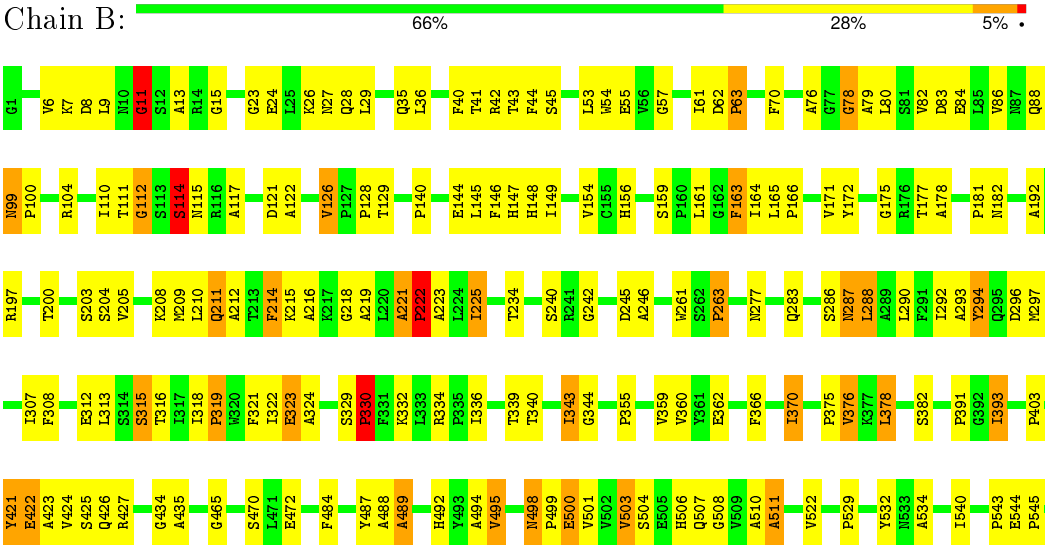
- Molecule 1: MAJOR INNER PROTEIN P1

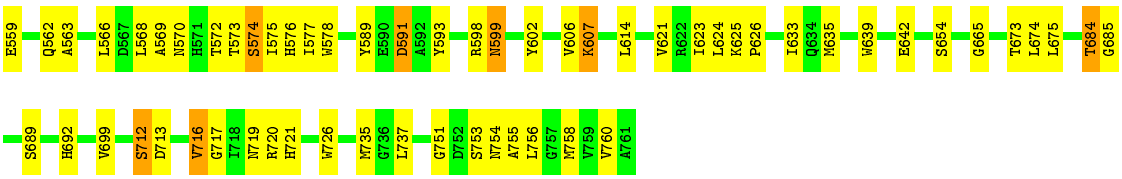
Chain A:



- Molecule 1: MAJOR INNER PROTEIN P1

Chain B:




E559
Q562
A563
L566
D567
L568
A569
N570
H571
T572
T573
S574
I575
H576
I577
W578
Y589
E590
D591
A592
Y593
R598
N599
Y602
V606
K607
L614
V621
R622
I623
L624
K625
P626
I633
Q634
M635
W639
E642
S654
G665
T673
L674
L675
T684
G685
S689
H692
V699
S712
D713
V716
G717
I718
N719
R720
H721
W726
M735
G736
L737
G751
D752
S753
N754
A755
L756
G757
M758
V759
V760
A761

4 Experimental information (i)

Property	Value	Source
Reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, I	Depositor
Number of images	Not provided	Depositor
Resolution determination method	Not provided	Depositor
CTF correction method	Not provided	Depositor
Microscope	FEI TECNAI F20	Depositor
Voltage (kV)	200	Depositor
Electron dose ($e^-/Å^2$)	Not provided	Depositor
Minimum defocus (nm)	Not provided	Depositor
Maximum defocus (nm)	Not provided	Depositor
Magnification	Not provided	Depositor
Image detector	Not provided	Depositor

5 Model quality

5.1 Standard geometry

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	2.08	33/3042 (1.1%)	1.70	53/3801 (1.4%)
1	B	2.43	81/3041 (2.7%)	1.82	66/3798 (1.7%)
All	All	2.26	114/6083 (1.9%)	1.76	119/7599 (1.6%)

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	31
1	B	0	42
All	All	0	73

The worst 5 of 114 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	465	GLY	CA-C	-7.93	1.39	1.51
1	B	719	ASN	N-CA	-7.11	1.32	1.46
1	B	635	MET	CA-C	-7.09	1.34	1.52
1	B	100	PRO	CA-C	-6.92	1.39	1.52
1	B	288	LEU	CA-C	-6.74	1.35	1.52

The worst 5 of 119 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	287	ASN	CA-C-O	-10.94	97.14	120.10
1	B	313	LEU	N-CA-C	-10.60	82.37	111.00
1	A	15	GLY	N-CA-C	-9.85	88.48	113.10
1	A	78	GLY	N-CA-C	-9.42	89.54	113.10
1	A	426	GLN	N-CA-C	-9.32	85.83	111.00

There are no chirality outliers.

5 of 73 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	129	THR	Peptide
1	A	130	ALA	Mainchain
1	A	161	LEU	Peptide
1	A	162	GLY	Peptide
1	A	35	GLN	Mainchain,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	3043	0	820	25	0
1	B	3043	0	817	25	0
All	All	6086	0	1637	50	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 6.

The worst 5 of 50 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:ARG:H	1:A:465:GLY:HA3	1.67	0.59
1:B:222:PRO:CA	1:B:225:ILE:H	2.17	0.58
1:A:504:SER:H	1:A:517:TYR:H	1.52	0.57
1:B:171:VAL:N	1:B:576:HIS:H	2.03	0.56
1:A:118:ILE:N	1:A:121:ASP:H	2.05	0.55

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	759/761 (100%)	561 (74%)	98 (13%)	100 (13%)	0 7
1	B	757/761 (100%)	559 (74%)	100 (13%)	98 (13%)	0 7
All	All	1516/1522 (100%)	1120 (74%)	198 (13%)	198 (13%)	1 7

5 of 198 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	13	ALA
1	A	16	LEU
1	A	17	THR
1	A	46	ALA
1	A	51	GLU

5.3.2 Protein sidechains (i)

There are no protein residues with a non-rotameric sidechain to report in this entry.

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.