



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 14, 2017 – 07:10 am GMT

PDB ID : 4LDD  
Title : Crystal Structure of Ebola virus VP40 Hexamer  
Authors : Bornholdt, Z.A.; Ableson, D.M.; Saphire, E.O.  
Deposited on : 2013-06-24  
Resolution : 3.50 Å(reported)

This is a wwPDB X-ray Structure Validation Summary Report for a publicly released PDB entry.

We welcome your comments at [validation@mail.wwpdb.org](mailto:validation@mail.wwpdb.org)

A user guide is available at

<http://wwpdb.org/validation/2016/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

---

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

MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.9-1692
EDS	:	trunk28620
Percentile statistics	:	20161228.v01 (using entries in the PDB archive December 28th 2016)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	recalc28949

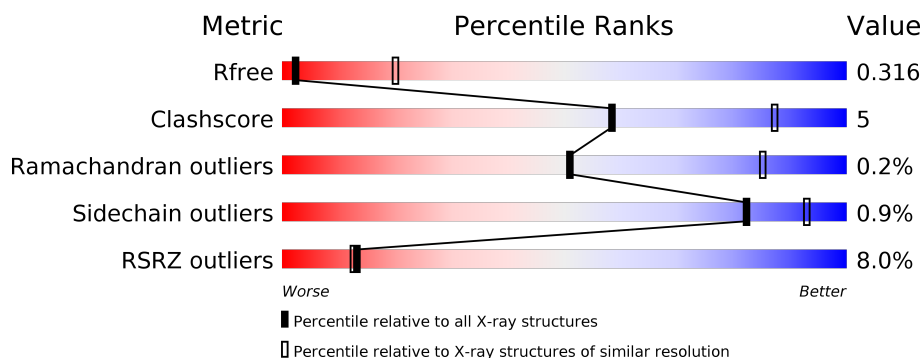
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 3.50 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))
$R_{free}$	100719	1195 (3.60-3.40)
Clashscore	112137	1322 (3.60-3.40)
Ramachandran outliers	110173	1283 (3.60-3.40)
Sidechain outliers	110143	1284 (3.60-3.40)
RSRZ outliers	101464	1226 (3.60-3.40)

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  $\geq 3$ , 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  $\leq 5\%$ . The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	297	<div> <div>3%</div> <div>42%</div> <div>54%</div> </div>
1	B	297	<div> <div>9%</div> <div>62%</div> <div>13%</div> <div>24%</div> </div>
1	C	297	<div> <div>2%</div> <div>42%</div> <div>5%</div> <div>53%</div> </div>

## 2 Entry composition

There is only 1 type of molecule in this entry. The entry contains 3883 atoms, of which 0 are hydrogens and 0 are deuteriums.

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 Matrix protein VP40.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	B	226	Total	C	N	O	S	0	0	0
			1755	1148	288	313	6			
1	C	139	Total	C	N	O	S	0	0	0
			1072	697	178	194	3			
1	A	137	Total	C	N	O	S	0	0	0
			1056	687	175	191	3			

There are 42 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	30	MET	-	EXPRESSION TAG	UNP Q05128
B	31	ALA	-	EXPRESSION TAG	UNP Q05128
B	32	HIS	-	EXPRESSION TAG	UNP Q05128
B	33	HIS	-	EXPRESSION TAG	UNP Q05128
B	34	HIS	-	EXPRESSION TAG	UNP Q05128
B	35	HIS	-	EXPRESSION TAG	UNP Q05128
B	36	HIS	-	EXPRESSION TAG	UNP Q05128
B	37	HIS	-	EXPRESSION TAG	UNP Q05128
B	38	VAL	-	EXPRESSION TAG	UNP Q05128
B	39	ASP	-	EXPRESSION TAG	UNP Q05128
B	40	ASP	-	EXPRESSION TAG	UNP Q05128
B	41	ASP	-	EXPRESSION TAG	UNP Q05128
B	42	ASP	-	EXPRESSION TAG	UNP Q05128
B	43	LYS	-	EXPRESSION TAG	UNP Q05128
C	30	MET	-	EXPRESSION TAG	UNP Q05128
C	31	ALA	-	EXPRESSION TAG	UNP Q05128
C	32	HIS	-	EXPRESSION TAG	UNP Q05128
C	33	HIS	-	EXPRESSION TAG	UNP Q05128
C	34	HIS	-	EXPRESSION TAG	UNP Q05128
C	35	HIS	-	EXPRESSION TAG	UNP Q05128
C	36	HIS	-	EXPRESSION TAG	UNP Q05128
C	37	HIS	-	EXPRESSION TAG	UNP Q05128
C	38	VAL	-	EXPRESSION TAG	UNP Q05128

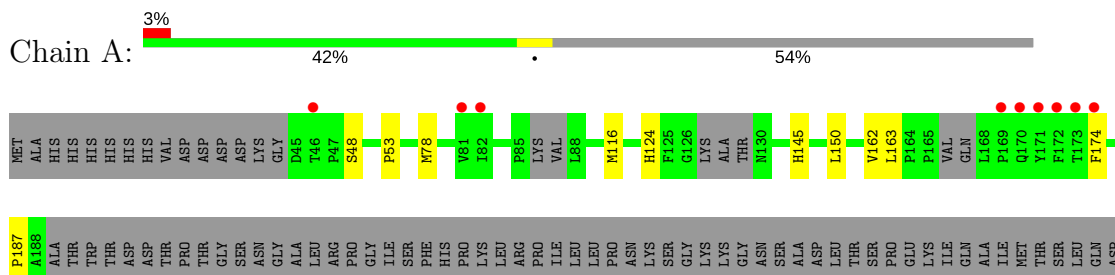
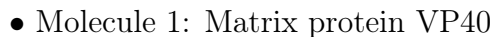
*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
C	39	ASP	-	EXPRESSION TAG	UNP Q05128
C	40	ASP	-	EXPRESSION TAG	UNP Q05128
C	41	ASP	-	EXPRESSION TAG	UNP Q05128
C	42	ASP	-	EXPRESSION TAG	UNP Q05128
C	43	LYS	-	EXPRESSION TAG	UNP Q05128
A	30	MET	-	EXPRESSION TAG	UNP Q05128
A	31	ALA	-	EXPRESSION TAG	UNP Q05128
A	32	HIS	-	EXPRESSION TAG	UNP Q05128
A	33	HIS	-	EXPRESSION TAG	UNP Q05128
A	34	HIS	-	EXPRESSION TAG	UNP Q05128
A	35	HIS	-	EXPRESSION TAG	UNP Q05128
A	36	HIS	-	EXPRESSION TAG	UNP Q05128
A	37	HIS	-	EXPRESSION TAG	UNP Q05128
A	38	VAL	-	EXPRESSION TAG	UNP Q05128
A	39	ASP	-	EXPRESSION TAG	UNP Q05128
A	40	ASP	-	EXPRESSION TAG	UNP Q05128
A	41	ASP	-	EXPRESSION TAG	UNP Q05128
A	42	ASP	-	EXPRESSION TAG	UNP Q05128
A	43	LYS	-	EXPRESSION TAG	UNP Q05128



- Molecule 1: Matrix protein VP40



PHE	ILE
LYS	THR
ILE	GLN
VAL	ASP
PRO	CYS
ILE	ASP
ASP	THR
PRO	CYS
THR	HIS
LYS	SER
ASN	PRO
ILE	ALA
MET	SER
GLY	LEU
ILE	PRO
GLU	ALA
VAL	VAL
PRO	ILE
GLU	GLU
THR	LYS
LEU	VAL
VAL	HIS
LYS	LYS
LEU	LYS
THR	THR
GLY	GLY
LYS	LYS
LYS	VAL
VAL	THR
THR	SER
LYS	LYS
ASN	GLY
GLY	GLN
PRO	PRO
ILE	ILE
ILE	ILE
PRO	PRO
VAL	VAL
LEU	LEU
LEU	PRO
PRO	LYS
LYS	TYR
TYR	ILE
ILE	GLY
GLY	LEU
LEU	ASP
ASP	PRO
PRO	VAL
VAL	ALA
ALA	PRO
PRO	GLY
GLY	ASP
ASP	LEU
LEU	THR
THR	MET
MET	VAL
VAL	

ILE
THR
GLN
ASP
CYS
ASP
THR
CYS
HIS
SER
PRO
ALA
SER
LEU
PRO
ALA
VAL
ILE
GLU
LYS

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 4 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	134.48Å 134.48Å 136.09Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	44.88 – 3.50 44.88 – 3.50	Depositor EDS
% Data completeness (in resolution range)	99.0 (44.88-3.50) 98.7 (44.88-3.50)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.26 (at 3.48Å)	Xtriage
Refinement program	PHENIX (phenix.refine: dev_1150)	Depositor
R, $R_{free}$	0.306 , 0.323 0.307 , 0.316	Depositor DCC
$R_{free}$ test set	823 reflections (5.09%)	DCC
Wilson B-factor (Å <sup>2</sup> )	126.8	Xtriage
Anisotropy	0.022	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.29 , 47.2	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.39$ , $\langle L^2 \rangle = 0.22$	Xtriage
Estimated twinning fraction	0.087 for -h,l,k 0.040 for -l,-k,-h	Xtriage
$F_o, F_c$ correlation	0.89	EDS
Total number of atoms	3883	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	68.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.00% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

## 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	0.22	0/1084	0.43	0/1484
1	B	0.22	0/1799	0.46	0/2461
1	C	0.22	0/1101	0.43	0/1508
All	All	0.22	0/3984	0.44	0/5453

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

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	1056	0	1069	8	0
1	B	1755	0	1826	26	0
1	C	1072	0	1088	11	0
All	All	3883	0	3983	43	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 43 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:265:GLU:HB3	1:B:269:HIS:HE1	1.31	0.95

*Continued on next page...*



*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:265:GLU:HB3	1:B:269:HIS:CE1	2.06	0.89
1:B:265:GLU:O	1:B:269:HIS:ND1	2.30	0.63
1:C:78:MET:HE1	1:A:187:PRO:HG2	1.80	0.62
1:B:242:THR:O	1:B:245:GLN:NE2	2.29	0.62

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 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	129/297 (43%)	127 (98%)	2 (2%)	0	100	100
1	B	214/297 (72%)	202 (94%)	11 (5%)	1 (0%)	32	73
1	C	133/297 (45%)	128 (96%)	5 (4%)	0	100	100
All	All	476/891 (53%)	457 (96%)	18 (4%)	1 (0%)	51	85

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	211	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 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	119/259 (46%)	119 (100%)	0	100	100
1	B	202/259 (78%)	198 (98%)	4 (2%)	60	84
1	C	121/259 (47%)	121 (100%)	0	100	100
All	All	442/777 (57%)	438 (99%)	4 (1%)	82	93

All (4) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	218	LEU
1	B	289	LEU
1	B	303	LEU
1	B	309	GLN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 12 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	309	GLN
1	C	64	HIS
1	A	64	HIS
1	B	257	ASN
1	C	170	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 [i](#)

There are no such residues in this entry.

## 5.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

## 6 Fit of model and data [i](#)

### 6.1 Protein, DNA and RNA chains [i](#)

In the following table, the column labelled ‘#RSRZ> 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95<sup>th</sup> percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled ‘Q< 0.9’ lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å <sup>2</sup> )	Q<0.9
1	A	137/297 (46%)	0.52	9 (6%) 19 16	19, 56, 106, 123	0
1	B	226/297 (76%)	0.70	26 (11%) 5 6	20, 77, 121, 140	0
1	C	139/297 (46%)	0.55	5 (3%) 43 37	18, 56, 111, 120	0
All	All	502/891 (56%)	0.61	40 (7%) 13 12	18, 62, 116, 140	0

The worst 5 of 40 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	172	PHE	5.2
1	B	88	LEU	4.9
1	A	171	TYR	4.8
1	A	173	THR	4.8
1	A	172	PHE	4.3

### 6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

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

### 6.3 Carbohydrates [i](#)

There are no carbohydrates in this entry.

### 6.4 Ligands [i](#)

There are no ligands in this entry.

## 6.5 Other polymers [i](#)

There are no such residues in this entry.