



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 15, 2026 – 01:50 PM UTC

PDB ID : 2R5H / pdb_00002r5h
Title : Pentamer structure of Major Capsid Protein L1 of Human Papilloma Virus type 16
Authors : Bishop, B.; Dasgupta, J.; Chen, X.S.
Deposited on : 2007-09-03
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

<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	:	4-5-2 with Phenix2.0
Xtriage (Phenix)	:	2.0
EDS	:	3.0
Percentile statistics	:	20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4	:	9.0.010 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.49

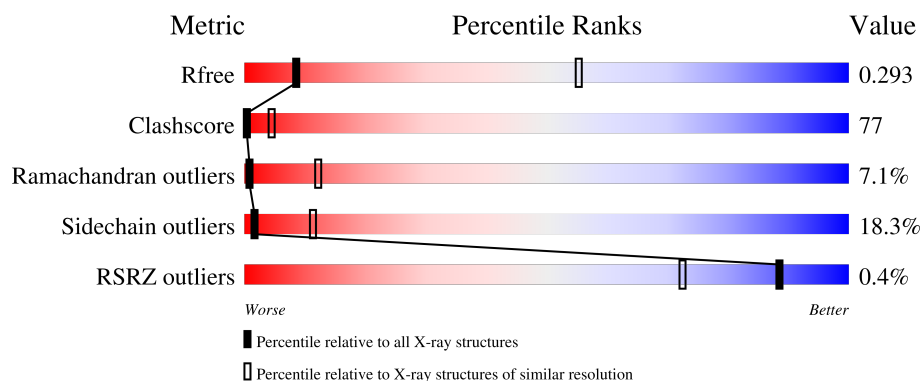
1 Overall quality at a glance

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}	180053	1131 (3.80-3.60)
Clashscore	190562	1171 (3.80-3.60)
Ramachandran outliers	187476	1129 (3.80-3.60)
Sidechain outliers	187428	1126 (3.80-3.60)
RSRZ outliers	180081	1130 (3.80-3.60)

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 $\leq 5\%$. 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	
1	B	424	
1	C	424	
1	D	424	
1	E	424	

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 mutation	UNP Q81007
A	181	GLN	ASN	engineered mutation	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 mutation	UNP Q81007
B	181	GLN	ASN	engineered mutation	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 mutation	UNP Q81007
C	181	GLN	ASN	engineered mutation	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 mutation	UNP Q81007
D	181	GLN	ASN	engineered mutation	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 mutation	UNP Q81007
E	181	GLN	ASN	engineered mutation	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 mutation	UNP Q81007
F	181	GLN	ASN	engineered mutation	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 mutation	UNP Q81007
G	181	GLN	ASN	engineered mutation	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 mutation	UNP Q81007
H	181	GLN	ASN	engineered mutation	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 mutation	UNP Q81007
I	181	GLN	ASN	engineered mutation	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 mutation	UNP Q81007
J	181	GLN	ASN	engineered mutation	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 mutation	UNP Q81007
K	181	GLN	ASN	engineered mutation	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 mutation	UNP Q81007
L	181	GLN	ASN	engineered mutation	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 mutation	UNP Q81007
M	181	GLN	ASN	engineered mutation	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 mutation	UNP Q81007
N	181	GLN	ASN	engineered mutation	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...

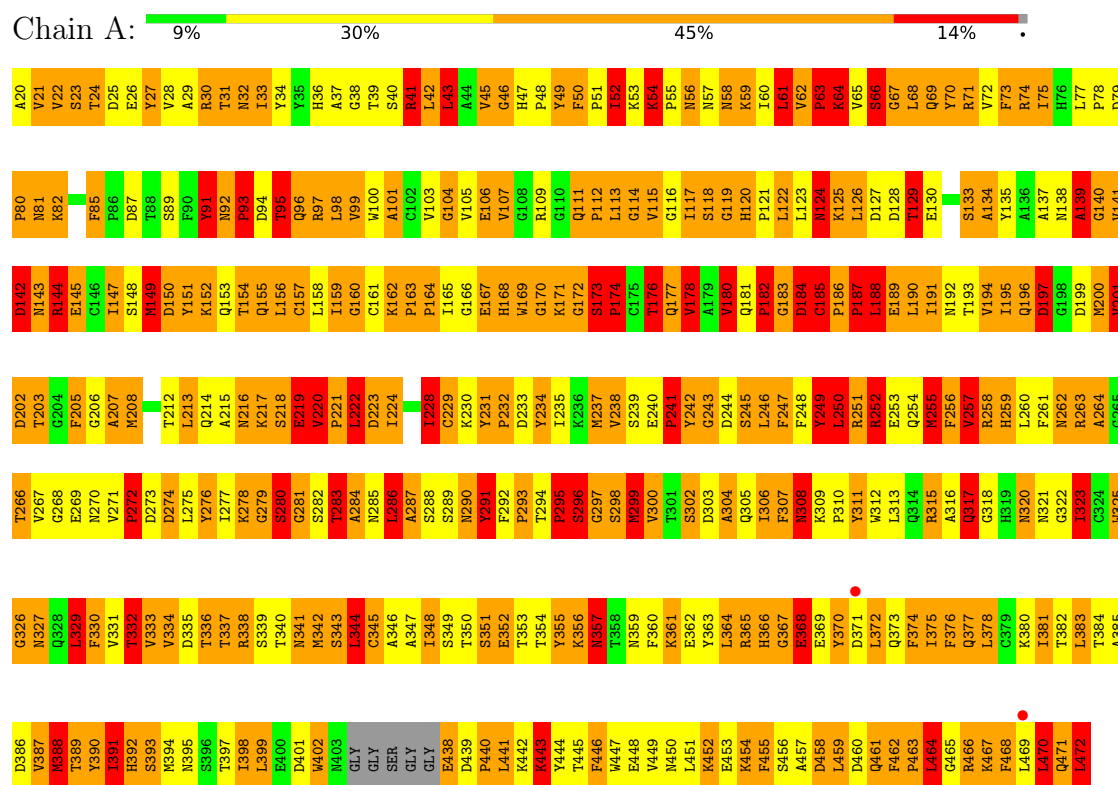
Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
O	177	GLN	ASN	engineered mutation	UNP Q81007
O	181	GLN	ASN	engineered mutation	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, 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 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



D202	N262	G322	T382	Q471
T203	R263	I323	L383	L472
G204	A264	G324	T384	
F205	G265	G325	A385	
G206	G266	G326	D386	
A207	V267	N327	V387	
M208	G268	G328	M388	
D209	E269	L329	T389	
F210	N270	F330	Y390	
T211	V271	V331	I391	
T212	T272	T332	H392	
L213	D273	V333	S393	
Q214	M274	V334	N394	
A215	L275	D335	N395	
N216	L276	T336	S396	
K217	I277	T337	T397	
S218	K278	I338	I398	
E219	G279	S339	I399	
V220	S280	T340	E400	
V221	G281	N341	D401	
L222	S282	M342	W402	
D223	T283	S343	N403	
L224	A284	L344	GLY	
G225	N285	C345	GLY	
T226	L286	A346	SER	
S227	A287	A347	GLY	
L228	S288	I348	GLY	
C229	S289	G349	E438	
K230	N290	T350	D439	
Y231	Y291	M351	P440	
P232	F292	E352	L441	
D233	T293	T353	K442	
Y234	T294	T354	K443	
L235	P295	Y355	Y444	
G236	S296	K356	T445	
M237	G297	N357	F446	
V238	S298	T358	W447	
S239	M299	N359	E448	
E240	V300	F360	V449	
P241	T301	K361	N450	
Y242	S302	E362	L451	
G243	D303	Y363	K452	
D244	L364	A365	E453	
S245	Q365	R366	K454	
L246	I366	G367	F455	
F247	E368	E369	S456	
Y249	K369	E370	D458	
L250	P310	D371	L459	
R251	Y311	D372	D460	
R252	W312	Q373	Q461	
E253	L313	F374	F462	
Q254	G314	F375	P463	
M255	R315	L376	L464	
F256	A316	F376	G465	
V257	Q317	Q377	R466	
R258	G318	L378	K467	
H259	H319	C379	F468	
L260	K320	K380	L469	
F261	N321	I381	L470	

• Molecule 1: Late major capsid protein L1

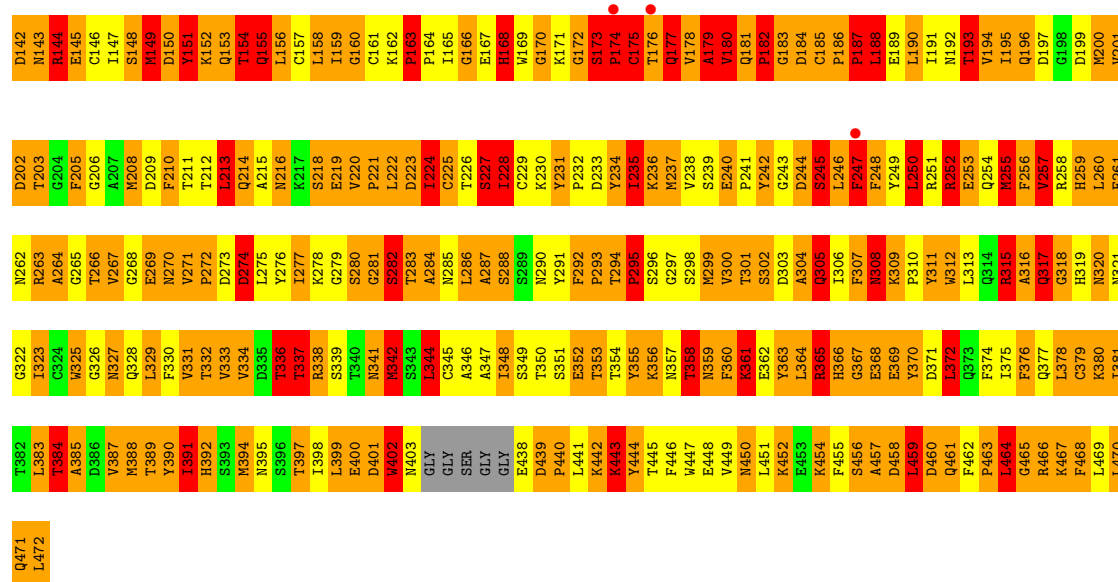
Chain C: 6% 30% 50% 12%

A20	P80	G140	H200	L260	N320	K380	L469
V21	N81	V141	V201	F261	N321	I381	L470
V22	K82	D142	D202	R262	G322	T382	Q471
T24	F83	N143	T203	R263	I323	L383	L472
D25	G84	R144	G204	A264	G324	T384	
E26	F85	E145	F205	G265	G325	A385	
Y27	P86	C146	G206	T266	G326	N386	
V28	D87	I147	A207	V267	N327	V387	
A29	T88	S148	R208	G268	Q328	M388	
R30	S89	M149	D209	E269	L329	T389	
T31	F90	D150	F210	N270	F330	Y390	
N32	Y91	Y151	T211	T212	V331	I391	
I33	N92	K152	L213	L214	T332	H392	
Y34	P93	Q153	Q214	Q215	V333	S393	
H36	D94	T154	A215	A216	V334	N394	
A37	T95	Q155	L156	L157	T336	N395	
G38	R97	L158	K217	K218	G336	S396	
T39	V99	I159	G278	G279	N337	I398	
S40	W100	G160	S280	S281	S338	L399	
R41	A101	C161	G281	G282	T340	E400	
L42	C102	K162	L222	L223	N341	W402	
L43	V103	P163	D223	D224	S343	M403	
A44	G104	P164	I224	I225	L344	GLY	
V45	V105	I165	C225	C226	C345	GLY	
G46	E106	G166	T226	T227	A346	SER	
H47	V107	E167	S227	S228	A347	GLY	
F48	G108	H168	I228	I229	I348	GLY	
Y49	R109	V169	C229	C230	S349	E438	
F50	Y110	G170	K230	K231	T350	D439	
P51	Q111	K171	Y231	Y232	S351	P440	
I52	L112	G172	F232	F233	E352	L441	
K53	L113	P173	D233	D234	T353	K442	
R54	G114	S174	Y234	Y235	T354	K443	
P55	V115	C175	I235	I236	Y355	Y444	
N56	G116	T176	K236	K237	K356	T445	
N57	I117	Q177	M237	M238	N357	F446	
N58	S118	V178	V238	V239	T358	W447	
K59	H119	A179	S298	S299	E448	E449	
L60	P120	V180	M299	M300	N359	N450	
L61	Q121	Q181	V300	V301	F360	L451	
L62	P122	P182	T301	T302	K361	V449	
R63	L123	G183	S302	S303	E362	N450	
K64	N124	D184	D244	D245	Y363	K452	
V65	K125	C185	Q305	Q306	L364	E453	
S66	L126	P186	L246	L247	R365	E454	
G67	D127	P187	F247	F248	H366	F455	
L68	L188	L189	F249	F250	G367	S456	
Q69	T129	E189	Y249	Y250	E368	A457	
Y70	E130	L190	L250	L251	E369	D458	
R71	I191	T191	R251	R252	Y370	L459	
F72	A132	N192	G252	G253	D371	D460	
F73	S133	T193	L253	L254	L372	Q461	
R74	A134	V194	Q254	Q255	Q373	F462	
I75	Y135	I195	R255	R256	F374	P463	
H76	A136	Q196	F256	F257	G375	L464	
L77	A137	D197	V257	V258	F376	G465	
P78	N138	G198	R258	R259	F377	R466	
D79	K139	L199	H259	H260	L378	R467	
D79	K139	L199	H259	H260	L378	R467	

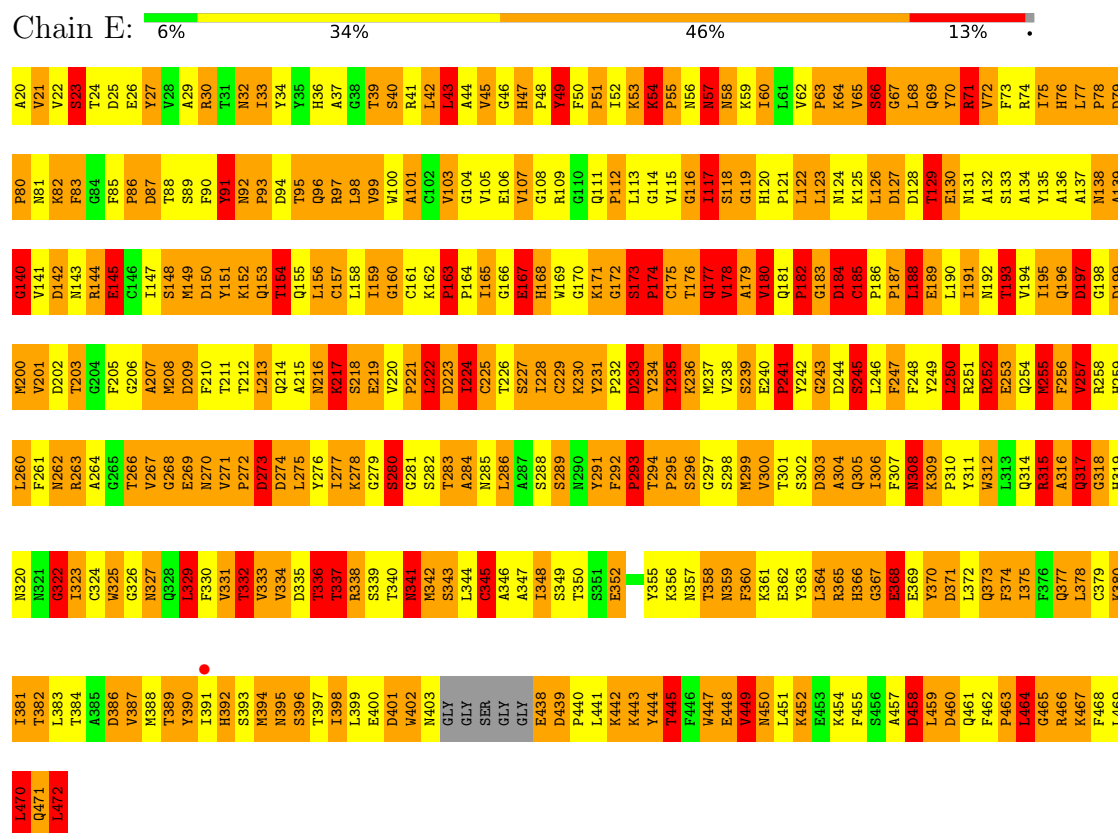
• Molecule 1: Late major capsid protein L1

Chain D: 6% 30% 49% 14%

A20	N81	S40	V105	E106	V107	R109	G110	P111	P112	K53	K54	P55	N56	N57	N58	K59	I60	L61	V62	P63	K64	V65	S66	G67	L68	Q69	Y70	R71	V72	F73	R74	I75	H76	L77	P78	G140	P80
V21	K82	R41	V106	E107	H47	R109	G110	P111	P112	K53	K54	P55	N56	N57	N58	K59	I60	L61	V62	P63	K64	V65	S66	G67	L68	Q69	Y70	R71	V72	F73	R74	I75	H76	L77	P78	G140	P80
V22	F83	L42	V107	E108	H48	R109	G110	P111	P112	K53	K54	P55	N56	N57	N58	K59	I60	L61	V62	P63	K64	V65	S66	G67	L68	Q69	Y70	R71	V72	F73	R74	I75	H76	L77	P78	G140	P80
T24	G84	L43	V108	E109	H49	R109	G110	P111	P112	K53	K54	P55	N56	N57	N58	K59	I60	L61	V62	P63	K64	V65	S66	G67	L68	Q69	Y70	R71	V72	F73	R74	I75	H76	L77	P78	G140	P80
D25	F85	L44	V109	E110	H50	R109	G110	P111	P112	K53	K54	P55	N56	N57	N58	K59	I60	L61	V62	P63	K64	V65	S66	G67	L68	Q69	Y70	R71	V72	F73	R74	I75	H76	L77	P78	G140	P80
E26	P86	L45	V110	E111	H51	R109	G110	P111	P112	K53	K54	P55	N56	N57	N58	K59	I60	L61	V62	P63	K64	V65	S66	G67	L68	Q69	Y70	R71	V72	F73	R74	I75	H76	L77	P78	G140	P80
Y27	T88	L46	V111	E112	H52	R109	G110	P111	P112	K53	K54	P55	N56	N57	N58	K59	I60	L61	V62	P63	K64	V65	S66	G67	L68	Q69	Y70	R71	V72	F73	R74	I75	H76	L77	P78	G140	P80
V28	S89	L47	V112	E113	H53	R109	G110	P111	P112	K53	K54	P55	N56	N57	N58	K59	I60	L61	V62	P63	K64	V65	S66	G67	L68	Q69	Y70	R71	V72	F73	R74	I75	H76	L77	P78	G140	P80
A29	F90	L48	V113	E114	H54	R109	G110	P111	P112	K53	K54	P55	N56	N57	N58	K59	I60	L61	V62	P63	K64	V65	S66	G67	L68	Q69	Y70	R71	V72	F73	R74	I75	H76	L77	P78	G140	P80
R30	Y91	L49	V114	E115	H55	R109	G110	P111	P112	K53	K54	P55	N56	N57	N58	K59	I60	L61	V62	P63	K64	V65	S66	G67	L68	Q69	Y70	R71	V72	F73	R74	I75	H76	L77	P78	G140	P80
T31	N92	L50	V115	E116	H56	R109	G110	P111	P112	K53	K54	P55	N56	N57	N58	K59	I60	L61	V62	P63	K64	V65	S66	G67	L68	Q69	Y70	R71	V72	F73	R74	I75	H76	L77	P78	G140	P80
N32	I33	L51	V116	E117	H57	R109	G110	P111	P112	K53	K54	P55	N56	N57	N58	K59	I60	L61	V62	P63	K64	V65	S66	G67	L68	Q69	Y70	R71	V72	F73	R74	I75	H76	L77	P78	G140	P80
I33	Y34	L52	V117	E118	H58	R109	G110	P111	P112	K53	K54	P55	N56	N57	N58	K59	I60	L61	V62	P63	K64	V65	S66	G67	L68	Q69	Y70	R71	V72	F73	R74	I75	H76	L77	P78	G140	P80
Y34	Q95	L53	V118	E119	H59	R109	G110	P111	P112	K53	K54	P55	N56	N57	N58	K59	I60	L61	V62	P63	K64	V65	S66	G67	L68	Q69	Y70	R71	V72	F73	R74	I75	H76	L77	P78	G140	P80
Q96	Y35	L54	V119	E120	H60	R109	G110	P111	P112	K53	K54	P55	N56	N57	N58	K59	I60	L61	V62	P63	K64	V65	S66	G67	L68	Q69	Y70	R71	V72	F73	R74	I75	H76	L77	P78	G140	P80
H36	R97	L55	V120	E121	H61	R109	G110	P111	P112	K53	K54	P55	N56	N57	N58	K59	I60	L61	V62	P63	K64	V65	S66	G67	L68	Q69	Y70	R71	V72	F73	R74	I75	H76	L77	P78	G140	P80
A37	A37	L56	V121	E122	H62	R109	G110	P111	P112	K53	K54	P55	N56	N57	N58	K59	I60	L61	V62	P63	K64	V65	S66	G67	L68	Q69	Y70	R71	V72	F73	R74	I75	H76	L77	P78	G140	P80

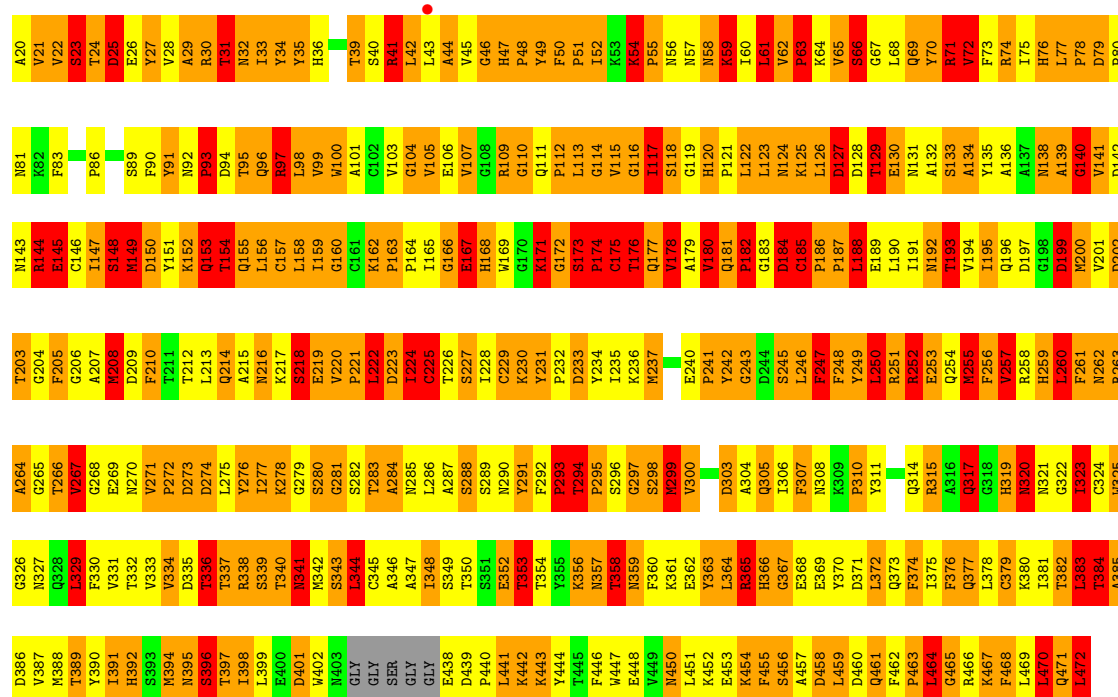


• Molecule 1: Late major capsid protein L1



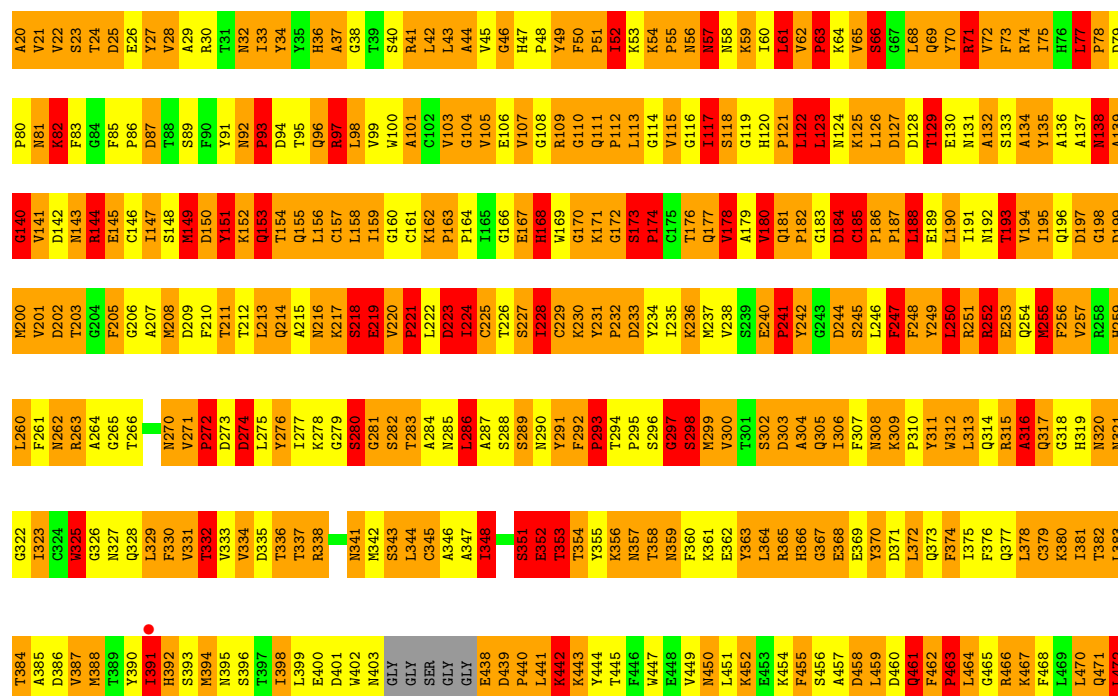
• Molecule 1: Late major capsid protein L1





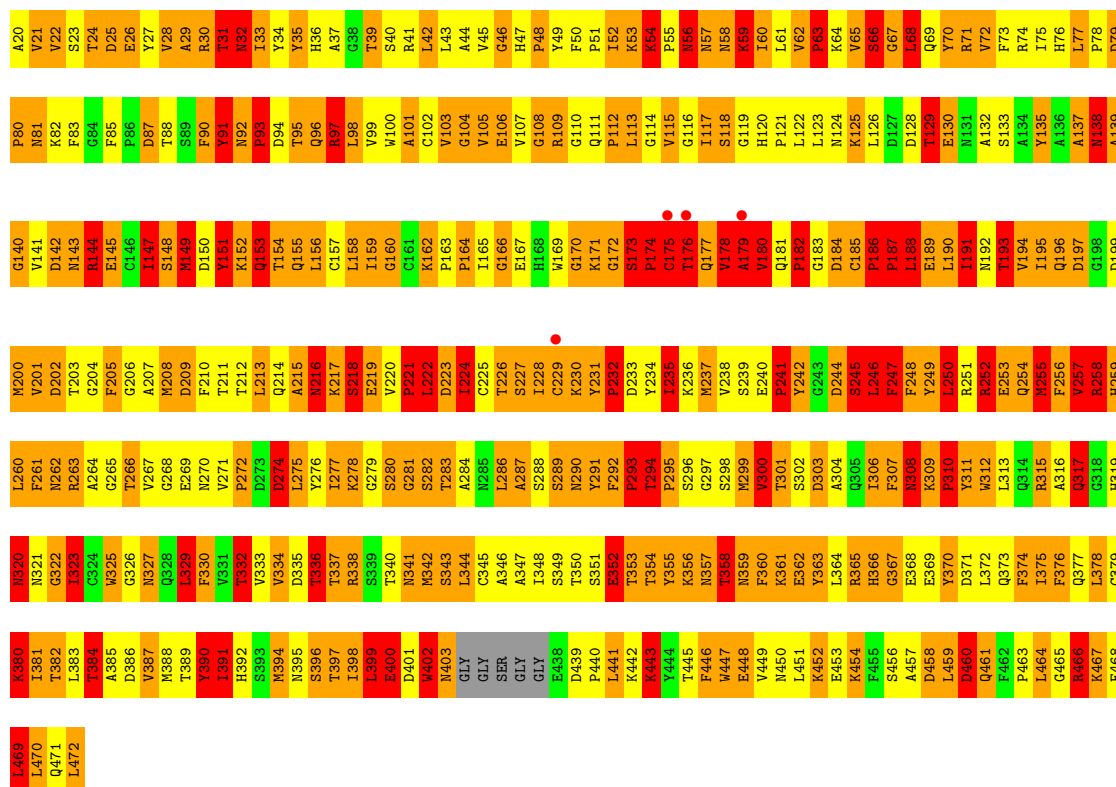
• Molecule 1: Late major capsid protein L1

Chain G: 7% 30% 48% 14%



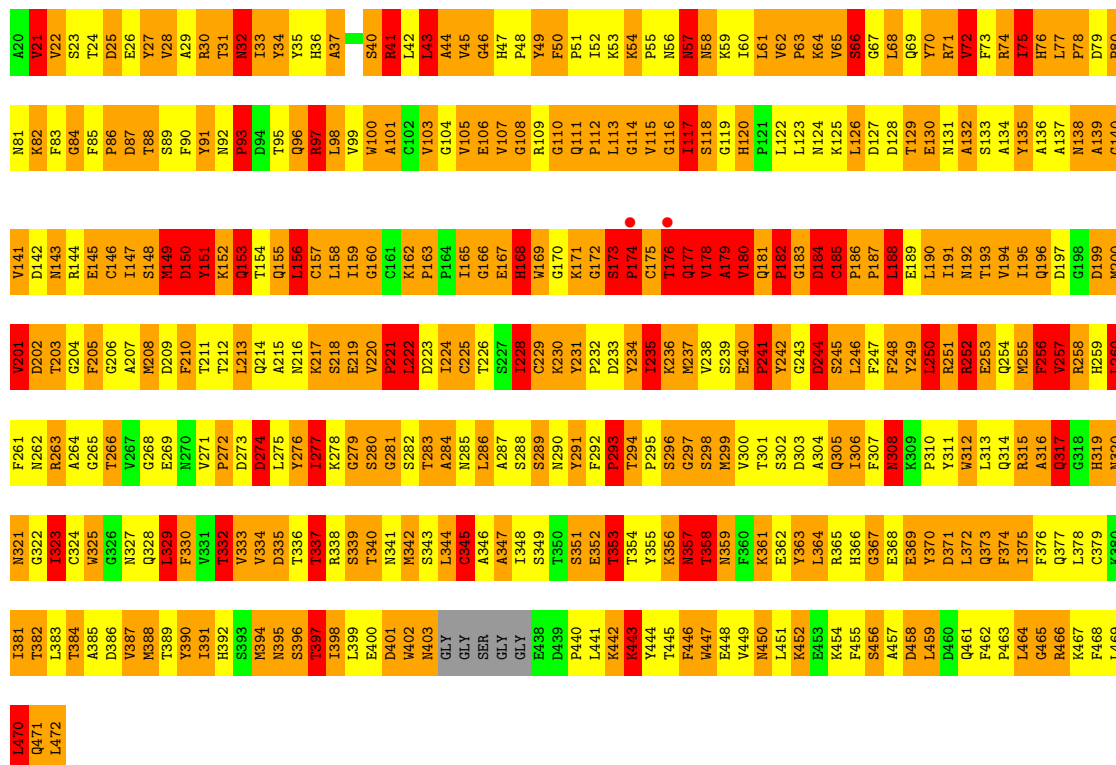
• Molecule 1: Late major capsid protein L1

Chain H: 6% 32% 43% 17%

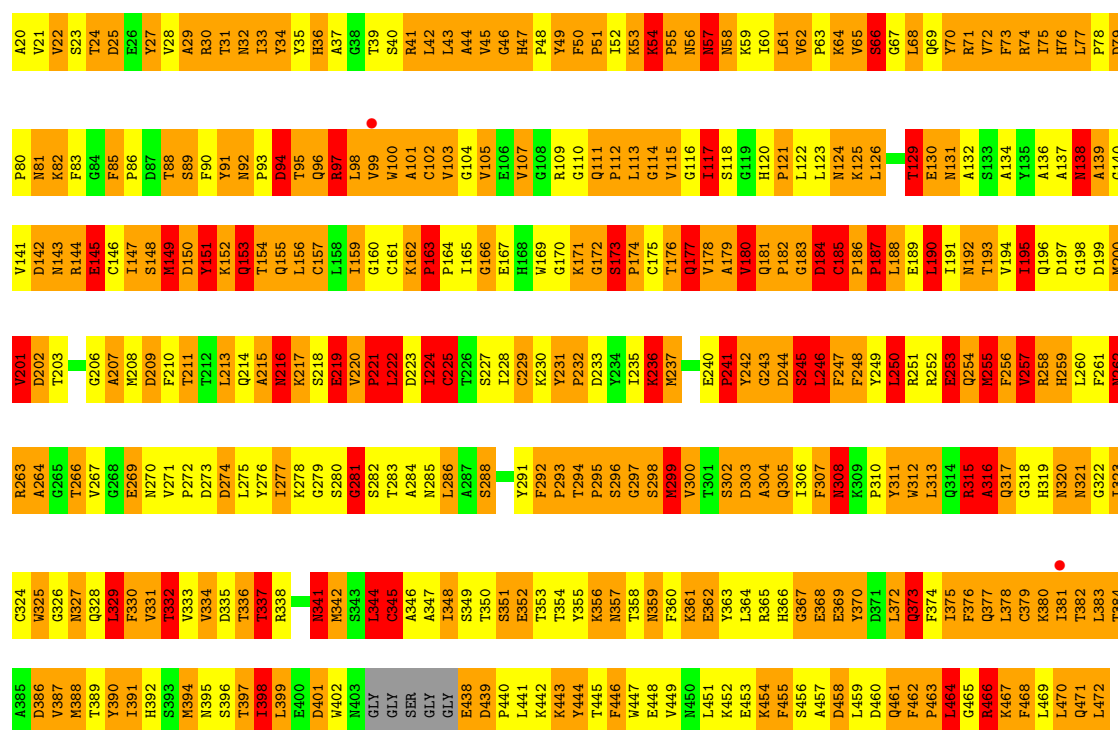


• Molecule 1: Late major capsid protein L1

Chain I: 6% 33% 47% 13%

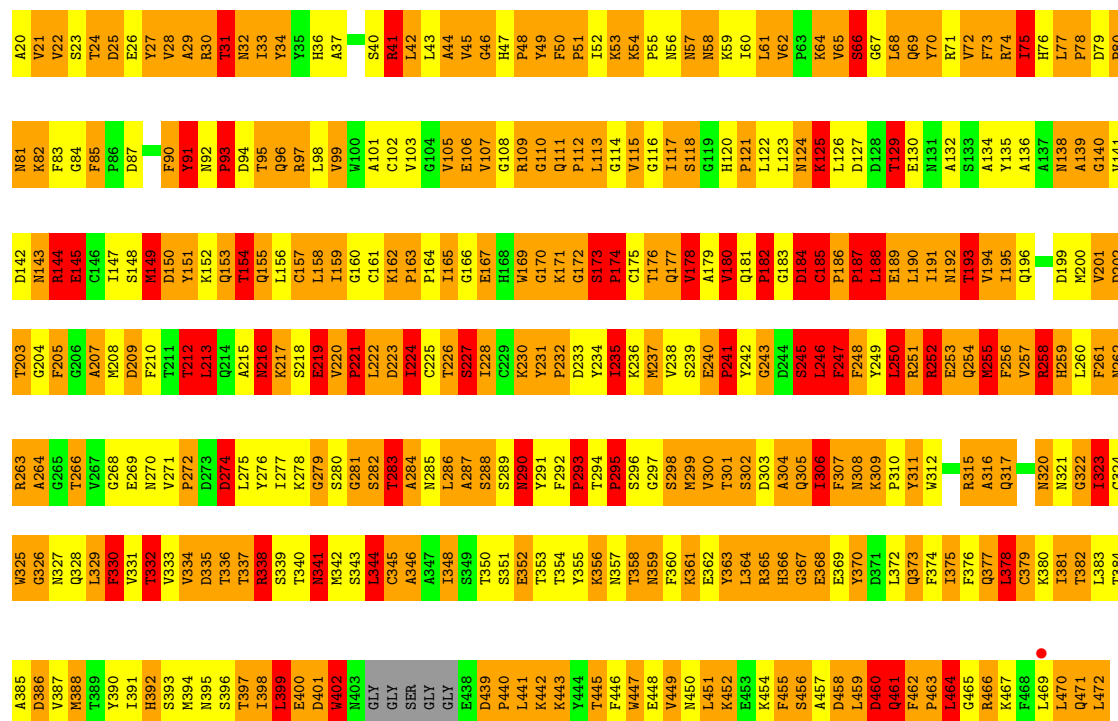


Chain J: 9% 31% 47% 12%

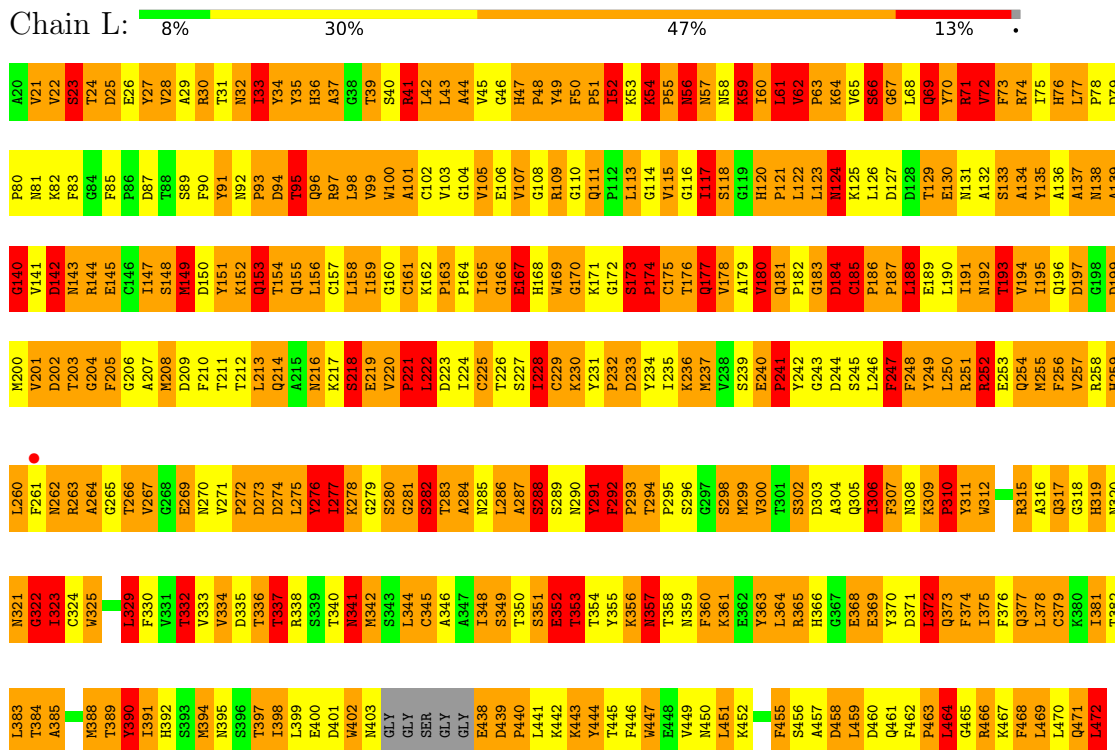


• Molecule 1: Late major capsid protein L1

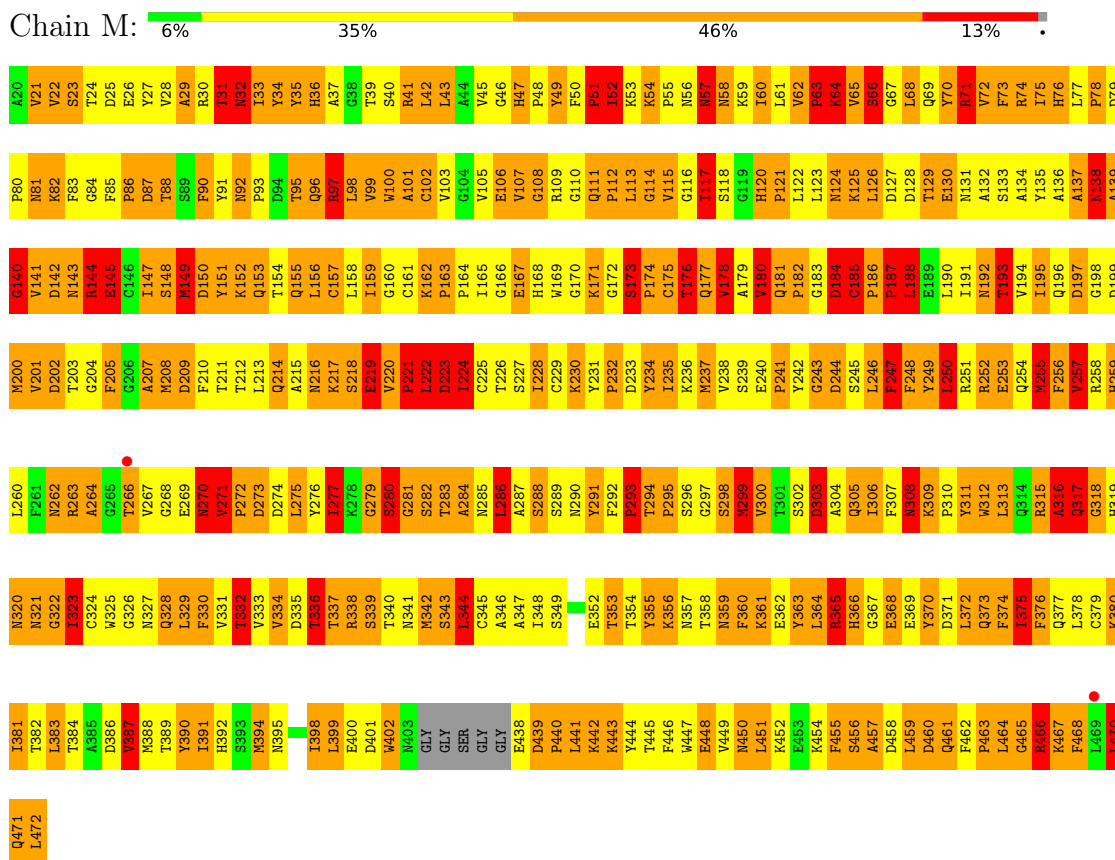
Chain K: 9% 31% 46% 13%



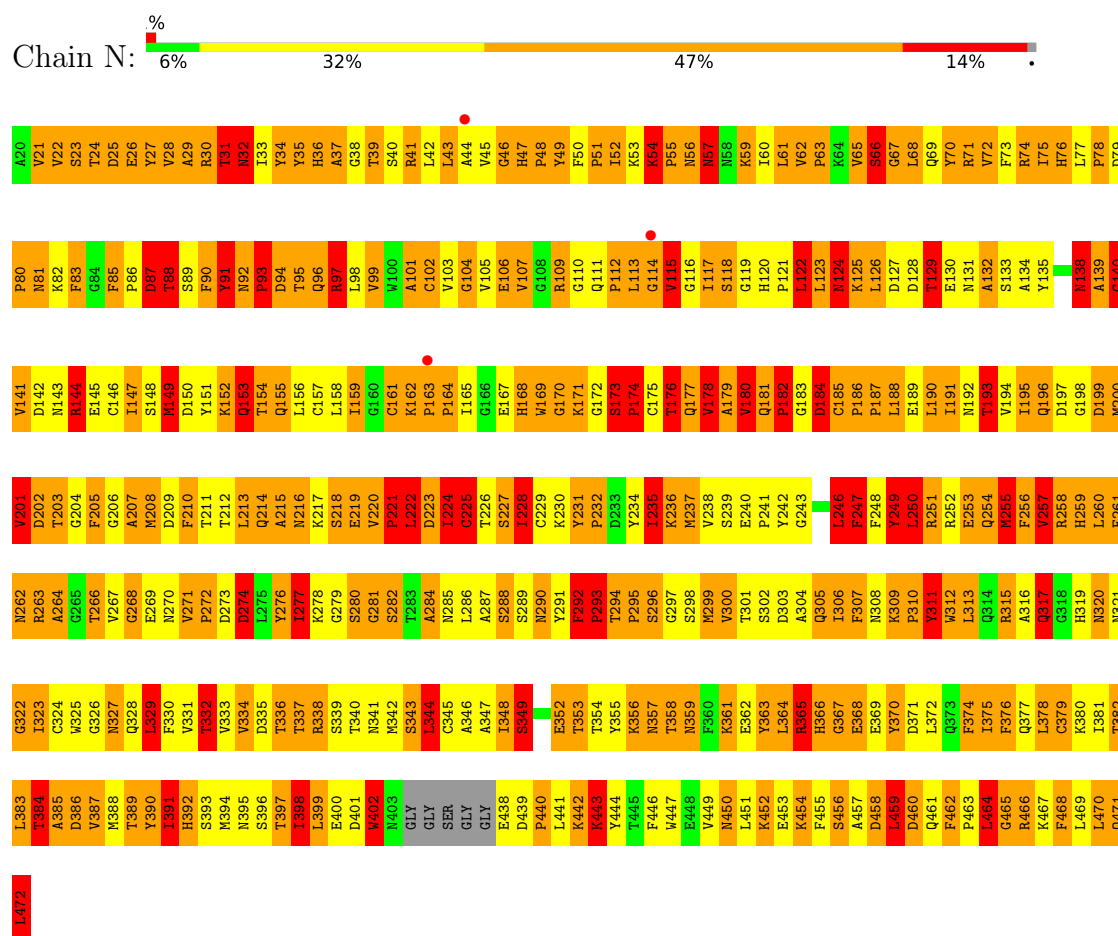
• Molecule 1: Late major capsid protein L1



• Molecule 1: Late major capsid protein L1



• Molecule 1: Late major capsid protein L1



[illegible]

4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	105.00Å 159.05Å 413.74Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	15.00 – 3.70 15.00 – 3.70	Depositor EDS
% Data completeness (in resolution range)	80.5 (15.00-3.70) 79.1 (15.00-3.70)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.70 (at 3.66Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, R_{free}	0.247 , 0.305 0.241 , 0.293	Depositor DCC
R_{free} test set	2971 reflections (4.03%)	wwPDB-VP
Wilson B-factor (Å ²)	92.7	Xtriage
Anisotropy	0.063	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.22 , 56.1	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.91	EDS
Total number of atoms	49650	wwPDB-VP
Average B, all atoms (Å ²)	80.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.88% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical 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 ⓘ

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	3.31	410/3395 (12.1%)	2.47	231/4616 (5.0%)
1	B	3.25	384/3395 (11.3%)	2.44	213/4616 (4.6%)
1	C	3.23	401/3395 (11.8%)	2.43	229/4616 (5.0%)
1	D	3.49	459/3395 (13.5%)	2.45	239/4616 (5.2%)
1	E	3.26	400/3395 (11.8%)	2.38	205/4616 (4.4%)
1	F	3.37	419/3395 (12.3%)	2.38	207/4616 (4.5%)
1	G	3.28	410/3395 (12.1%)	2.37	203/4616 (4.4%)
1	H	3.35	416/3395 (12.3%)	2.43	210/4616 (4.5%)
1	I	3.27	392/3395 (11.5%)	2.42	216/4616 (4.7%)
1	J	3.29	402/3395 (11.8%)	2.47	209/4616 (4.5%)
1	K	3.27	407/3395 (12.0%)	2.41	232/4616 (5.0%)
1	L	3.26	401/3395 (11.8%)	2.50	235/4616 (5.1%)
1	M	3.22	375/3395 (11.0%)	2.52	237/4616 (5.1%)
1	N	3.28	393/3395 (11.6%)	2.48	243/4616 (5.3%)
1	O	3.18	368/3395 (10.8%)	2.25	171/4616 (3.7%)
All	All	3.29	6037/50925 (11.9%)	2.43	3280/69240 (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	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 6037 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	E	471	GLN	CA-C	20.37	1.65	1.52
1	H	139	ALA	CA-CB	19.73	1.78	1.53
1	H	178	VAL	C-O	19.33	1.47	1.24
1	K	245	SER	CB-OG	19.26	1.80	1.42
1	F	139	ALA	CA-CB	19.14	1.79	1.52

The worst 5 of 3280 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	L	185	CYS	CA-C-N	21.35	135.03	119.66
1	L	185	CYS	C-N-CA	21.35	135.03	119.66
1	B	202	ASP	N-CA-C	-17.40	87.56	110.53
1	E	185	CYS	CA-C-N	17.27	132.26	119.66
1	E	185	CYS	C-N-CA	17.27	132.26	119.66

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

5.2 Too-close contacts

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	580	0
1	B	3310	0	3213	569	0
1	C	3310	0	3213	543	0
1	D	3310	0	3213	570	11
1	E	3310	0	3213	528	5
1	F	3310	0	3213	576	5
1	G	3310	0	3213	608	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	550	5
1	I	3310	0	3213	589	26
1	J	3310	0	3213	569	26
1	K	3310	0	3213	553	0
1	L	3310	0	3213	544	0
1	M	3310	0	3213	555	5
1	N	3310	0	3213	551	11
1	O	3310	0	3213	555	0
All	All	49650	0	48195	7582	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 77.

The worst 5 of 7582 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:CG1	1:D:115:VAL:CB	1.75	1.65
1:B:82:LYS:CB	1:B:82:LYS:CG	1.75	1.64
1:I:88:THR:CB	1:I:88:THR:CG2	1.76	1.64
1:F:176:THR:CB	1:F:176:THR:CG2	1.74	1.64
1:O:331:VAL:CB	1:O:331:VAL:CG1	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

5.3 Torsion angles

5.3.1 Protein backbone

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	12
1	B	415/424 (98%)	327 (79%)	64 (15%)	24 (6%)	1	15
1	C	415/424 (98%)	322 (78%)	63 (15%)	30 (7%)	1	12
1	D	415/424 (98%)	309 (74%)	76 (18%)	30 (7%)	1	12
1	E	415/424 (98%)	321 (77%)	66 (16%)	28 (7%)	1	13
1	F	415/424 (98%)	325 (78%)	63 (15%)	27 (6%)	1	14
1	G	415/424 (98%)	310 (75%)	71 (17%)	34 (8%)	0	9
1	H	415/424 (98%)	322 (78%)	59 (14%)	34 (8%)	0	9
1	I	415/424 (98%)	321 (77%)	64 (15%)	30 (7%)	1	12
1	J	415/424 (98%)	316 (76%)	65 (16%)	34 (8%)	0	9
1	K	415/424 (98%)	322 (78%)	66 (16%)	27 (6%)	1	14
1	L	415/424 (98%)	321 (77%)	67 (16%)	27 (6%)	1	14
1	M	415/424 (98%)	319 (77%)	72 (17%)	24 (6%)	1	15
1	N	415/424 (98%)	319 (77%)	67 (16%)	29 (7%)	1	12
1	O	415/424 (98%)	325 (78%)	59 (14%)	31 (8%)	1	11
All	All	6225/6360 (98%)	4795 (77%)	991 (16%)	439 (7%)	1	12

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

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	367/368 (100%)	297 (81%)	70 (19%)	1	10
1	B	367/368 (100%)	292 (80%)	75 (20%)	1	8
1	C	367/368 (100%)	296 (81%)	71 (19%)	1	9
1	D	367/368 (100%)	302 (82%)	65 (18%)	2	12
1	E	367/368 (100%)	297 (81%)	70 (19%)	1	10
1	F	367/368 (100%)	296 (81%)	71 (19%)	1	9
1	G	367/368 (100%)	306 (83%)	61 (17%)	2	14
1	H	367/368 (100%)	298 (81%)	69 (19%)	1	10
1	I	367/368 (100%)	305 (83%)	62 (17%)	2	13
1	J	367/368 (100%)	308 (84%)	59 (16%)	2	14
1	K	367/368 (100%)	296 (81%)	71 (19%)	1	9
1	L	367/368 (100%)	300 (82%)	67 (18%)	2	11
1	M	367/368 (100%)	298 (81%)	69 (19%)	1	10
1	N	367/368 (100%)	295 (80%)	72 (20%)	1	9
1	O	367/368 (100%)	309 (84%)	58 (16%)	2	15
All	All	5505/5520 (100%)	4495 (82%)	1010 (18%)	2	11

5 of 1010 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	G	293	PRO
1	N	66	SER
1	I	156	LEU
1	M	464	LEU
1	N	402	TRP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 267 such sidechains are listed below:

Mol	Chain	Res	Type
1	N	32	ASN
1	N	153	GLN
1	O	305	GLN
1	F	259	HIS
1	F	155	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 oligosaccharides in this entry.

5.6 Ligand geometry [i](#)

There are no ligands in this entry.

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

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%)	-0.35	2 (0%) 87 69	40, 77, 111, 122	0
1	B	419/424 (98%)	-0.36	1 (0%) 91 80	36, 76, 109, 131	0
1	C	419/424 (98%)	-0.36	0 100 100	44, 79, 110, 121	0
1	D	419/424 (98%)	-0.27	5 (1%) 76 51	37, 78, 109, 130	0
1	E	419/424 (98%)	-0.40	1 (0%) 91 80	41, 79, 114, 128	0
1	F	419/424 (98%)	-0.31	1 (0%) 91 80	40, 80, 111, 131	0
1	G	419/424 (98%)	-0.36	1 (0%) 91 80	46, 83, 112, 135	0
1	H	419/424 (98%)	-0.34	4 (0%) 79 55	34, 78, 111, 136	0
1	I	419/424 (98%)	-0.34	2 (0%) 87 69	48, 82, 113, 132	0
1	J	419/424 (98%)	-0.28	2 (0%) 87 69	43, 79, 111, 134	0
1	K	419/424 (98%)	-0.38	1 (0%) 91 80	35, 76, 109, 129	0
1	L	419/424 (98%)	-0.36	1 (0%) 91 80	38, 73, 108, 137	0
1	M	419/424 (98%)	-0.26	2 (0%) 87 69	43, 75, 108, 133	0
1	N	419/424 (98%)	-0.28	3 (0%) 84 63	45, 80, 112, 128	0
1	O	419/424 (98%)	-0.39	2 (0%) 87 69	45, 82, 112, 133	0
All	All	6285/6360 (98%)	-0.33	28 (0%) 88 72	34, 78, 111, 137	0

The worst 5 of 28 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	I	176	THR	5.9
1	H	179	ALA	5.0
1	I	174	PRO	4.4
1	N	114	GLY	4.4
1	N	163	PRO	3.3

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

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

6.3 Carbohydrates [i](#)

There are no oligosaccharides in this entry.

6.4 Ligands [i](#)

There are no ligands in this entry.

6.5 Other polymers [i](#)

There are no such residues in this entry.