wwPDB X-ray Structure Validation Summary Report (i)

Sep 5, 2023 – 10:13 AM EDT

PDB ID	:	3VBU
Title	:	Crystal structure of empty human Enterovirus 71 particle
Authors	:	Wang, X.; Peng, W.; Ren, J.; Hu, Z.; Xu, J.; Lou, Z.; Li, X.; Yin, W.; Shen, X.; Porta, C.; Walter, T.S.; Evans, G.; Axford, D.; Owen, R.; Rowlands, D.J.; Wang, J.; Stuart, D.I.; Fry, E.E.; Rao, Z.
Deposited on	:	2012-01-02
Resolution	:	4.00 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*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) 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 (1)) were used in the production of this report:

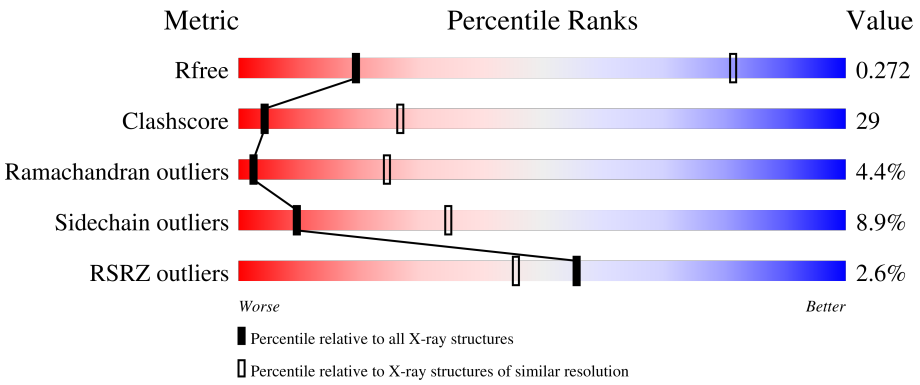
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.35
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 4.00 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)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	1087 (4.30-3.70)
Clashscore	141614	1148 (4.30-3.70)
Ramachandran outliers	138981	1108 (4.30-3.70)
Sidechain outliers	138945	1099 (4.30-3.70)
RSRZ outliers	127900	1028 (4.34-3.66)

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 of 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 <=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	225	44% 44% 9% .
2	B	237	4% 48% 46% 5%
3	C	239	3% 55% 37% 8%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 5397 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 Genome Polyprotein, capsid protein VP1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	218	1724	1107	290	316	11	0	0	0

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	225	MET	CYS	SEE REMARK 999	UNP B2ZUN0

- Molecule 2 is a protein called Genome Polyprotein, capsid protein VP0.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	237	1834	1179	301	346	8	0	0	0

- Molecule 3 is a protein called Genome Polyprotein, capsid protein VP3.

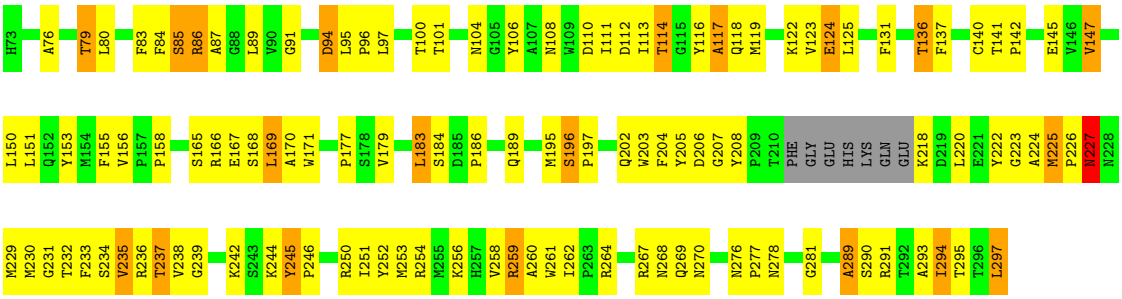
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	C	239	1839	1182	306	340	11	0	0	0

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 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.

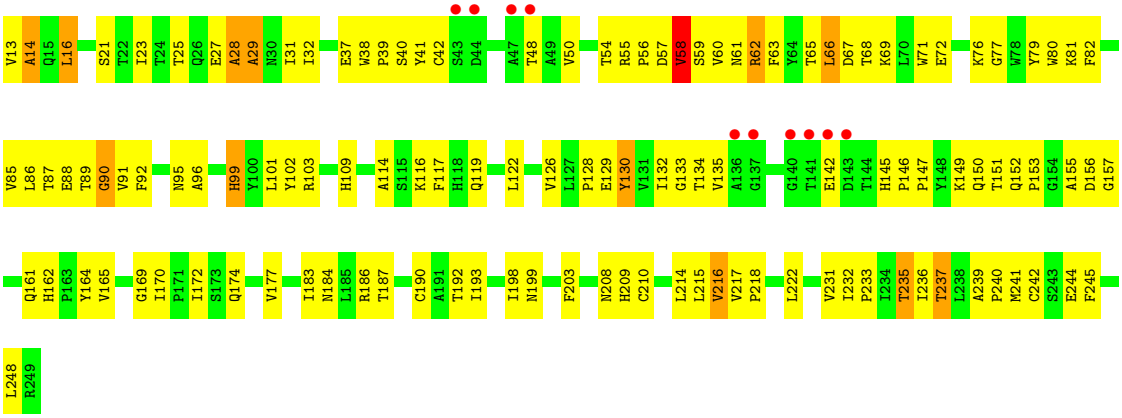
• Molecule 1: Genome Polyprotein, capsid protein VP1

Chain A: 44% 44% 9%



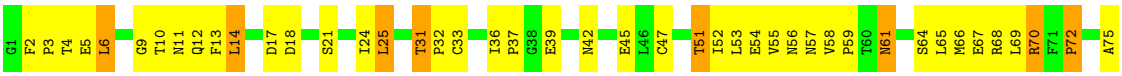
• Molecule 2: Genome Polyprotein, capsid protein VP0

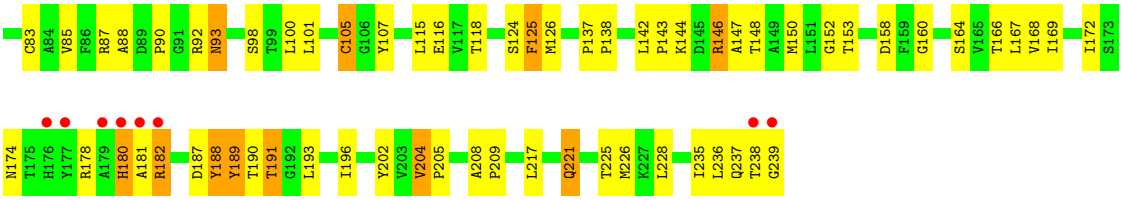
Chain B: 4% 48% 46% 5%



• Molecule 3: Genome Polyprotein, capsid protein VP3

Chain C: 3% 55% 37% 8%




C83
A84
V85
F86
R87
A88
D89
P90
G91
R92
N93
S98
T99
L100
L101
C105
G106
Y107
L115
E116
V117
T118
S124
F125
M126
P137
P138
L142
P143
K144
D145
R146
A147
T148
A149
M150
L151
G152
T153
D158
F159
G160
S164
V165
T166
L167
V168
I169
I172
S173
N174
T175
H176
Y177
R178
A179
H180
A181
R182
D187
Y188
Y189
T190
T191
G192
L193
I196
Y202
V203
V204
P205
A208
P209
L217
Q221
T225
M226
K227
L228
I235
L236
Q237
T238
G239

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 42 3 2	Depositor
Cell constants a, b, c, α, β, γ	355.90Å 355.90Å 355.90Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	49.84 – 4.00 49.84 – 4.00	Depositor EDS
% Data completeness (in resolution range)	98.8 (49.84-4.00) 98.8 (49.84-4.00)	Depositor EDS
R_{merge}	0.39	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.18 (at 4.00Å)	Xtriage
Refinement program	CNS	Depositor
R, R_{free}	0.272 , 0.278 0.284 , 0.272	Depositor DCC
R_{free} test set	657 reflections (1.02%)	wwPDB-VP
Wilson B-factor (Å^2)	92.2	Xtriage
Anisotropy	0.000	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.27 , 96.7	EDS
L-test for twinning[2]	$< \|L\| > = 0.38$, $< L^2 > = 0.20$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.79	EDS
Total number of atoms	5397	wwPDB-VP
Average B, all atoms (Å^2)	111.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 1.38% of the height of the origin peak. No significant pseudotranslation is detected.*

[1]Intensities estimated from amplitudes.

[2]Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

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	0.53	0/1776	0.78	0/2419
2	B	0.56	0/1889	0.75	1/2592 (0.0%)
3	C	0.50	0/1891	0.73	0/2586
All	All	0.53	0/5556	0.75	1/7597 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	248	LEU	CA-CB-CG	5.64	128.28	115.30

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	1724	0	1688	118	0
2	B	1834	0	1774	123	0
3	C	1839	0	1814	123	0
All	All	5397	0	5276	312	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 29.

The worst 5 of 312 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:281:GLY:HA3	2:B:135:VAL:HG13	1.14	1.09
1:A:229:MET:HE2	1:A:231:GLY:H	1.16	1.07
1:A:156:VAL:HG22	1:A:232:THR:HB	1.42	0.99
2:B:134:THR:HG23	2:B:146:PRO:HG3	1.47	0.96
2:B:164:TYR:HA	3:C:66:MET:HE3	1.49	0.94

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	214/225 (95%)	188 (88%)	15 (7%)	11 (5%)	2 22
2	B	235/237 (99%)	196 (83%)	31 (13%)	8 (3%)	3 30
3	C	237/239 (99%)	202 (85%)	24 (10%)	11 (5%)	2 24
All	All	686/701 (98%)	586 (85%)	70 (10%)	30 (4%)	2 24

5 of 30 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	85	SER
2	B	28	ALA
2	B	29	ALA
2	B	58	VAL
3	C	93	ASN

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	187/193 (97%)	170 (91%)	17 (9%)	9 33
2	B	201/201 (100%)	186 (92%)	15 (8%)	13 41
3	C	199/199 (100%)	179 (90%)	20 (10%)	7 29
All	All	587/593 (99%)	535 (91%)	52 (9%)	9 34

5 of 52 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	142	GLU
3	C	25	LEU
3	C	221	GLN
2	B	216	VAL
3	C	6	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 17 such sidechains are listed below:

Mol	Chain	Res	Type
3	C	154	HIS
3	C	180	HIS
2	B	145	HIS
2	B	150	GLN
2	B	162	HIS

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 monosaccharides 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 (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, 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	218/225 (96%)	-0.46	0 100 100	79, 97, 150, 207	0
2	B	237/237 (100%)	-0.09	10 (4%) 36 29	81, 111, 194, 237	0
3	C	239/239 (100%)	-0.44	8 (3%) 46 37	80, 99, 180, 232	0
All	All	694/701 (99%)	-0.33	18 (2%) 56 46	79, 100, 180, 237	0

The worst 5 of 18 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	140	GLY	4.9
3	C	182	ARG	3.9
2	B	143	ASP	3.1
2	B	136	ALA	2.9
2	B	142	GLU	2.9

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 monosaccharides 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.