wwPDB EM Map/Model Validation Report (i)

Apr 10, 2016 – 01:56 PM BST

PDB ID	:	4BTG
EMDB ID:	:	EMD-2364
Title	:	Coordinates of the bacteriophage phi6 capsid subunits (P1A and P1B) fitted into the cryoEM reconstruction of the procapsid at 4.4 A resolution
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	:	2013-06-17
Resolution	:	4.40 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

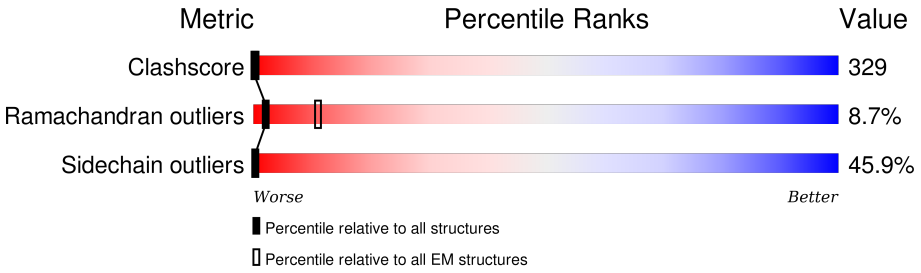
MolProbity	:	4.02b-467
Mogul	:	unknown
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP)	:	trunk27241

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 4.40 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	761	5% 40% 41% 14%
1	B	761	. 48% 38% 10%

2 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 11840 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	S	0	0
			5920	3741	1048	1109	22		
1	B	761	Total	C	N	O	S	0	0
			5920	3741	1048	1109	22		

There are 2 discrepancies between the modelled and reference sequences:

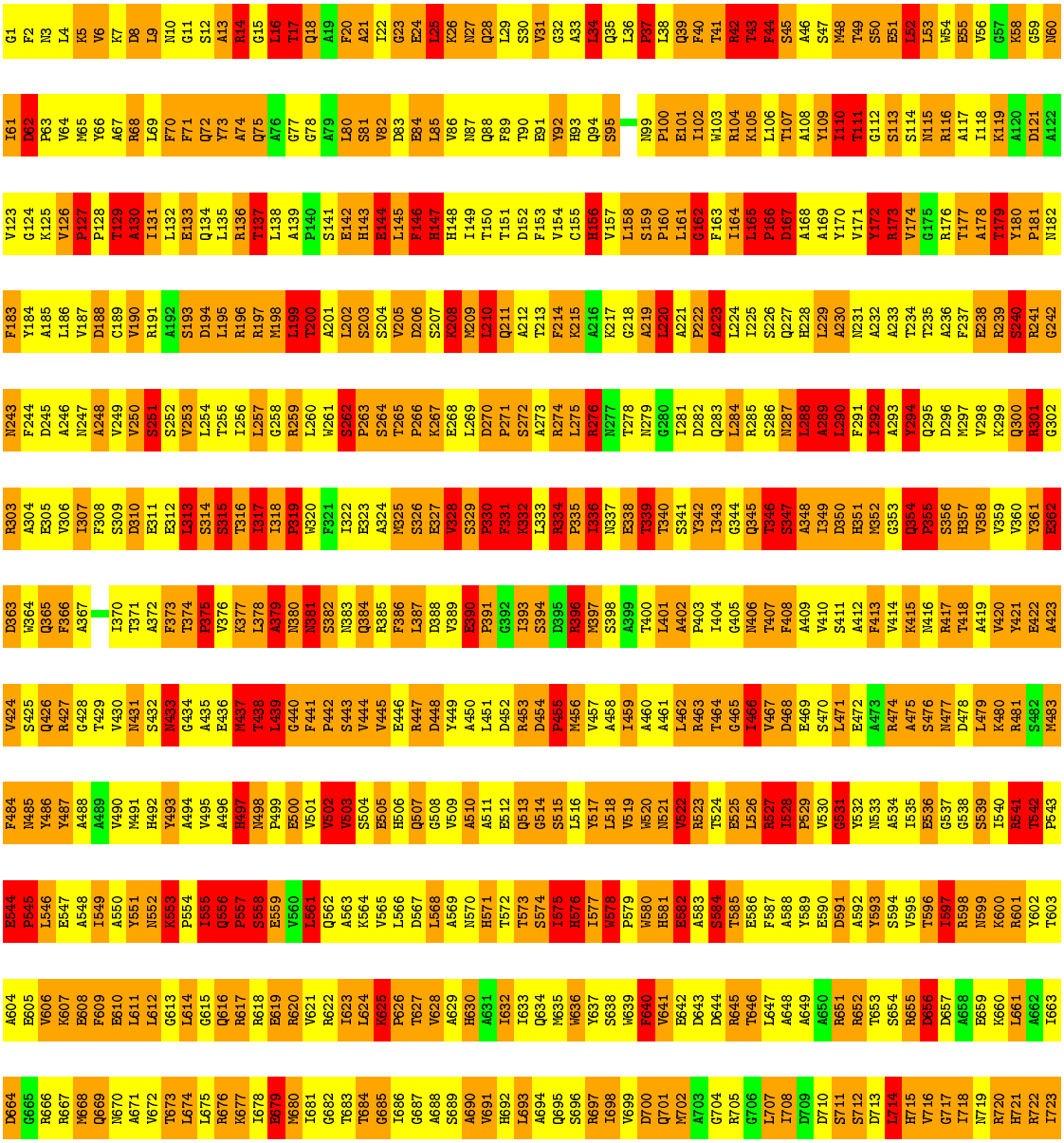
Chain	Residue	Modelled	Actual	Comment	Reference
A	1	GLY	-	EXPRESSION TAG	UNP P11126
B	1	GLY	-	EXPRESSION TAG	UNP P11126

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.

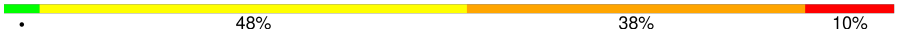
- Molecule 1: MAJOR INNER PROTEIN P1

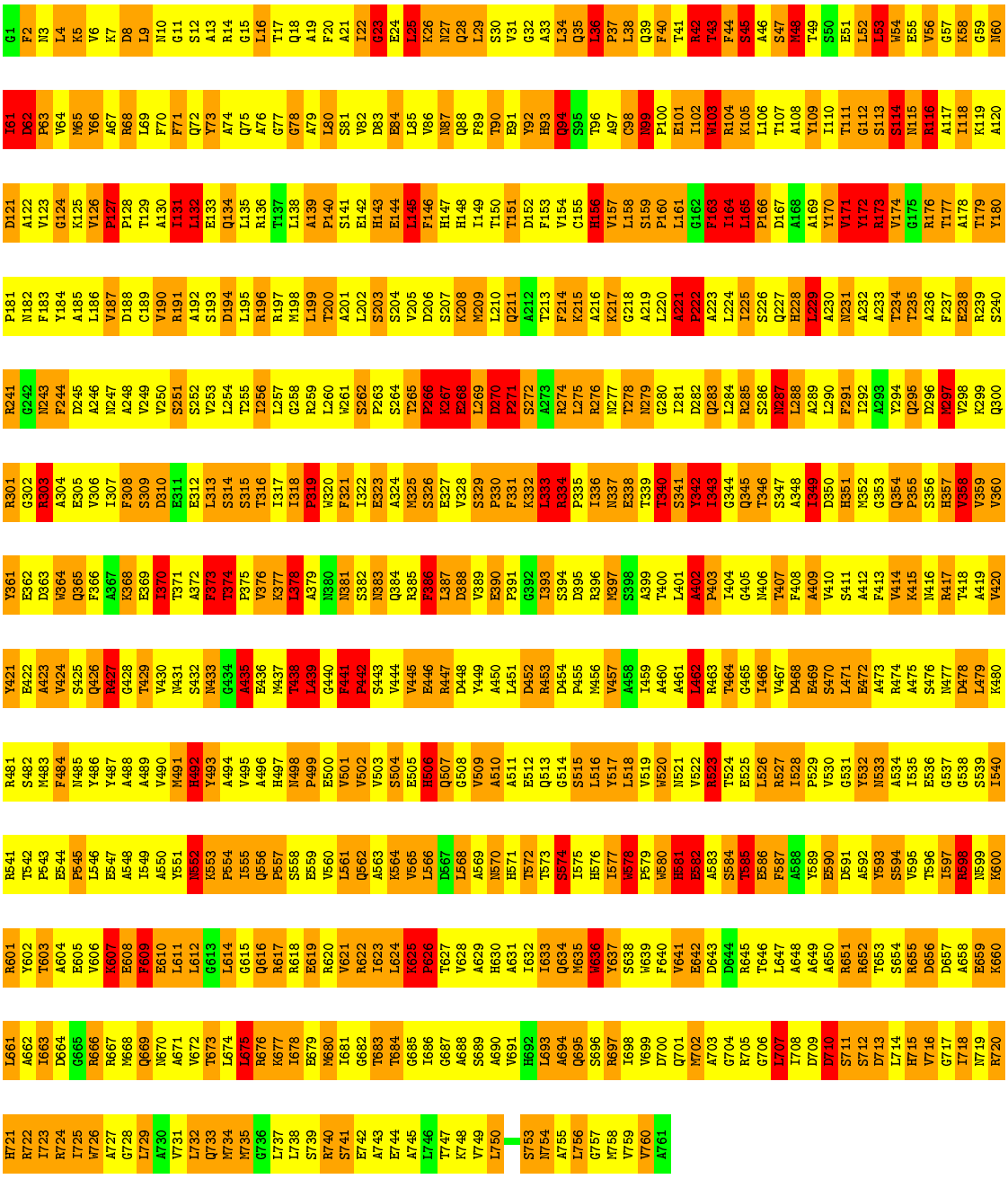
Chain A: 5% 40% 41% 14%



- Molecule 1: MAJOR INNER PROTEIN P1

Chain B:





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	PARTICLES FROM EACH MICROGRAPH	Depositor
Microscope	OTHER	Depositor
Voltage (kV)	300	Depositor
Electron dose ($e^-/Å^2$)	15	Depositor
Minimum defocus (nm)	800	Depositor
Maximum defocus (nm)	2200	Depositor
Magnification	46000	Depositor
Image detector	KODAK SO163 FILM	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	1.22	32/6038 (0.5%)	1.56	144/8200 (1.8%)
1	B	0.96	27/6039 (0.4%)	1.41	88/8203 (1.1%)
All	All	1.10	59/12077 (0.5%)	1.49	232/16403 (1.4%)

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	67
1	B	0	45
All	All	0	112

The worst 5 of 59 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	466	ILE	C-N	-36.28	0.50	1.34
1	B	127	PRO	C-N	30.25	1.91	1.34
1	A	437	MET	C-N	11.92	1.61	1.34
1	A	75	GLN	CA-CB	-11.01	1.29	1.53
1	B	124	GLY	C-N	9.05	1.54	1.34

The worst 5 of 232 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	441	PHE	C-N-CD	-46.69	17.89	120.60
1	A	542	THR	C-N-CD	15.13	160.17	128.40
1	A	290	LEU	CB-CG-CD1	-13.07	88.79	111.00
1	A	16	LEU	CB-CG-CD1	-12.81	89.22	111.00
1	A	739	SER	O-C-N	-11.26	104.68	122.70

There are no chirality outliers.

5 of 112 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	23	GLY	Peptide
1	A	25	LEU	Peptide
1	A	34	LEU	Peptide
1	A	52	LEU	Peptide
1	A	60	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	5920	0	5870	3751	0
1	B	5920	0	5883	4050	0
All	All	11840	0	11753	7764	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 329.

The worst 5 of 7764 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:61:ILE:HG22	1:A:156:HIS:CD2	1.18	1.68
1:A:732:LEU:CD1	1:A:738:LEU:HD22	1.23	1.67
1:B:373:PHE:CD2	1:B:583:ALA:HB1	1.31	1.66
1:B:670:ASN:HD21	1:B:749:VAL:CG1	1.07	1.65
1:B:156:HIS:CD2	1:B:158:LEU:H	1.14	1.65

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	755/761 (99%)	575 (76%)	109 (14%)	71 (9%)	1 16
1	B	757/761 (100%)	591 (78%)	105 (14%)	61 (8%)	1 18
All	All	1512/1522 (99%)	1166 (77%)	214 (14%)	132 (9%)	2 17

5 of 132 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	34	LEU
1	A	61	ILE
1	A	119	LYS
1	A	123	VAL
1	A	144	GLU

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	629/629 (100%)	341 (54%)	288 (46%)	0 0
1	B	629/629 (100%)	340 (54%)	289 (46%)	0 0
All	All	1258/1258 (100%)	681 (54%)	577 (46%)	0 0

5 of 577 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	698	ILE
1	B	91	GLU
1	B	642	GLU
1	A	714	LEU
1	B	26	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 53 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	72	GLN
1	B	182	ASN
1	B	581	HIS
1	B	75	GLN
1	B	88	GLN

5.3.3 RNA

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates

There are no carbohydrates in this entry.

5.6 Ligand geometry

There are no ligands in this entry.

5.7 Other polymers

There are no such residues in this entry.

5.8 Polymer linkage issues

There are no chain breaks in this entry.