wwPDB X-ray Structure Validation Summary Report (i)

Feb 10, 2024 – 02:03 PM EST

PDB ID	:	2TBV
Title	:	STRUCTURE OF TOMATO BUSHY STUNT VIRUS. V. COAT PROTEIN SEQUENCE DETERMINATION AND ITS STRUCTURAL IMPLICATIONS
Authors	:	Harrison, S.C.
Deposited on	:	1984-06-22
Resolution	:	2.90 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
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) 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 (i)) were used in the production of this report:

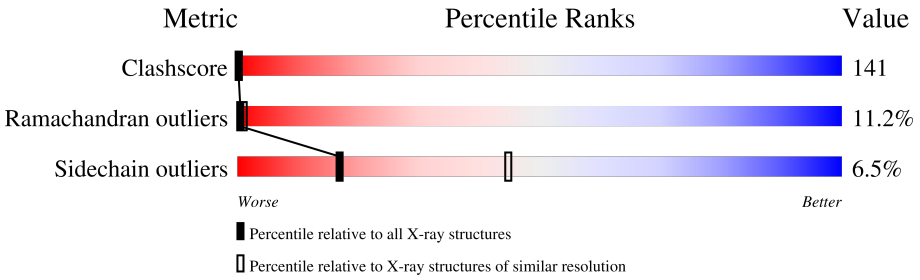
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

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 2.90 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	2172 (2.90-2.90)
Ramachandran outliers	138981	2115 (2.90-2.90)
Sidechain outliers	138945	2117 (2.90-2.90)

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 >=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 <=5%

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	387	8% 38% 21% 7% 26%
1	B	387	7% 37% 23% 7% 26%
1	C	387	8% 43% 24% 7% 17%

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 6648 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 TOMATO BUSHY STUNT VIRUS.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	287	Total	C	N	O	S	0	3	1
			2136	1351	360	420	5			
1	B	287	Total	C	N	O	S	0	2	1
			2130	1348	359	418	5			
1	C	321	Total	C	N	O	S	0	3	0
			2376	1502	406	462	6			

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	102	GLY	SER	conflict	UNP P11795
A	107	SER	GLY	conflict	UNP P11795
B	102	GLY	SER	conflict	UNP P11795
B	107	SER	GLY	conflict	UNP P11795
C	102	GLY	SER	conflict	UNP P11795
C	107	SER	GLY	conflict	UNP P11795

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

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

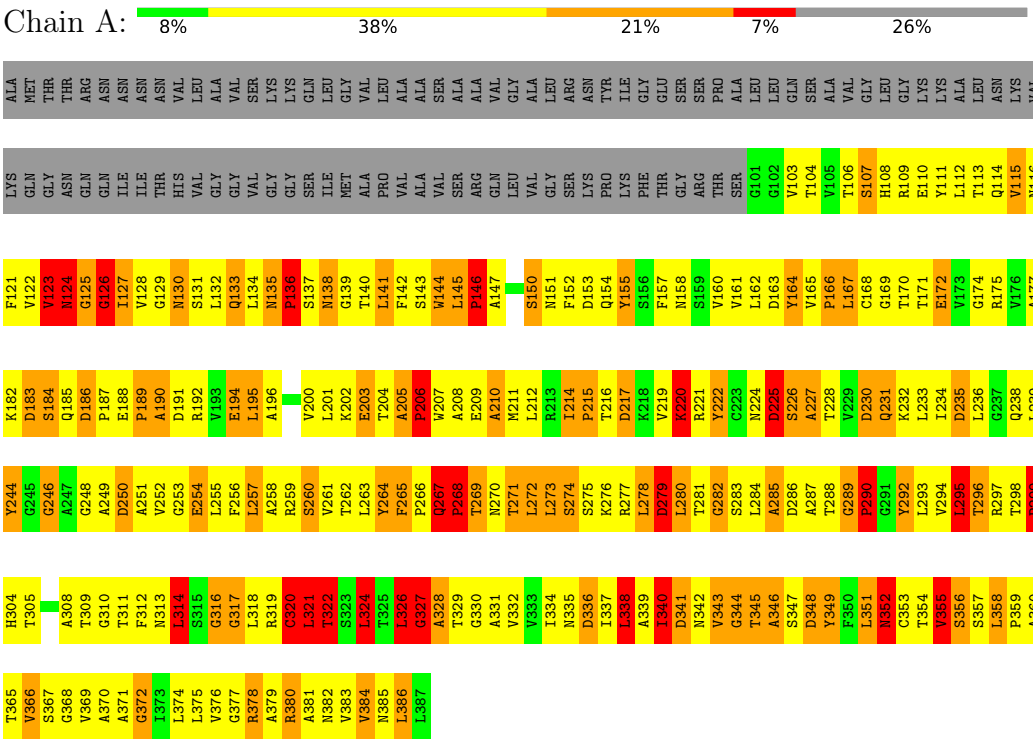
3 Residue-property plots (i)

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.

Note EDS was not executed.

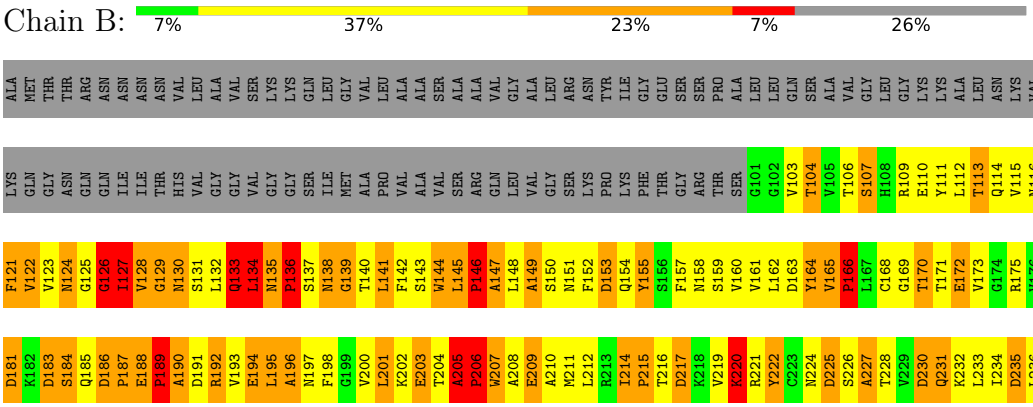
• Molecule 1: TOMATO BUSHY STUNT VIRUS

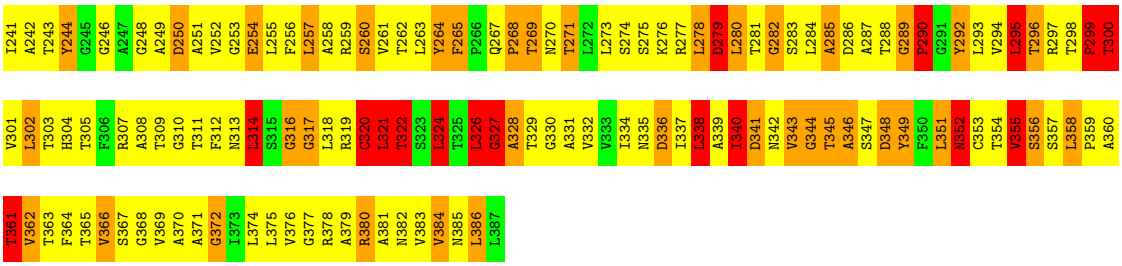
Chain A:



• Molecule 1: TOMATO BUSHY STUNT VIRUS

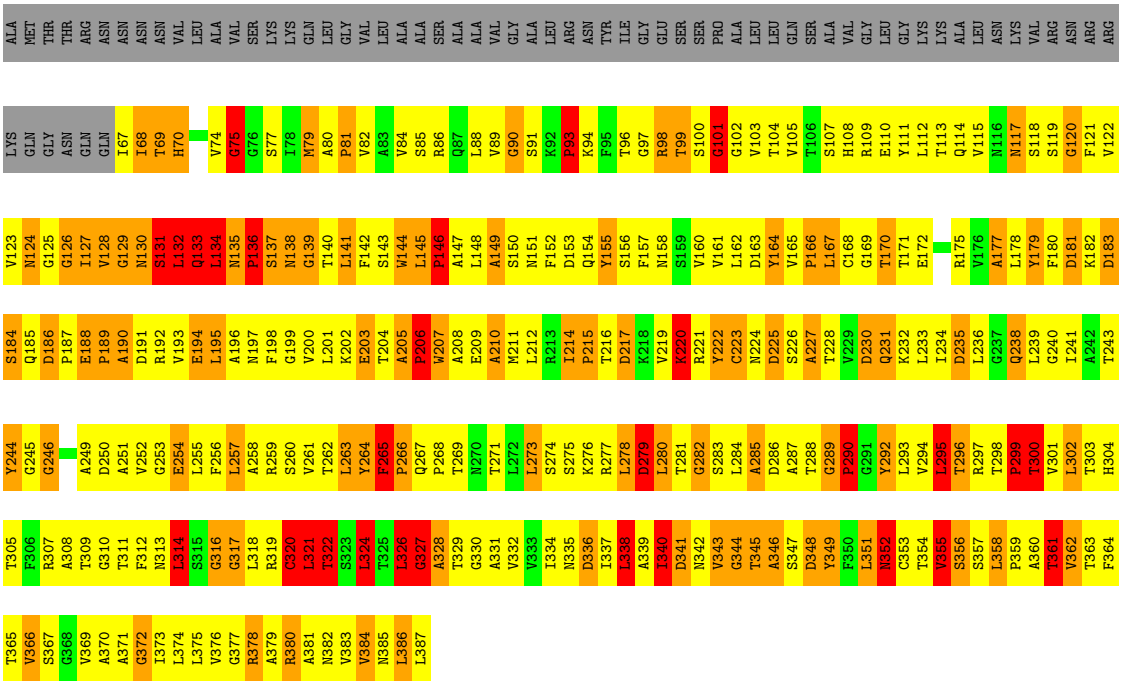
Chain B:





• Molecule 1: TOMATO BUSHY STUNT VIRUS

Chain C: 8% 43% 24% 7% 17%



4 Data and refinement statistics (i)

Xtriage (Phenix) and EDS were not executed - this section is therefore incomplete.

Property	Value	Source
Space group	I 2 3	Depositor
Cell constants a, b, c, α, β, γ	383.20Å 383.20Å 383.20Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	(Not available) – 2.90	Depositor
% Data completeness (in resolution range)	(Not available) ((Not available)-2.90)	Depositor
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
Refinement program	unknown	Depositor
R, R_{free}	(Not available) , (Not available)	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	6648	wwPDB-VP
Average B, all atoms (Å^2)	20.0	wwPDB-VP

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	1.53	29/2165 (1.3%)	2.19	119/2956 (4.0%)
1	B	1.52	29/2171 (1.3%)	2.16	123/2964 (4.1%)
1	C	1.47	32/2409 (1.3%)	2.12	125/3286 (3.8%)
All	All	1.50	90/6745 (1.3%)	2.16	367/9206 (4.0%)

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	2	35
1	B	3	28
1	C	3	32
All	All	8	95

The worst 5 of 90 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	138	ASN	N-CA	-24.59	0.97	1.46
1	C	138	ASN	N-CA	-24.56	0.97	1.46
1	A	138	ASN	N-CA	-24.55	0.97	1.46
1	B	137	SER	N-CA	-17.72	1.10	1.46
1	A	137	SER	N-CA	-17.70	1.10	1.46

The worst 5 of 367 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	267	GLN	C-N-CD	-22.78	70.48	120.60
1	B	165	VAL	C-N-CD	-19.72	77.22	120.60

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	C	145	LEU	C-N-CD	-18.03	80.94	120.60
1	A	135	ASN	CA-C-O	-15.93	86.65	120.10
1	C	135	ASN	CA-C-O	-15.91	86.70	120.10

5 of 8 chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
1	A	146	PRO	CA
1	A	189	PRO	CA
1	B	146	PRO	CA
1	B	166	PRO	CA
1	B	189	PRO	CA

5 of 95 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	119	SER	Peptide
1	A	123	VAL	Mainchain
1	A	124	ASN	Peptide
1	A	126	GLY	Mainchain,Peptide
1	A	127	ILE	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	2136	0	2111	647	0
1	B	2130	0	2111	596	2
1	C	2376	0	2374	695	19
2	A	3	0	0	0	0
2	B	1	0	0	0	0
2	C	2	0	0	0	0
All	All	6648	0	6596	1878	19

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 141.

The worst 5 of 1878 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:267:GLN:HG3	1:A:268:PRO:CD	1.18	1.59
1:B:122:VAL:N	1:B:122:VAL:CA	1.68	1.56
1:B:107:SER:N	1:B:107:SER:CA	1.70	1.53
1:C:223:CYS:CA	1:C:223:CYS:C	1.75	1.52
1:A:172:GLU:N	1:A:172:GLU:CA	1.70	1.51

The worst 5 of 19 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:C:340:ILE:CD1	1:C:348:ASP:CG[2_555]	1.00	1.20
1:C:340:ILE:CD1	1:C:348:ASP:OD2[2_555]	1.13	1.07
1:C:340:ILE:CG1	1:C:348:ASP:OD2[2_555]	1.24	0.96
1:C:340:ILE:CD1	1:C:348:ASP:OD1[2_555]	1.60	0.60
1:C:373:ILE:CD1	1:C:386:LEU:N[2_555]	1.67	0.53

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	288/387 (74%)	208 (72%)	49 (17%)	31 (11%)	0 1
1	B	287/387 (74%)	212 (74%)	43 (15%)	32 (11%)	0 1
1	C	322/387 (83%)	240 (74%)	45 (14%)	37 (12%)	0 1
All	All	897/1161 (77%)	660 (74%)	137 (15%)	100 (11%)	0 1

5 of 100 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	124	ASN
1	A	130	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	139	GLY
1	A	166	PRO
1	A	170	THR

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	232/309 (75%)	217 (94%)	15 (6%)	17	45
1	B	233/309 (75%)	217 (93%)	16 (7%)	15	41
1	C	258/309 (84%)	242 (94%)	16 (6%)	18	47
All	All	723/927 (78%)	676 (94%)	47 (6%)	17	45

5 of 47 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	338	LEU
1	C	189	PRO
1	B	351	LEU
1	C	81	PRO
1	C	211	MET

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 27 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	304	HIS
1	C	108	HIS
1	C	304	HIS
1	B	352	ASN
1	C	114	GLN

5.3.3 RNA (i)

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 monosaccharides in this entry.

5.6 Ligand geometry

Of 6 ligands modelled in this entry, 6 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

EDS was not executed - this section is therefore empty.

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

EDS was not executed - this section is therefore empty.

6.3 Carbohydrates

EDS was not executed - this section is therefore empty.

6.4 Ligands

EDS was not executed - this section is therefore empty.

6.5 Other polymers

EDS was not executed - this section is therefore empty.