



wwPDB X-ray Structure Validation Summary Report

Mar 31, 2014 – 05:35 PM BST

PDB ID : 4H6F
Title : Crystal Structure of Rift Valley Fever Virus Nucleocapsid Protein Hexamer Bound to Single-stranded RNA. This entry contains three out of six hexamers bound to RNA.
Authors : Raymond, D.D.; Smith, J.L.
Deposited on : 2012-09-19
Resolution : 3.40 Å(reported)

This is a wwPDB validation summary report for a publicly released PDB entry.
We welcome your comments at validation@mail.wwpdb.org
A user guide is available at <http://wwpdb.org/ValidationPDFNotes.html>

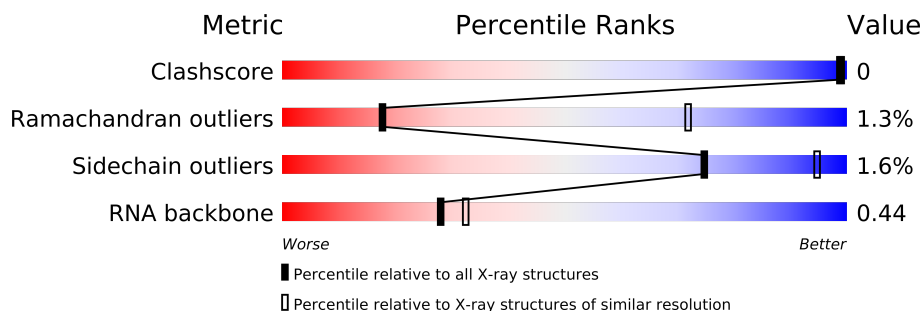
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity : 4.02b-467
Mogul : 1.15 2013
Xtriage (Phenix) : dev-1323
EDS : **NOT EXECUTED**
Percentile statistics : 21963
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP) : stable23004

1 Overall quality at a glance

The reported resolution of this entry is 3.40 Å.













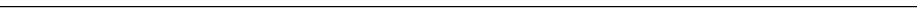


Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
Clashscore	79885	1214 (3.50-3.30)
Ramachandran outliers	78287	1177 (3.50-3.30)
Sidechain outliers	78261	1177 (3.50-3.30)
RNA backbone	1838	1002 (4.02-2.76)






The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	245	
1	B	245	
1	C	245	
1	D	245	
1	E	245	
1	F	245	
1	G	245	
1	H	245	
1	I	245	
1	J	245	
1	K	245	
1	L	245	
1	M	245	
1	N	245	
1	O	245	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	P	245	
1	Q	245	
1	R	245	
2	a	36	
2	g	36	
2	m	36	

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 36036 atoms, of which 0 are hydrogen and 0 are deuterium.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, 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 Nucleocapsid protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	241	Total	C	N	O	S	0	0	0
			1882	1188	339	343	12			
1	B	241	Total	C	N	O	S	0	0	0
			1882	1188	339	343	12			
1	C	241	Total	C	N	O	S	0	0	0
			1882	1188	339	343	12			
1	D	241	Total	C	N	O	S	0	0	0
			1882	1188	339	343	12			
1	E	241	Total	C	N	O	S	0	0	0
			1882	1188	339	343	12			
1	F	241	Total	C	N	O	S	0	0	0
			1882	1188	339	343	12			
1	G	241	Total	C	N	O	S	0	0	0
			1882	1188	339	343	12			
1	H	241	Total	C	N	O	S	0	0	0
			1882	1188	339	343	12			
1	I	241	Total	C	N	O	S	0	0	0
			1882	1188	339	343	12			
1	J	241	Total	C	N	O	S	0	0	0
			1882	1188	339	343	12			
1	K	241	Total	C	N	O	S	0	0	0
			1882	1188	339	343	12			
1	L	241	Total	C	N	O	S	0	0	0
			1882	1188	339	343	12			
1	M	241	Total	C	N	O	S	0	0	0
			1882	1188	339	343	12			
1	N	241	Total	C	N	O	S	0	0	0
			1882	1188	339	343	12			
1	O	241	Total	C	N	O	S	0	0	0
			1882	1188	339	343	12			
1	P	241	Total	C	N	O	S	0	0	0
			1882	1188	339	343	12			

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	Q	241	Total	C	N	O	S	0	0	0
			1882	1188	339	343	12			
1	R	241	Total	C	N	O	S	0	0	0
			1882	1188	339	343	12			

- Molecule 2 is a RNA chain called 35-mer poly(U) RNA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	a	36	Total	C	N	O	P	0	0	0
			720	324	72	288	36			
2	g	36	Total	C	N	O	P	0	0	0
			720	324	72	288	36			
2	m	36	Total	C	N	O	P	0	0	0
			720	324	72	288	36			

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 and electron density. 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. A red dot above a residue indicates a poor fit to the electron density ($RSRZ > 2$). 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.

Note EDS was not executed.

- Molecule 1: Nucleocapsid protein

Chain A: 



- Molecule 1: Nucleocapsid protein

Chain B: 



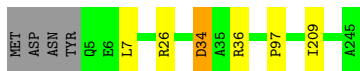
- Molecule 1: Nucleocapsid protein

Chain C: 



- Molecule 1: Nucleocapsid protein

Chain D: 



- Molecule 1: Nucleocapsid protein

Chain E: 



- Molecule 1: Nucleocapsid protein

Chain F: 



- Molecule 1: Nucleocapsid protein

Chain G: 



- Molecule 1: Nucleocapsid protein

Chain H: 



- Molecule 1: Nucleocapsid protein

Chain I: 



- Molecule 1: Nucleocapsid protein

Chain J: 



- Molecule 1: Nucleocapsid protein

Chain K: 



- Molecule 1: Nucleocapsid protein

Chain L: 



- Molecule 1: Nucleocapsid protein

Chain M: 



- Molecule 1: Nucleocapsid protein

Chain N: 



- Molecule 1: Nucleocapsid protein

Chain O: 



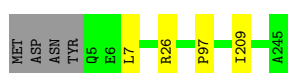
- Molecule 1: Nucleocapsid protein

Chain P: 



- Molecule 1: Nucleocapsid protein

Chain Q: 



- Molecule 1: Nucleocapsid protein

Chain R: 



- Molecule 2: 35-mer poly(U) RNA

Chain a: 



- Molecule 2: 35-mer poly(U) RNA

Chain g: 



- Molecule 2: 35-mer poly(U) RNA

Chain m: 



4 Data and refinement statistics

EDS was not executed - this section will therefore be incomplete.

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α , β , γ	91.66Å 173.33Å 172.90Å 119.95° 99.34° 90.12°	Depositor
Resolution (Å)	57.42 – 3.40	Depositor
% Data completeness (in resolution range)	94.7 (57.42-3.40)	Depositor
R_{merge}	(Not available)	Depositor
R_{sym}	0.19	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.83 (at 3.40Å)	Xtriage
Refinement program	BUSTER 2.10.0	Depositor
R, R_{free}	0.225 , 0.240	Depositor
Wilson B-factor (Å ²)	58.7	Xtriage
Anisotropy	0.679	Xtriage
Estimated twinning fraction	0.316 for -h,k,-k-l	Xtriage
L-test for twinning	$\langle L \rangle = 0.36$, $\langle L^2 \rangle = 0.19$	Xtriage
Outliers	0 of 117932 reflections	Xtriage
Total number of atoms	36036	wwPDB-VP
Average B, all atoms (Å ²)	71.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 21.66 % of the origin peak, indicating pseudo translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo translational symmetry is equal to 6.6968e-03.*

¹Intensities estimated from amplitudes.

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	0.40	0/1918	0.53	0/2586
1	B	0.40	0/1918	0.53	0/2586
1	C	0.40	0/1918	0.53	0/2586
1	D	0.39	0/1918	0.54	0/2586
1	E	0.40	0/1918	0.53	0/2586
1	F	0.40	0/1918	0.53	0/2586
1	G	0.40	0/1918	0.54	0/2586
1	H	0.39	0/1918	0.53	0/2586
1	I	0.40	0/1918	0.53	0/2586
1	J	0.40	0/1918	0.56	0/2586
1	K	0.39	0/1918	0.53	0/2586
1	L	0.40	0/1918	0.53	0/2586
1	M	0.39	0/1918	0.53	0/2586
1	N	0.40	0/1918	0.53	0/2586
1	O	0.40	0/1918	0.53	0/2586
1	P	0.40	0/1918	0.53	0/2586
1	Q	0.40	0/1918	0.53	0/2586
1	R	0.40	0/1918	0.53	0/2586
2	a	1.10	0/786	0.92	0/1200
2	g	1.11	0/786	0.94	0/1200
2	m	1.12	0/786	0.92	0/1200
All	All	0.48	0/36882	0.57	0/50148

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

5.2 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 hydrogens

added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	1882	0	23	0	0
1	B	1882	0	23	0	0
1	C	1882	0	23	0	0
1	D	1882	0	23	0	0
1	E	1882	0	23	0	0
1	F	1882	0	23	0	0
1	G	1882	0	23	0	0
1	H	1882	0	23	0	0
1	I	1882	0	23	0	0
1	J	1882	0	23	1	0
1	K	1882	0	23	0	0
1	L	1882	0	23	0	0
1	M	1882	0	23	0	0
1	N	1882	0	23	0	0
1	O	1882	0	23	1	0
1	P	1882	0	23	0	0
1	Q	1882	0	23	0	0
1	R	1882	0	23	0	0
2	a	720	0	366	0	0
2	g	720	0	366	0	0
2	m	720	0	366	0	0
All	All	36036	0	1512	2	0

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 0.

All (2) close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:O:197:THR:O	1:O:201:ASN:ND2	2.50	0.45
1:J:34:ASP:O	1:J:36:ARG:N	2.54	0.41

There are no symmetry-related clashes.

5.3 Torsion angles

5.3.1 Protein backbone ⓘ

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution. 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	239/245 (98%)	227 (95%)	8 (3%)	4 (2%)	14	66
1	B	239/245 (98%)	227 (95%)	8 (3%)	4 (2%)	14	66
1	C	239/245 (98%)	227 (95%)	8 (3%)	4 (2%)	14	66
1	D	239/245 (98%)	225 (94%)	11 (5%)	3 (1%)	18	72
1	E	239/245 (98%)	225 (94%)	12 (5%)	2 (1%)	27	81
1	F	239/245 (98%)	227 (95%)	9 (4%)	3 (1%)	18	72
1	G	239/245 (98%)	227 (95%)	9 (4%)	3 (1%)	18	72
1	H	239/245 (98%)	227 (95%)	10 (4%)	2 (1%)	27	81
1	I	239/245 (98%)	229 (96%)	7 (3%)	3 (1%)	18	72
1	J	239/245 (98%)	224 (94%)	12 (5%)	3 (1%)	18	72
1	K	239/245 (98%)	229 (96%)	7 (3%)	3 (1%)	18	72
1	L	239/245 (98%)	224 (94%)	12 (5%)	3 (1%)	18	72
1	M	239/245 (98%)	229 (96%)	8 (3%)	2 (1%)	27	81
1	N	239/245 (98%)	224 (94%)	11 (5%)	4 (2%)	14	66
1	O	239/245 (98%)	227 (95%)	10 (4%)	2 (1%)	27	81
1	P	239/245 (98%)	225 (94%)	10 (4%)	4 (2%)	14	66
1	Q	239/245 (98%)	227 (95%)	10 (4%)	2 (1%)	27	81
1	R	239/245 (98%)	226 (95%)	9 (4%)	4 (2%)	14	66
All	All	4302/4410 (98%)	4076 (95%)	171 (4%)	55 (1%)	18	72

5 of 55 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	J	35	ALA
1	A	32	GLY
1	D	34	ASP
1	P	31	GLN
1	A	35	ALA

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 X-ray entries followed by that with respect to entries of similar resolution. 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	194/198 (98%)	191 (98%)	3 (2%)	76	95
1	B	194/198 (98%)	191 (98%)	3 (2%)	76	95
1	C	194/198 (98%)	192 (99%)	2 (1%)	85	96
1	D	194/198 (98%)	190 (98%)	4 (2%)	66	93
1	E	194/198 (98%)	190 (98%)	4 (2%)	66	93
1	F	194/198 (98%)	189 (97%)	5 (3%)	59	91
1	G	194/198 (98%)	191 (98%)	3 (2%)	76	95
1	H	194/198 (98%)	190 (98%)	4 (2%)	66	93
1	I	194/198 (98%)	191 (98%)	3 (2%)	76	95
1	J	194/198 (98%)	190 (98%)	4 (2%)	66	93
1	K	194/198 (98%)	189 (97%)	5 (3%)	59	91
1	L	194/198 (98%)	192 (99%)	2 (1%)	85	96
1	M	194/198 (98%)	191 (98%)	3 (2%)	76	95
1	N	194/198 (98%)	191 (98%)	3 (2%)	76	95
1	O	194/198 (98%)	192 (99%)	2 (1%)	85	96
1	P	194/198 (98%)	192 (99%)	2 (1%)	85	96
1	Q	194/198 (98%)	192 (99%)	2 (1%)	85	96
1	R	194/198 (98%)	191 (98%)	3 (2%)	76	95
All	All	3492/3564 (98%)	3435 (98%)	57 (2%)	75	95

5 of 57 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	H	26	ARG
1	J	7	LEU
1	Q	7	LEU
1	H	36	ARG
1	I	7	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no

such sidechains identified.

5.3.3 RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
2	a	30/36 (83%)	13 (43%)	0
2	g	30/36 (83%)	13 (43%)	0
2	m	30/36 (83%)	12 (40%)	0
All	All	90/108 (83%)	38 (42%)	0

5 of 38 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
2	a	4	U
2	a	5	U
2	a	10	U
2	a	11	U
2	a	14	U

There are no RNA pucker outliers to report.

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.

6 Fit of model and data ⓘ

6.1 Protein, DNA and RNA chains ⓘ

EDS was not executed - this section will therefore be empty.

6.2 Non-standard residues in protein, DNA, RNA chains ⓘ

EDS was not executed - this section will therefore be empty.

6.3 Carbohydrates ⓘ

EDS was not executed - this section will therefore be empty.

6.4 Ligands ⓘ

EDS was not executed - this section will therefore be empty.

6.5 Other polymers ⓘ

EDS was not executed - this section will therefore be empty.