WORLDWIDE PDB

PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report ⓘ

Jun 17, 2024 – 02:46 AM EDT

|		
PDB ID	:	5MJV
Title	:	Rebuild and re-refined model for Human Parechovirus 1
Authors	:	Shakeel, S.; Dykeman, E.C.; White, S.J.; Ora, A.; Cockburn, J.J.B.; Butcher, S.J.; Stockley, P.G.; Twarock, R.
Deposited on	:	2016-12-01
Resolution	:	3.09 Å(reported)

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 3.09 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.
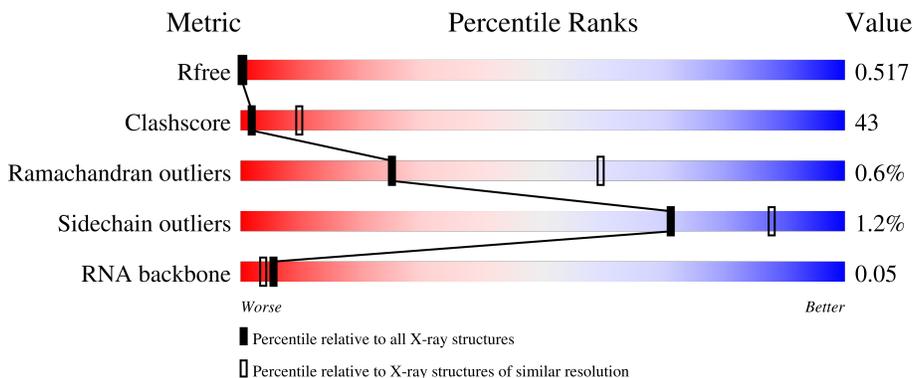
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	1094 (3.10-3.10)
Clashscore	141614	1184 (3.10-3.10)
Ramachandran outliers	138981	1141 (3.10-3.10)
Sidechain outliers	138945	1141 (3.10-3.10)
RNA backbone	3102	1116 (3.40-2.80)

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%

Mol	Chain	Length	Quality of chain
1	A	234	34% · 46% · 18%
2	B	253	47% · 46% · 6%
3	C	289	47% · 41% · 11%
4	D	6	50% · 17% · 33%

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 5486 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 SUBUNIT VP1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	192	Total 1507	C 971	N 250	O 279	S 7	0	0	0

- Molecule 2 is a protein called CAPSID SUBUNIT VP3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	239	Total 1864	C 1185	N 320	O 349	S 10	0	0	0

- Molecule 3 is a protein called CAPSID SUBUNIT VP0.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	257	Total 1995	C 1265	N 332	O 392	S 6	0	0	0

- Molecule 4 is a RNA chain called RNA (5'-R(*GP*UP*UP*UP*UP*U)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	D	6	Total 120	C 55	N 15	O 45	P 5	0	0	0

3 Residue-property plots ⓘ

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. 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 SUBUNIT VP1

Chain A:
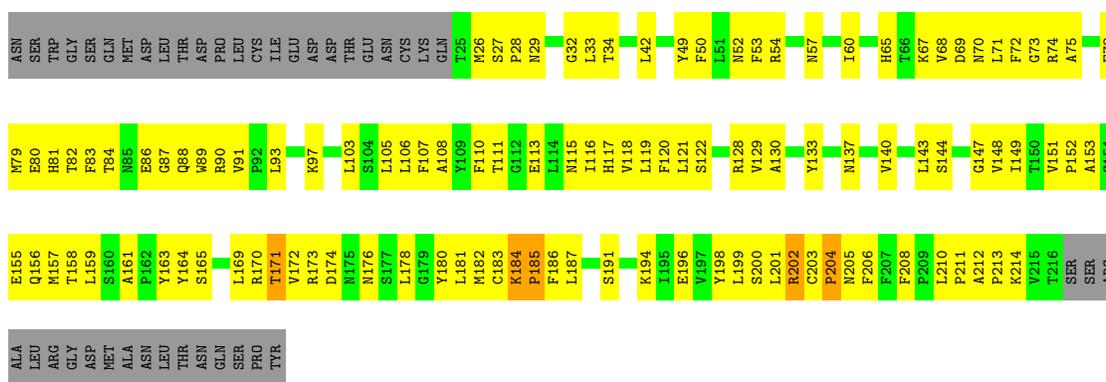
• Molecule 2: CAPSID SUBUNIT VP3

Chain B:
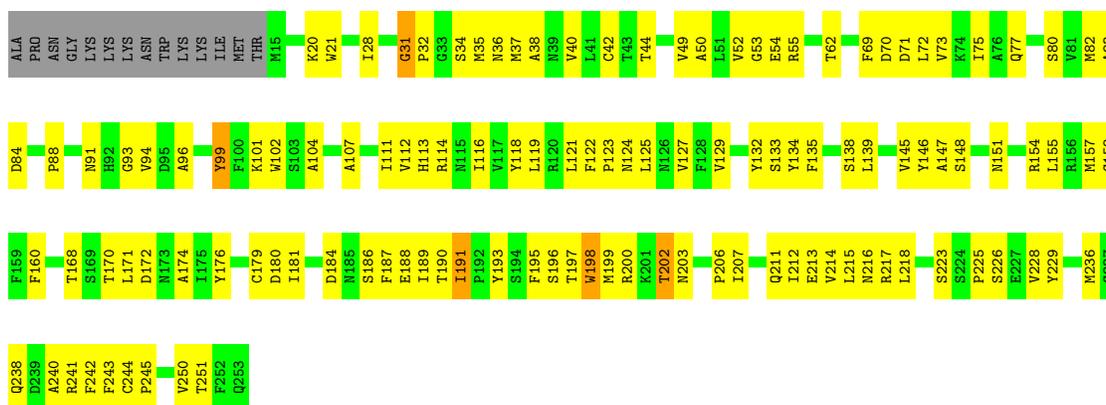
• Molecule 3: CAPSID SUBUNIT VP0

Chain C:

P93 H102 S103 M104 I108 T109 N114 P115 L118 L121 E122 T123 H124 E125 W126 T127 W130 A131 R132 G133 Y134 Q135 I136 T137 H138 V139 E140 L141 P142 K143 V144 F145 W146 D147 H148 Q149 Y154 R158 Y159 F160 A161 A162 V163 F167 Q170 V175 T179 A180

G181 S182 A183 V186 Y187 E188 P189 K190 P191 L199 E200 F201 G202 A203 F204 T205 N206 L207 P208 H209 V210 L211 M212 N213 L214 A215 E216 T217 A220 D221 L222 Y226 V227 A228 D229 V233 K234 T235 D236 S237 S238 D239 L240 G241 Q242 L243 K244 V247 W248 T249 P250 L251 S252 I253

P254 T255 G256 N259 Q260 V261 D262 V263 L266 G267 L272 D273 F274 Q275 N276 P277 R278 V279 F280 A281 Q282 D283 V284 N285 I286 Y287 D288 ASN

• Molecule 4: RNA (5'-R(*GP*UP*UP*UP*UP*U)-3')

Chain D:

50%   17%   33%

G1 U2 U3 U4 U5 U6

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 63 2 2	Depositor
Cell constants a, b, c, α, β, γ	399.50Å 399.50Å 332.86Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	345.98 – 3.09 63.93 – 3.09	Depositor EDS
% Data completeness (in resolution range)	82.7 (345.98-3.09) 82.8 (63.93-3.09)	Depositor EDS
R_{merge}	0.35	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.32 (at 3.07Å)	Xtriage
Refinement program	CNS 1.3, REFMAC 5.8.0135	Depositor
R, R_{free}	0.258 , 0.261 0.517 , 0.517	Depositor DCC
R_{free} test set	11563 reflections (4.96%)	wwPDB-VP
Wilson B-factor (Å^2)	46.8	Xtriage
Anisotropy	0.012	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.35 , 49.6	EDS
L-test for twinning[2]	$< |L| > = 0.45, < L^2 > = 0.28$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.20	EDS
Total number of atoms	5486	wwPDB-VP
Average B, all atoms (Å^2)	42.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 1.47% 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	A	0.45	0/1551	0.72	5/2113 (0.2%)
2	B	0.53	2/1909 (0.1%)	0.64	2/2592 (0.1%)
3	C	0.37	0/2048	0.61	0/2810
4	D	1.85	3/132 (2.3%)	1.28	2/203 (1.0%)
All	All	0.53	5/5640 (0.1%)	0.68	9/7718 (0.1%)

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	D	5	U	O3'-P	-9.23	1.50	1.61
4	D	2	U	O3'-P	-8.16	1.51	1.61
4	D	3	U	O3'-P	-7.73	1.51	1.61
2	B	99	TYR	CE1-CZ	-6.88	1.29	1.38
2	B	99	TYR	CE2-CZ	-5.88	1.30	1.38

The worst 5 of 9 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	D	5	U	C2'-C3'-O3'	-7.70	92.57	109.50
1	A	203	CYS	N-CA-C	-6.69	92.94	111.00
1	A	173	ARG	NE-CZ-NH1	-6.28	117.16	120.30
1	A	202	ARG	N-CA-C	5.89	126.91	111.00
1	A	184	LYS	C-N-CD	5.53	140.01	128.40

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	1507	0	1444	217	0
2	B	1864	0	1814	162	0
3	C	1995	0	1901	135	0
4	D	120	0	63	10	0
All	All	5486	0	5222	457	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 43.

The worst 5 of 457 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:86:GLU:HG2	1:A:185:PRO:O	1.29	1.26
1:A:158:THR:OG1	2:B:35:MET:CE	1.85	1.24
1:A:78:PHE:HB2	1:A:199:LEU:CD1	1.72	1.19
1:A:158:THR:OG1	2:B:35:MET:HE2	1.41	1.17
2:B:28:ILE:HD11	2:B:32:PRO:CB	1.78	1.13

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	190/234 (81%)	176 (93%)	12 (6%)	2 (1%)	14	46
2	B	237/253 (94%)	226 (95%)	11 (5%)	0	100	100
3	C	255/289 (88%)	231 (91%)	22 (9%)	2 (1%)	19	54
All	All	682/776 (88%)	633 (93%)	45 (7%)	4 (1%)	25	59

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	C	37	GLY
1	A	185	PRO
3	C	210	VAL
1	A	204	PRO

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	166/209 (79%)	164 (99%)	2 (1%)	71	88
2	B	204/220 (93%)	200 (98%)	4 (2%)	55	80
3	C	222/252 (88%)	221 (100%)	1 (0%)	88	94
All	All	592/681 (87%)	585 (99%)	7 (1%)	71	88

5 of 7 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	196	SER
2	B	198	TRP
3	C	35	GLU
2	B	202	THR
2	B	195	PHE

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	213	ASN
3	C	276	ASN
2	B	216	ASN
3	C	39	ASN
3	C	82	GLN

5.3.3 RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
4	D	6/6 (100%)	4 (66%)	2 (33%)

All (4) RNA backbone outliers are listed below:

Mol	Chain	Res	Type
4	D	3	U
4	D	4	U
4	D	5	U
4	D	6	U

All (2) RNA pucker outliers are listed below:

Mol	Chain	Res	Type
4	D	1	G
4	D	5	U

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 monosaccharides 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.