



wwPDB X-ray Structure Validation Summary Report ⓘ

Jul 25, 2017 – 11:45 PM EDT

PDB ID : 2R5H
Title : Pentamer structure of Major Capsid Protein L1 of Human Papilloma Virus type 16
Authors : Bishop, B.; Dasgupta, J.; Chen, X.S.
Deposited on : unknown
Resolution : 3.70 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 ⓘ symbol.

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20029824
Percentile statistics	:	20161228.v01 (using entries in the PDB archive December 28th 2016)
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)	:	rb-20029824

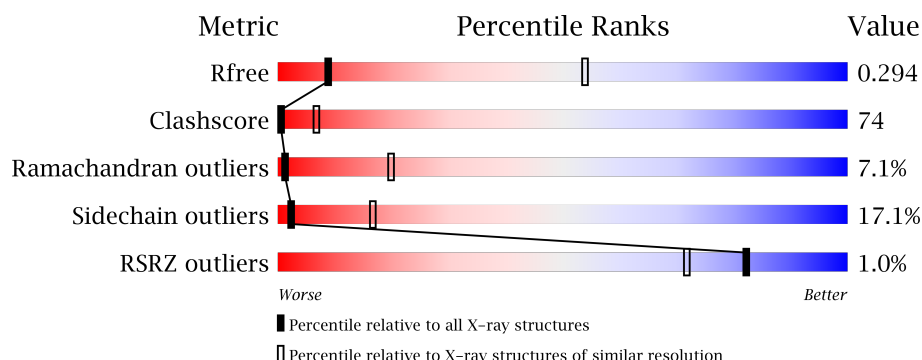
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.70 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.





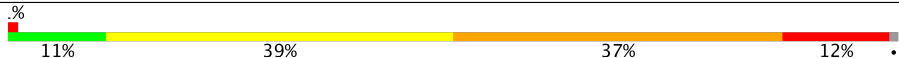
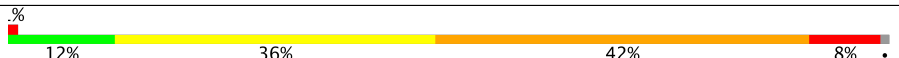
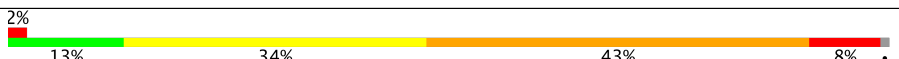
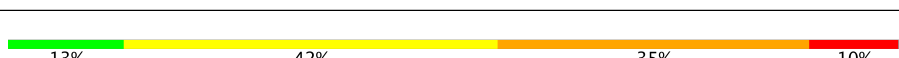
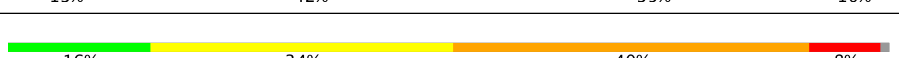



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	100719	1234 (3.90-3.50)
Clashscore	112137	1377 (3.90-3.50)
Ramachandran outliers	110173	1323 (3.90-3.50)
Sidechain outliers	110143	1320 (3.90-3.50)
RSRZ outliers	101464	1262 (3.90-3.50)

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 $\leq 5\%$. 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	424	<div> <div>15%</div> <div>36%</div> <div>37%</div> <div>11%</div> <div>.</div> </div>
1	B	424	<div> <div>12%</div> <div>36%</div> <div>42%</div> <div>9%</div> <div>.</div> </div>
1	C	424	<div> <div>11%</div> <div>42%</div> <div>36%</div> <div>10%</div> <div>.</div> </div>
1	D	424	<div> <div>9%</div> <div>39%</div> <div>41%</div> <div>9%</div> <div>.</div> </div>
1	E	424	<div> <div>12%</div> <div>40%</div> <div>37%</div> <div>10%</div> <div>.</div> </div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	424	
1	G	424	
1	H	424	
1	I	424	
1	J	424	
1	K	424	
1	L	424	
1	M	424	
1	N	424	
1	O	424	

2 Entry composition [i](#)

There is only 1 type of molecule in this entry. The entry contains 49650 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 Late major capsid protein L1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	419	Total	C	N	O	S	0	0	0
			3310	2108	554	628	20			
1	B	419	Total	C	N	O	S	0	0	0
			3310	2108	554	628	20			
1	C	419	Total	C	N	O	S	0	0	0
			3310	2108	554	628	20			
1	D	419	Total	C	N	O	S	0	0	0
			3310	2108	554	628	20			
1	E	419	Total	C	N	O	S	0	0	0
			3310	2108	554	628	20			
1	F	419	Total	C	N	O	S	0	0	0
			3310	2108	554	628	20			
1	G	419	Total	C	N	O	S	0	0	0
			3310	2108	554	628	20			
1	H	419	Total	C	N	O	S	0	0	0
			3310	2108	554	628	20			
1	I	419	Total	C	N	O	S	0	0	0
			3310	2108	554	628	20			
1	J	419	Total	C	N	O	S	0	0	0
			3310	2108	554	628	20			
1	K	419	Total	C	N	O	S	0	0	0
			3310	2108	554	628	20			
1	L	419	Total	C	N	O	S	0	0	0
			3310	2108	554	628	20			
1	M	419	Total	C	N	O	S	0	0	0
			3310	2108	554	628	20			
1	N	419	Total	C	N	O	S	0	0	0
			3310	2108	554	628	20			
1	O	419	Total	C	N	O	S	0	0	0
			3310	2108	554	628	20			

There are 135 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	20	ALA	-	EXPRESSION TAG	UNP Q81007
A	177	GLN	ASN	ENGINEERED	UNP Q81007
A	181	GLN	ASN	ENGINEERED	UNP Q81007
A	404	GLY	-	LINKER	UNP Q81007
A	405	GLY	-	LINKER	UNP Q81007
A	406	SER	-	LINKER	UNP Q81007
A	407	GLY	-	LINKER	UNP Q81007
A	408	GLY	-	LINKER	UNP Q81007
A	472	LEU	-	EXPRESSION TAG	UNP Q81007
B	20	ALA	-	EXPRESSION TAG	UNP Q81007
B	177	GLN	ASN	ENGINEERED	UNP Q81007
B	181	GLN	ASN	ENGINEERED	UNP Q81007
B	404	GLY	-	LINKER	UNP Q81007
B	405	GLY	-	LINKER	UNP Q81007
B	406	SER	-	LINKER	UNP Q81007
B	407	GLY	-	LINKER	UNP Q81007
B	408	GLY	-	LINKER	UNP Q81007
B	472	LEU	-	EXPRESSION TAG	UNP Q81007
C	20	ALA	-	EXPRESSION TAG	UNP Q81007
C	177	GLN	ASN	ENGINEERED	UNP Q81007
C	181	GLN	ASN	ENGINEERED	UNP Q81007
C	404	GLY	-	LINKER	UNP Q81007
C	405	GLY	-	LINKER	UNP Q81007
C	406	SER	-	LINKER	UNP Q81007
C	407	GLY	-	LINKER	UNP Q81007
C	408	GLY	-	LINKER	UNP Q81007
C	472	LEU	-	EXPRESSION TAG	UNP Q81007
D	20	ALA	-	EXPRESSION TAG	UNP Q81007
D	177	GLN	ASN	ENGINEERED	UNP Q81007
D	181	GLN	ASN	ENGINEERED	UNP Q81007
D	404	GLY	-	LINKER	UNP Q81007
D	405	GLY	-	LINKER	UNP Q81007
D	406	SER	-	LINKER	UNP Q81007
D	407	GLY	-	LINKER	UNP Q81007
D	408	GLY	-	LINKER	UNP Q81007
D	472	LEU	-	EXPRESSION TAG	UNP Q81007
E	20	ALA	-	EXPRESSION TAG	UNP Q81007
E	177	GLN	ASN	ENGINEERED	UNP Q81007
E	181	GLN	ASN	ENGINEERED	UNP Q81007
E	404	GLY	-	LINKER	UNP Q81007
E	405	GLY	-	LINKER	UNP Q81007
E	406	SER	-	LINKER	UNP Q81007
E	407	GLY	-	LINKER	UNP Q81007

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
E	408	GLY	-	LINKER	UNP Q81007
E	472	LEU	-	EXPRESSION TAG	UNP Q81007
F	20	ALA	-	EXPRESSION TAG	UNP Q81007
F	177	GLN	ASN	ENGINEERED	UNP Q81007
F	181	GLN	ASN	ENGINEERED	UNP Q81007
F	404	GLY	-	LINKER	UNP Q81007
F	405	GLY	-	LINKER	UNP Q81007
F	406	SER	-	LINKER	UNP Q81007
F	407	GLY	-	LINKER	UNP Q81007
F	408	GLY	-	LINKER	UNP Q81007
F	472	LEU	-	EXPRESSION TAG	UNP Q81007
G	20	ALA	-	EXPRESSION TAG	UNP Q81007
G	177	GLN	ASN	ENGINEERED	UNP Q81007
G	181	GLN	ASN	ENGINEERED	UNP Q81007
G	404	GLY	-	LINKER	UNP Q81007
G	405	GLY	-	LINKER	UNP Q81007
G	406	SER	-	LINKER	UNP Q81007
G	407	GLY	-	LINKER	UNP Q81007
G	408	GLY	-	LINKER	UNP Q81007
G	472	LEU	-	EXPRESSION TAG	UNP Q81007
H	20	ALA	-	EXPRESSION TAG	UNP Q81007
H	177	GLN	ASN	ENGINEERED	UNP Q81007
H	181	GLN	ASN	ENGINEERED	UNP Q81007
H	404	GLY	-	LINKER	UNP Q81007
H	405	GLY	-	LINKER	UNP Q81007
H	406	SER	-	LINKER	UNP Q81007
H	407	GLY	-	LINKER	UNP Q81007
H	408	GLY	-	LINKER	UNP Q81007
H	472	LEU	-	EXPRESSION TAG	UNP Q81007
I	20	ALA	-	EXPRESSION TAG	UNP Q81007
I	177	GLN	ASN	ENGINEERED	UNP Q81007
I	181	GLN	ASN	ENGINEERED	UNP Q81007
I	404	GLY	-	LINKER	UNP Q81007
I	405	GLY	-	LINKER	UNP Q81007
I	406	SER	-	LINKER	UNP Q81007
I	407	GLY	-	LINKER	UNP Q81007
I	408	GLY	-	LINKER	UNP Q81007
I	472	LEU	-	EXPRESSION TAG	UNP Q81007
J	20	ALA	-	EXPRESSION TAG	UNP Q81007
J	177	GLN	ASN	ENGINEERED	UNP Q81007
J	181	GLN	ASN	ENGINEERED	UNP Q81007
J	404	GLY	-	LINKER	UNP Q81007

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
J	405	GLY	-	LINKER	UNP Q81007
J	406	SER	-	LINKER	UNP Q81007
J	407	GLY	-	LINKER	UNP Q81007
J	408	GLY	-	LINKER	UNP Q81007
J	472	LEU	-	EXPRESSION TAG	UNP Q81007
K	20	ALA	-	EXPRESSION TAG	UNP Q81007
K	177	GLN	ASN	ENGINEERED	UNP Q81007
K	181	GLN	ASN	ENGINEERED	UNP Q81007
K	404	GLY	-	LINKER	UNP Q81007
K	405	GLY	-	LINKER	UNP Q81007
K	406	SER	-	LINKER	UNP Q81007
K	407	GLY	-	LINKER	UNP Q81007
K	408	GLY	-	LINKER	UNP Q81007
K	472	LEU	-	EXPRESSION TAG	UNP Q81007
L	20	ALA	-	EXPRESSION TAG	UNP Q81007
L	177	GLN	ASN	ENGINEERED	UNP Q81007
L	181	GLN	ASN	ENGINEERED	UNP Q81007
L	404	GLY	-	LINKER	UNP Q81007
L	405	GLY	-	LINKER	UNP Q81007
L	406	SER	-	LINKER	UNP Q81007
L	407	GLY	-	LINKER	UNP Q81007
L	408	GLY	-	LINKER	UNP Q81007
L	472	LEU	-	EXPRESSION TAG	UNP Q81007
M	20	ALA	-	EXPRESSION TAG	UNP Q81007
M	177	GLN	ASN	ENGINEERED	UNP Q81007
M	181	GLN	ASN	ENGINEERED	UNP Q81007
M	404	GLY	-	LINKER	UNP Q81007
M	405	GLY	-	LINKER	UNP Q81007
M	406	SER	-	LINKER	UNP Q81007
M	407	GLY	-	LINKER	UNP Q81007
M	408	GLY	-	LINKER	UNP Q81007
M	472	LEU	-	EXPRESSION TAG	UNP Q81007
N	20	ALA	-	EXPRESSION TAG	UNP Q81007
N	177	GLN	ASN	ENGINEERED	UNP Q81007
N	181	GLN	ASN	ENGINEERED	UNP Q81007
N	404	GLY	-	LINKER	UNP Q81007
N	405	GLY	-	LINKER	UNP Q81007
N	406	SER	-	LINKER	UNP Q81007
N	407	GLY	-	LINKER	UNP Q81007
N	408	GLY	-	LINKER	UNP Q81007
N	472	LEU	-	EXPRESSION TAG	UNP Q81007
O	20	ALA	-	EXPRESSION TAG	UNP Q81007

Continued on next page...

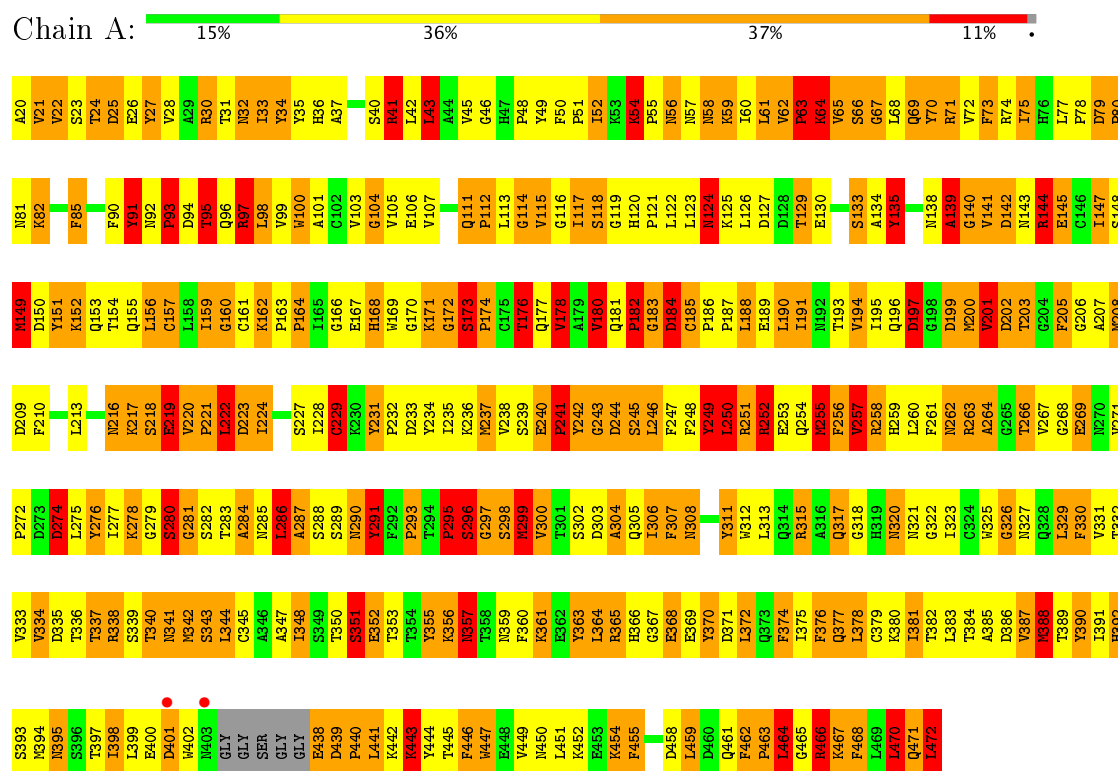
Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
O	177	GLN	ASN	ENGINEERED	UNP Q81007
O	181	GLN	ASN	ENGINEERED	UNP Q81007
O	404	GLY	-	LINKER	UNP Q81007
O	405	GLY	-	LINKER	UNP Q81007
O	406	SER	-	LINKER	UNP Q81007
O	407	GLY	-	LINKER	UNP Q81007
O	408	GLY	-	LINKER	UNP Q81007
O	472	LEU	-	EXPRESSION TAG	UNP Q81007

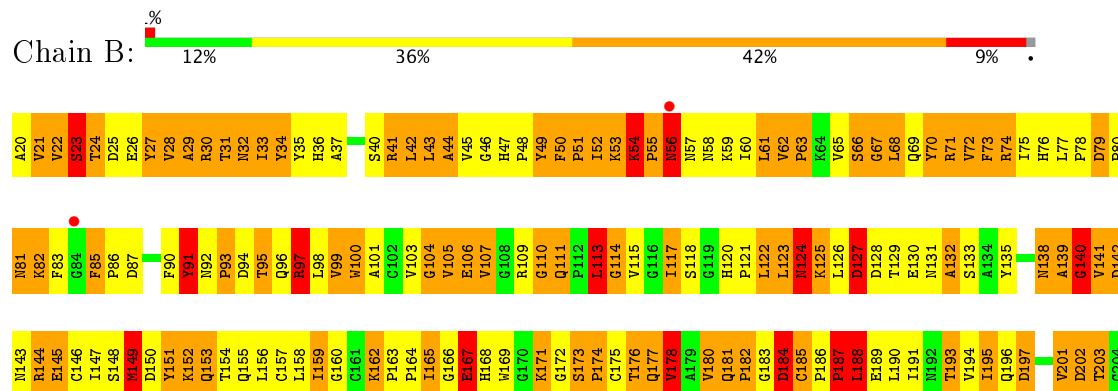
3 Residue-property plots

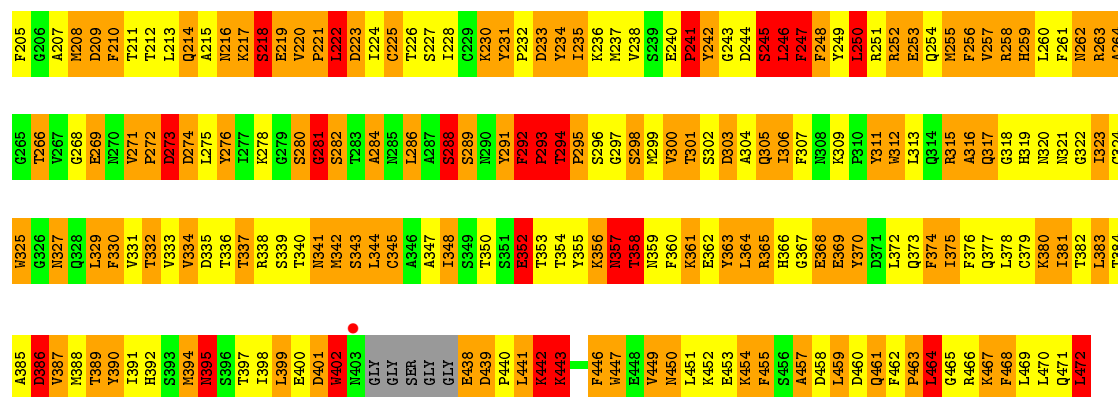
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are 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: Late major capsid protein L1

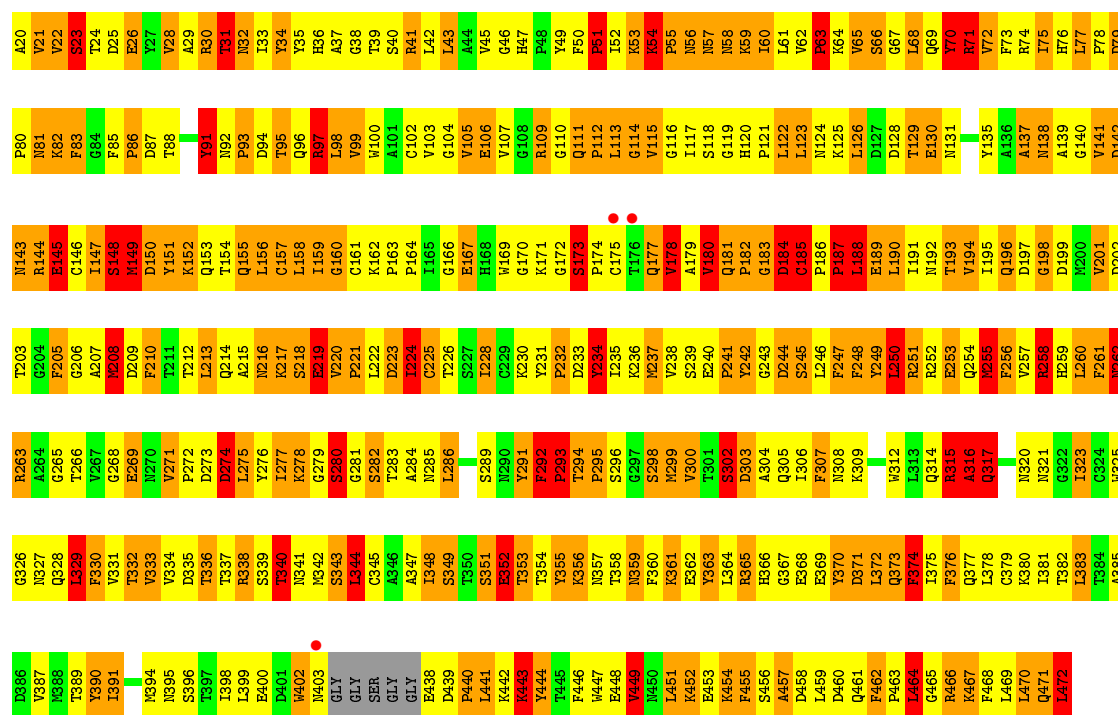


• Molecule 1: Late major capsid protein L1

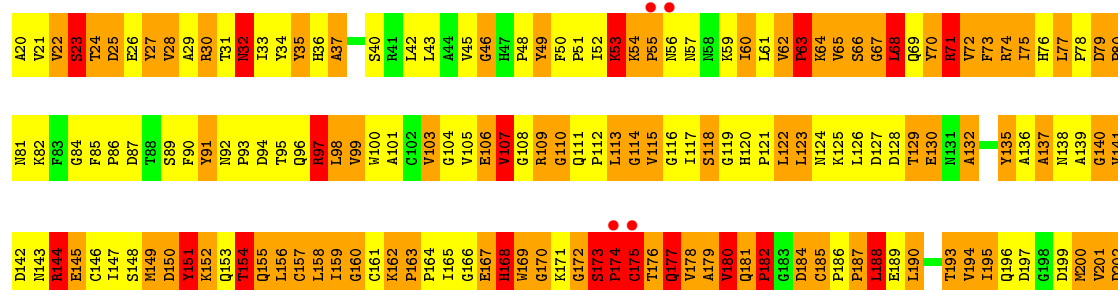


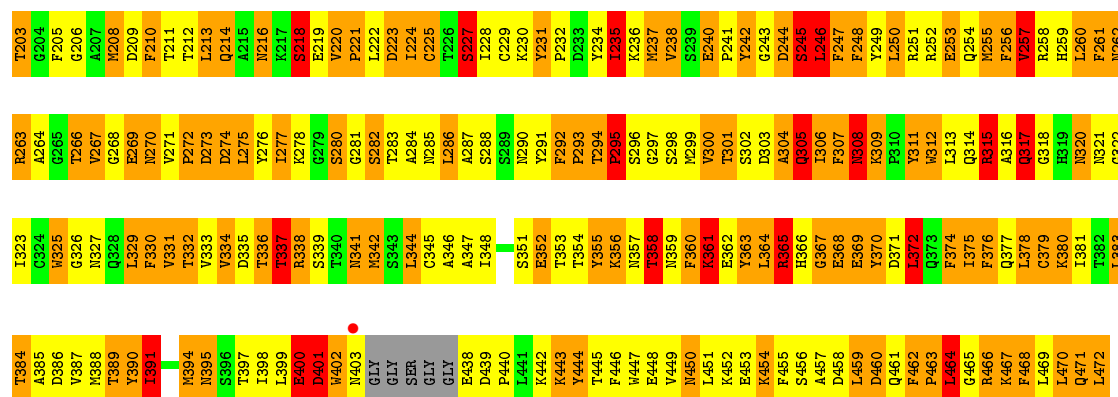


• Molecule 1: Late major capsid protein L1

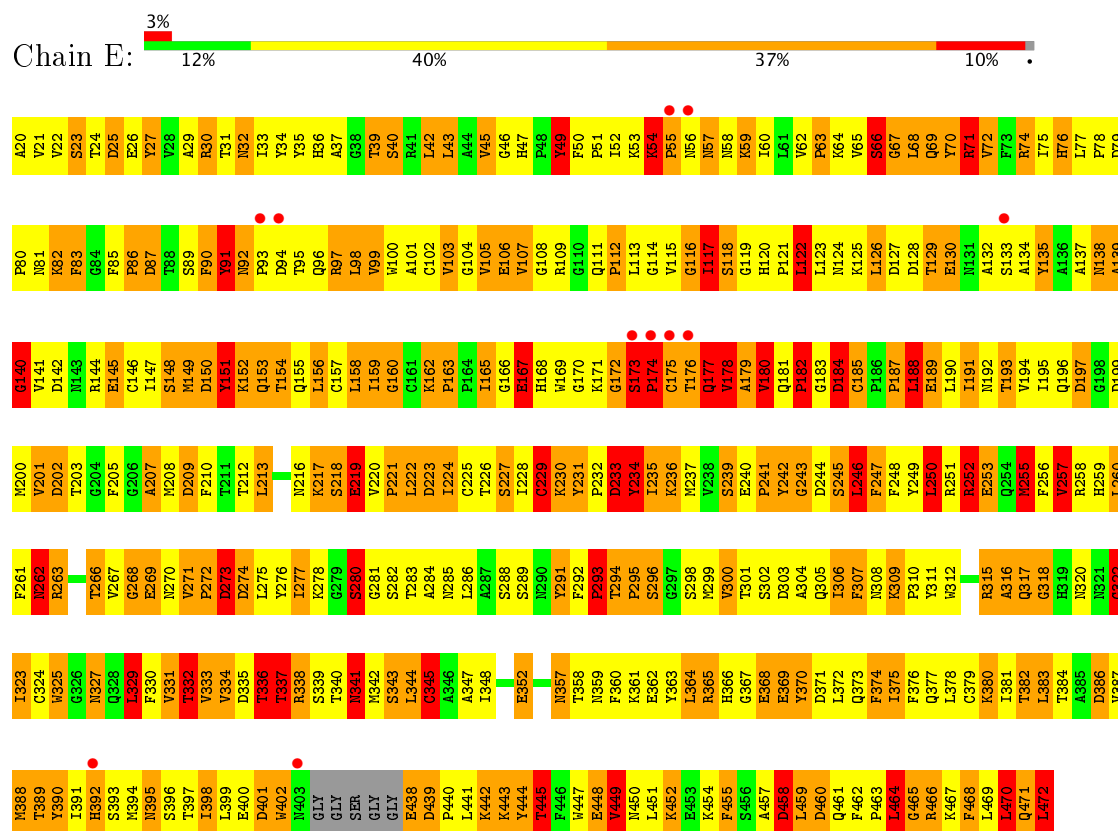


• Molecule 1: Late major capsid protein L1

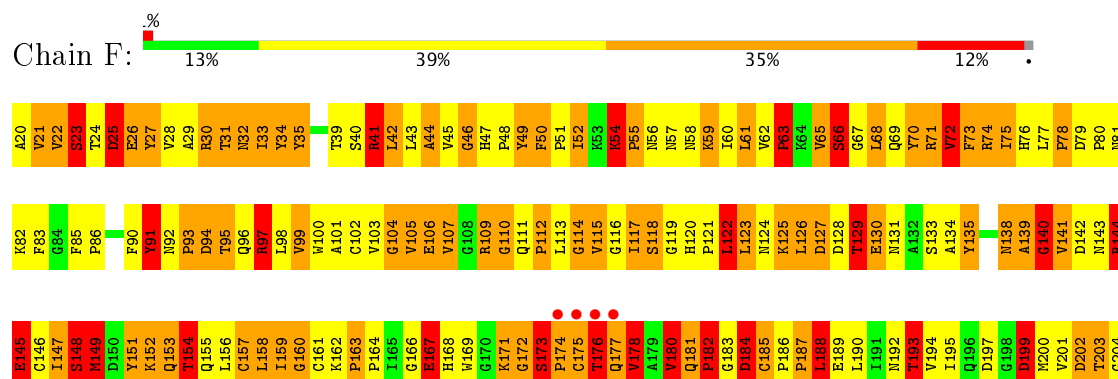


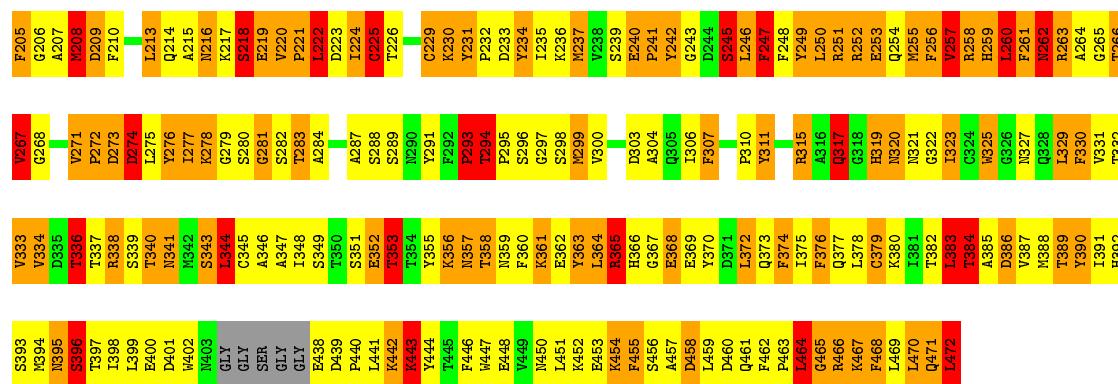


• Molecule 1: Late major capsid protein L1

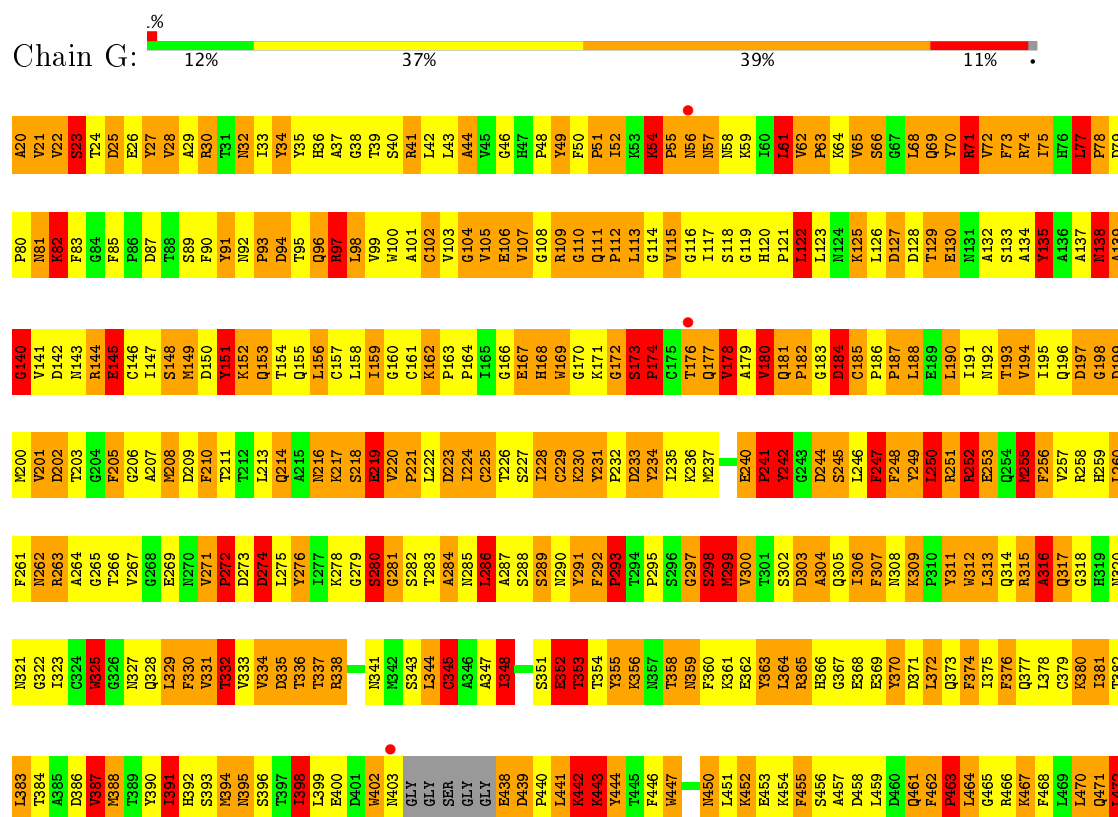


• Molecule 1: Late major capsid protein L1

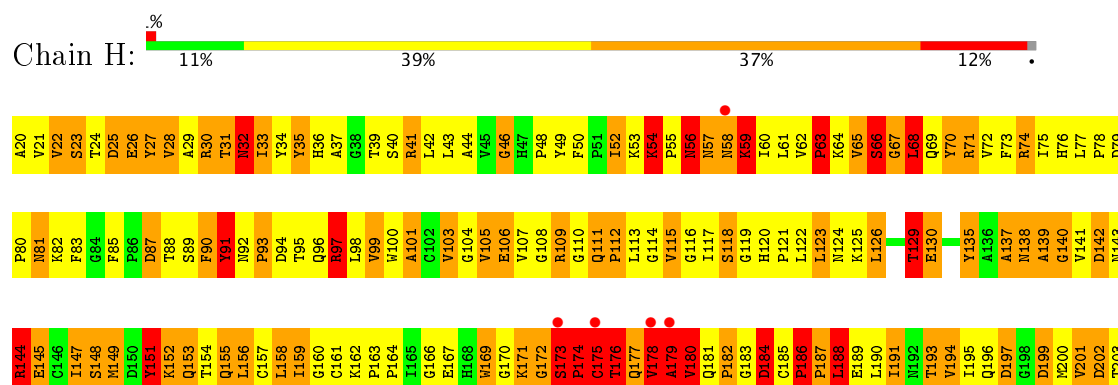


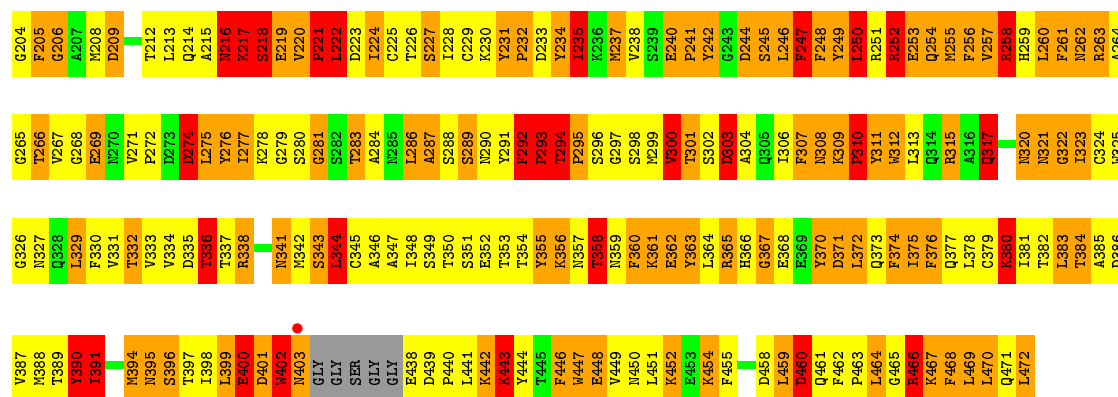


• Molecule 1: Late major capsid protein L1

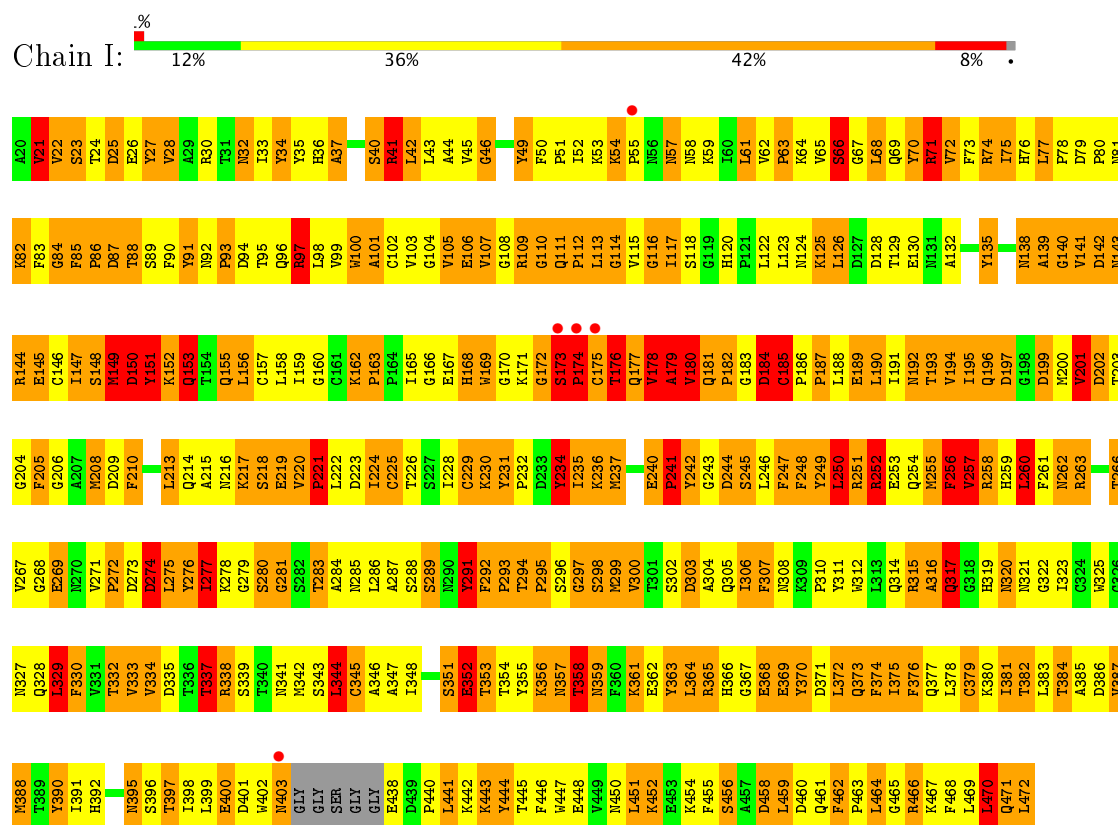


• Molecule 1: Late major capsid protein L1

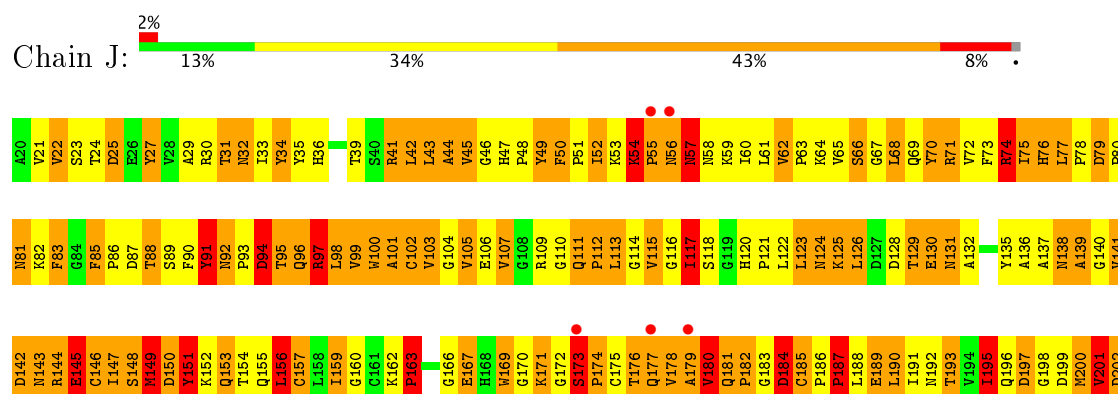


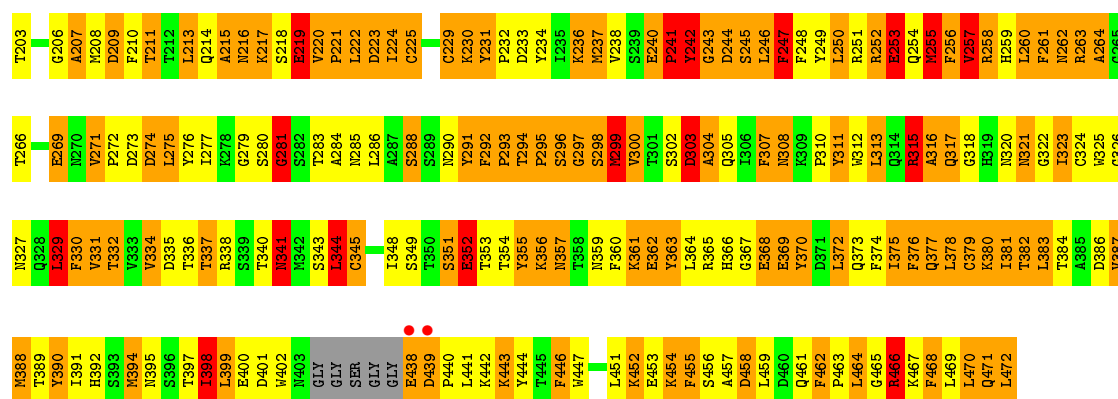


• Molecule 1: Late major capsid protein L1



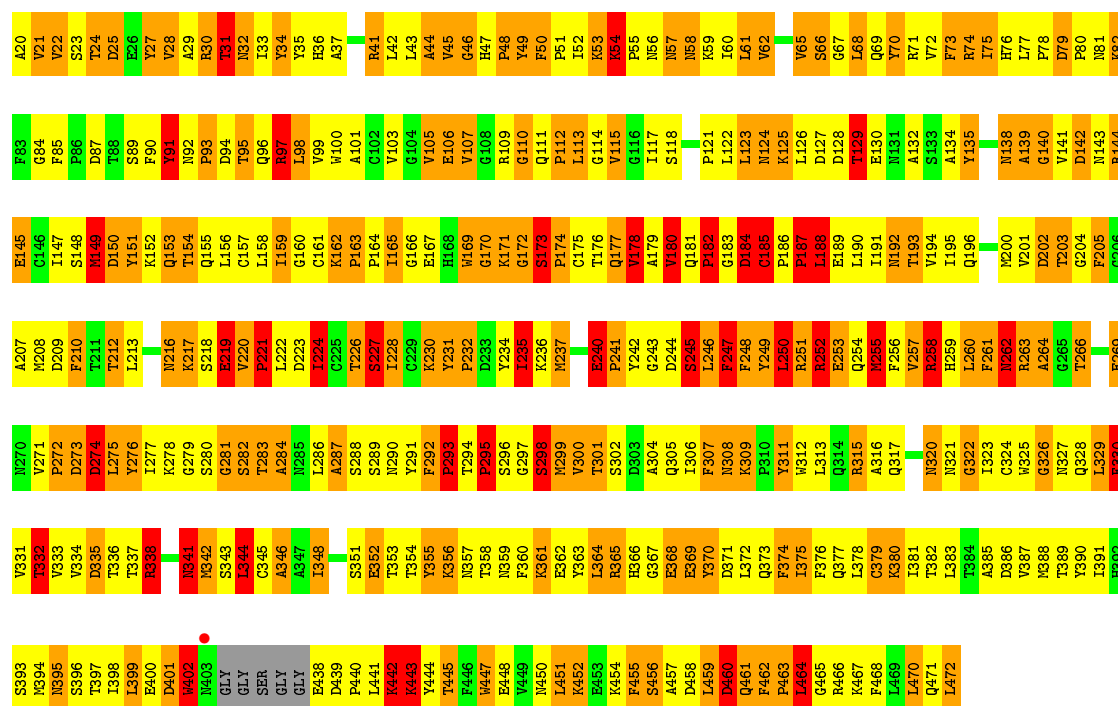
• Molecule 1: Late major capsid protein L1





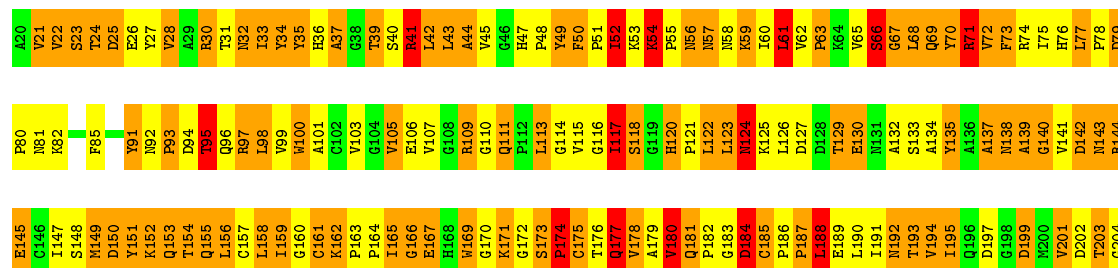
• Molecule 1: Late major capsid protein L1

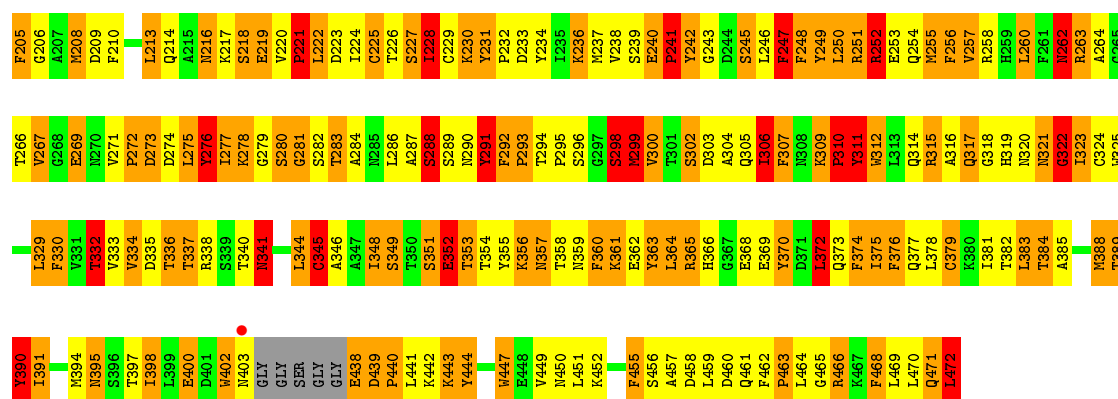
Chain K: 13% 42% 35% 10%



• Molecule 1: Late major capsid protein L1

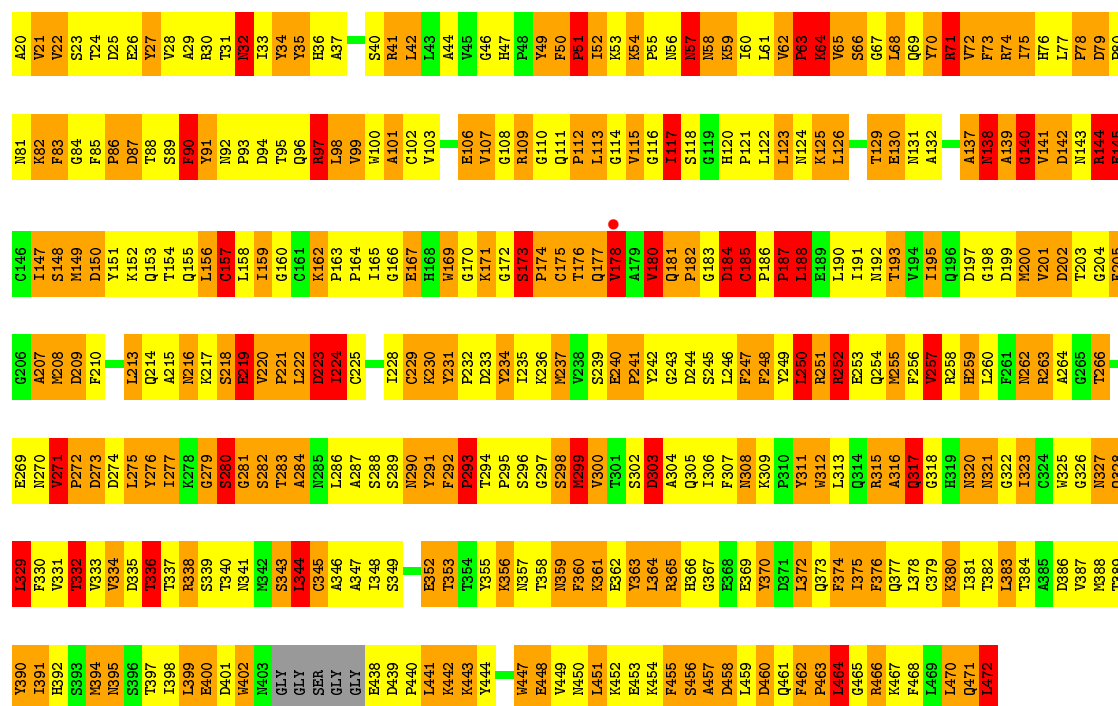
Chain L: 16% 34% 40% 8%





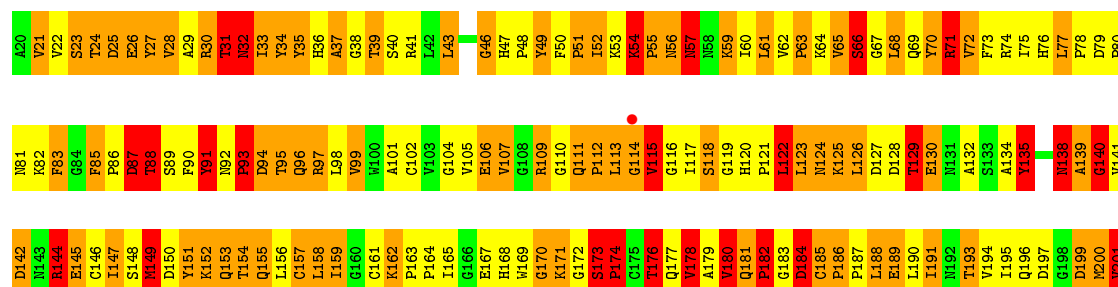
• Molecule 1: Late major capsid protein L1

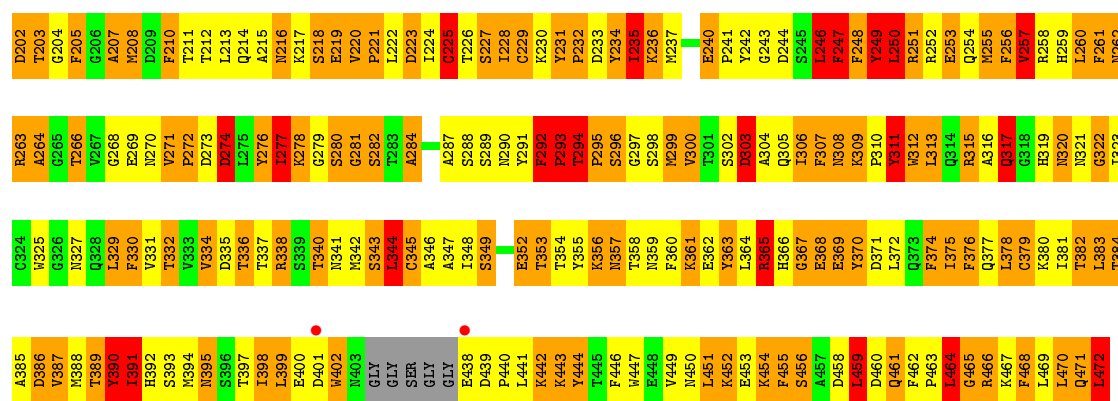
Chain M: 12% 40% 38% 9%



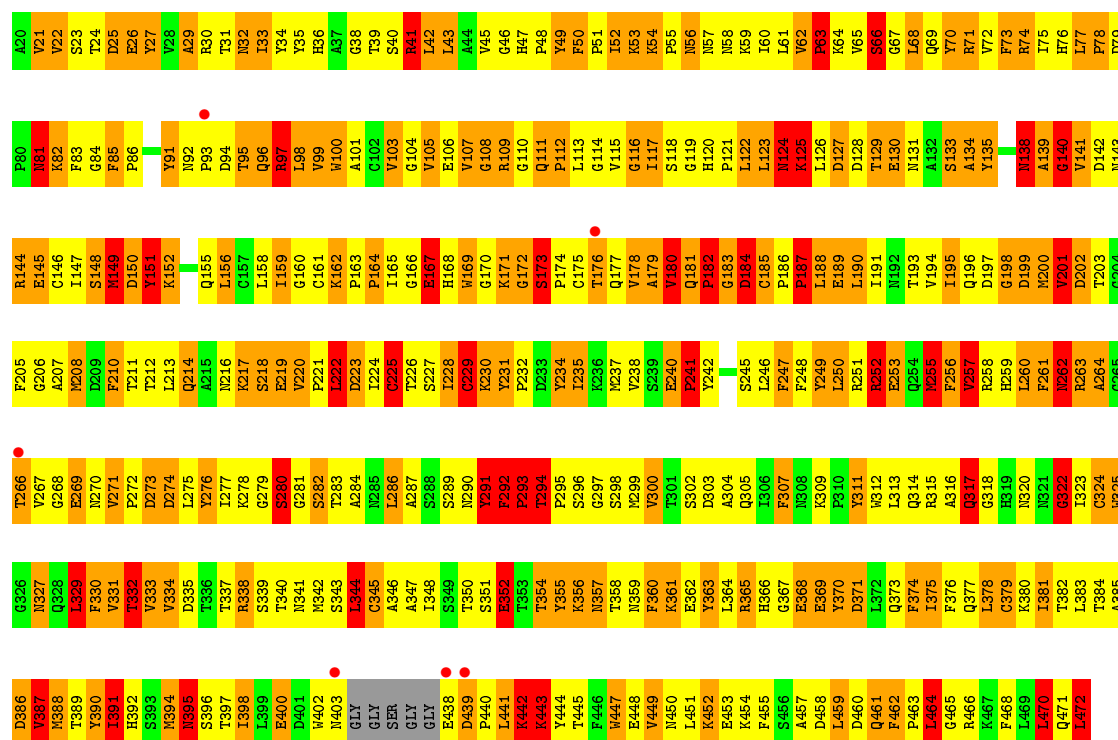
• Molecule 1: Late major capsid protein L1

Chain N: 11% 36% 41% 11%





• Molecule 1: Late major capsid protein L1



GLOBAL-STATISTICS INFOmissingINFO

4 Model quality

4.1 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	2.76	242/3395 (7.1%)	2.11	115/4616 (2.5%)
1	B	2.74	260/3395 (7.7%)	2.03	110/4616 (2.4%)
1	C	2.76	241/3395 (7.1%)	2.14	110/4616 (2.4%)
1	D	2.98	305/3395 (9.0%)	2.15	130/4616 (2.8%)
1	E	2.76	244/3395 (7.2%)	2.02	113/4616 (2.4%)
1	F	2.85	287/3395 (8.5%)	2.09	128/4616 (2.8%)
1	G	2.82	264/3395 (7.8%)	2.07	121/4616 (2.6%)
1	H	2.87	287/3395 (8.5%)	2.09	133/4616 (2.9%)
1	I	2.82	257/3395 (7.6%)	2.08	120/4616 (2.6%)
1	J	2.78	267/3395 (7.9%)	2.09	123/4616 (2.7%)
1	K	2.82	258/3395 (7.6%)	2.04	97/4616 (2.1%)
1	L	2.82	243/3395 (7.2%)	2.07	124/4616 (2.7%)
1	M	2.72	230/3395 (6.8%)	2.14	129/4616 (2.8%)
1	N	2.76	269/3395 (7.9%)	2.12	136/4616 (2.9%)
1	O	2.73	266/3395 (7.8%)	1.99	90/4616 (1.9%)
All	All	2.80	3920/50925 (7.7%)	2.08	1779/69240 (2.6%)

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	C	0	3
1	D	0	1
1	E	0	1
1	F	0	1
1	G	0	1
1	L	0	3
1	M	0	1
1	N	0	1
1	O	0	1
All	All	0	13

The worst 5 of 3920 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	K	245	SER	CB-OG	29.56	1.80	1.42
1	D	438	GLU	CD-OE2	20.41	1.48	1.25
1	D	438	GLU	CD-OE1	18.10	1.45	1.25
1	I	356	LYS	CE-NZ	17.09	1.91	1.49
1	D	178	VAL	CB-CG2	16.71	1.88	1.52

The worst 5 of 1779 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	71	ARG	NE-CZ-NH2	23.80	132.20	120.30
1	F	365	ARG	NE-CZ-NH1	21.93	131.26	120.30
1	C	144	ARG	NE-CZ-NH1	21.91	131.25	120.30
1	I	258	ARG	NE-CZ-NH1	20.20	130.40	120.30
1	D	251	ARG	NE-CZ-NH2	19.49	130.04	120.30

There are no chirality outliers.

5 of 13 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	C	292	PHE	Sidechain
1	C	390	TYR	Sidechain
1	C	70	TYR	Sidechain
1	D	294	THR	Mainchain
1	E	234	TYR	Sidechain

4.2 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	3310	0	3213	561	0
1	B	3310	0	3213	550	0
1	C	3310	0	3213	527	0
1	D	3310	0	3213	540	11
1	E	3310	0	3213	509	5
1	F	3310	0	3213	553	5
1	G	3310	0	3213	579	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	H	3310	0	3213	520	5
1	I	3310	0	3213	568	26
1	J	3310	0	3213	543	26
1	K	3310	0	3213	526	0
1	L	3310	0	3213	530	0
1	M	3310	0	3213	537	5
1	N	3310	0	3213	540	11
1	O	3310	0	3213	531	0
All	All	49650	0	48195	7270	47

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 74.

The worst 5 of 7270 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:D:115:VAL:CB	1:D:115:VAL:CG1	1.75	1.65
1:B:82:LYS:CG	1:B:82:LYS:CB	1.75	1.64
1:I:88:THR:CG2	1:I:88:THR:CB	1.76	1.64
1:F:176:THR:CG2	1:F:176:THR:CB	1.74	1.64
1:O:331:VAL:CG1	1:O:331:VAL:CB	1.75	1.63

The worst 5 of 47 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:E:177:GLN:OE1	1:M:87:ASP:N[3_645]	1.49	0.71
1:E:177:GLN:OE1	1:M:86:PRO:C[3_645]	1.62	0.58
1:I:177:GLN:CG	1:J:95:THR:C[3_645]	1.66	0.54
1:D:175:CYS:O	1:N:88:THR:CG2[3_545]	1.72	0.48
1:D:177:GLN:N	1:N:88:THR:CA[3_545]	1.73	0.47

4.3 Torsion angles [i](#)

4.3.1 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	415/424 (98%)	316 (76%)	69 (17%)	30 (7%)	1	19
1	B	415/424 (98%)	327 (79%)	64 (15%)	24 (6%)	2	24
1	C	415/424 (98%)	322 (78%)	63 (15%)	30 (7%)	1	19
1	D	415/424 (98%)	309 (74%)	76 (18%)	30 (7%)	1	19
1	E	415/424 (98%)	321 (77%)	66 (16%)	28 (7%)	1	21
1	F	415/424 (98%)	325 (78%)	63 (15%)	27 (6%)	1	22
1	G	415/424 (98%)	310 (75%)	71 (17%)	34 (8%)	1	15
1	H	415/424 (98%)	322 (78%)	59 (14%)	34 (8%)	1	15
1	I	415/424 (98%)	321 (77%)	64 (15%)	30 (7%)	1	19
1	J	415/424 (98%)	316 (76%)	65 (16%)	34 (8%)	1	15
1	K	415/424 (98%)	322 (78%)	66 (16%)	27 (6%)	1	22
1	L	415/424 (98%)	321 (77%)	67 (16%)	27 (6%)	1	22
1	M	415/424 (98%)	319 (77%)	72 (17%)	24 (6%)	2	24
1	N	415/424 (98%)	319 (77%)	67 (16%)	29 (7%)	1	20
1	O	415/424 (98%)	325 (78%)	59 (14%)	31 (8%)	1	18
All	All	6225/6360 (98%)	4795 (77%)	991 (16%)	439 (7%)	1	20

5 of 439 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	140	GLY
1	A	174	PRO
1	A	180	VAL
1	A	299	MET
1	A	303	ASP

4.3.2 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	367/368 (100%)	304 (83%)	63 (17%)	2	16
1	B	367/368 (100%)	302 (82%)	65 (18%)	2	14
1	C	367/368 (100%)	300 (82%)	67 (18%)	2	13
1	D	367/368 (100%)	308 (84%)	59 (16%)	3	20
1	E	367/368 (100%)	299 (82%)	68 (18%)	2	12
1	F	367/368 (100%)	305 (83%)	62 (17%)	2	17
1	G	367/368 (100%)	303 (83%)	64 (17%)	2	15
1	H	367/368 (100%)	302 (82%)	65 (18%)	2	14
1	I	367/368 (100%)	310 (84%)	57 (16%)	3	22
1	J	367/368 (100%)	311 (85%)	56 (15%)	3	22
1	K	367/368 (100%)	300 (82%)	67 (18%)	2	13
1	L	367/368 (100%)	312 (85%)	55 (15%)	3	23
1	M	367/368 (100%)	305 (83%)	62 (17%)	2	17
1	N	367/368 (100%)	296 (81%)	71 (19%)	1	11
1	O	367/368 (100%)	307 (84%)	60 (16%)	3	19
All	All	5505/5520 (100%)	4564 (83%)	941 (17%)	2	17

5 of 941 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	G	292	PHE
1	I	149	MET
1	N	357	ASN
1	G	352	GLU
1	H	232	PRO

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 238 such sidechains are listed below:

Mol	Chain	Res	Type
1	G	341	ASN
1	I	168	HIS
1	N	395	ASN
1	G	395	ASN
1	H	262	ASN

4.3.3 RNA [i](#)

There are no RNA molecules in this entry.

4.4 Non-standard residues in protein, DNA, RNA chains [i](#)

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

4.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

4.6 Ligand geometry [i](#)

There are no ligands in this entry.

4.7 Other polymers [i](#)

There are no such residues in this entry.

4.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

5 Fit of model and data ⓘ

5.1 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, 95th 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	419/424 (98%)	-0.52	2 (0%) 90 85	40, 77, 111, 122	0
1	B	419/424 (98%)	-0.53	3 (0%) 87 80	36, 76, 109, 131	0
1	C	419/424 (98%)	-0.48	3 (0%) 87 80	44, 79, 110, 121	0
1	D	419/424 (98%)	-0.47	5 (1%) 79 68	37, 78, 109, 130	0
1	E	419/424 (98%)	-0.45	11 (2%) 56 44	41, 79, 114, 128	0
1	F	419/424 (98%)	-0.45	4 (0%) 82 72	40, 80, 111, 131	0
1	G	419/424 (98%)	-0.43	3 (0%) 87 80	46, 83, 112, 135	0
1	H	419/424 (98%)	-0.48	6 (1%) 75 64	34, 78, 111, 136	0
1	I	419/424 (98%)	-0.43	5 (1%) 79 68	48, 82, 113, 132	0
1	J	419/424 (98%)	-0.42	7 (1%) 70 58	43, 79, 111, 134	0
1	K	419/424 (98%)	-0.53	1 (0%) 94 92	35, 76, 109, 129	0
1	L	419/424 (98%)	-0.56	1 (0%) 94 92	38, 73, 108, 137	0
1	M	419/424 (98%)	-0.54	1 (0%) 94 92	43, 75, 108, 133	0
1	N	419/424 (98%)	-0.45	3 (0%) 87 80	45, 80, 112, 128	0
1	O	419/424 (98%)	-0.42	6 (1%) 75 64	45, 82, 112, 133	0
All	All	6285/6360 (98%)	-0.48	61 (0%) 82 72	34, 78, 111, 137	0

The worst 5 of 61 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	L	403	ASN	6.3
1	H	179	ALA	5.6
1	O	403	ASN	5.5
1	I	174	PRO	5.5
1	O	176	THR	5.0

5.2 Non-standard residues in protein, DNA, RNA chains [i](#)

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

5.3 Carbohydrates [i](#)

There are no carbohydrates in this entry.

5.4 Ligands [i](#)

There are no ligands in this entry.

5.5 Other polymers [i](#)

There are no such residues in this entry.