wwPDB X-ray Structure Validation Summary Report (i)

Feb 1, 2016 – 05:46 PM GMT

PDB ID	:	4JGZ
Title	:	Crystal structure of human coxsackievirus A16 uncoating intermediate (space group I222)
Authors	:	Ren, J.; Wang, X.; Hu, Z.; Gao, Q.; Sun, Y.; Li, X.; Porta, C.; Walter, T.S.; Gilbert, R.J.; Zhao, Y.; Axford, D.; Williams, M.; McAuley, K.; Rowlands, D.J.; Yin, W.; Wang, J.; Stuart, D.I.; Rao, Z.; Fry, E.E.
Deposited on	:	2013-03-04
Resolution	:	3.00 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
http://wwpdb.org/validation/2016/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:

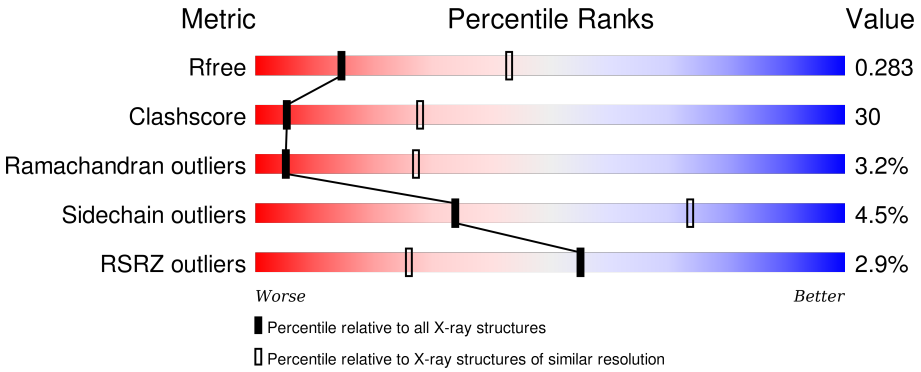
MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20026688
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	trunk26865

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 3.00 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}	91344	1578 (3.00-3.00)
Clashscore	102246	1912 (3.00-3.00)
Ramachandran outliers	100387	1853 (3.00-3.00)
Sidechain outliers	100360	1856 (3.00-3.00)
RSRZ outliers	91569	1592 (3.00-3.00)

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. 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	297	% 35% 39% • 23%
2	B	254	5% 46% 45% • 6%
3	C	242	% 48% 44% • 5%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 5468 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 Polyprotein, capsid protein VP1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	228	1830	1170	312	335	13	0	0	0

- Molecule 2 is a protein called Polyprotein, capsid protein VP2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	239	1844	1186	305	343	10	0	0	0

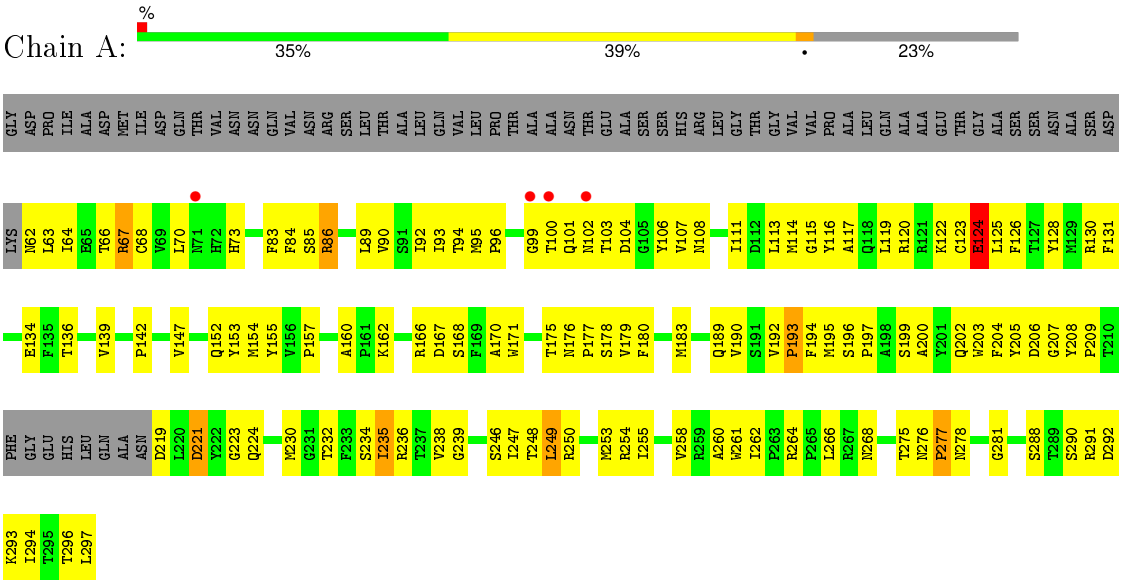
- Molecule 3 is a protein called Polyprotein, capsid protein VP3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	C	231	1794	1156	291	336	11	0	0	0

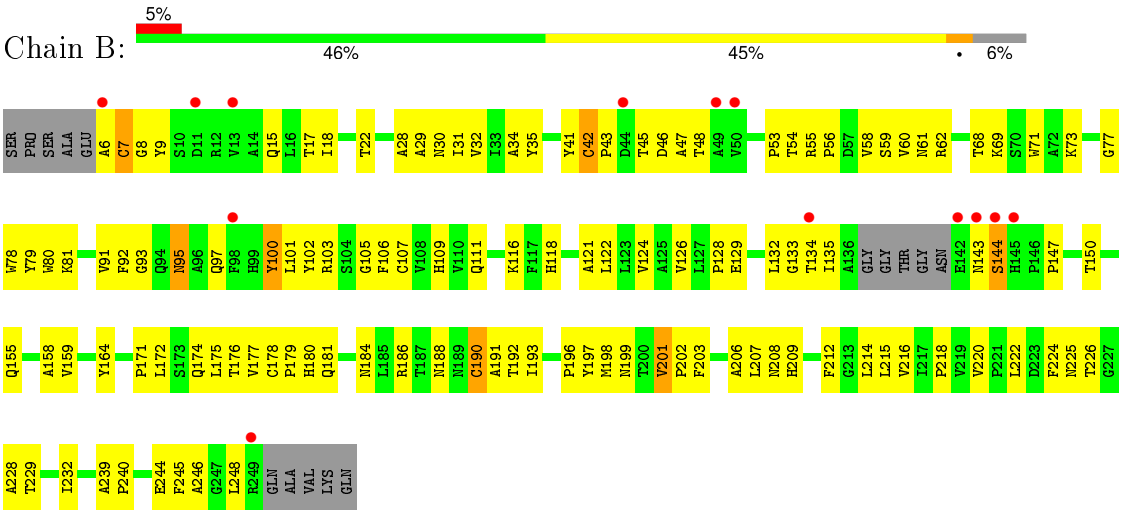
3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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.

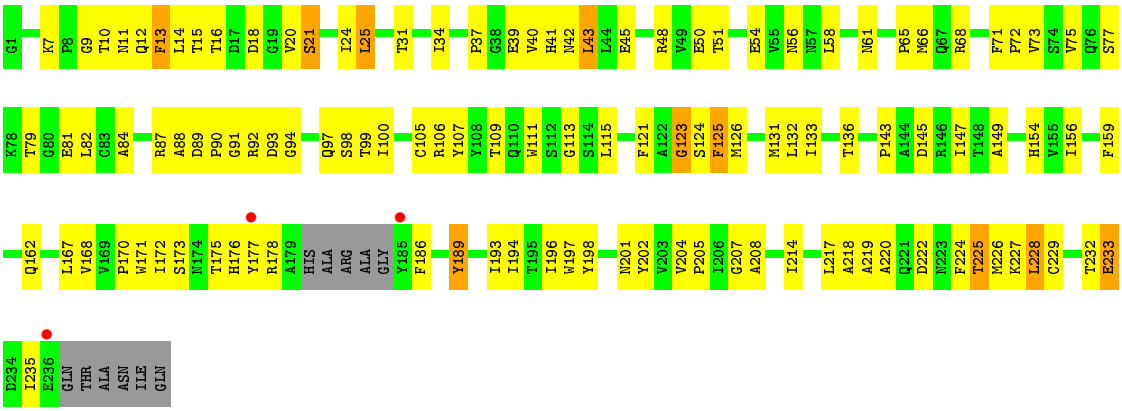
- Molecule 1: Polyprotein, capsid protein VP1



- Molecule 2: Polyprotein, capsid protein VP2



- Molecule 3: Polyprotein, capsid protein VP3


G1 K7 P8 G9 T10 N11 Q12 F13 L14 T15 T16 D17 D18 G19 V20 S21 I24 L25 T31 I34 P37 G38 E39 V40 H41 N42 L43 L44 E45 R48 V49 E50 T51 E54 V55 N56 N57 L58 N61 P65 M66 Q67 R68 F71 P72 V73 S74 V75 Q76 S77
K78 T79 G80 E81 L82 C83 A84 R87 A88 D89 P90 G91 R92 D93 G94 Q97 S98 T99 I100 C105 R106 Y107 Y108 T109 Q110 W111 S112 G113 S114 L115 F121 A122 G123 S124 F125 M126 M131 L132 I133 T136 P143 A144 D145 R146 I147 T148 A149 H154 V155 I156 F159
Q162 L167 V168 V169 P170 W171 I172 S173 N174 T175 H176 Y177 R178 A179 HIS ALA ARG ALA GLY Y185 F186 Y189 I193 I194 T195 I196 W197 Y198 N201 Y202 V203 V204 P205 I206 G207 A208 I214 L217 A218 A219 A220 Q221 D222 N223 F224 T225 M226 K227 L228 C229 T232 E233
D234 I235 E236 GLN THR ALA ASN ILE GLN

4 Data and refinement statistics (i)

Property	Value	Source
Space group	I 2 2 2	Depositor
Cell constants a, b, c, α, β, γ	338.40Å 360.20Å 366.00Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	49.46 – 3.00 49.46 – 3.00	Depositor EDS
% Data completeness (in resolution range)	24.5 (49.46-3.00) 20.0 (49.46-3.00)	Depositor EDS
R_{merge}	0.46	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.16 (at 3.01Å)	Xtriage
Refinement program	CNS 1.3	Depositor
R, R_{free}	0.286 , 0.311 0.288 , 0.283	Depositor DCC
R_{free} test set	5344 reflections (5.21%)	DCC
Wilson B-factor (Å^2)	45.6	Xtriage
Anisotropy	0.069	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	(Not available) , (Not available)	EDS
Estimated twinning fraction	0.147 for -h,-l,-k	Xtriage
L-test for twinning[2]	$< \|L\| > = 0.33$, $< L^2 > = 0.16$	Xtriage
Outliers	1 of 108008 reflections (0.001%)	Xtriage
F_o,F_c correlation	0.76	EDS
Total number of atoms	5468	wwPDB-VP
Average B, all atoms (Å^2)	58.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.81% 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, 0.375 respectively for untwinned datasets, and 0.333, 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	0.39	0/1883	0.64	0/2564
2	B	0.38	0/1898	0.65	0/2608
3	C	0.40	0/1845	0.67	0/2528
All	All	0.39	0/5626	0.65	0/7700

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	1830	0	1792	127	0
2	B	1844	0	1786	127	0
3	C	1794	0	1759	121	0
All	All	5468	0	5337	326	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 30.

The worst 5 of 326 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)
2:B:181:GLN:HG2	2:B:191:ALA:HB1	1.38	1.02

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:189:GLN:HG2	3:C:21:SER:HB3	1.45	0.98
2:B:134:THR:HG22	2:B:135:ILE:H	1.32	0.93
1:A:139:VAL:HG12	1:A:249:LEU:HB2	1.49	0.91
3:C:65:PRO:O	3:C:68:ARG:HG3	1.75	0.87

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	224/297 (75%)	193 (86%)	22 (10%)	9 (4%)	4	21
2	B	235/254 (92%)	191 (81%)	36 (15%)	8 (3%)	5	25
3	C	227/242 (94%)	196 (86%)	26 (12%)	5 (2%)	8	38
All	All	686/793 (86%)	580 (84%)	84 (12%)	22 (3%)	5	27

5 of 22 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	193	PRO
2	B	58	VAL
2	B	144	SER
1	A	124	GLU
2	B	29	ALA

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	204/257 (79%)	196 (96%)	8 (4%)	39 77
2	B	202/212 (95%)	195 (96%)	7 (4%)	43 80
3	C	197/204 (97%)	185 (94%)	12 (6%)	23 61
All	All	603/673 (90%)	576 (96%)	27 (4%)	34 74

5 of 27 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	155	GLN
3	C	13	PHE
3	C	225	THR
2	B	190	CYS
1	A	235	ILE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 21 such sidechains are listed below:

Mol	Chain	Res	Type
2	B	15	GLN
2	B	94	GLN
2	B	225	ASN
1	A	278	ASN
3	C	11	ASN

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 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

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^{th} 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	228/297 (76%)	-0.47	4 (1%) 71 43	14, 48, 125, 173	0
2	B	239/254 (94%)	-0.25	13 (5%) 29 11	23, 52, 157, 194	0
3	C	231/242 (95%)	-0.55	3 (1%) 79 53	16, 42, 105, 139	0
All	All	698/793 (88%)	-0.42	20 (2%) 55 26	14, 46, 140, 194	0

The worst 5 of 20 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	99	GLY	4.1
2	B	144	SER	4.0
3	C	236	GLU	3.7
1	A	71	ASN	3.5
2	B	143	ASN	3.0

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

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

6.3 Carbohydrates

There are no carbohydrates in this entry.

6.4 Ligands

There are no ligands in this entry.

6.5 Other polymers (i)

There are no such residues in this entry.