



# Full wwPDB EM Validation Report ⓘ

Feb 24, 2025 – 06:00 PM EST

PDB ID : 2FTE  
Title : Bacteriophage HK97 Expansion Intermediate IV  
Authors : Gan, L.; Speir, J.A.; Conway, J.F.; Lander, G.; Cheng, N.; Firek, B.A.; Hendrix, R.W.; Duda, R.L.; Liljas, L.; Johnson, J.E.  
Deposited on : 2006-01-24  
Resolution : 12.00 Å (reported)  
Based on initial model : 2FSY

This is a Full wwPDB EM Validation Report for a publicly released PDB/EMDB entry.

We welcome your comments at [validation@mail.wwpdb.org](mailto: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 ⓘ 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 ⓘ](#)) were used in the production of this report:

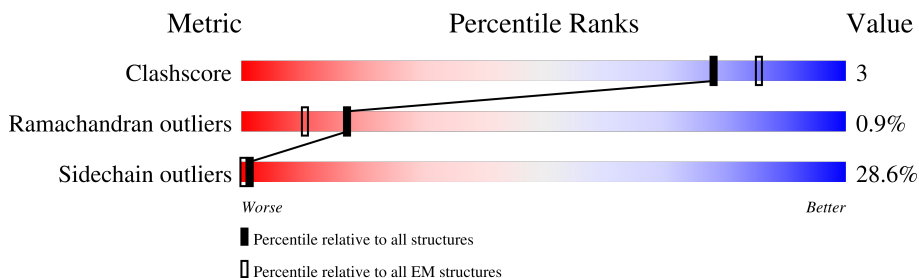
MolProbity : 4.02b-467  
Percentile statistics : 20231227.v01 (using entries in the PDB archive December 27th 2023)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.41.4

# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:  
*ELECTRON MICROSCOPY*

The reported resolution of this entry is 12.00 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	210492	15764
Ramachandran outliers	207382	16835
Sidechain outliers	206894	16415

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  $\geq 3$ , 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  $\leq 5\%$ .

Mol	Chain	Length	Quality of chain
1	A	282	
1	B	282	
1	C	282	
1	D	282	
1	E	282	
1	F	282	
1	G	282	

## 2 Entry composition

There is only 1 type of molecule in this entry. The entry contains 9387 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.

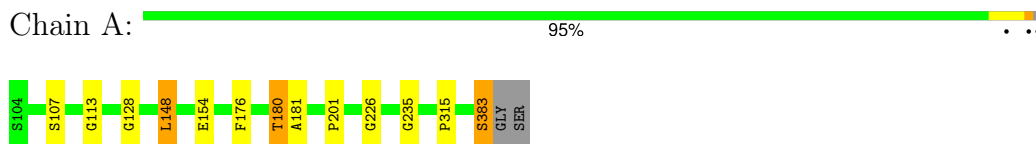
- Molecule 1 is a protein called major capsid protein.

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
1	A	280	1379	819	280	280	0	0
1	B	280	1378	818	280	280	0	0
1	C	280	1379	819	280	280	0	0
1	D	280	1379	819	280	280	0	0
1	E	280	1379	819	280	280	0	0
1	F	264	1301	773	264	264	0	0
1	G	242	1192	708	242	242	0	0

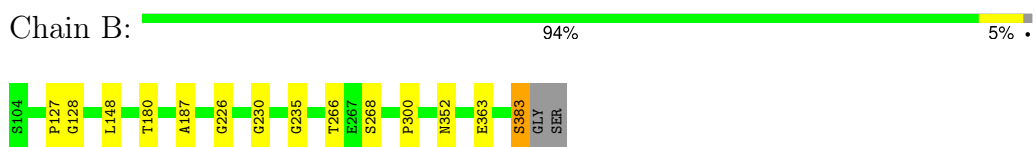
### 3 Residue-property plots [i](#)

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.

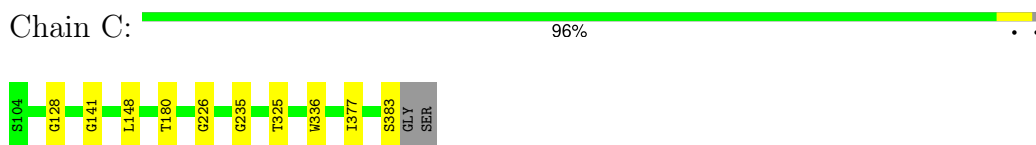
- Molecule 1: major capsid protein



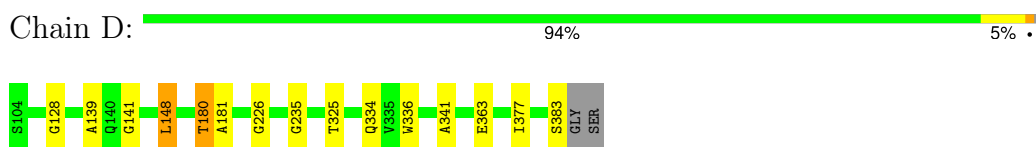
- Molecule 1: major capsid protein



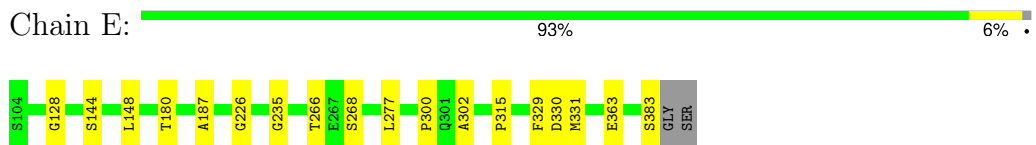
- Molecule 1: major capsid protein



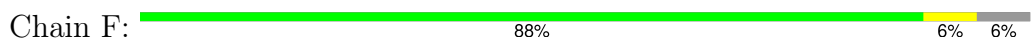
- Molecule 1: major capsid protein



- Molecule 1: major capsid protein



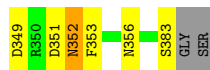
- Molecule 1: major capsid protein





- Molecule 1: major capsid protein

Chain G: 80% 5% • 14%



## 4 Data and refinement statistics

Xtrriage (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, $\alpha$ , $\beta$ , $\gamma$	1.00Å 1.00Å 1.00Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	(Not available) – 12.00	Depositor
% Data completeness (in resolution range)	(Not available) ((Not available)-12.00)	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.	Xtrriage
Total number of atoms	9387	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	404.0	wwPDB-VP

## 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  > 5$	RMSZ	# $ Z  > 5$
1	A	2.76	4/1378 (0.3%)	0.80	4/1915 (0.2%)
1	B	2.69	4/1377 (0.3%)	0.74	4/1913 (0.2%)
1	C	4.17	4/1378 (0.3%)	0.80	1/1915 (0.1%)
1	D	4.25	4/1378 (0.3%)	0.87	5/1915 (0.3%)
1	E	4.26	4/1378 (0.3%)	0.88	5/1915 (0.3%)
1	F	4.35	3/1300 (0.2%)	0.86	3/1807 (0.2%)
1	G	4.59	2/1190 (0.2%)	0.88	1/1652 (0.1%)
All	All	3.92	25/9379 (0.3%)	0.83	23/13032 (0.2%)

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	1
1	D	0	1
All	All	0	2

All (25) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	G	383	SER	CB-OG	157.42	3.46	1.42
1	E	383	SER	CB-OG	157.32	3.46	1.42
1	D	383	SER	CB-OG	156.24	3.45	1.42
1	F	383	SER	CB-OG	156.24	3.45	1.42
1	C	383	SER	CB-OG	154.22	3.42	1.42
1	A	383	SER	CB-OG	99.77	2.71	1.42
1	B	383	SER	CB-OG	98.93	2.70	1.42
1	D	180	THR	C-N	17.36	1.74	1.34
1	A	180	THR	C-N	17.17	1.73	1.34
1	G	180	THR	C-N	14.55	1.67	1.34
1	A	128	GLY	C-N	11.04	1.59	1.34

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	F	148	LEU	C-N	10.88	1.59	1.34
1	D	148	LEU	C-N	10.21	1.57	1.34
1	E	180	THR	C-N	9.53	1.55	1.34
1	E	128	GLY	C-N	8.71	1.54	1.34
1	B	128	GLY	C-N	8.23	1.52	1.34
1	A	148	LEU	C-N	7.47	1.51	1.34
1	D	128	GLY	C-N	7.43	1.51	1.34
1	C	128	GLY	C-N	7.01	1.50	1.34
1	C	180	THR	C-N	6.81	1.49	1.34
1	C	148	LEU	C-N	6.19	1.48	1.34
1	F	180	THR	C-N	5.92	1.47	1.34
1	E	148	LEU	C-N	5.82	1.47	1.34
1	B	148	LEU	C-N	5.48	1.46	1.34
1	B	180	THR	C-N	5.20	1.46	1.34

All (23) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	E	383	SER	CA-CB-OG	-28.02	35.55	111.20
1	G	383	SER	CA-CB-OG	-28.00	35.59	111.20
1	F	383	SER	CA-CB-OG	-27.75	36.28	111.20
1	D	383	SER	CA-CB-OG	-27.75	36.28	111.20
1	C	383	SER	CA-CB-OG	-27.39	37.23	111.20
1	A	383	SER	CA-CB-OG	-18.24	61.95	111.20
1	B	383	SER	CA-CB-OG	-18.11	62.31	111.20
1	A	128	GLY	O-C-N	13.55	144.39	122.70
1	A	128	GLY	CA-C-N	-10.85	93.34	117.20
1	B	128	GLY	O-C-N	10.65	139.74	122.70
1	A	128	GLY	C-N-CA	-8.77	99.78	121.70
1	D	148	LEU	O-C-N	-8.62	108.91	122.70
1	E	128	GLY	O-C-N	8.50	136.29	122.70
1	B	128	GLY	CA-C-N	-8.04	99.51	117.20
1	E	128	GLY	CA-C-N	-6.86	102.10	117.20
1	B	128	GLY	C-N-CA	-6.79	104.73	121.70
1	D	180	THR	O-C-N	6.76	133.52	122.70
1	D	180	THR	CA-C-N	-6.72	102.40	117.20
1	E	128	GLY	C-N-CA	-5.78	107.26	121.70
1	F	127	PRO	N-CA-CB	5.65	110.08	103.30
1	F	122	PRO	N-CA-CB	5.62	110.05	103.30
1	E	148	LEU	O-C-N	-5.26	114.29	122.70
1	D	180	THR	C-N-CA	-5.04	109.09	121.70



There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	148	LEU	Mainchain
1	D	148	LEU	Mainchain

## 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	659	6	0
1	B	1378	0	659	3	0
1	C	1379	0	660	4	0
1	D	1379	0	659	9	0
1	E	1379	0	660	6	0
1	F	1301	0	621	5	0
1	G	1192	0	569	5	0
All	All	9387	0	4487	37	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 3.

All (37) 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:G:180:THR:C	1:G:181:ALA:N	1.67	1.43
1:A:180:THR:C	1:A:181:ALA:N	1.73	1.42
1:D:180:THR:C	1:D:181:ALA:N	1.74	1.38
1:D:180:THR:CA	1:D:181:ALA:N	2.54	0.70
1:D:180:THR:C	1:D:181:ALA:CA	2.61	0.69
1:B:226:GLY:H	1:B:235:GLY:HA3	1.61	0.66
1:A:226:GLY:H	1:A:235:GLY:HA3	1.64	0.62
1:C:226:GLY:HA3	1:C:235:GLY:H	1.69	0.57
1:E:187:ALA:HB2	1:E:363:GLU:HA	1.87	0.56
1:A:107:SER:O	1:A:113:GLY:HA3	2.07	0.55
1:C:141:GLY:O	1:C:336:TRP:HA	2.07	0.54
1:D:141:GLY:O	1:D:336:TRP:HA	2.08	0.53

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:226:GLY:H	1:D:235:GLY:HA3	1.76	0.50
1:F:226:GLY:H	1:F:235:GLY:HA3	1.75	0.50
1:F:343:VAL:HA	1:F:361:LEU:O	2.12	0.50
1:D:139:ALA:HB3	1:D:334:GLN:CB	2.41	0.50
1:G:139:ALA:HB3	1:G:334:GLN:CB	2.42	0.49
1:E:226:GLY:H	1:E:235:GLY:HA3	1.77	0.49
1:A:180:THR:CA	1:A:181:ALA:N	2.70	0.48
1:E:266:THR:C	1:E:268:SER:H	2.17	0.48
1:A:180:THR:C	1:A:181:ALA:CA	2.74	0.46
1:D:341:ALA:HA	1:D:363:GLU:O	2.16	0.46
1:C:325:THR:HA	1:C:377:ILE:O	2.16	0.46
1:C:226:GLY:H	1:C:235:GLY:HA3	1.81	0.46
1:A:154:GLU:H	1:A:176:PHE:HA	1.82	0.45
1:F:130:ARG:C	1:F:132:LEU:H	2.19	0.45
1:G:151:VAL:HA	1:G:178:LYS:HA	1.99	0.44
1:G:352:ASN:O	1:G:356:ASN:N	2.51	0.43
1:D:325:THR:HA	1:D:377:ILE:O	2.18	0.42
1:F:187:ALA:HB2	1:F:363:GLU:HA	2.01	0.42
1:B:266:THR:C	1:B:268:SER:H	2.23	0.42
1:E:302:ALA:HA	1:F:310:GLY:N	2.36	0.41
1:E:329:PHE:C	1:E:331:MET:H	2.24	0.41
1:G:351:ASP:O	1:G:353:PHE:N	2.54	0.41
1:D:226:GLY:N	1:D:235:GLY:HA3	2.35	0.40
1:B:187:ALA:HB2	1:B:363:GLU:HA	2.04	0.40
1:E:277:LEU:O	1:E:315:PRO:HA	2.22	0.40

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	278/282 (99%)	249 (90%)	27 (10%)	2 (1%)	19 57

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	B	278/282 (99%)	252 (91%)	22 (8%)	4 (1%)	9	41
1	C	278/282 (99%)	258 (93%)	20 (7%)	0	100	100
1	D	278/282 (99%)	254 (91%)	24 (9%)	0	100	100
1	E	278/282 (99%)	246 (88%)	29 (10%)	3 (1%)	12	47
1	F	262/282 (93%)	237 (90%)	22 (8%)	3 (1%)	12	47
1	G	238/282 (84%)	204 (86%)	29 (12%)	5 (2%)	5	30
All	All	1890/1974 (96%)	1700 (90%)	173 (9%)	17 (1%)	17	52

All (17) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	E	300	PRO
1	F	352	ASN
1	G	352	ASN
1	G	154	GLU
1	E	330	ASP
1	G	224	LEU
1	A	201	PRO
1	B	127	PRO
1	B	230	GLY
1	G	349	ASP
1	B	352	ASN
1	E	144	SER
1	F	240	ALA
1	F	300	PRO
1	G	230	GLY
1	A	315	PRO
1	B	300	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.

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	1/231 (0%)	0	1 (100%)	0	0
1	B	1/231 (0%)	0	1 (100%)	0	0
1	C	1/231 (0%)	1 (100%)	0	100	100
1	D	1/231 (0%)	1 (100%)	0	100	100
1	E	1/231 (0%)	1 (100%)	0	100	100
1	F	1/231 (0%)	1 (100%)	0	100	100
1	G	1/231 (0%)	1 (100%)	0	100	100
All	All	7/1617 (0%)	5 (71%)	2 (29%)	1	2

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

Mol	Chain	Res	Type
1	A	383	SER
1	B	383	SER

Sometimes 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 oligosaccharides 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	D	1
1	A	1
1	G	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	D	180:THR	C	181:ALA	N	1.74
1	A	180:THR	C	181:ALA	N	1.73
1	G	180:THR	C	181:ALA	N	1.67