



# wwPDB X-ray Structure Validation Summary Report

Feb 13, 2017 – 12:44 am GMT

PDB ID : 3LQE  
Title : X-Ray Structure of the Murine Norovirus (MNV)-1 Capsid Protein Protruding (P) Domain  
Authors : Rubin, J.R.; Stuckey, J.A.  
Deposited on : 2010-02-09  
Resolution : 2.00 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](mailto:validation@mail.wwpdb.org)

A user guide is available at

<http://wwpdb.org/validation/2016/XrayValidationReportHelp>

with specific help available everywhere you see the  symbol.

---

The following versions of software and data (see [references](#) ) were used in the production of this report:

MolProbity : 4.02b-467  
Xtriage (Phenix) : 1.9-1692  
EDS : trunk28620  
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)  
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) : recalc28949

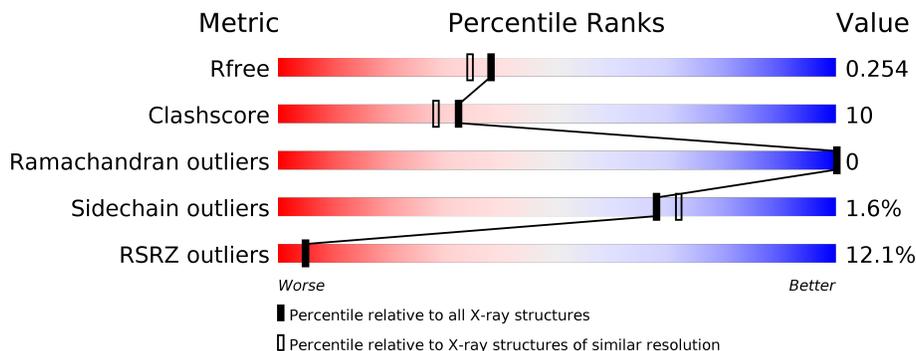
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.00 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))
$R_{free}$	100719	6609 (2.00-2.00)
Clashscore	112137	7775 (2.00-2.00)
Ramachandran outliers	110173	7679 (2.00-2.00)
Sidechain outliers	110143	7678 (2.00-2.00)
RSRZ outliers	101464	6696 (2.00-2.00)

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  $\geq 3$ , 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  $\leq 5\%$ . 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	320	

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 2530 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.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	313	2404	1542	397	457	8	0	0	0

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	222	SER	-	EXPRESSION TAG	UNP Q80J94
A	223	ASN	-	EXPRESSION TAG	UNP Q80J94
A	224	ALA	-	EXPRESSION TAG	UNP Q80J94
A	225	THR	-	EXPRESSION TAG	UNP Q80J94
A	226	ILE	-	EXPRESSION TAG	UNP Q80J94
A	227	TYR	-	EXPRESSION TAG	UNP Q80J94
A	296	GLU	LYS	ENGINEERED	UNP Q80J94
A	541	GLN	-	EXPRESSION TAG	UNP Q80J94

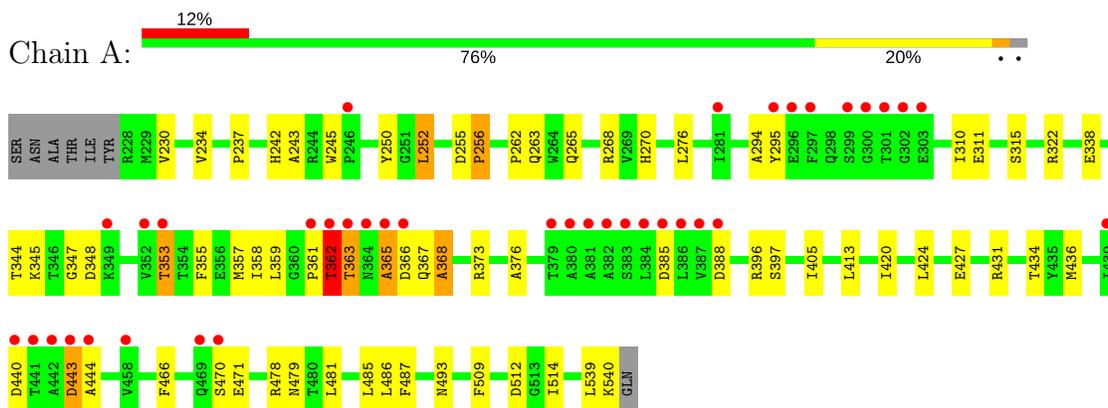
- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	126	Total	O	0	0
			126	126		

### 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 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: Capsid protein



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	86.26Å 105.22Å 84.05Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	32.84 – 2.00 32.84 – 2.00	Depositor EDS
% Data completeness (in resolution range)	93.3 (32.84-2.00) 93.3 (32.84-2.00)	Depositor EDS
$R_{merge}$	0.07	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	4.58 (at 2.00Å)	Xtrriage
Refinement program	REFMAC 5.5.0072	Depositor
R, $R_{free}$	0.215 , 0.256 0.214 , 0.254	Depositor DCC
$R_{free}$ test set	1253 reflections (5.41%)	DCC
Wilson B-factor (Å <sup>2</sup> )	27.2	Xtrriage
Anisotropy	0.064	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 45.9	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	2530	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	28.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.88% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  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	1.23	8/2467 (0.3%)	1.14	19/3376 (0.6%)

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	3

The worst 5 of 8 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	256	PRO	N-CD	10.85	1.63	1.47
1	A	405	ILE	C-N	-5.88	1.23	1.34
1	A	487	PHE	CE2-CZ	5.69	1.48	1.37
1	A	250	TYR	CD1-CE1	5.50	1.47	1.39
1	A	243	ALA	CA-CB	5.41	1.63	1.52

The worst 5 of 19 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	362	THR	C-N-CA	-11.63	92.63	121.70
1	A	362	THR	CB-CA-C	9.77	137.98	111.60
1	A	444	ALA	CB-CA-C	8.92	123.48	110.10
1	A	471	GLU	N-CA-C	8.51	133.97	111.00
1	A	481	LEU	CA-CB-CG	6.67	130.64	115.30

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	362	THR	Peptide
1	A	365	ALA	Peptide
1	A	385	ASP	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	2404	0	2359	47	0
2	A	126	0	0	1	0
All	All	2530	0	2359	47	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 10.

The worst 5 of 47 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:362:THR:HG23	1:A:363:THR:OG1	1.45	1.14
1:A:365:ALA:HB1	1:A:366:ASP:CB	1.79	1.12
1:A:440:ASP:HB3	1:A:443:ASP:HB2	1.18	1.12
1:A:365:ALA:CB	1:A:366:ASP:HB2	1.85	1.07
1:A:365:ALA:HB1	1:A:366:ASP:HB2	1.02	0.99

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	311/320 (97%)	298 (96%)	13 (4%)	0	100 100

There are no Ramachandran outliers to report.

### 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	257/263 (98%)	253 (98%)	4 (2%)	68 72

All (4) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	237	PRO
1	A	256	PRO
1	A	353	THR
1	A	363	THR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	263	GLN
1	A	526	GLN

### 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.

## 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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	313/320 (97%)	0.53	38 (12%) <b>5</b> <b>5</b>	11, 25, 55, 63	18 (5%)

The worst 5 of 38 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	442	ALA	14.2
1	A	364	ASN	10.1
1	A	382	ALA	8.7
1	A	379	THR	6.4
1	A	381	ALA	6.4

### 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 carbohydrates 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.