



# Full wwPDB X-ray Structure Validation Report ⓘ

May 23, 2020 – 01:07 am BST

PDB ID : 4LZI  
Title : Characterization of Solanum tuberosum Multicystatin and Significance of Core Domains  
Authors : Nissen, M.S.; Kumar, G.N.; Green, A.R.; Knowles, N.R.; Kang, C.  
Deposited on : 2013-07-31  
Resolution : 2.20 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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

<https://www.wwpdb.org/validation/2017/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)	:	1.13
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

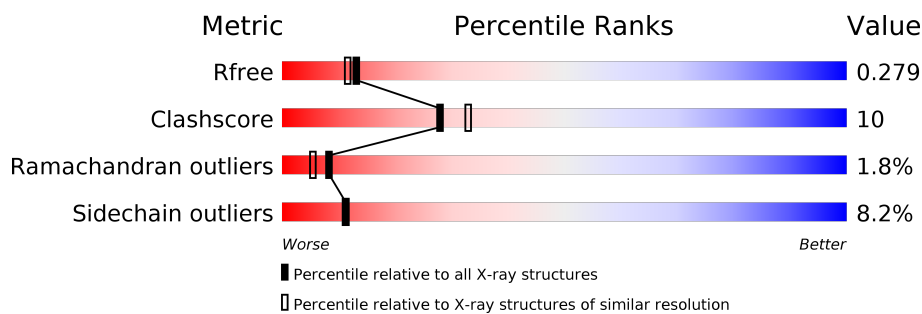
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.20 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	4898 (2.20-2.20)
Clashscore	141614	5594 (2.20-2.20)
Ramachandran outliers	138981	5503 (2.20-2.20)
Sidechain outliers	138945	5504 (2.20-2.20)

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 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  $\leq 5\%$ .

Mol	Chain	Length	Quality of chain
1	A	282	

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 2458 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 Multicystatin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	277	Total	C	N	O	S	0	1	0
			2252	1475	361	412	4			

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	2	GLY	-	EXPRESSION TAG	UNP P37842

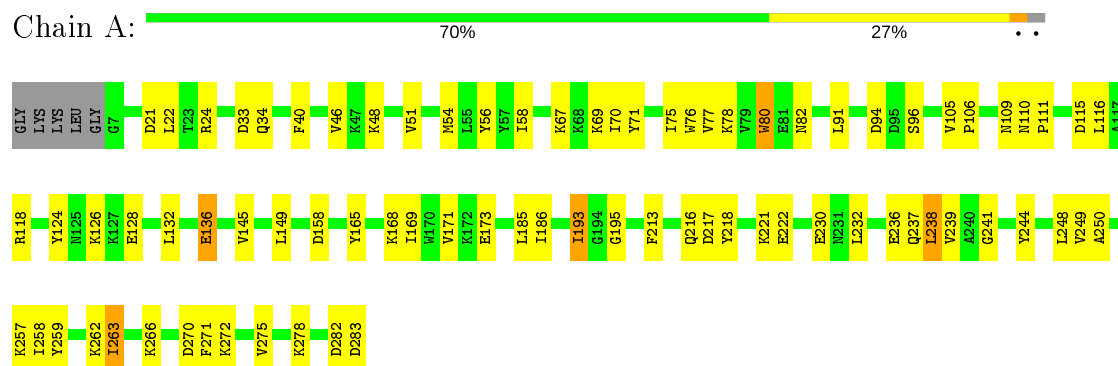
- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	206	Total	O	0	0
			206	206		

### 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: Multicystatin



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 43 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	65.94Å 65.94Å 191.85Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	25.12 – 2.20 25.12 – 1.49	Depositor EDS
% Data completeness (in resolution range)	89.2 (25.12-2.20) 61.5 (25.12-1.49)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.03 (at 1.49Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.8.2_1309)	Depositor
R, $R_{free}$	0.219 , 0.279 0.219 , 0.279	Depositor DCC
$R_{free}$ test set	2000 reflections (4.59%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	9.1	Xtriage
Anisotropy	0.209	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.31 , 56.8	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.32$ , $\langle L^2 \rangle = 0.16$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.89	EDS
Total number of atoms	2458	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	21.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.28% 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

### 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	0.48	0/2310	0.62	0/3123

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	2252	0	2217	44	0
2	A	206	0	0	17	1
All	All	2458	0	2217	44	1

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.

All (44) 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:193:ILE:O	2:A:476:HOH:O	1.89	0.89
1:A:33:ASP:OD2	2:A:324:HOH:O	1.98	0.80
1:A:230:GLU:OE2	2:A:302:HOH:O	2.04	0.73
1:A:270:ASP:OD1	2:A:347:HOH:O	2.07	0.73

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:33:ASP:OD2	2:A:400:HOH:O	2.12	0.68
1:A:111:PRO:O	2:A:322:HOH:O	2.11	0.67
1:A:34:GLN:NE2	2:A:395:HOH:O	2.23	0.66
1:A:168:LYS:NZ	2:A:356:HOH:O	2.27	0.66
1:A:257:LYS:HE2	1:A:283:ASP:HA	1.78	0.65
1:A:76:TRP:CZ3	1:A:173:GLU:HG3	2.32	0.64
1:A:115:ASP:OD1	1:A:118:ARG:NH2	2.34	0.61
1:A:282:ASP:OD2	2:A:335:HOH:O	2.17	0.58
1:A:263:ILE:HD11	1:A:272:LYS:HG2	1.86	0.58
1:A:46:VAL:HG22	1:A:58:ILE:HG12	1.86	0.57
1:A:217:ASP:OD1	2:A:353:HOH:O	2.17	0.56
1:A:218:TYR:CE1	1:A:222:GLU:HG3	2.41	0.56
1:A:54[B]:MET:SD	2:A:360:HOH:O	2.58	0.55
1:A:78:LYS:HA	2:A:454:HOH:O	2.08	0.53
1:A:48:LYS:HG2	2:A:360:HOH:O	2.09	0.53
1:A:109:ASN:ND2	2:A:379:HOH:O	2.32	0.51
1:A:136:GLU:OE1	2:A:303:HOH:O	2.19	0.51
1:A:262:LYS:HD3	1:A:275:VAL:HG23	1.91	0.51
1:A:266:LYS:HB2	1:A:271:PHE:HB3	1.94	0.49
1:A:48:LYS:CG	2:A:360:HOH:O	2.60	0.48
1:A:250:ALA:HB3	1:A:259:TYR:CE1	2.48	0.48
1:A:195:GLY:O	1:A:237:GLN:NE2	2.46	0.48
1:A:69:LYS:HD3	1:A:71:TYR:OH	2.13	0.48
1:A:186:ILE:HG21	1:A:258:ILE:HD12	1.96	0.47
1:A:238:LEU:HD13	1:A:241:GLY:O	2.14	0.47
1:A:22:LEU:HB3	1:A:75:ILE:HD12	1.97	0.46
1:A:78:LYS:HD2	1:A:80:TRP:CZ2	2.51	0.46
1:A:158:ASP:HA	2:A:384:HOH:O	2.15	0.46
1:A:124:TYR:CE1	1:A:128:GLU:HG3	2.51	0.45
1:A:24:ARG:CZ	1:A:40:PHE:HD2	2.30	0.44
1:A:236:GLU:HG2	1:A:244:TYR:CD1	2.53	0.44
1:A:70:ILE:O	1:A:91:LEU:HD12	2.17	0.44
1:A:213:PHE:O	1:A:216:GLN:HG3	2.18	0.44
1:A:78:LYS:HD2	1:A:80:TRP:CH2	2.54	0.43
1:A:173:GLU:HG2	1:A:173:GLU:H	1.67	0.42
1:A:48:LYS:HG3	1:A:56:TYR:CE1	2.54	0.41
1:A:132:LEU:HD21	1:A:165:TYR:CE2	2.55	0.41
1:A:116:LEU:HA	1:A:116:LEU:HD23	1.82	0.41
1:A:278:LYS:HB2	1:A:278:LYS:HE2	1.85	0.41
1:A:105:VAL:HA	1:A:106:PRO:HD3	1.90	0.40

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the sym-

metry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:A:401:HOH:O	2:A:415:HOH:O[3_454]	2.09	0.11

## 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	276/282 (98%)	256 (93%)	15 (5%)	5 (2%)	<b>8</b> <b>5</b>

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	94	ASP
1	A	239	VAL
1	A	51	VAL
1	A	145	VAL
1	A	193	ILE

### 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	234/244 (96%)	215 (92%)	19 (8%)	<b>11</b> <b>12</b>

All (19) residues with a non-rotameric sidechain are listed below:



Mol	Chain	Res	Type
1	A	21	ASP
1	A	67	LYS
1	A	77	VAL
1	A	80	TRP
1	A	82	ASN
1	A	96	SER
1	A	110	ASN
1	A	126	LYS
1	A	136	GLU
1	A	149	LEU
1	A	169	ILE
1	A	171	VAL
1	A	185	LEU
1	A	221	LYS
1	A	232	LEU
1	A	238	LEU
1	A	248	LEU
1	A	249	VAL
1	A	263	ILE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

### 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

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.