wwPDB EM Validation Summary Report (i)

Feb 12, 2024 – 12:02 AM EST

PDB ID	:	3DDX
Title	:	HK97 bacteriophage capsid Expansion Intermediate-II model
Authors	:	Lee, K.K.; Gan, L.; Conway, J.F.; Hendrix, R.W.; Steven, A.C.; Johnson, J.E.
Deposited on	:	2008-06-06
Resolution	:	Not provided
Based on initial model	:	1OHG

This is a wwPDB EM 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.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

The following versions of software and data (see references (i)) were used in the production of this report:

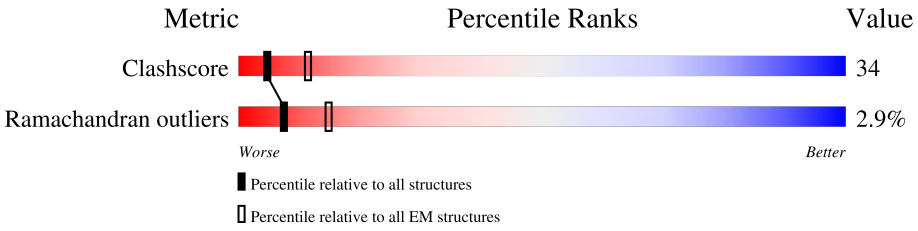
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

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 unknown.

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	158937	4297
Ramachandran outliers	154571	4023

The table below summarises the geometric issues observed across the polymeric chains and their fit to the map. The red, orange, yellow and green segments of the bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria respectively. 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	282	82% 17% •
1	B	282	76% 22% ••
1	C	282	80% 18% ••
1	D	282	78% 19% ••
1	E	282	84% 14% ••
1	F	282	79% 20% ••
1	G	282	88% 12% •

2 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 9653 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 capsid protein.

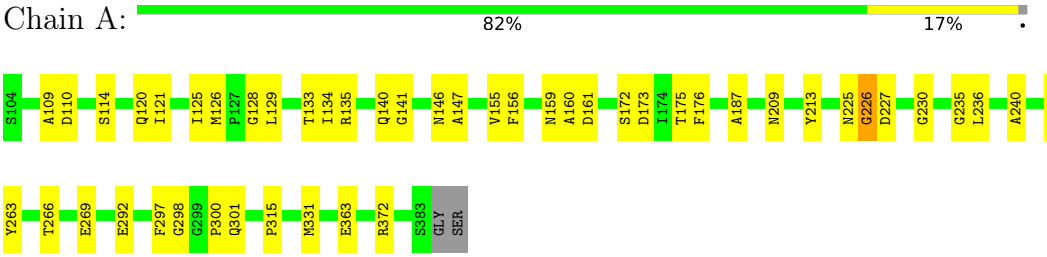
Mol	Chain	Residues	Atoms				AltConf	Trace
1	A	280	Total	C	N	O	0	0
			1379	819	280	280		
1	B	280	Total	C	N	O	0	0
			1379	819	280	280		
1	C	280	Total	C	N	O	0	0
			1379	819	280	280		
1	D	280	Total	C	N	O	0	0
			1379	819	280	280		
1	E	280	Total	C	N	O	0	0
			1379	819	280	280		
1	F	280	Total	C	N	O	0	0
			1379	819	280	280		
1	G	280	Total	C	N	O	0	0
			1379	819	280	280		

3 Residue-property plots (i)

These plots are drawn for all protein, RNA, DNA and oligosaccharide 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.

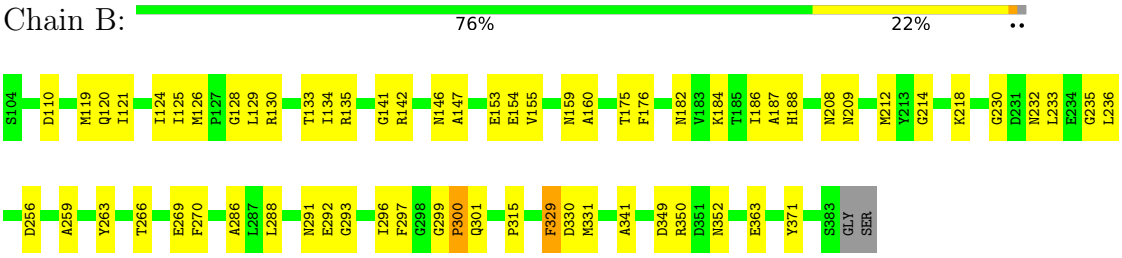
• Molecule 1: Major capsid protein

Chain A:



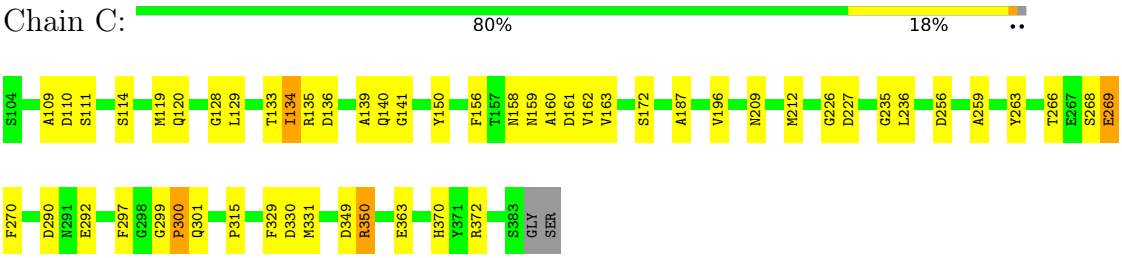
• Molecule 1: Major capsid protein

Chain B:



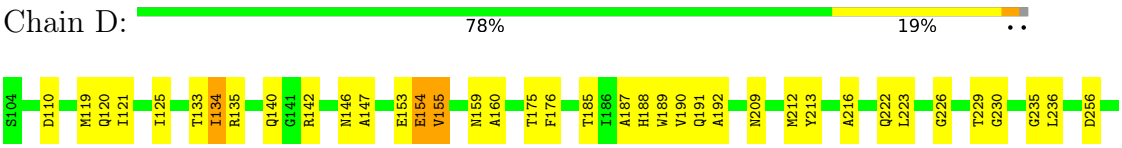
• Molecule 1: Major capsid protein

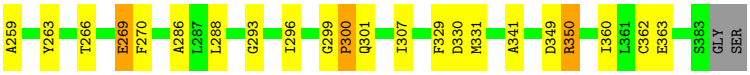
Chain C:



• Molecule 1: Major capsid protein

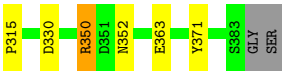
Chain D:





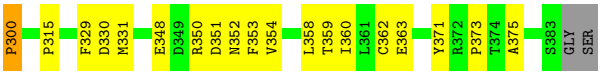
• Molecule 1: Major capsid protein

Chain E:



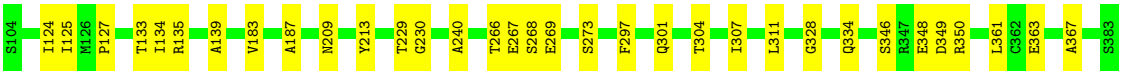
• Molecule 1: Major capsid protein

Chain F:



• Molecule 1: Major capsid protein

Chain G:



4 Data and refinement statistics (i)

Xtriage (Phenix) and EDS failed to run properly - this section is therefore incomplete.

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α, β, γ	1.00Å 1.00Å 1.00Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	(Not available) – (Not available)	Depositor
% Data completeness (in resolution range)	(Not available) ((Not available)-(Not available))	Depositor
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
Refinement program	unknown	Depositor
R, R_{free}	(Not available) , (Not available)	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	9653	wwPDB-VP
Average B, all atoms (Å^2)	100.0	wwPDB-VP

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	#$\|Z\|$ >5	RMSZ	#$\|Z\|$ >5
1	A	0.41	0/1375	0.71	1/1906 (0.1%)
1	B	0.42	0/1375	0.70	0/1906
1	C	0.44	0/1375	0.71	0/1906
1	D	0.60	1/1376 (0.1%)	0.97	3/1909 (0.2%)
1	E	0.44	0/1375	0.72	1/1906 (0.1%)
1	F	0.42	0/1375	0.72	0/1906
1	G	0.42	0/1375	0.71	0/1906
All	All	0.45	1/9626 (0.0%)	0.75	5/13345 (0.0%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	D	154	GLU	C-N	14.65	1.67	1.34

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	D	154	GLU	O-C-N	-21.77	87.86	122.70
1	D	154	GLU	C-N-CA	-15.39	83.23	121.70
1	D	154	GLU	CA-C-N	-14.38	85.57	117.20
1	A	298	GLY	N-CA-C	5.28	126.30	113.10
1	E	298	GLY	N-CA-C	5.18	126.05	113.10

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts

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	1379	0	653	132	0
1	B	1379	0	654	165	0
1	C	1379	0	651	137	0
1	D	1379	0	649	115	0
1	E	1379	0	649	93	0
1	F	1379	0	650	150	0
1	G	1379	0	653	29	0
All	All	9653	0	4559	481	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 34.

The worst 5 of 481 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:B:110:ASP:CB	1:C:209:ASN:HA	1.30	1.60
1:E:128:GLY:CA	1:F:270:PHE:CB	1.80	1.59
1:B:120:GLN:CB	1:C:141:GLY:HA3	1.21	1.58
1:A:160:ALA:CB	1:F:187:ALA:HB1	1.30	1.56
1:E:128:GLY:HA2	1:F:270:PHE:CB	1.30	1.56

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	272/282 (96%)	232 (85%)	34 (12%)	6 (2%)	6	6
1	B	272/282 (96%)	223 (82%)	42 (15%)	7 (3%)	5	5
1	C	272/282 (96%)	231 (85%)	32 (12%)	9 (3%)	4	4
1	D	274/282 (97%)	228 (83%)	37 (14%)	9 (3%)	4	4
1	E	272/282 (96%)	233 (86%)	32 (12%)	7 (3%)	5	5

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
1	F	272/282 (96%)	225 (83%)	39 (14%)	8 (3%)	4 4
1	G	272/282 (96%)	227 (84%)	36 (13%)	9 (3%)	4 4
All	All	1906/1974 (97%)	1599 (84%)	252 (13%)	55 (3%)	7 4

5 of 55 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	300	PRO
1	C	134	ILE
1	C	300	PRO
1	D	155	VAL
1	D	229	THR

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 monosaccharides 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)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
1	G	3
1	A	3
1	C	3
1	D	3
1	B	3
1	E	3
1	F	3

The worst 5 of 21 chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	G	125:ILE	C	126:MET	N	9.60
1	A	154:GLU	C	155:VAL	N	8.52
1	A	175:THR	C	176:PHE	N	8.34
1	C	175:THR	C	176:PHE	N	7.85
1	D	125:ILE	C	126:MET	N	6.89