



# wwPDB X-ray Structure Validation Summary Report ⓘ

Sep 27, 2023 – 10:09 PM EDT

PDB ID : 1SMV  
Title : PRIMARY STRUCTURE OF SESBANIA MOSAIC VIRUS COAT PROTEIN: ITS IMPLICATIONS TO THE ASSEMBLY AND ARCHITECTURE OF THE VIRUS  
Authors : Bhuvaneshwari, M.; Murthy, M.R.N.  
Deposited on : 1995-06-16  
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](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 types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

---

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 : 2.35.1  
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) : 2.35.1

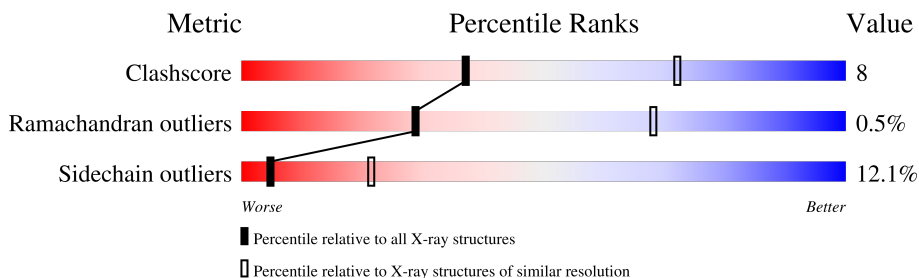
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.00 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))
Clashscore	141614	2416 (3.00-3.00)
Ramachandran outliers	138981	2333 (3.00-3.00)
Sidechain outliers	138945	2336 (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 of 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	266	
1	B	266	
1	C	266	

## 2 Entry composition i

There are 2 unique types of molecules in this entry. The entry contains 4516 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 SESBANIA MOSAIC VIRUS COAT PROTEIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	196	1437	911	234	284	8	0	0	0
1	B	196	1452	924	234	287	7	0	0	0
1	C	222	1623	1028	269	316	10	0	0	0

There are 42 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	0	ILE	LYS	CONFLICT	UNP Q9EB06
A	36	PRO	GLN	CONFLICT	UNP Q9EB06
A	38	ILE	THR	CONFLICT	UNP Q9EB06
A	66	ALA	GLY	CONFLICT	UNP Q9EB06
A	?	-	THR	DELETION	UNP Q9EB06
A	72	CYS	SER	CONFLICT	UNP Q9EB06
A	75	THR	SER	CONFLICT	UNP Q9EB06
A	105	ASP	ALA	CONFLICT	UNP Q9EB06
A	149	LYS	GLN	CONFLICT	UNP Q9EB06
A	173	ASN	GLY	CONFLICT	UNP Q9EB06
A	174	SER	THR	CONFLICT	UNP Q9EB06
A	191	GLU	LYS	CONFLICT	UNP Q9EB06
A	218	ASP	PRO	CONFLICT	UNP Q9EB06
A	244	ASP	CYS	CONFLICT	UNP Q9EB06
B	0	ILE	LYS	CONFLICT	UNP Q9EB06
B	36	PRO	GLN	CONFLICT	UNP Q9EB06
B	38	ILE	THR	CONFLICT	UNP Q9EB06
B	66	ALA	GLY	CONFLICT	UNP Q9EB06
B	?	-	THR	DELETION	UNP Q9EB06
B	72	CYS	SER	CONFLICT	UNP Q9EB06
B	75	THR	SER	CONFLICT	UNP Q9EB06
B	105	ASP	ALA	CONFLICT	UNP Q9EB06
B	149	LYS	GLN	CONFLICT	UNP Q9EB06

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
B	173	ASN	GLY	CONFLICT	UNP Q9EB06
B	174	SER	THR	CONFLICT	UNP Q9EB06
B	191	GLU	LYS	CONFLICT	UNP Q9EB06
B	218	ASP	PRO	CONFLICT	UNP Q9EB06
B	244	ASP	CYS	CONFLICT	UNP Q9EB06
C	0	ILE	LYS	CONFLICT	UNP Q9EB06
C	36	PRO	GLN	CONFLICT	UNP Q9EB06
C	38	ILE	THR	CONFLICT	UNP Q9EB06
C	66	ALA	GLY	CONFLICT	UNP Q9EB06
C	?	-	THR	DELETION	UNP Q9EB06
C	72	CYS	SER	CONFLICT	UNP Q9EB06
C	75	THR	SER	CONFLICT	UNP Q9EB06
C	105	ASP	ALA	CONFLICT	UNP Q9EB06
C	149	LYS	GLN	CONFLICT	UNP Q9EB06
C	173	ASN	GLY	CONFLICT	UNP Q9EB06
C	174	SER	THR	CONFLICT	UNP Q9EB06
C	191	GLU	LYS	CONFLICT	UNP Q9EB06
C	218	ASP	PRO	CONFLICT	UNP Q9EB06
C	244	ASP	CYS	CONFLICT	UNP Q9EB06

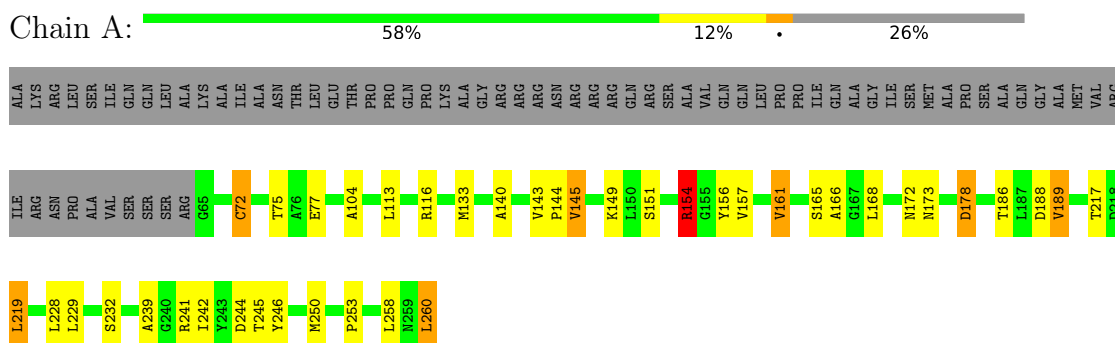
- Molecule 2 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	3	Total Ca 3 3	0	0
2	C	1	Total Ca 1 1	0	0

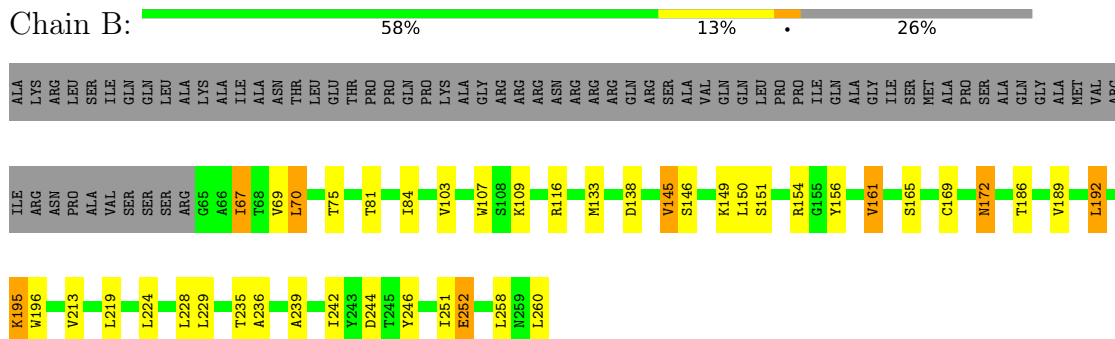
### 3 Residue-property plots

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.

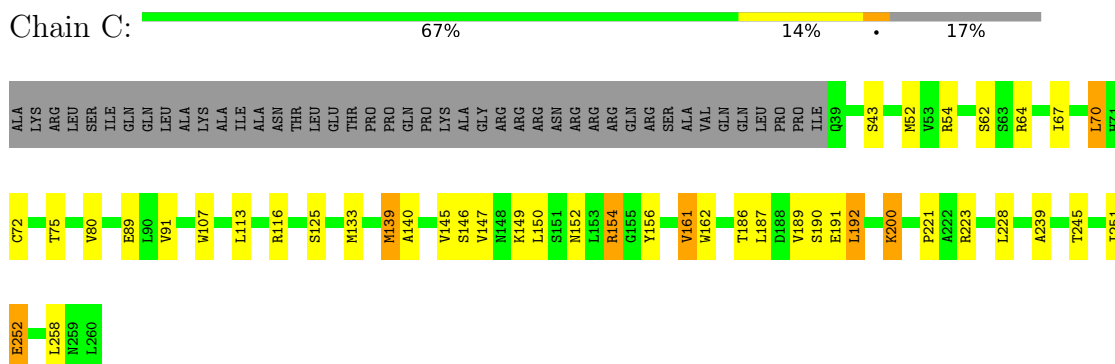
- Molecule 1: SESBANIA MOSAIC VIRUS COAT PROTEIN



- Molecule 1: SESBANIA MOSAIC VIRUS COAT PROTEIN



- Molecule 1: SESBANIA MOSAIC VIRUS COAT PROTEIN



## 4 Data and refinement statistics

Property	Value	Source
Space group	R 3	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	291.46Å 291.46Å 291.46Å 61.95° 61.95° 61.95°	Depositor
Resolution (Å)	10.00 – 3.00 51.40 – 2.90	Depositor EDS
% Data completeness (in resolution range)	70.0 (10.00-3.00) 66.4 (51.40-2.90)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$	-	Xtrriage
Refinement program	X-PLOR 3.1	Depositor
R, $R_{free}$	0.227 , (Not available) 0.699 , (Not available)	Depositor DCC
$R_{free}$ test set	No test flags present.	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	(Not available)	Xtrriage
Anisotropy	(Not available)	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 64.8	EDS
L-test for twinning <sup>1</sup>	$\langle  L  \rangle =$ (Not available), $\langle L^2 \rangle =$ (Not available)	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.18	EDS
Total number of atoms	4516	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	11.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *(Not available)*

<sup>1</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](#)

Bond lengths and bond angles in the following residue types are not validated in this section: CA

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	0.62	0/1468	0.84	2/2012 (0.1%)
1	B	0.59	0/1483	0.83	1/2033 (0.0%)
1	C	0.60	0/1656	0.81	2/2268 (0.1%)
All	All	0.60	0/4607	0.82	5/6313 (0.1%)

There are no bond length outliers.

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	70	LEU	CA-CB-CG	-7.71	97.56	115.30
1	C	70	LEU	CA-CB-CG	-7.02	99.15	115.30
1	A	154	ARG	NE-CZ-NH2	-5.64	117.48	120.30
1	C	70	LEU	CB-CG-CD2	-5.58	101.52	111.00
1	A	219	LEU	CA-CB-CG	5.07	126.96	115.30

There are no chirality outliers.

There are no planarity outliers.

### 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	1437	0	1369	24	0
1	B	1452	0	1416	26	0
1	C	1623	0	1586	26	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	A	3	0	0	0	0
2	C	1	0	0	0	0
All	All	4516	0	4371	69	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 8.

The worst 5 of 69 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:145:VAL:HG23	1:A:149:LYS:HE3	1.42	1.00
1:A:72:CYS:SG	1:A:245:THR:HG23	2.08	0.93
1:A:149:LYS:HD3	1:B:258:LEU:HD21	1.47	0.92
1:B:116:ARG:HB2	1:B:186:THR:HG22	1.52	0.91
1:C:72:CYS:SG	1:C:245:THR:HG23	2.16	0.85

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	194/266 (73%)	180 (93%)	12 (6%)	2 (1%)	15	53
1	B	194/266 (73%)	184 (95%)	9 (5%)	1 (0%)	29	68
1	C	220/266 (83%)	211 (96%)	9 (4%)	0	100	100
All	All	608/798 (76%)	575 (95%)	30 (5%)	3 (0%)	29	68

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	173	ASN

Continued on next page...



*Continued from previous page...*

Mol	Chain	Res	Type
1	B	161	VAL
1	A	189	VAL

### 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	150/220 (68%)	131 (87%)	19 (13%)	4	19
1	B	157/220 (71%)	140 (89%)	17 (11%)	6	26
1	C	172/220 (78%)	150 (87%)	22 (13%)	4	19
All	All	479/660 (73%)	421 (88%)	58 (12%)	5	21

5 of 58 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	192	LEU
1	C	228	LEU
1	B	252	GLU
1	C	200	LYS
1	C	161	VAL

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

Mol	Chain	Res	Type
1	B	172	ASN
1	C	49	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 monosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

Of 4 ligands modelled in this entry, 4 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

## 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

### 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.