



## wwPDB/EMDataBank EM Map/Model Validation Summary Report ⓘ

Feb 20, 2018 – 09:41 pm GMT

PDB ID : 6F9C  
EMDB ID: : EMD-4198  
Title : Model of the Rift Valley fever virus glycoprotein hexamer type 1  
Authors : Halldorsson, S.; Bowden, T.A.; Huiskonen, J.T.  
Deposited on : 2017-12-14  
Resolution : 8.00 Å(reported)

This is a wwPDB/EMDataBank EM Map/Model Validation Summary 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.

---

MolProbity : 4.02b-467  
Percentile statistics : 20171227.v01 (using entries in the PDB archive December 27th 2017)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)  
Validation Pipeline (wwPDB-VP) : trunk30686

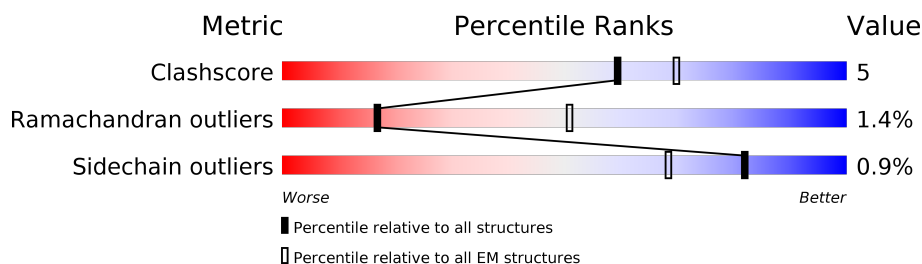
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*ELECTRON MICROSCOPY*

The reported resolution of this entry is 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	136279	1886
Ramachandran outliers	132675	1663
Sidechain outliers	132484	1531

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

Mol	Chain	Length	Quality of chain
1	A	316	81% 14% 5%
1	C	316	81% 14% • 5%
1	E	316	82% 13% • 5%
1	G	316	82% 13% • 5%
1	I	316	78% 16% • 5%
1	K	316	83% 12% 5%
2	B	431	86% 14%
2	D	431	82% 18%
2	F	431	83% 17%

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain	
2	H	431		16%
2	J	431		16%
2	L	431		17%

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 33048 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 Glycoprotein.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	301	Total	C	N	O	S	0	0
			2284	1437	400	441	6		
1	C	301	Total	C	N	O	S	0	0
			2284	1437	400	441	6		
1	E	301	Total	C	N	O	S	0	0
			2284	1437	400	441	6		
1	G	301	Total	C	N	O	S	0	0
			2284	1437	400	441	6		
1	I	301	Total	C	N	O	S	0	0
			2284	1437	400	441	6		
1	K	301	Total	C	N	O	S	0	0
			2284	1437	400	441	6		

- Molecule 2 is a protein called Glycoprotein.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	B	431	Total	C	N	O	S	0	0
			3224	2006	561	652	5		
2	D	431	Total	C	N	O	S	0	0
			3224	2006	561	652	5		
2	F	431	Total	C	N	O	S	0	0
			3224	2006	561	652	5		
2	H	431	Total	C	N	O	S	0	0
			3224	2006	561	652	5		
2	J	431	Total	C	N	O	S	0	0
			3224	2006	561	652	5		
2	L	431	Total	C	N	O	S	0	0
			3224	2006	561	652	5		

There are 18 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	688	ASP	-	expression tag	UNP A2T072
B	689	PRO	-	expression tag	UNP A2T072

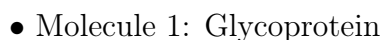
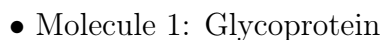
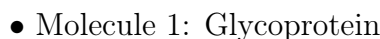
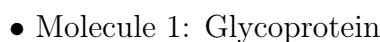
*Continued on next page...*

*Continued from previous page...*

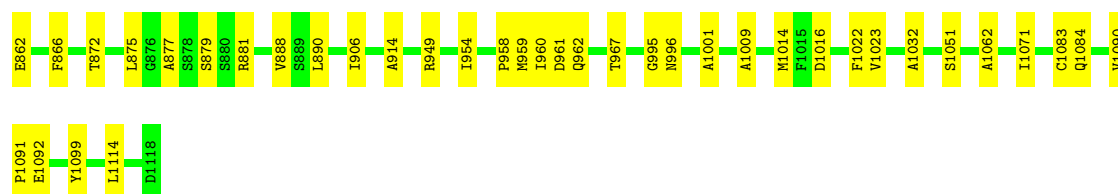
Chain	Residue	Modelled	Actual	Comment	Reference
B	690	GLY	-	expression tag	UNP A2T072
D	688	ASP	-	expression tag	UNP A2T072
D	689	PRO	-	expression tag	UNP A2T072
D	690	GLY	-	expression tag	UNP A2T072
F	688	ASP	-	expression tag	UNP A2T072
F	689	PRO	-	expression tag	UNP A2T072
F	690	GLY	-	expression tag	UNP A2T072
H	688	ASP	-	expression tag	UNP A2T072
H	689	PRO	-	expression tag	UNP A2T072
H	690	GLY	-	expression tag	UNP A2T072
J	688	ASP	-	expression tag	UNP A2T072
J	689	PRO	-	expression tag	UNP A2T072
J	690	GLY	-	expression tag	UNP A2T072
L	688	ASP	-	expression tag	UNP A2T072
L	689	PRO	-	expression tag	UNP A2T072
L	690	GLY	-	expression tag	UNP A2T072



- Molecule 1: Glycoprotein

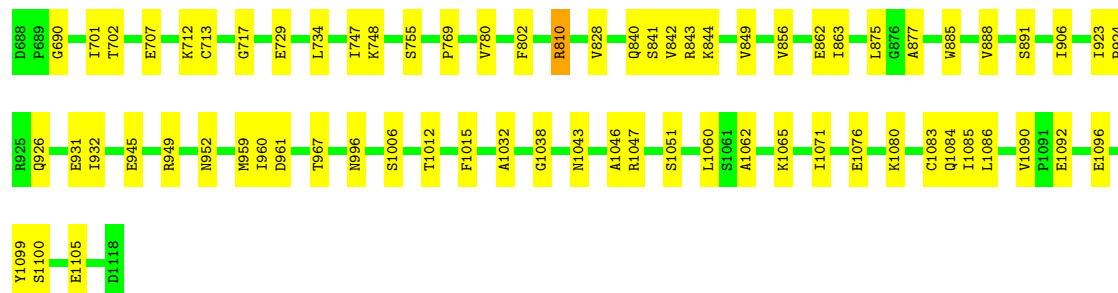






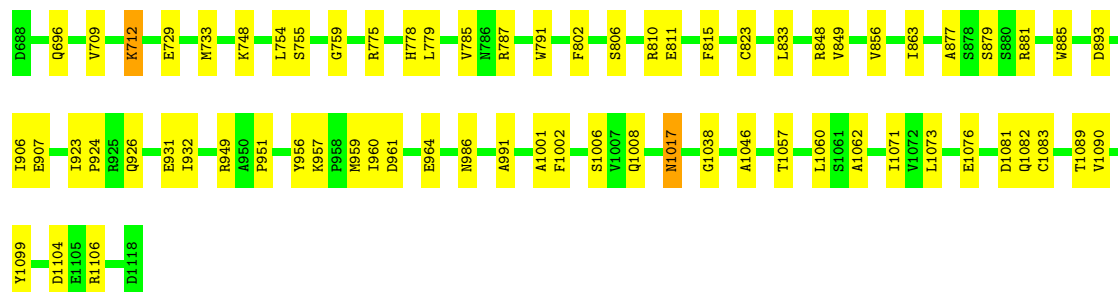
• Molecule 2: Glycoprotein

Chain H: 84% 16%



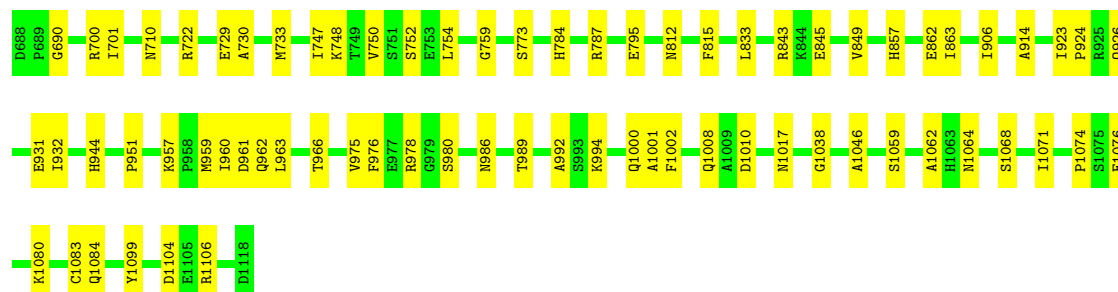
• Molecule 2: Glycoprotein

Chain J: 84% 16%



• Molecule 2: Glycoprotein

Chain L: 83% 17%





## 4 Experimental information

Property	Value	Source
Reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, C6	Depositor
Number of particles used	55710	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	PHASE FLIPPING AND AMPLITUDE CORRECTION	Depositor
Microscope	FEI TECNAI F30	Depositor
Voltage (kV)	300	Depositor
Electron dose ( $e^-/\text{\AA}^2$ )	22	Depositor
Minimum defocus (nm)	Not provided	Depositor
Maximum defocus (nm)	Not provided	Depositor
Magnification	Not provided	Depositor
Image detector	GATAN K2 SUMMIT (4k x 4k)	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	0.24	0/2333	0.41	0/3136
1	C	0.24	0/2333	0.41	0/3136
1	E	0.24	0/2333	0.40	0/3136
1	G	0.24	0/2333	0.40	0/3136
1	I	0.24	0/2333	0.40	0/3136
1	K	0.24	0/2333	0.40	0/3136
2	B	0.24	0/3284	0.42	0/4431
2	D	0.24	0/3284	0.43	0/4431
2	F	0.24	0/3284	0.43	0/4431
2	H	0.24	0/3284	0.42	0/4431
2	J	0.24	0/3284	0.42	0/4431
2	L	0.24	0/3284	0.42	0/4431
All	All	0.24	0/33702	0.41	0/45402

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

### 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	2284	0	2205	24	0
1	C	2284	0	2205	25	0
1	E	2284	0	2205	24	0
1	G	2284	0	2205	26	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	I	2284	0	2205	25	0
1	K	2284	0	2205	17	0
2	B	3224	0	3071	29	0
2	D	3224	0	3071	37	0
2	F	3224	0	3071	39	0
2	H	3224	0	3071	38	0
2	J	3224	0	3071	34	0
2	L	3224	0	3071	35	0
All	All	33048	0	31656	334	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 5.

The worst 5 of 334 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:412:LYS:HD2	1:A:426:THR:HG21	1.76	0.68
1:A:261:ASN:HD21	1:A:265:LYS:HB2	1.60	0.66
2:L:843:ARG:HG3	2:L:845:GLU:H	1.61	0.65
1:E:262:ASN:HD22	1:E:326:VAL:HG22	1.60	0.65
2:L:1064:ASN:HD21	2:L:1068:SER:H	1.44	0.65

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	295/316 (93%)	250 (85%)	43 (15%)	2 (1%)	24	67
1	C	295/316 (93%)	259 (88%)	33 (11%)	3 (1%)	17	60
1	E	295/316 (93%)	258 (88%)	32 (11%)	5 (2%)	10	49

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	G	295/316 (93%)	255 (86%)	36 (12%)	4 (1%)	12	52
1	I	295/316 (93%)	246 (83%)	41 (14%)	8 (3%)	5	38
1	K	295/316 (93%)	255 (86%)	35 (12%)	5 (2%)	10	49
2	B	429/431 (100%)	375 (87%)	49 (11%)	5 (1%)	14	56
2	D	429/431 (100%)	384 (90%)	39 (9%)	6 (1%)	12	52
2	F	429/431 (100%)	370 (86%)	54 (13%)	5 (1%)	14	56
2	H	429/431 (100%)	376 (88%)	49 (11%)	4 (1%)	19	61
2	J	429/431 (100%)	380 (89%)	42 (10%)	7 (2%)	11	50
2	L	429/431 (100%)	377 (88%)	47 (11%)	5 (1%)	14	56
All	All	4344/4482 (97%)	3785 (87%)	500 (12%)	59 (1%)	17	52

5 of 59 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	1083	CYS
1	E	281	CYS
2	F	1083	CYS
1	G	239	CYS
1	I	456	CYS

### 5.3.2 Protein sidechains ⓘ

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	241/271 (89%)	238 (99%)	3 (1%)	74	87
1	C	241/271 (89%)	240 (100%)	1 (0%)	92	95
1	E	241/271 (89%)	239 (99%)	2 (1%)	83	92
1	G	241/271 (89%)	240 (100%)	1 (0%)	92	95
1	I	241/271 (89%)	239 (99%)	2 (1%)	83	92
1	K	241/271 (89%)	239 (99%)	2 (1%)	83	92
2	B	347/371 (94%)	344 (99%)	3 (1%)	81	90

Continued on next page...

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
2	D	347/371 (94%)	345 (99%)	2 (1%)	87	93
2	F	347/371 (94%)	343 (99%)	4 (1%)	74	87
2	H	347/371 (94%)	343 (99%)	4 (1%)	74	87
2	J	347/371 (94%)	345 (99%)	2 (1%)	87	93
2	L	347/371 (94%)	343 (99%)	4 (1%)	74	87
All	All	3528/3852 (92%)	3498 (99%)	30 (1%)	82	90

5 of 30 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	F	844	LYS
2	H	810	ARG
2	L	733	MET
1	G	366	ASN
2	H	844	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 29 such sidechains are listed below:

Mol	Chain	Res	Type
1	G	377	HIS
2	H	829	ASN
2	L	944	HIS
1	G	438	ASN
2	H	996	ASN

### 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 [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.