wwPDB X-ray Structure Validation Summary Report (i)

May 25, 2020 – 03:36 pm BST

PDB ID	:	4GMP
Title	:	Crystal structure of enterovirus 71 strain 1095 procapsid
Authors	:	Yoder, J.D.; Hafenstein, S.
Deposited on	:	2012-08-16
Resolution	:	3.90 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 following versions of software and data (see references (1)) were used in the production of this report:

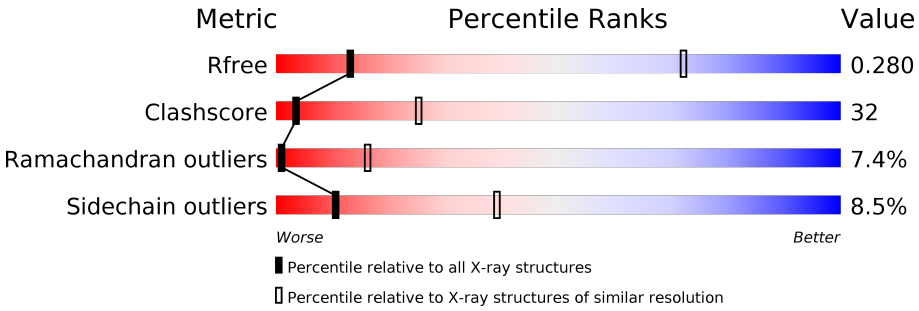
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

1 Overall quality at a glance (i)

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

The reported resolution of this entry is 3.90 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	1002 (4.14-3.66)
Clashscore	141614	1004 (4.12-3.68)
Ramachandran outliers	138981	1021 (4.14-3.66)
Sidechain outliers	138945	1014 (4.14-3.66)

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 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 <=5%

Mol	Chain	Length	Quality of chain
1	0	323	31% 37% 5% 27%
2	1	297	33% 34% 6% • 27%
3	3	242	44% 44% 10% ••

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 5395 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 capsid protein VP0.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	0	237	1833	1179	301	345	8	0	0	0

- Molecule 2 is a protein called capsid protein VP1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	1	218	1717	1101	289	316	11	0	0	0

- Molecule 3 is a protein called capsid protein VP3.

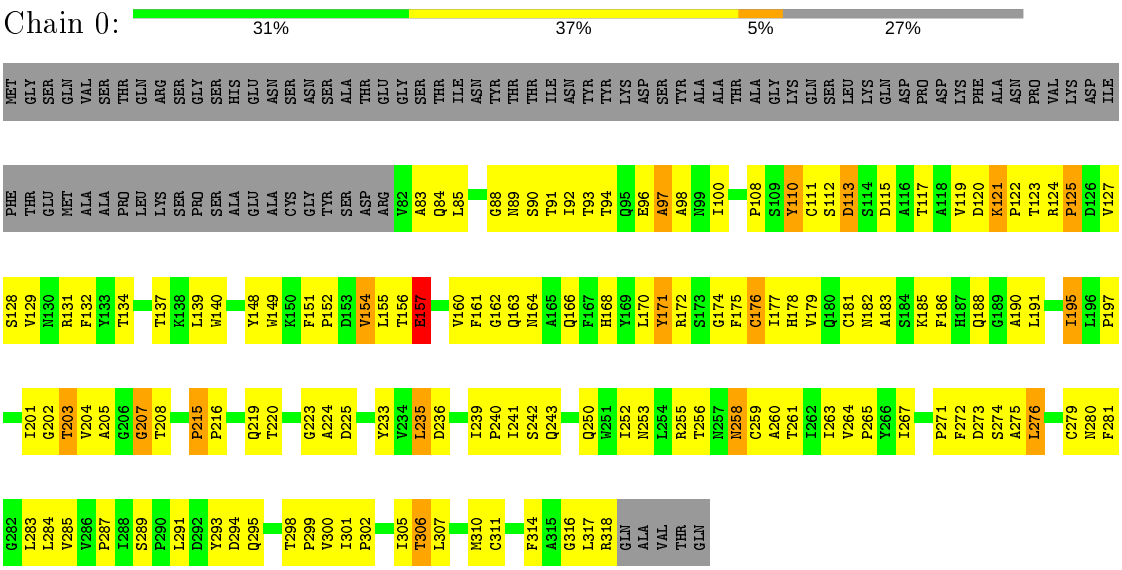
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	3	240	1845	1188	304	342	11	0	0	0

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 the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry. 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. 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.

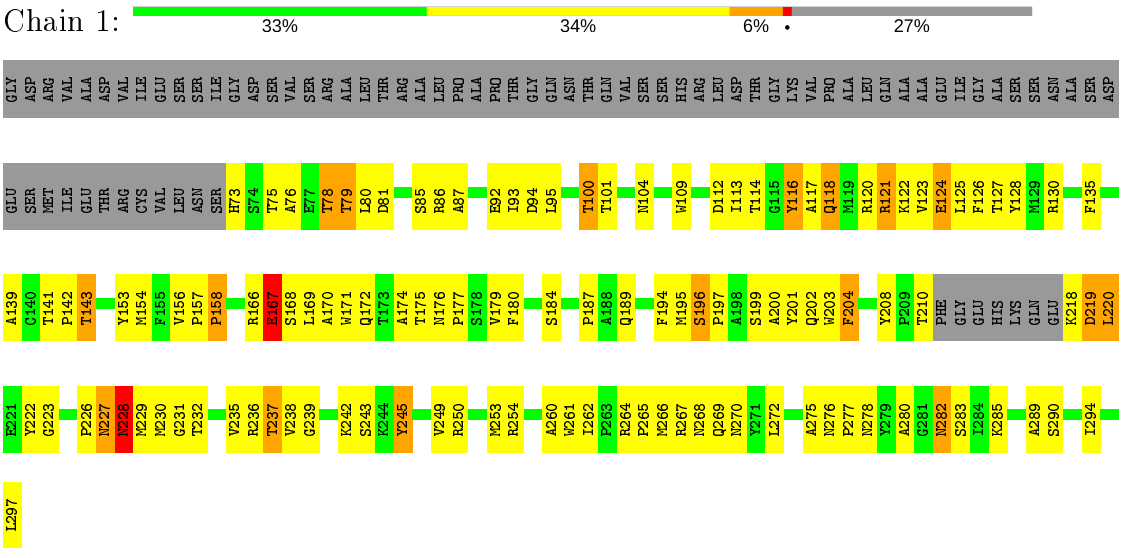
- Molecule 1: capsid protein VP0

Chain 0:



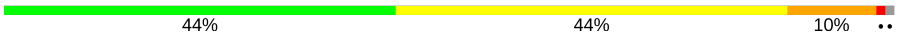
- Molecule 2: capsid protein VP1

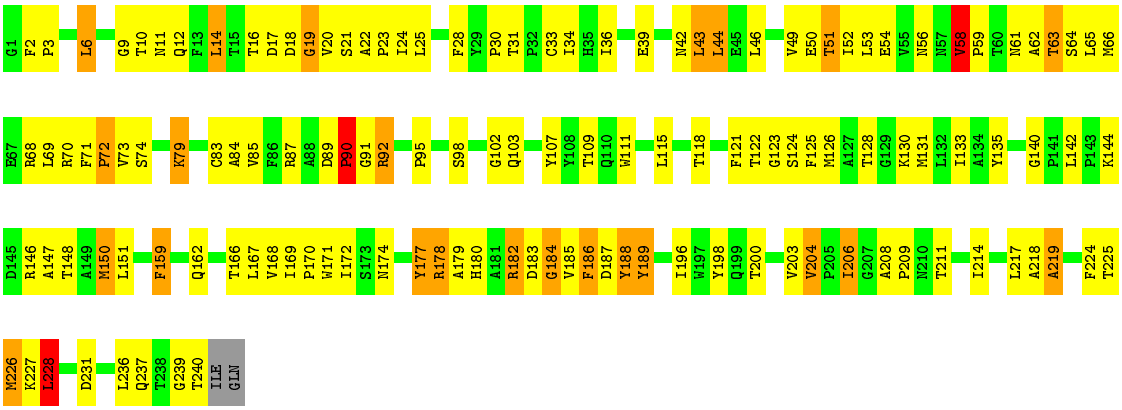
Chain 1:



- Molecule 3: capsid protein VP3

Chain 3:


44%
44%
10%


G1 F2 P3 L6 G9 T10 N11 Q12 F13 L14 T15 T16 D17 D18 G19 V20 S21 A22 P23 I24 L25 F28 Y29 P30 T31 P32 C33 I34 H35 I36 E39 N42 L43 L44 E45 L46 V49 E50 T51 I52 L53 E54 V55 N56 N57 Y58 P59 T60 N61 A62 T63 S64 L65 M66
E67 R68 L69 R70 F71 P72 V73 S74 K79 C83 A84 Y85 F86 R87 A88 D89 P90 G91 R92 P95 S98 G102 Q103 Y107 Y108 T109 Q110 W111 L115 T118 F121 T122 G123 S124 F125 M126 A127 T128 G129 K130 M131 L132 I133 A134 Y135 G140 P141 L142 P143 K144
D145 R146 A147 T148 A149 M150 L151 F159 Q162 T166 L167 V168 I169 P170 W171 I172 S173 N174 Y177 R178 A179 H180 A181 R182 D183 G184 V185 F186 D187 Y188 Y189 I196 W197 Y198 Q199 T200 V203 V204 P205 I206 G207 A208 P209 N210 T211 I214 L217 A218 A219 F224 T225
M226 K227 L228 D231 L236 Q237 T238 G239 T240 ILE GLN

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 42 3 2	Depositor
Cell constants a, b, c, α, β, γ	350.25Å 350.25Å 350.25Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	49.53 – 3.90 49.53 – 3.90	Depositor EDS
% Data completeness (in resolution range)	99.8 (49.53-3.90) 99.9 (49.53-3.90)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.81 (at 3.88Å)	Xtriage
Refinement program	CNS 1.3	Depositor
R, R_{free}	0.269 , 0.285 0.264 , 0.280	Depositor DCC
R_{free} test set	3398 reflections (5.08%)	wwPDB-VP
Wilson B-factor (Å^2)	109.8	Xtriage
Anisotropy	0.000	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.32 , 86.2	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.71	EDS
Total number of atoms	5395	wwPDB-VP
Average B, all atoms (Å^2)	125.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 1.44% 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

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	0	0.44	0/1888	0.67	0/2591
2	1	0.46	0/1769	0.68	0/2411
3	3	0.49	0/1897	0.71	0/2596
All	All	0.47	0/5554	0.69	0/7598

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	0	1833	0	1774	124	3
2	1	1717	0	1674	125	8
3	3	1845	0	1822	134	1
All	All	5395	0	5270	341	9

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 32.

The worst 5 of 341 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)
3:3:20:VAL:HG22	3:3:21:SER:H	1.17	1.06

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:1:229:MET:HE2	2:1:231:GLY:H	1.24	1.02
3:3:206:ILE:HD12	3:3:206:ILE:H	1.23	1.01
3:3:109:THR:HB	3:3:228:LEU:HB3	1.52	0.91
3:3:58:VAL:HG23	3:3:59:PRO:HD3	1.50	0.90

The worst 5 of 9 symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:0:208:THR:CG2	2:1:283:SER:OG[13_454]	1.16	1.04
1:0:208:THR:CG2	2:1:283:SER:CB[13_454]	1.57	0.63
2:1:282:ASN:ND2	2:1:283:SER:N[13_454]	1.68	0.52
3:3:63:THR:CG2	3:3:63:THR:CG2[13_454]	1.71	0.49
2:1:282:ASN:O	2:1:282:ASN:CB[13_454]	1.75	0.45

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	0	235/323 (73%)	175 (74%)	42 (18%)	18 (8%)	1 16
2	1	214/297 (72%)	171 (80%)	27 (13%)	16 (8%)	1 16
3	3	238/242 (98%)	187 (79%)	34 (14%)	17 (7%)	1 17
All	All	687/862 (80%)	533 (78%)	103 (15%)	51 (7%)	1 16

5 of 51 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	0	97	ALA
1	0	112	SER
1	0	127	VAL
2	1	169	LEU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	1	196	SER

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	0	201/272 (74%)	188 (94%)	13 (6%)	17 46
2	1	186/250 (74%)	168 (90%)	18 (10%)	8 32
3	3	200/202 (99%)	181 (90%)	19 (10%)	8 33
All	All	587/724 (81%)	537 (92%)	50 (8%)	10 39

5 of 50 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	1	179	VAL
2	1	228	ASN
3	3	204	VAL
2	1	184	SER
2	1	204	PHE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 11 such sidechains are listed below:

Mol	Chain	Res	Type
2	1	152	GLN
2	1	176	ASN
3	3	48	GLN
1	0	214	HIS
2	1	278	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

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

Unable to reproduce the depositors R factor - this section is therefore empty.

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

Unable to reproduce the depositors R factor - this section is therefore empty.

6.3 Carbohydrates

Unable to reproduce the depositors R factor - this section is therefore empty.

6.4 Ligands

Unable to reproduce the depositors R factor - this section is therefore empty.

6.5 Other polymers

Unable to reproduce the depositors R factor - this section is therefore empty.