wwPDB EM Map/Model Validation Report (i)

Apr 10, 2016 – 01:54 PM BST

PDB ID	:	3ZX9
EMDB ID:	:	EMD-1864
Title	:	Cryo-EM reconstruction of native and expanded Turnip Crinkle virus
Authors	:	Bakker, S.E.; Robottom, J.; Pearson, A.R.; Stockley, P.G.; Ranson, N.A.
Deposited on	:	2011-08-08
Resolution	:	17.00 Å(reported)

This is a wwPDB EM Map/Model Validation Report for a publicly released PDB/EMDB entry.
For rigid body fitted models, validation errors reported here could
stem from errors in the original structure(s) used in the fitting.
We welcome your comments at validation@mail.wwpdb.org
A user guide is available at
http://wwpdb.org/validation/2016/EMValidationReportHelp

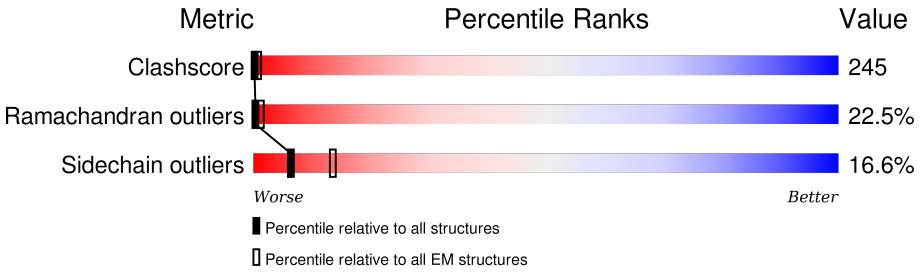
MolProbity	:	4.02b-467
Mogul	:	unknown
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP)	:	trunk27241

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
ELECTRON MICROSCOPY

The reported resolution of this entry is 17.00 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)	EM structures (#Entries)
Clashscore	114402	924
Ramachandran outliers	111179	726
Sidechain outliers	111093	686

The table below summarises the geometric issues observed across the polymeric chains. The red, orange, yellow and green segments on the 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%

Mol	Chain	Length	Quality of chain
1	A	347	6% 35% 24% 12% 23%
1	B	347	6% 35% 25% 11% 23%
1	C	347	• 40% 28% 13% 15%

2 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 6284 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, 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.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	267	Total	C	N	O	S	0	0
			2024	1280	343	396	5		
1	B	267	Total	C	N	O	S	0	0
			2024	1280	343	396	5		
1	C	295	Total	C	N	O	S	0	0
			2236	1409	386	436	5		

There are 15 discrepancies between the modelled and reference sequences:

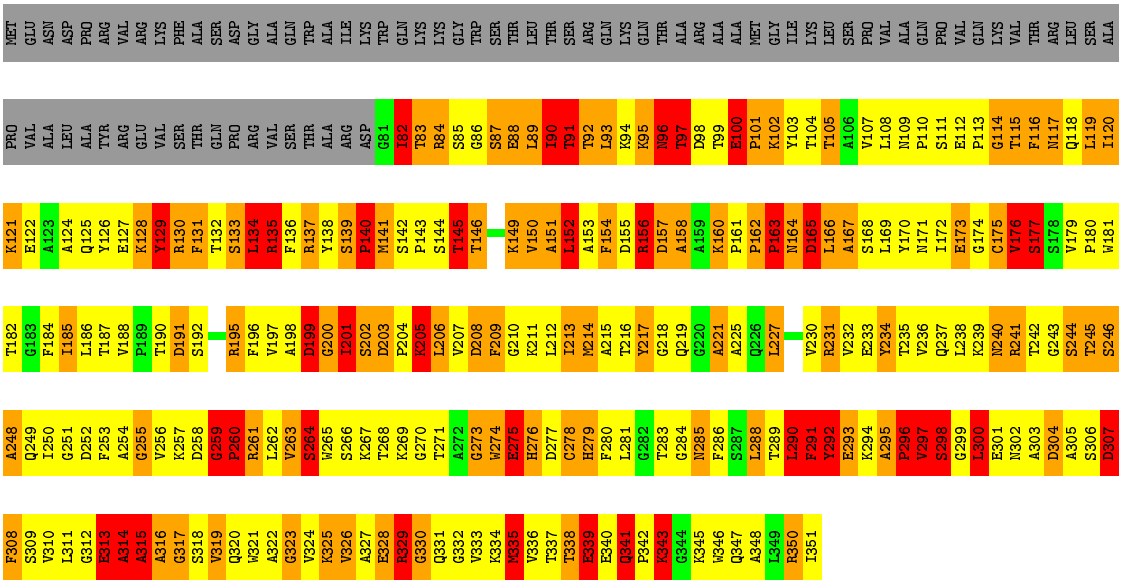
Chain	Residue	Modelled	Actual	Comment	Reference
A	.	-	ASN	DELETION	UNP P06663
A	.	-	ASP	DELETION	UNP P06663
A	.	-	ALA	DELETION	UNP P06663
A	.	-	ASP	DELETION	UNP P06663
A	346	TRP	LEU	VARIANT	UNP P06663
B	.	-	ASN	DELETION	UNP P06663
B	.	-	ASP	DELETION	UNP P06663
B	.	-	ALA	DELETION	UNP P06663
B	.	-	ASP	DELETION	UNP P06663
B	346	TRP	LEU	VARIANT	UNP P06663
C	.	-	ASN	DELETION	UNP P06663
C	.	-	ASP	DELETION	UNP P06663
C	.	-	ALA	DELETION	UNP P06663
C	.	-	ASP	DELETION	UNP P06663
C	346	TRP	LEU	VARIANT	UNP P06663

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 errors 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.

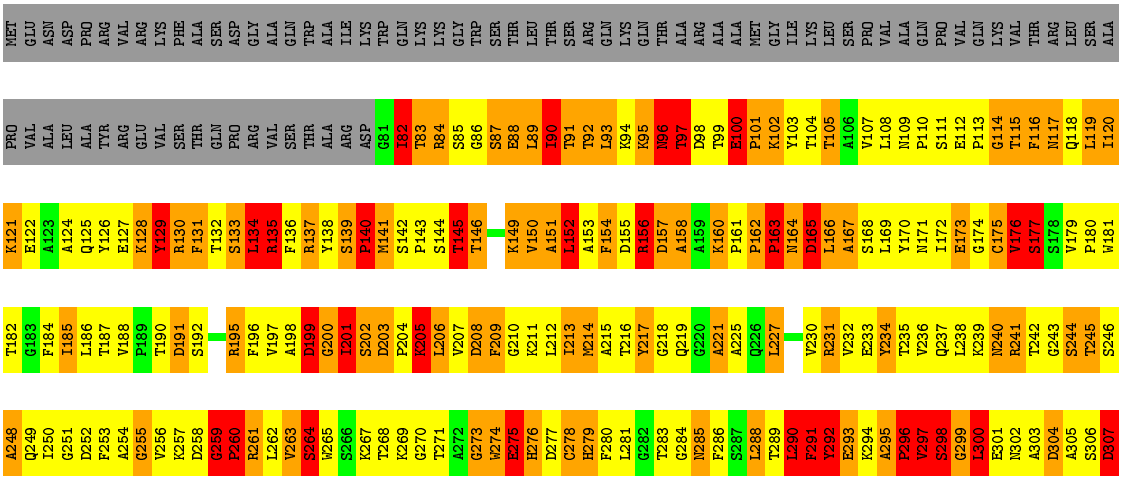
- Molecule 1: CAPSID PROTEIN

Chain A:



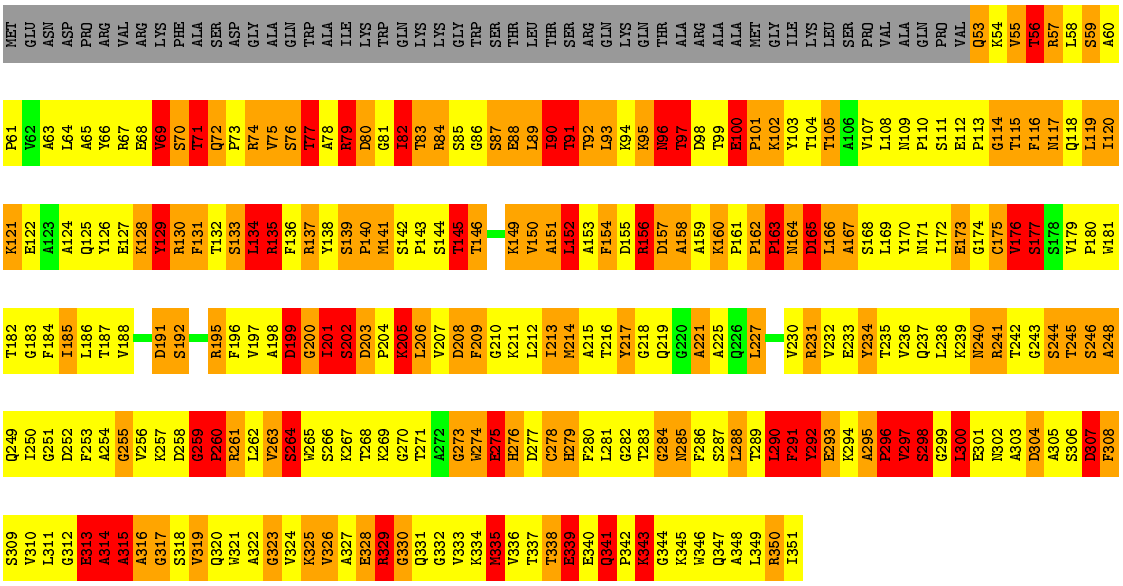
- Molecule 1: CAPSID PROTEIN

Chain B:



- Molecule 1: CAPSID PROTEIN

Chain C:



4 Experimental information (i)

Property	Value	Source
Reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, I	Depositor
Number of images	Not provided	Depositor
Resolution determination method	Not provided	Depositor
CTF correction method	PHASE-FLIPPING EACH PARTICLE	Depositor
Microscope	FEI TECNAI F20	Depositor
Voltage (kV)	200	Depositor
Electron dose ($e^-/Å^2$)	15	Depositor
Minimum defocus (nm)	1000	Depositor
Maximum defocus (nm)	3500	Depositor
Magnification	50000	Depositor
Image detector	KODAK SO163 FILM	Depositor

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	#$\|Z\|$ >2	RMSZ	#$\|Z\|$ >2
1	A	2.49	42/2068 (2.0%)	2.48	130/2804 (4.6%)
1	B	2.49	42/2068 (2.0%)	2.48	130/2804 (4.6%)
1	C	2.51	56/2282 (2.5%)	2.46	145/3096 (4.7%)
All	All	2.50	140/6418 (2.2%)	2.47	405/8704 (4.7%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	8
1	B	0	8
1	C	0	10
All	All	0	26

The worst 5 of 140 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	201	ILE	C-N	-42.35	0.36	1.34
1	C	201	ILE	C-N	-42.34	0.36	1.34
1	A	201	ILE	C-N	-42.34	0.36	1.34
1	B	93	LEU	N-CA	-41.10	0.64	1.46
1	C	93	LEU	N-CA	-41.09	0.64	1.46

The worst 5 of 405 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	201	ILE	O-C-N	-59.03	28.25	122.70
1	A	201	ILE	O-C-N	-59.02	28.27	122.70
1	C	201	ILE	O-C-N	-58.96	28.37	122.70
1	A	82	ILE	O-C-N	-16.84	95.75	122.70
1	C	82	ILE	O-C-N	-16.83	95.78	122.70

There are no chirality outliers.

5 of 26 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	100	GLU	Mainchain
1	A	177	SER	Mainchain
1	A	201	ILE	Mainchain
1	A	260	PRO	Mainchain
1	A	92	THR	Peptide

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	2024	0	1964	947	0
1	B	2024	0	1963	941	0
1	C	2236	0	2181	1146	0
All	All	6284	0	6108	3034	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 245.

The worst 5 of 3034 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:138:TYR:CB	1:A:184:PHE:CE1	1.75	1.66
1:C:138:TYR:CB	1:C:184:PHE:CE1	1.75	1.65
1:C:138:TYR:HB2	1:C:184:PHE:CZ	1.15	1.64
1:A:138:TYR:HB2	1:A:184:PHE:CZ	1.15	1.64
1:B:101:PRO:CD	1:B:166:LEU:CD1	1.75	1.64

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 PDB entries followed by that with respect to all EM entries.

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	263/347 (76%)	144 (55%)	60 (23%)	59 (22%)	0 2
1	B	263/347 (76%)	144 (55%)	60 (23%)	59 (22%)	0 2
1	C	291/347 (84%)	161 (55%)	64 (22%)	66 (23%)	0 2
All	All	817/1041 (78%)	449 (55%)	184 (22%)	184 (22%)	0 2

5 of 184 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	83	THR
1	A	87	SER
1	A	90	ILE
1	A	95	LYS
1	A	96	ASN

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 PDB entries followed by that with respect to all EM entries.

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	215/282 (76%)	178 (83%)	37 (17%)	2 17
1	B	215/282 (76%)	179 (83%)	36 (17%)	3 19
1	C	238/282 (84%)	200 (84%)	38 (16%)	3 21
All	All	668/846 (79%)	557 (83%)	111 (17%)	6 19

5 of 111 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	152	LEU
1	B	217	TYR
1	C	264	SER
1	B	154	PHE
1	B	192	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 36 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	164	ASN
1	B	240	ASN
1	C	249	GLN
1	B	226	GLN
1	B	285	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 (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.