wwPDB X-ray Structure Validation Summary Report (i)

Feb 1, 2016 – 12:52 AM GMT

PDB ID	:	2BTV
Title	:	ATOMIC MODEL FOR BLUETONGUE VIRUS (BTV) CORE
Authors	:	Grimes, J.M.; Burroughs, J.N.; Gouet, P.; Diprose, J.M.; Malby, R.; Zientras, S.; Mertens, P.P.C.; Stuart, D.I.
Deposited on	:	1998-09-05
Resolution	:	3.50 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 (i)) were used in the production of this report:

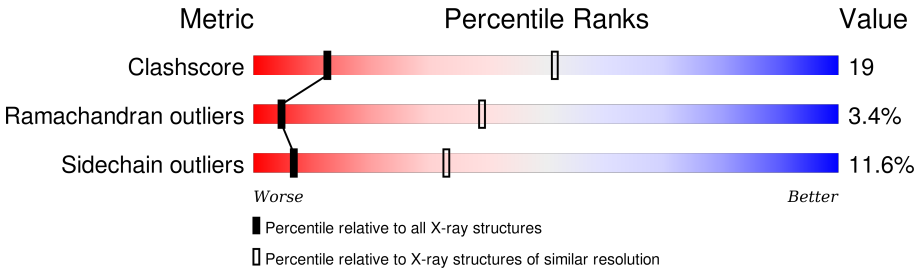
MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	1.9-1692
EDS	:	FAILED
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation 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 3.50 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	1157 (3.60-3.40)
Ramachandran outliers	100387	1120 (3.60-3.40)
Sidechain outliers	100360	1121 (3.60-3.40)

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.

Note EDS failed to run properly.

Mol	Chain	Length	Quality of chain
1	A	901	37% 42% 13% • 6%
1	B	901	45% 41% 12% • •
2	C	349	69% 29% •
2	D	349	74% 23% •
2	E	349	67% 28% 5%
2	F	349	64% 30% 5%
2	G	349	66% 29% •

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	H	349	69% 27%
2	I	349	64% 31% 5%
2	J	349	66% 30%
2	P	349	69% 27%
2	Q	349	65% 30% 5%
2	R	349	64% 30% 5%
2	S	349	61% 35% 5%
2	T	349	68% 27% 5%

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 49061 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 PROTEIN (VP3 CORE PROTEIN).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	845	Total	C	N	O	S	0	0	0
			6824	4357	1181	1248	38			
1	B	885	Total	C	N	O	S	0	0	0
			7150	4561	1238	1311	40			

There are 10 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	77	GLU	ALA	CONFLICT	UNP P56582
A	213	ILE	PHE	CONFLICT	UNP P56582
A	220	LEU	PHE	CONFLICT	UNP P56582
A	772	VAL	ILE	CONFLICT	UNP P56582
A	773	ARG	GLY	CONFLICT	UNP P56582
B	77	GLU	ALA	CONFLICT	UNP P56582
B	213	ILE	PHE	CONFLICT	UNP P56582
B	220	LEU	PHE	CONFLICT	UNP P56582
B	772	VAL	ILE	CONFLICT	UNP P56582
B	773	ARG	GLY	CONFLICT	UNP P56582

- Molecule 2 is a protein called PROTEIN (VP7 CORE PROTEIN).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	P	349	Total	C	N	O	S	0	0	0
			2699	1706	478	492	23			
2	C	349	Total	C	N	O	S	0	0	0
			2699	1706	478	492	23			
2	D	349	Total	C	N	O	S	0	0	0
			2699	1706	478	492	23			
2	Q	349	Total	C	N	O	S	0	0	0
			2699	1706	478	492	23			
2	E	349	Total	C	N	O	S	0	0	0
			2699	1706	478	492	23			

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	F	349	Total	C	N	O	S	0	0	0
			2699	1706	478	492	23			
2	R	349	Total	C	N	O	S	0	0	0
			2699	1706	478	492	23			
2	G	349	Total	C	N	O	S	0	0	0
			2699	1706	478	492	23			
2	H	349	Total	C	N	O	S	0	0	0
			2699	1706	478	492	23			
2	S	349	Total	C	N	O	S	0	0	0
			2699	1706	478	492	23			
2	I	349	Total	C	N	O	S	0	0	0
			2699	1706	478	492	23			
2	J	349	Total	C	N	O	S	0	0	0
			2699	1706	478	492	23			
2	T	349	Total	C	N	O	S	0	0	0
			2699	1706	478	492	23			

There are 13 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
P	278	TRP	GLY	CONFLICT	UNP P18259
C	278	TRP	GLY	CONFLICT	UNP P18259
D	278	TRP	GLY	CONFLICT	UNP P18259
Q	278	TRP	GLY	CONFLICT	UNP P18259
E	278	TRP	GLY	CONFLICT	UNP P18259
F	278	TRP	GLY	CONFLICT	UNP P18259
R	278	TRP	GLY	CONFLICT	UNP P18259
G	278	TRP	GLY	CONFLICT	UNP P18259
H	278	TRP	GLY	CONFLICT	UNP P18259
S	278	TRP	GLY	CONFLICT	UNP P18259
I	278	TRP	GLY	CONFLICT	UNP P18259
J	278	TRP	GLY	CONFLICT	UNP P18259
T	278	TRP	GLY	CONFLICT	UNP P18259

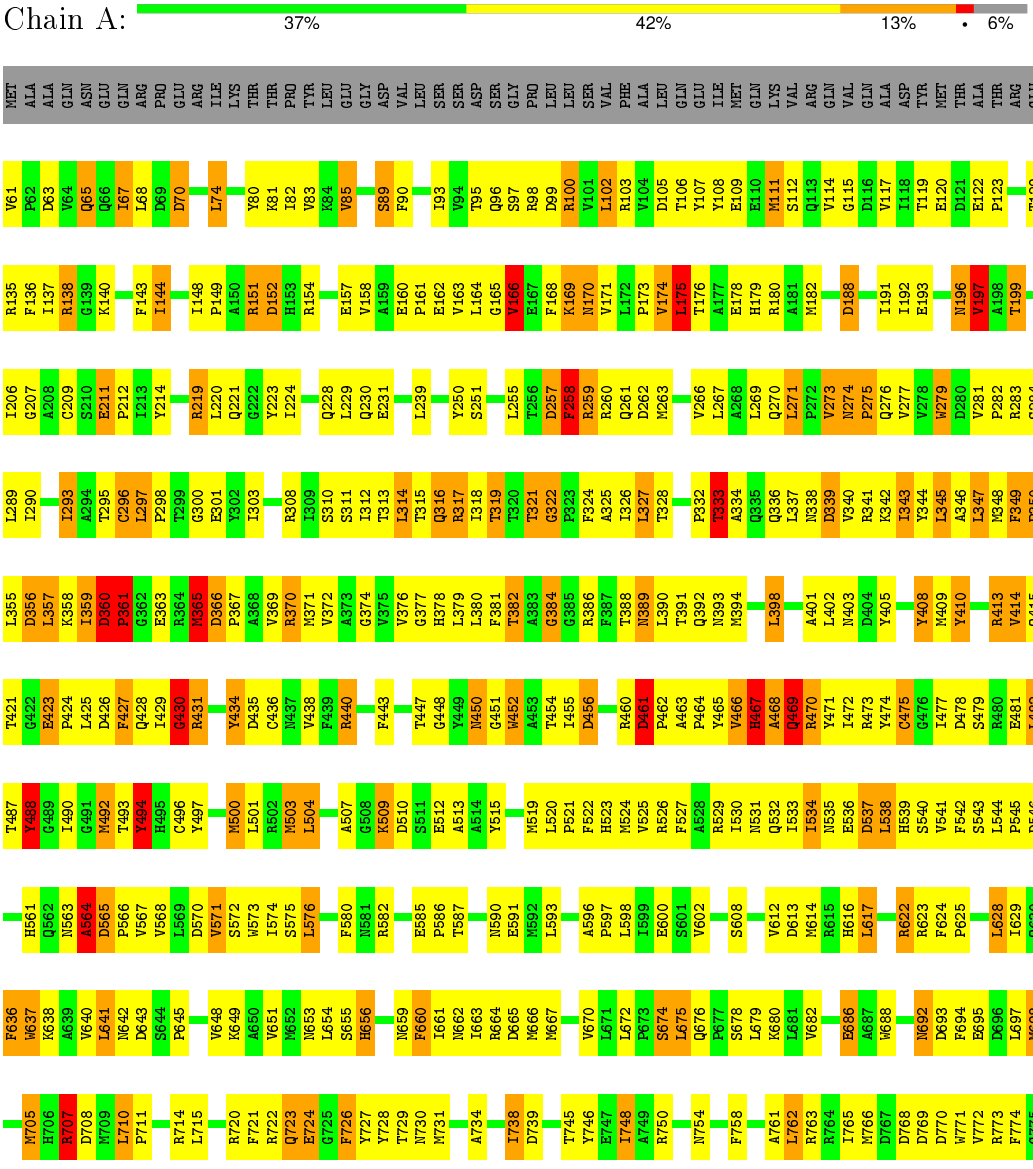
3 Residue-property plots (i)

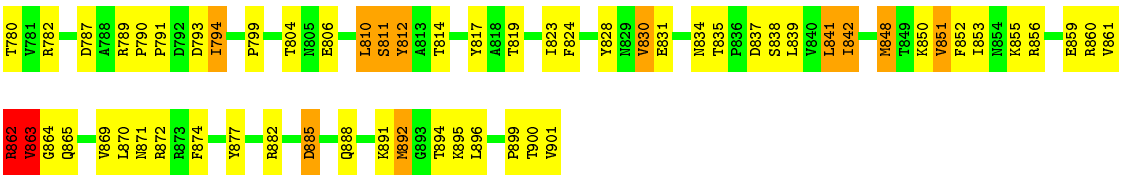
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according 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.

Note EDS failed to run properly.

- Molecule 1: PROTEIN (VP3 CORE PROTEIN)

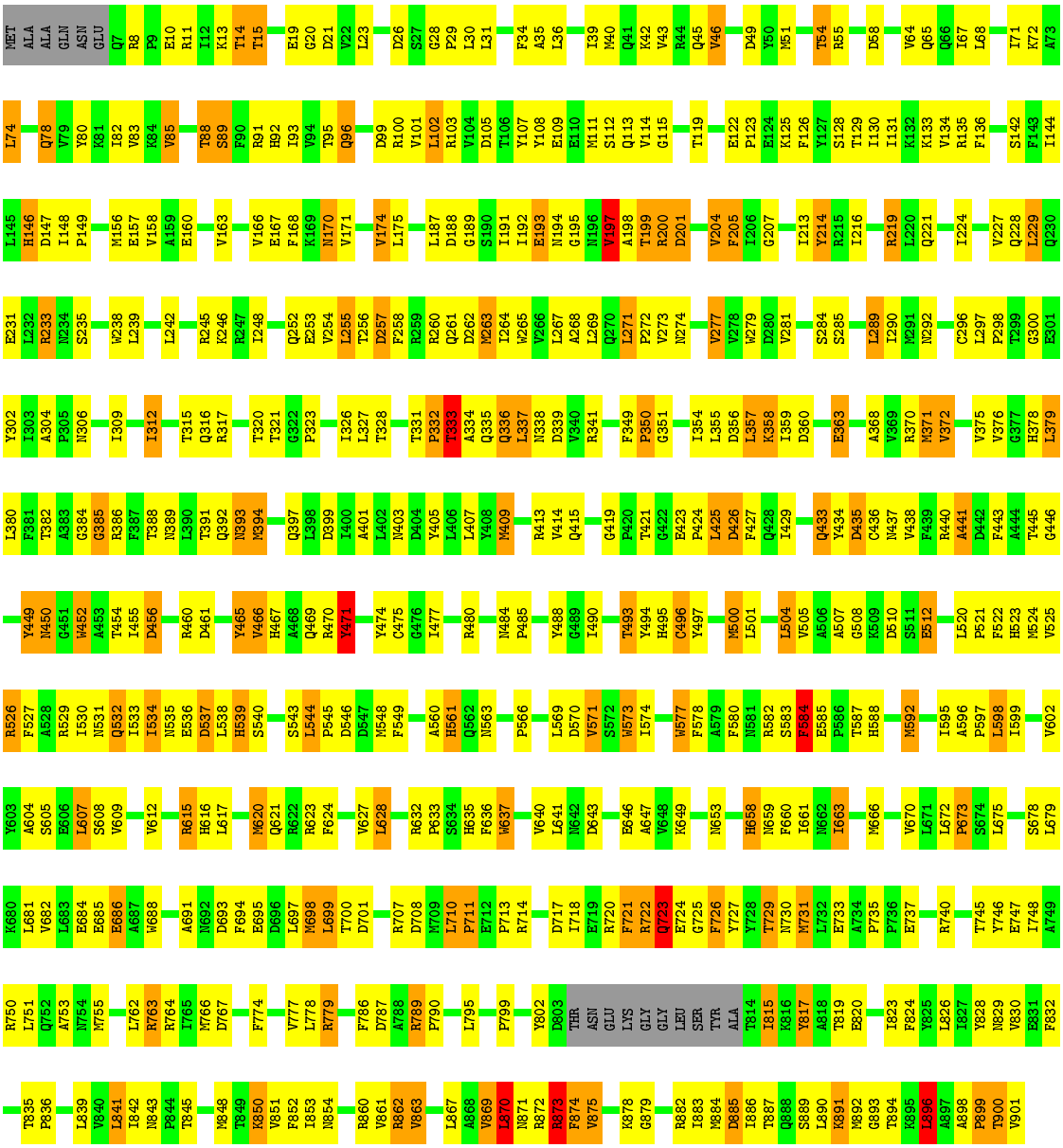
Chain A:





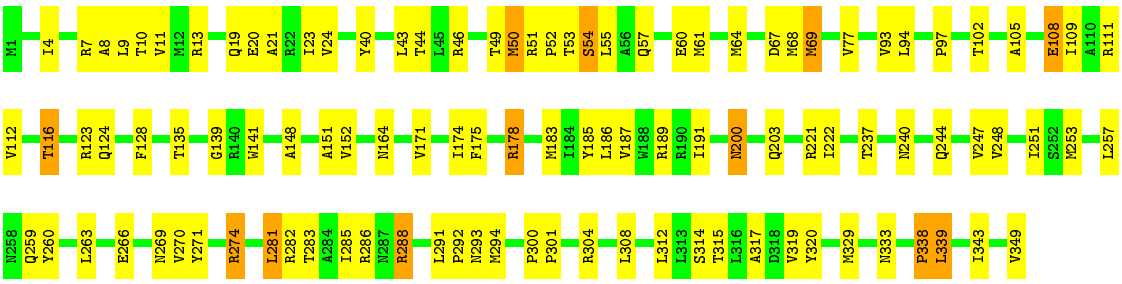
• Molecule 1: PROTEIN (VP3 CORE PROTEIN)

Chain B:



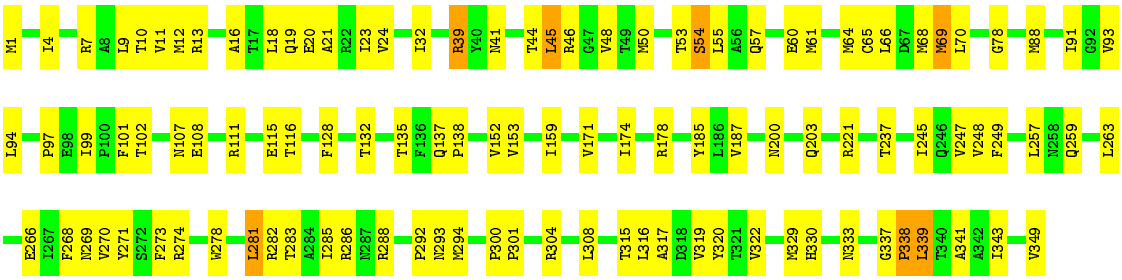
• Molecule 2: PROTEIN (VP7 CORE PROTEIN)

Chain P:



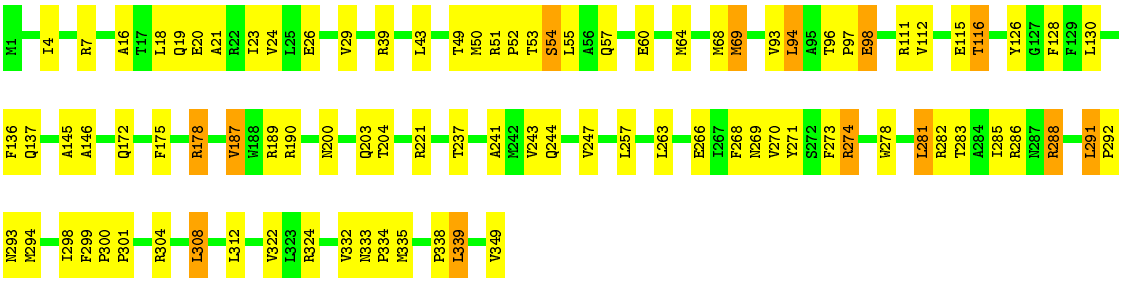
• Molecule 2: PROTEIN (VP7 CORE PROTEIN)

Chain C:



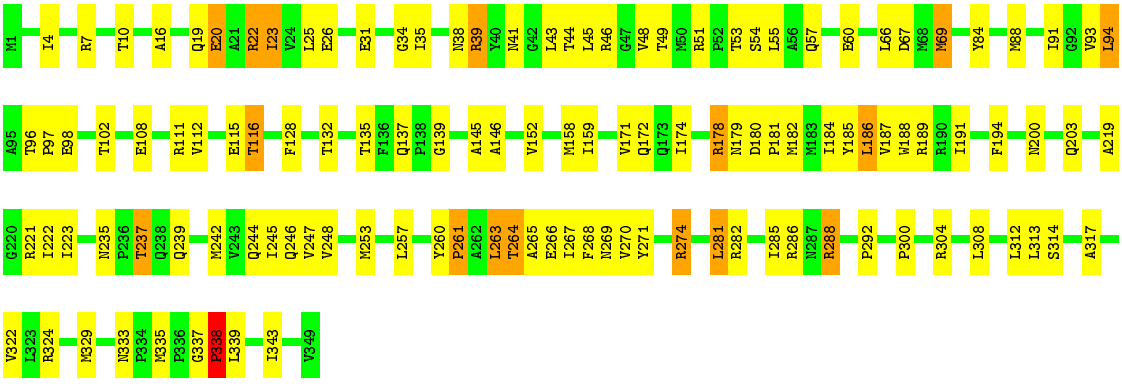
• Molecule 2: PROTEIN (VP7 CORE PROTEIN)

Chain D:



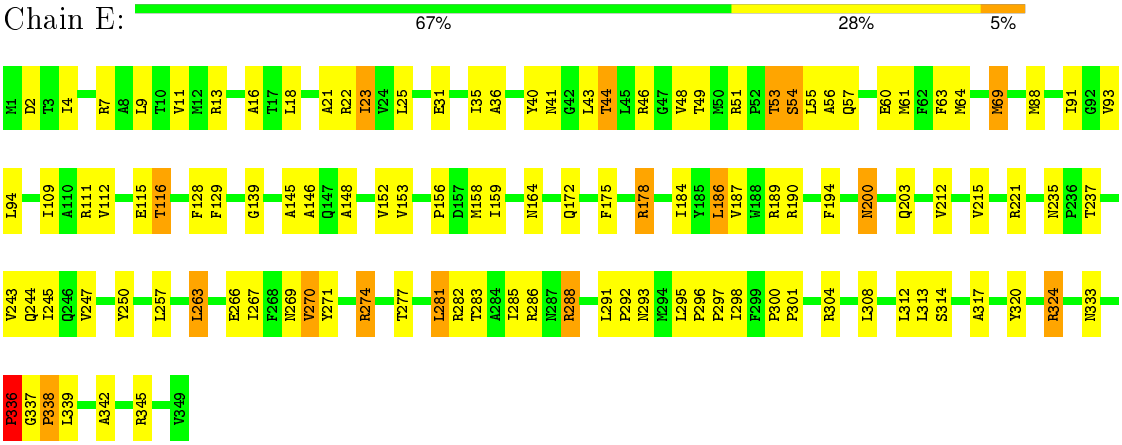
• Molecule 2: PROTEIN (VP7 CORE PROTEIN)

Chain Q:



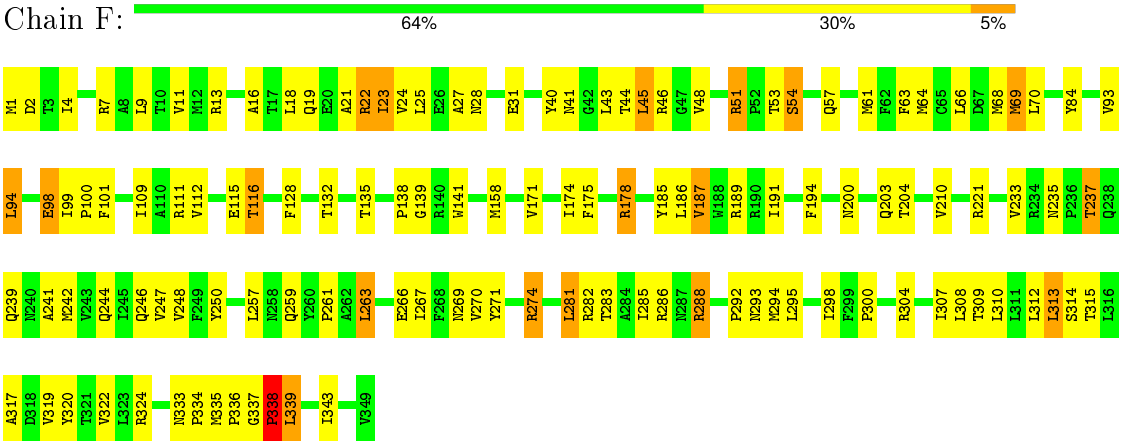
• Molecule 2: PROTEIN (VP7 CORE PROTEIN)

Chain E:



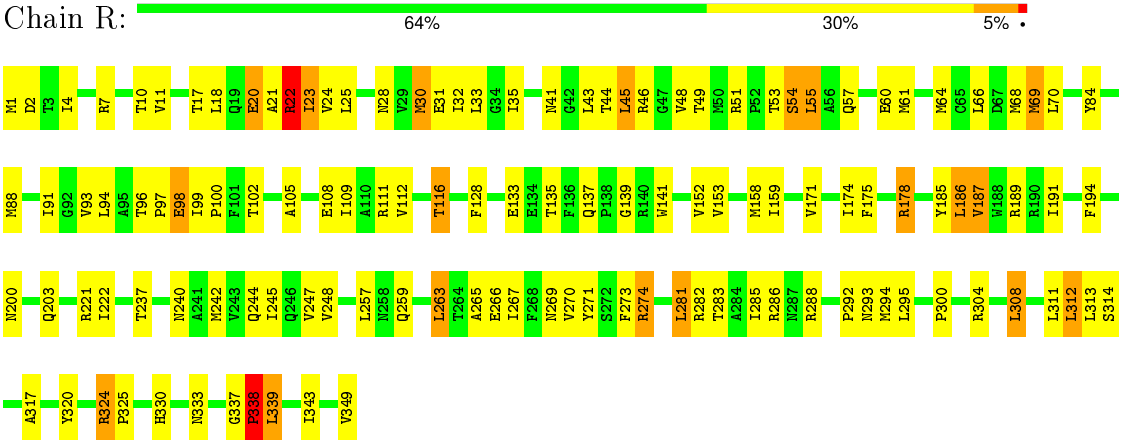
• Molecule 2: PROTEIN (VP7 CORE PROTEIN)

Chain F:



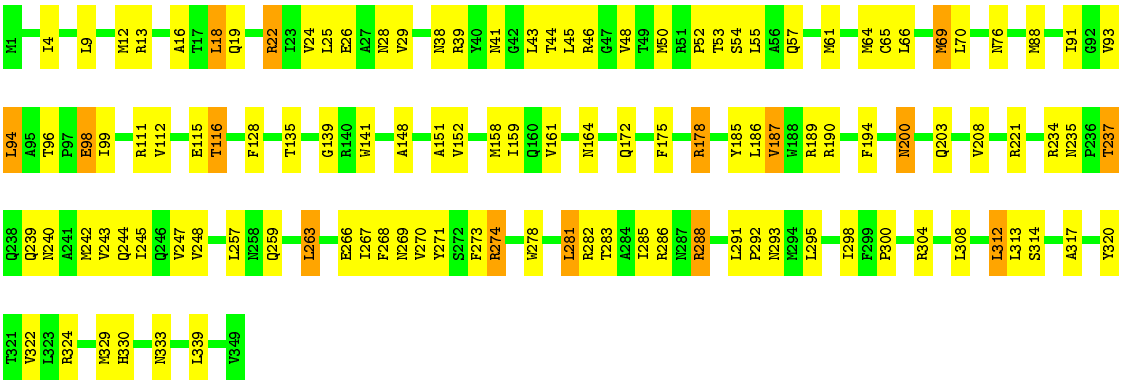
• Molecule 2: PROTEIN (VP7 CORE PROTEIN)

Chain R:



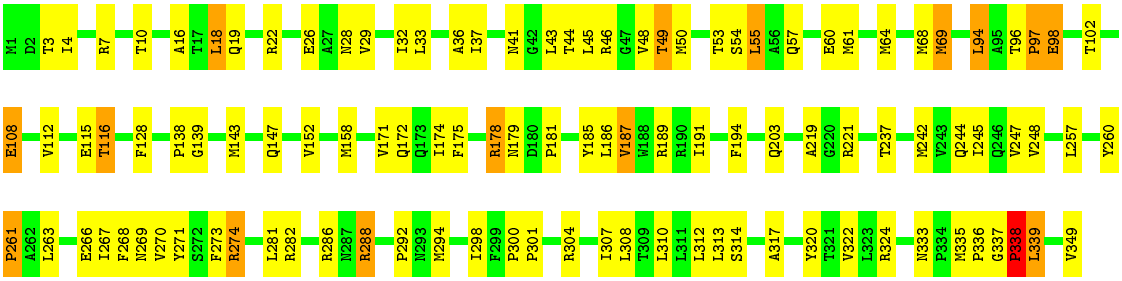
• Molecule 2: PROTEIN (VP7 CORE PROTEIN)

Chain G:



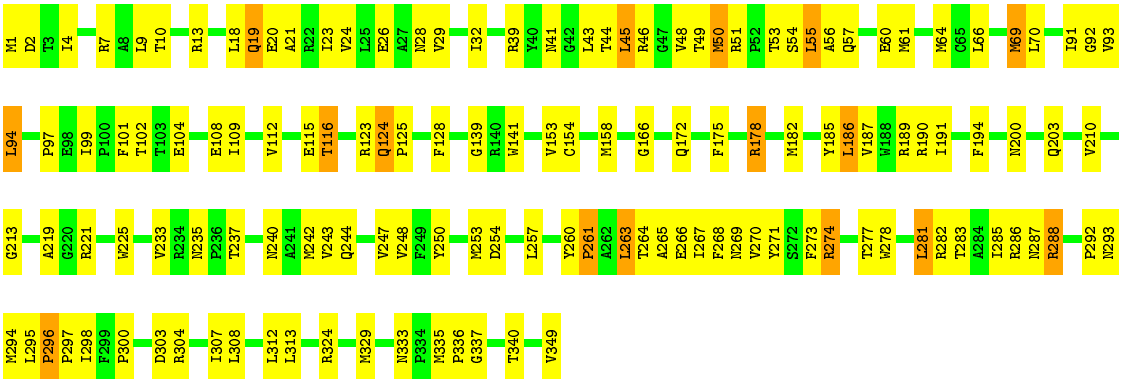
• Molecule 2: PROTEIN (VP7 CORE PROTEIN)

Chain H:



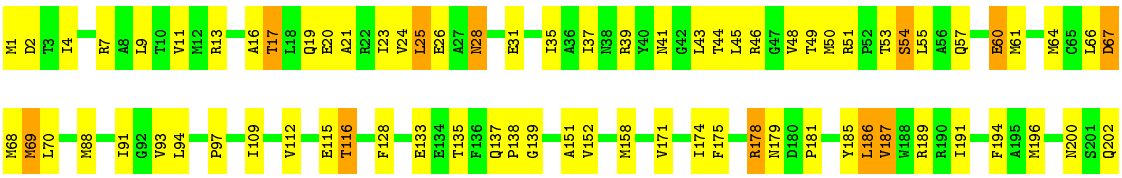
• Molecule 2: PROTEIN (VP7 CORE PROTEIN)

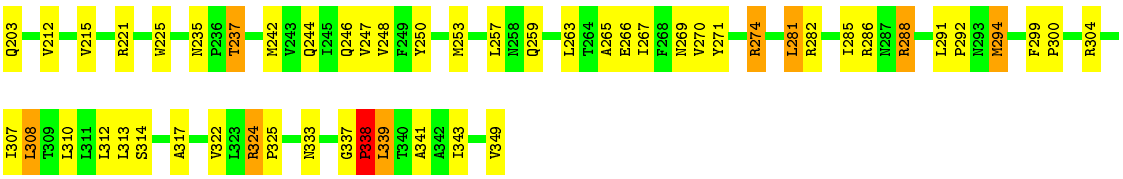
Chain S:



• Molecule 2: PROTEIN (VP7 CORE PROTEIN)

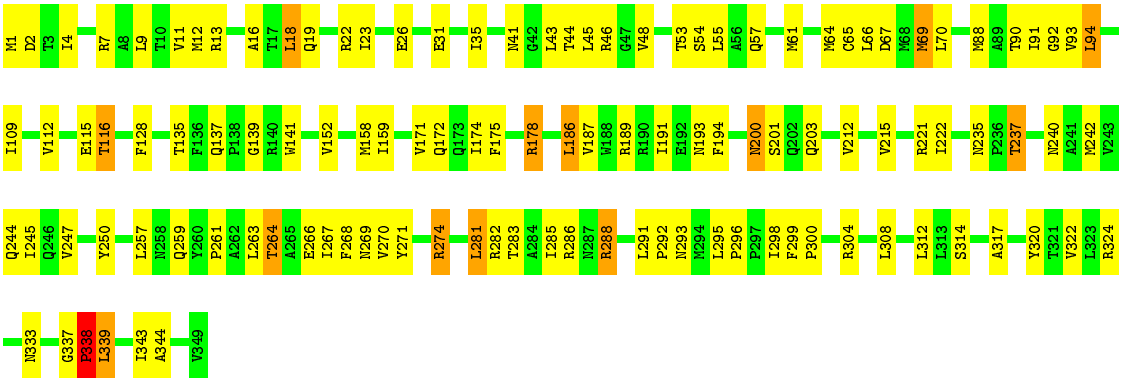
Chain I:





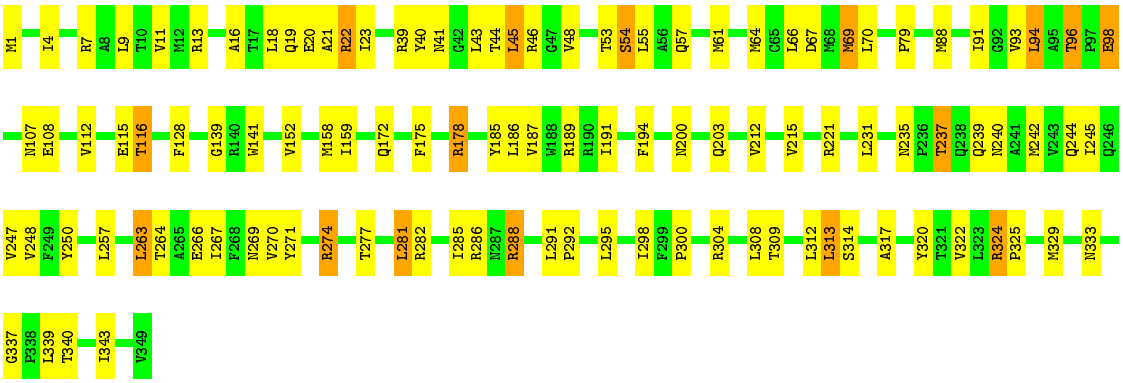
- Molecule 2: PROTEIN (VP7 CORE PROTEIN)

Chain J:



- Molecule 2: PROTEIN (VP7 CORE PROTEIN)

Chain T:



4 Data and refinement statistics (i)

EDS failed to run properly - this section will therefore be incomplete.

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, α, β, γ	795.60Å 821.80Å 753.30Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	100.00 – 3.50	Depositor
% Data completeness (in resolution range)	54.0 (100.00-3.50)	Depositor
R_{merge}	0.23	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	4.83 (at 3.49Å)	Xtriage
Refinement program	X-PLOR 3.1	Depositor
R, R_{free}	0.266 , (Not available)	Depositor
Wilson B-factor (Å^2)	48.5	Xtriage
Anisotropy	0.015	Xtriage
Estimated twinning fraction	0.078 for k,h,-l	Xtriage
L-test for twinning[2]	$< \|L\| > = 0.38$, $< L^2 > = 0.21$	Xtriage
Outliers	0 of 3524397 reflections	Xtriage
Total number of atoms	49061	wwPDB-VP
Average B, all atoms (Å^2)	58.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 1.12% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] 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

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	0.46	0/6978	0.80	7/9479 (0.1%)
1	B	0.46	0/7308	0.78	3/9925 (0.0%)
2	C	0.33	0/2757	0.59	0/3756
2	D	0.32	0/2757	0.58	0/3756
2	E	0.35	0/2757	0.63	0/3756
2	F	0.36	0/2757	0.64	0/3756
2	G	0.38	0/2757	0.64	0/3756
2	H	0.37	0/2757	0.64	0/3756
2	I	0.37	0/2757	0.64	0/3756
2	J	0.36	0/2757	0.64	0/3756
2	P	0.33	0/2757	0.59	0/3756
2	Q	0.37	0/2757	0.67	0/3756
2	R	0.37	0/2757	0.65	1/3756 (0.0%)
2	S	0.39	0/2757	0.66	0/3756
2	T	0.37	0/2757	0.64	0/3756
All	All	0.39	0/50127	0.68	11/68232 (0.0%)

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	0	1
1	B	0	1
All	All	0	2

There are no bond length outliers.

The worst 5 of 11 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	467	HIS	N-CA-C	-6.11	94.52	111.00

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	564	ALA	N-CA-C	-5.88	95.12	111.00
1	A	271	LEU	N-CA-C	5.82	126.71	111.00
1	A	360	ASP	N-CA-C	5.41	125.61	111.00
2	R	20	GLU	N-CA-C	-5.38	96.48	111.00

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	494	TYR	Sidechain
1	B	471	TYR	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	6824	0	6804	484	0
1	B	7150	0	7137	456	0
2	C	2699	0	2697	66	0
2	D	2699	0	2697	58	0
2	E	2699	0	2697	82	0
2	F	2699	0	2697	89	0
2	G	2699	0	2697	90	0
2	H	2699	0	2697	77	0
2	I	2699	0	2697	104	0
2	J	2699	0	2697	89	0
2	P	2699	0	2697	66	0
2	Q	2699	0	2697	95	0
2	R	2699	0	2697	94	0
2	S	2699	0	2697	102	0
2	T	2699	0	2697	71	0
All	All	49061	0	49002	1904	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 19.

The worst 5 of 1904 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)
2:I:97:PRO:HD3	2:J:22:ARG:HH21	1.27	0.99
1:B:358:LYS:HB2	1:B:570:ASP:HB3	1.49	0.94
2:S:282:ARG:HD3	2:S:286:ARG:NH1	1.83	0.94
2:T:257:LEU:HD12	2:T:300:PRO:HB3	1.47	0.94
1:A:219:ARG:HG3	1:A:686:GLU:HG3	1.46	0.94

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	843/901 (94%)	674 (80%)	109 (13%)	60 (7%)	1 17
1	B	881/901 (98%)	715 (81%)	110 (12%)	56 (6%)	2 20
2	C	347/349 (99%)	313 (90%)	29 (8%)	5 (1%)	14 58
2	D	347/349 (99%)	316 (91%)	25 (7%)	6 (2%)	11 54
2	E	347/349 (99%)	296 (85%)	40 (12%)	11 (3%)	5 40
2	F	347/349 (99%)	305 (88%)	34 (10%)	8 (2%)	8 48
2	G	347/349 (99%)	303 (87%)	39 (11%)	5 (1%)	14 58
2	H	347/349 (99%)	302 (87%)	37 (11%)	8 (2%)	8 48
2	I	347/349 (99%)	300 (86%)	40 (12%)	7 (2%)	9 51
2	J	347/349 (99%)	296 (85%)	45 (13%)	6 (2%)	11 54
2	P	347/349 (99%)	312 (90%)	28 (8%)	7 (2%)	9 51
2	Q	347/349 (99%)	300 (86%)	39 (11%)	8 (2%)	8 48
2	R	347/349 (99%)	299 (86%)	39 (11%)	9 (3%)	7 45
2	S	347/349 (99%)	296 (85%)	40 (12%)	11 (3%)	5 40
2	T	347/349 (99%)	306 (88%)	35 (10%)	6 (2%)	11 54
All	All	6235/6339 (98%)	5333 (86%)	689 (11%)	213 (3%)	5 39

5 of 213 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	106	THR
1	A	166	VAL
1	A	169	LYS
1	A	175	LEU
1	A	197	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	744/793 (94%)	592 (80%)	152 (20%)	1 8
1	B	782/793 (99%)	644 (82%)	138 (18%)	2 13
2	C	284/284 (100%)	263 (93%)	21 (7%)	17 56
2	D	284/284 (100%)	262 (92%)	22 (8%)	16 54
2	E	284/284 (100%)	259 (91%)	25 (9%)	12 48
2	F	284/284 (100%)	260 (92%)	24 (8%)	13 49
2	G	284/284 (100%)	258 (91%)	26 (9%)	11 45
2	H	284/284 (100%)	258 (91%)	26 (9%)	11 45
2	I	284/284 (100%)	260 (92%)	24 (8%)	13 49
2	J	284/284 (100%)	261 (92%)	23 (8%)	15 52
2	P	284/284 (100%)	262 (92%)	22 (8%)	16 54
2	Q	284/284 (100%)	258 (91%)	26 (9%)	11 45
2	R	284/284 (100%)	259 (91%)	25 (9%)	12 48
2	S	284/284 (100%)	260 (92%)	24 (8%)	13 49
2	T	284/284 (100%)	256 (90%)	28 (10%)	10 41
All	All	5218/5278 (99%)	4612 (88%)	606 (12%)	7 33

5 of 606 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	817	TYR
2	D	98	GLU
2	J	115	GLU
1	B	873	ARG
2	P	288	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 147 such sidechains are listed below:

Mol	Chain	Res	Type
2	Q	193	ASN
2	F	200	ASN
2	J	200	ASN
2	Q	244	GLN
2	E	244	GLN

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

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

EDS failed to run properly - this section will therefore be empty.

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

EDS failed to run properly - this section will therefore be empty.

6.3 Carbohydrates

EDS failed to run properly - this section will therefore be empty.

6.4 Ligands

EDS failed to run properly - this section will therefore be empty.

6.5 Other polymers

EDS failed to run properly - this section will therefore be empty.