wwPDB X-ray Structure Validation Summary Report (i)

Feb 1, 2016 – 08:32 PM GMT

PDB ID	:	4SBV
Title	:	The REFINEMENT OF SOUTHERN BEAN MOSAIC VIRUS IN RECIPROCAL SPACE
Authors	:	Rossmann, M.G.
Deposited on	:	1985-04-01
Resolution	:	2.80 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:

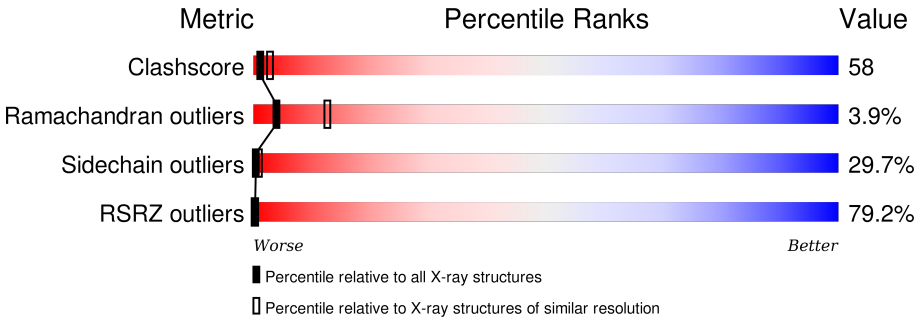
MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20026688
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	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 2.80 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	102246	2827 (2.80-2.80)
Ramachandran outliers	100387	2782 (2.80-2.80)
Sidechain outliers	100360	2784 (2.80-2.80)
RSRZ outliers	91569	2404 (2.80-2.80)

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	260	55% 16% 35% 17% 9% 23%
1	B	260	66% 8% 34% 24% 11% 23%
1	C	260	68% 7% 43% 26% 10% 15%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 4723 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 SOUTHERN BEAN MOSAIC VIRUS COAT PROTEIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	199	Total	C	N	O	S	0	0	0
			1506	956	249	292	9			
1	B	199	Total	C	N	O	S	0	0	0
			1506	956	249	292	9			
1	C	222	Total	C	N	O	S	0	0	0
			1674	1062	281	319	12			

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

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

- Molecule 3 is water.

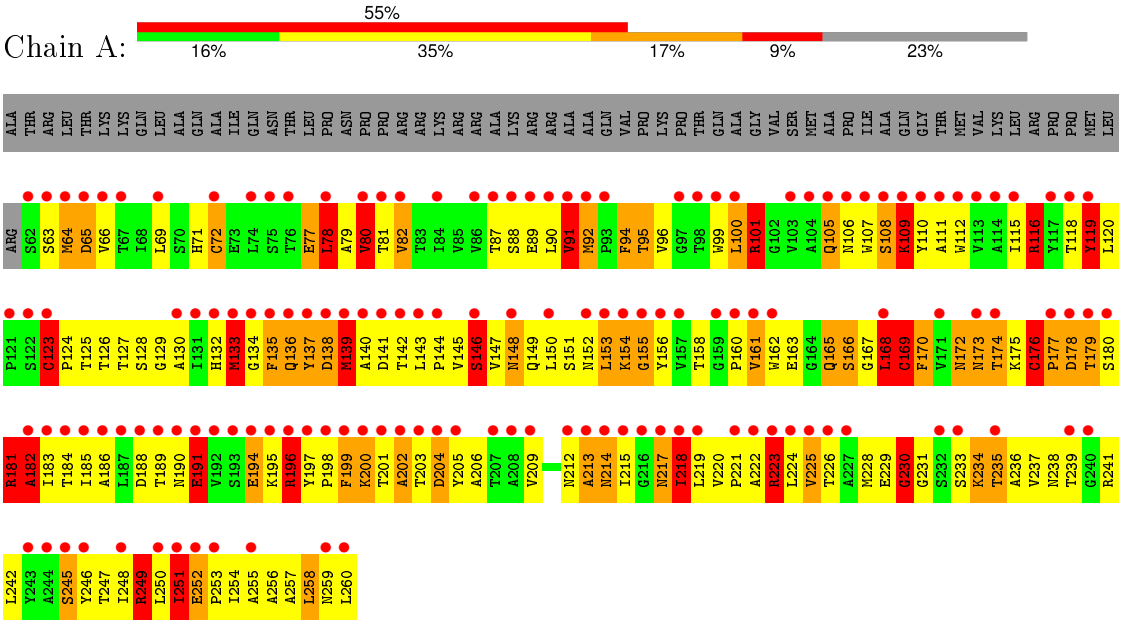
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	15	Total	O	0	0
			15	15		
3	B	10	Total	O	0	0
			10	10		
3	C	9	Total	O	0	0
			9	9		

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 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.

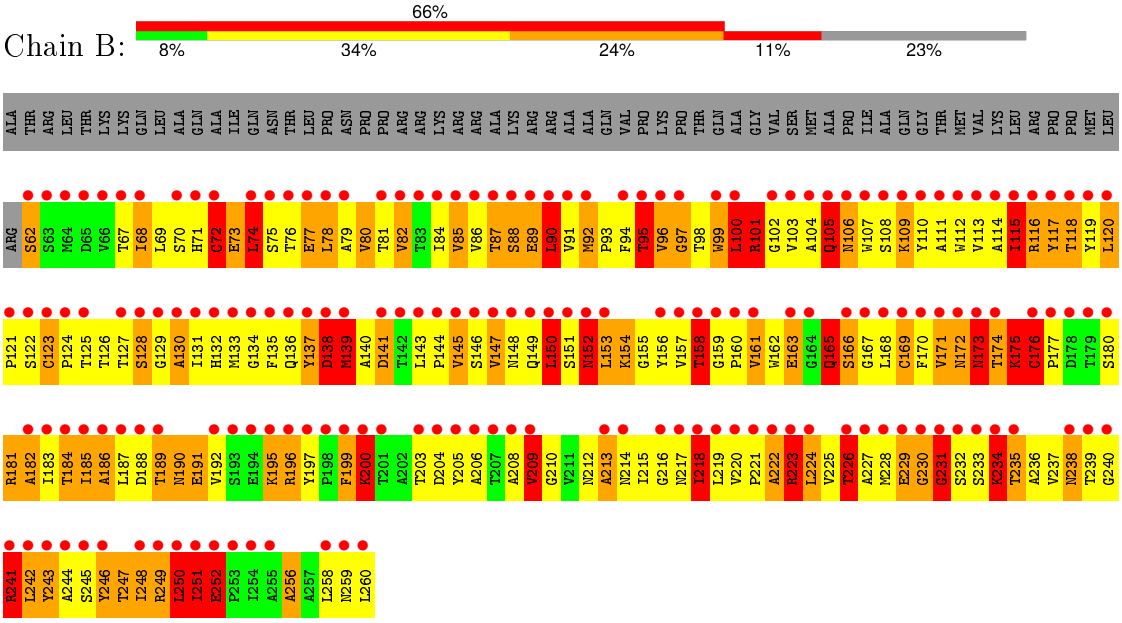
- Molecule 1: SOUTHERN BEAN MOSAIC VIRUS COAT PROTEIN

Chain A:



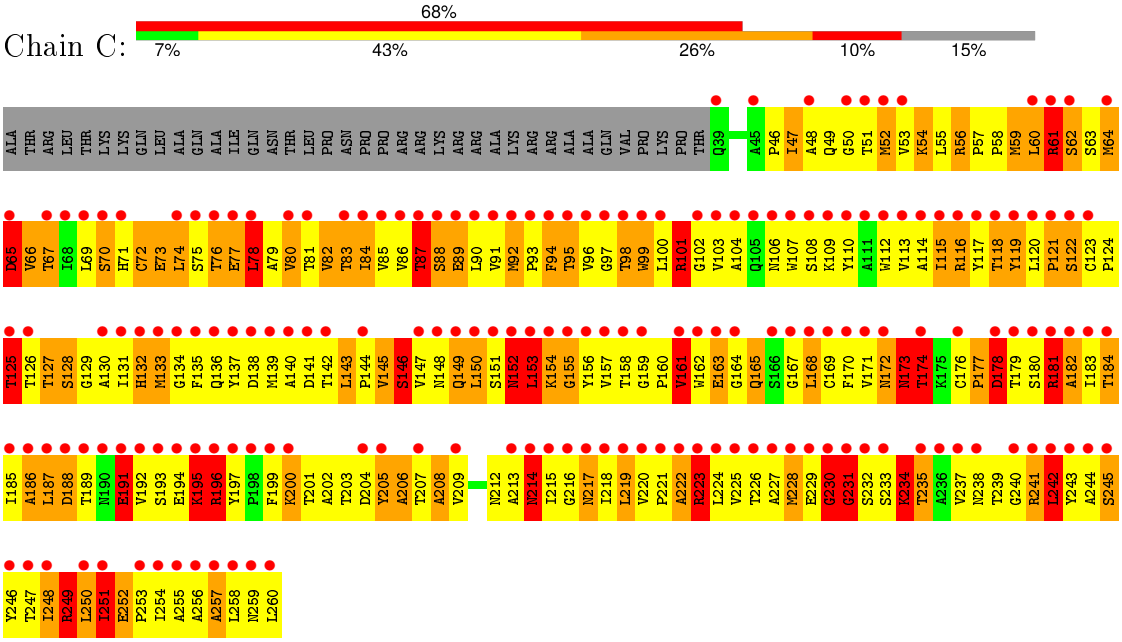
- Molecule 1: SOUTHERN BEAN MOSAIC VIRUS COAT PROTEIN

Chain B:



• Molecule 1: SOUTHERN BEAN MOSAIC VIRUS COAT PROTEIN

Chain C:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	H 3 2	Depositor
Cell constants a, b, c, α, β, γ	334.30Å 334.30Å 757.50Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	(Not available) – 2.80 142.18 – 2.80	Depositor EDS
% Data completeness (in resolution range)	(Not available) ((Not available)-2.80) 75.4 (142.18-2.80)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$	-	Xtriage
Refinement program	unknown	Depositor
R, R_{free}	0.254 , (Not available) 0.515 , (Not available)	Depositor DCC
R_{free} test set	No test flags present.	DCC
Wilson B-factor (Å^2)	30.6	Xtriage
Anisotropy	0.521	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.48 , 923.1	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning[1]	$< \|L\| > = 0.39$, $< L^2 > = 0.22$	Xtriage
Outliers	10 of 298615 reflections (0.003%)	Xtriage
F_o,F_c correlation	0.22	EDS
Total number of atoms	4723	wwPDB-VP
Average B, all atoms (Å^2)	24.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 1.93% of the height of the origin peak. No significant pseudotranslation is detected.*

[1]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

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	1.47	8/1537 (0.5%)	2.64	103/2104 (4.9%)
1	B	1.43	7/1537 (0.5%)	2.46	106/2104 (5.0%)
1	C	1.48	11/1708 (0.6%)	2.55	96/2335 (4.1%)
All	All	1.46	26/4782 (0.5%)	2.55	305/6543 (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	3
1	B	0	3
1	C	0	4
All	All	0	10

The worst 5 of 26 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	252	GLU	N-CA	10.24	1.66	1.46
1	C	155	GLY	N-CA	9.10	1.59	1.46
1	B	231	GLY	N-CA	-8.86	1.32	1.46
1	C	231	GLY	N-CA	-8.81	1.32	1.46
1	A	230	GLY	N-CA	8.31	1.58	1.46

The worst 5 of 305 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	196	ARG	NE-CZ-NH1	31.30	135.95	120.30
1	C	101	ARG	NE-CZ-NH1	26.03	133.32	120.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	109	LYS	CA-CB-CG	25.72	169.98	113.40
1	A	196	ARG	NE-CZ-NH2	-18.85	110.87	120.30
1	A	125	THR	N-CA-CB	17.78	144.08	110.30

There are no chirality outliers.

5 of 10 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	116	ARG	Sidechain
1	A	223	ARG	Sidechain
1	A	249	ARG	Sidechain
1	B	101	ARG	Sidechain
1	B	241	ARG	Sidechain

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	1506	0	1504	207	87
1	B	1506	0	1504	182	1771
1	C	1674	0	1691	187	1943
2	A	1	0	0	0	0
2	B	1	0	0	0	0
2	C	1	0	0	0	0
3	A	15	0	0	1	1
3	B	10	0	0	0	6
3	C	9	0	0	0	15
All	All	4723	0	4699	540	2046

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 58.

The worst 5 of 540 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:C:212:ASN:ND2	1:C:214:ASN:H	1.42	1.15

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:145:VAL:H	1:C:149:GLN:NE2	1.46	1.11
1:C:161:VAL:HG22	1:C:239:THR:CG2	1.82	1.08
1:C:161:VAL:HG22	1:C:239:THR:HG21	1.09	1.07
1:C:131:ILE:HG13	1:C:228:MET:HE2	1.34	1.06

The worst 5 of 2046 symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:185:ILE:CG1	1:C:118:THR:N[3_555]	0.11	2.09
1:B:130:ALA:CB	1:C:136:GLN:N[3_555]	0.18	2.02
1:B:120:LEU:C	1:C:156:TYR:CE1[3_555]	0.23	1.97
1:B:74:LEU:CB	1:C:130:ALA:CB[3_555]	0.25	1.95
1:B:160:PRO:O	1:C:153:LEU:C[3_555]	0.29	1.91

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	197/260 (76%)	160 (81%)	27 (14%)	10 (5%)	2 8
1	B	197/260 (76%)	163 (83%)	29 (15%)	5 (2%)	7 24
1	C	220/260 (85%)	182 (83%)	29 (13%)	9 (4%)	3 11
All	All	614/780 (79%)	505 (82%)	85 (14%)	24 (4%)	4 12

5 of 24 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	189	THR
1	B	231	GLY
1	C	219	LEU
1	C	230	GLY

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	C	231	GLY

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 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	167/217 (77%)	121 (72%)	46 (28%)	0 1
1	B	167/217 (77%)	116 (70%)	51 (30%)	0 1
1	C	185/217 (85%)	128 (69%)	57 (31%)	0 1
All	All	519/651 (80%)	365 (70%)	154 (30%)	0 1

5 of 154 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	153	LEU
1	B	209	VAL
1	C	195	LYS
1	B	158	THR
1	B	176	CYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 26 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	152	ASN
1	B	173	ASN
1	C	212	ASN
1	B	165	GLN
1	B	172	ASN

5.3.3 RNA

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates

There are no carbohydrates in this entry.

5.6 Ligand geometry

Of 3 ligands modelled in this entry, 3 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

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	199/260 (76%)	2.57	142 (71%)	0	0	12, 22, 60, 83	0
1	B	199/260 (76%)	3.14	172 (86%)	0	0	10, 21, 50, 65	0
1	C	222/260 (85%)	2.93	177 (79%)	0	0	8, 21, 37, 60	0
All	All	620/780 (79%)	2.88	491 (79%)	0	0	8, 21, 46, 83	0

The worst 5 of 491 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	222	ALA	7.5
1	B	240	GLY	7.3
1	C	70	SER	6.7
1	A	193	SER	6.7
1	C	65	ASP	6.6

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

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors

of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors(Å^2)	Q<0.9
2	CA	C	261	1/1	0.82	0.22	-3.49	20,20,20,20	0
2	CA	B	261	1/1	0.93	0.32	-3.77	16,16,16,16	0
2	CA	A	261	1/1	0.82	0.21	-4.75	27,27,27,27	0

6.5 Other polymers (i)

There are no such residues in this entry.