



Full wwPDB X-ray Structure Validation Report ⓘ

May 25, 2020 – 08:34 am BST

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 Full wwPDB X-ray Structure Validation Report for a publicly released PDB entry.

We welcome your comments at validation@mail.wwpdb.org

A user guide is available at

<https://www.wwpdb.org/validation/2017/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)	:	1.13
EDS	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

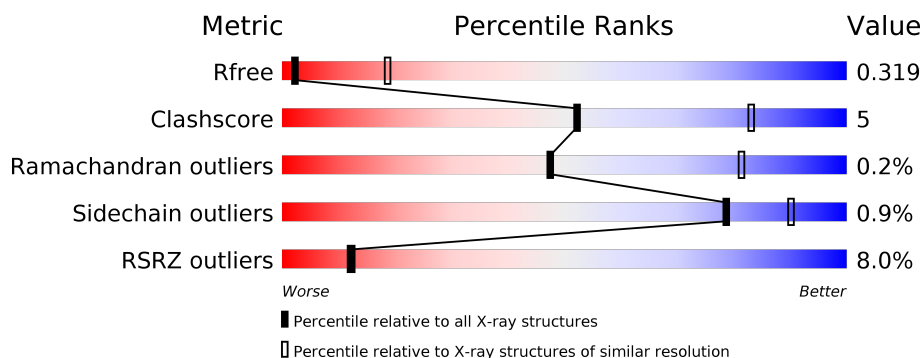
1 Overall quality at a glance

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}	130704	1659 (3.60-3.40)
Clashscore	141614	1036 (3.58-3.42)
Ramachandran outliers	138981	1005 (3.58-3.42)
Sidechain outliers	138945	1006 (3.58-3.42)
RSRZ outliers	127900	1559 (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 ≥ 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 $\leq 5\%$. 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

These plots are drawn for all protein, RNA and DNA 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 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 ($\text{RSRZ} > 2$). 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.

- Chain B:
-
- Sequence logo for Chain B. The y-axis represents information content in bits (0.00 to 0.15). The x-axis shows positions 1 to 300. The color scale at the top indicates conservation levels: 9% (red), 62% (green), 13% (yellow), and 24% (grey).
- Key residues (positions and amino acids) are highlighted in the sequence logo:
- Position 1: MET
 - Position 2: ALA
 - Position 3: HIS
 - Position 4: HIS
 - Position 5: HIS
 - Position 6: HIS
 - Position 7: HIS
 - Position 8: HIS
 - Position 9: VAL
 - Position 10: ASP
 - Position 11: ASP
 - Position 12: ASP
 - Position 13: ASP
 - Position 14: LYS
 - Position 15: GLY
 - Position 16: D45
 - Position 17: S48
 - Position 18: P53
 - Position 19: V81
 - Position 20: I82
 - Position 21: S83
 - Position 22: G84
 - Position 23: P85
 - Position 24: K86
 - Position 25: V87
 - Position 26: L88
 - Position 27: M89
 - Position 28: Q90
 - Position 29: I91
 - Position 30: I92
 - Position 31: P93
 - Position 32: I94
 - Position 33: M116
 - Position 34: L117
 - Position 35: A118
 - Position 36: H124
 - Position 37: F125
 - Position 38: GLY
 - Position 39: LYS
 - Position 40: ALA
 - Position 41: T129
 - Position 42: V133
 - Position 43: H145
 - Position 44: L150
 - Position 45: A156
 - Position 46: F161
 - Position 47: V162
 - Position 48: L163
 - Position 49: V174

- [illegible]

- Chain A:
-
- 3% 42% 54%
- MET ALA HIS HIS HIS HIS HIS VAL ASP ASP ASP LYS GLY D45 T46 P47 S48 P63 M78 T81 T82 P85 LYS VAL L88 M116 H124 F125 F126 LYS LYS ALA THR H130 H145 L150 V162 L163 P164 P165 VAL GLN L168 P169 Q170 Y171 F172 F173 T174 F174

PHE	LYS	ILE	VAL	PRO	CYS	ASP	THR	CYS	HIS	SER	ASN	MET	GLY	ILE	GLU	VAL	PRO	GLU	THR	LEU	VAL	HIS	LYS	LEU	THR	GLY	LYS	VAL	THR	SER	LYS	ASN	GLY	GLN	PRO	ILE	ILE	PRO	VAL	LEU	LEU	PRO	PRO	LYS	TYR	ILE	GLY	LEU	ASP	PRO	VAL	ALA	PRO	GLY	ASP	LEU	THR	MET	VAL
ILE	THR	GLN	ASP	CYS	ASP	THR	CYS	HIS	SER	PRO	ALA	ALA	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, α , β , γ	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
$\langle I/\sigma(I) \rangle$ ¹	2.26 (at 3.48Å)	Xtriage
Refinement program	PHENIX (phenix.refine: dev_1150)	Depositor
R, R_{free}	0.306 , 0.323 0.313 , 0.319	Depositor DCC
R_{free} test set	823 reflections (5.09%)	wwPDB-VP
Wilson B-factor (Å ²)	126.8	Xtriage
Anisotropy	0.022	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.29 , 47.2	EDS
L-test for twinning ²	$\langle L \rangle = 0.39$, $\langle L^2 \rangle = 0.22$	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 ²)	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.*

¹Intensities estimated from amplitudes.

²Theoretical 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.

All (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
1:B:253:ASP:OD1	1:B:255:THR:OG1	2.18	0.56
1:C:124:HIS:HB2	1:C:174:PHE:HE2	1.73	0.54
1:A:124:HIS:HB2	1:A:174:PHE:HE2	1.73	0.53
1:B:124:HIS:HB2	1:B:174:PHE:HE2	1.74	0.53
1:C:48:SER:HB3	1:C:174:PHE:HB2	1.92	0.52
1:A:48:SER:HB3	1:A:174:PHE:HB2	1.92	0.52
1:B:48:SER:HB3	1:B:174:PHE:HB2	1.92	0.51
1:A:145:HIS:CE1	1:A:150:LEU:HD12	2.45	0.51
1:C:145:HIS:CE1	1:C:150:LEU:HD12	2.46	0.50
1:B:145:HIS:CE1	1:B:150:LEU:HD12	2.47	0.49
1:B:156:ALA:O	1:B:214:ARG:NH2	2.47	0.48
1:B:210:HIS:CG	1:B:211:PRO:HD2	2.50	0.47
1:A:124:HIS:HB2	1:A:174:PHE:CE2	2.50	0.46
1:C:124:HIS:HB2	1:C:174:PHE:CE2	2.49	0.46
1:B:124:HIS:HB2	1:B:174:PHE:CE2	2.50	0.46
1:A:53:PRO:HG2	1:A:116:MET:HA	1.99	0.45
1:C:82:ILE:HG22	1:C:87:VAL:HG12	1.99	0.45
1:B:262:GLU:OE1	1:B:282:GLN:NE2	2.49	0.45
1:C:162:VAL:HG23	1:C:163:LEU:HG	1.99	0.45
1:B:162:VAL:HG23	1:B:163:LEU:HG	1.99	0.44
1:B:53:PRO:HG2	1:B:116:MET:HA	1.99	0.44
1:C:53:PRO:HG2	1:C:116:MET:HA	1.99	0.44
1:B:82:ILE:HG22	1:B:87:VAL:HG12	1.99	0.44
1:A:162:VAL:HG23	1:A:163:LEU:HG	1.99	0.44
1:B:184:GLN:HG2	1:B:258:ILE:HD12	1.99	0.43
1:B:216:ILE:HD12	1:B:216:ILE:H	1.84	0.43
1:B:282:GLN:HA	1:B:283:PRO:HD2	1.81	0.42
1:B:257:ASN:HD22	1:B:289:LEU:HB3	1.85	0.42
1:B:251:PRO:HA	1:B:259:MET:HA	2.02	0.42
1:B:207:ILE:O	1:B:216:ILE:HB	2.19	0.42
1:B:208:SER:HB2	1:B:215:PRO:HA	2.02	0.41
1:B:287:VAL:HG13	1:B:289:LEU:HB2	2.03	0.41
1:B:89:MET:HB3	1:B:92:ILE:HD11	2.03	0.41
1:B:86:LYS:HB2	1:B:86:LYS:HE3	1.85	0.41
1:B:184:GLN:HA	1:B:185:PRO:HD3	1.81	0.41
1:C:187:PRO:HG2	1:A:78:MET:HE1	2.03	0.40
1:C:186:LEU:HA	1:C:187:PRO:HD2	1.93	0.40

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:86:LYS:HE3	1:C:86:LYS:HB2	1.84	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 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%)	29	68
1	C	133/297 (45%)	128 (96%)	5 (4%)	0	100	100
All	All	476/891 (53%)	457 (96%)	18 (4%)	1 (0%)	47	81

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%)	55	79

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	C	121/259 (47%)	121 (100%)	0	100	100
All	All	442/777 (57%)	438 (99%)	4 (1%)	78	90

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. All (12) such sidechains are listed below:

Mol	Chain	Res	Type
1	B	64	HIS
1	B	91	GLN
1	B	170	GLN
1	B	238	GLN
1	B	257	ASN
1	B	309	GLN
1	C	64	HIS
1	C	91	GLN
1	C	170	GLN
1	A	64	HIS
1	A	91	GLN
1	A	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 [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.

6 Fit of model and data ⓘ

6.1 Protein, DNA and RNA chains ⓘ

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, 95th 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 ²)	Q<0.9
1	A	137/297 (46%)	0.52	9 (6%) 18 17	19, 56, 106, 123	0
1	B	226/297 (76%)	0.70	26 (11%) 4 5	20, 77, 121, 140	0
1	C	139/297 (46%)	0.55	5 (3%) 42 38	18, 56, 111, 120	0
All	All	502/891 (56%)	0.61	40 (7%) 12 12	18, 62, 116, 140	0

All (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	4.7
1	A	172	PHE	4.3
1	B	283	PRO	4.3
1	C	83	SER	4.1
1	B	87	VAL	3.7
1	B	133	VAL	3.6
1	B	82	ILE	3.5
1	B	294	GLY	3.5
1	C	170	GLN	3.4
1	B	161	PHE	3.4
1	B	171	TYR	3.3
1	B	252	ILE	3.3
1	B	174	PHE	3.2
1	B	81	VAL	3.1
1	A	46	THR	3.0
1	A	169	PRO	2.9
1	B	83	SER	2.9
1	B	124	HIS	2.9
1	B	162	VAL	2.8
1	C	132	LEU	2.8

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	118	ALA	2.8
1	C	160	GLU	2.7
1	B	267	LEU	2.7
1	B	247	PHE	2.7
1	B	85	PRO	2.6
1	B	90	LYS	2.6
1	B	89	MET	2.6
1	A	81	VAL	2.5
1	C	159	GLN	2.3
1	B	94	ILE	2.3
1	B	269	HIS	2.3
1	B	263	VAL	2.2
1	B	86	LYS	2.2
1	A	82	ILE	2.1
1	A	174	PHE	2.1
1	A	170	GLN	2.1
1	B	84	GLY	2.1

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.