wwPDB EM Validation Summary Report (i)

Jun 10, 2025 – 04:56 PM JST

PDB ID	:	7BQX / pdb_00007bqx
EMDB ID	:	EMD-30157
Title	:	Epstein-Barr virus, C5 portal vertex
Authors	:	Li, Z.; Yu, X.
Deposited on	:	2020-03-25
Resolution	:	4.20 Å(reported)

This is a wwPDB EM 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/EMValidationReportHelp
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:

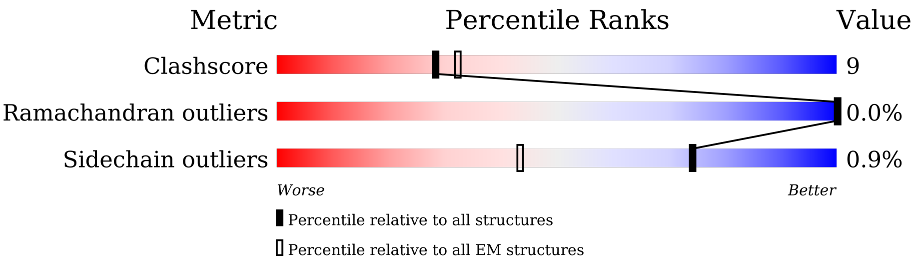
EMDB validation analysis	:	0.0.1.dev118
MolProbity	:	4-5-2 with Phenix2.0rc1
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
MapQ	:	1.9.13
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.43.1

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
ELECTRON MICROSCOPY

The reported resolution of this entry is 4.20 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)	EM structures (#Entries)
Clashscore	210492	15764
Ramachandran outliers	207382	16835
Sidechain outliers	206894	16415

The table below summarises the geometric issues observed across the polymeric chains and their fit to the map. The red, orange, yellow and green segments of the 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% The upper red bar (where present) indicates the fraction of residues that have poor fit to the EM map (all-atom inclusion < 40%). The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	5	364	30% / 59% 19% 22%
1	e	364	6% / 74% 13% 12%
2	6	301	35% / 76% 23% •
2	7	301	33% / 72% 23% 5%
2	f	301	14% / 72% 24% •
2	g	301	16% / 75% 21% •
3	G	570	8% / 12% • 85%
3	K	570	6% / 10% • 88%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
4	C	507	36% 51% 17% 32%
5	B	3149	99%
5	O	3149	99%
6	2	176	38% 34% 7% 58%
6	Y	176	23% 32% 10% 58%
6	Z	176	26% 31% 11% 58%
6	y	176	27% 30% 11% 58%
7	S	1381	16% 73% 20% 6%
7	T	1381	13% 74% 20% 6%
7	W	1381	15% 76% 20%
7	x	1381	15% 72% 20% 7%

2 Entry composition (i)

There are 7 unique types of molecules in this entry. The entry contains 61847 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, 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 Triplex capsid protein 1.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
1	5	284	2199	1421	374	397	7	0	0
1	e	319	2505	1608	444	446	7	0	0

- Molecule 2 is a protein called Triplex capsid protein 2.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
2	6	297	2329	1496	385	430	18	0	0
2	7	287	2262	1452	375	417	18	0	0
2	f	290	2279	1466	378	419	16	0	0
2	g	290	2272	1464	376	414	18	0	0

- Molecule 3 is a protein called Capsid vertex component 2.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
3	K	68	563	343	113	104	3	0	0
3	G	84	675	412	135	125	3	0	0

- Molecule 4 is a protein called Capsid vertex component 1.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
4	C	346	2705	1733	482	478	12	0	0

- Molecule 5 is a protein called Large tegument protein deneddylase.

Mol	Chain	Residues	Atoms				AltConf	Trace
5	B	35	Total 301	C 192	N 61	O 48	0	0
5	O	29	Total 252	C 163	N 49	O 40	0	0

- Molecule 6 is a protein called Small capsomere-interacting protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
6	Y	74	Total 621	C 394	N 114	O 112	S 1	0	0
6	Z	74	Total 621	C 394	N 114	O 112	S 1	0	0
6	2	74	Total 621	C 394	N 114	O 112	S 1	0	0
6	y	74	Total 621	C 394	N 114	O 112	S 1	0	0

- Molecule 7 is a protein called Major capsid protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
7	S	1292	Total 10166	C 6462	N 1758	O 1889	S 57	0	0
7	T	1305	Total 10258	C 6515	N 1780	O 1904	S 59	0	0
7	W	1330	Total 10456	C 6648	N 1808	O 1941	S 59	0	0
7	x	1290	Total 10141	C 6438	N 1760	O 1884	S 59	0	0

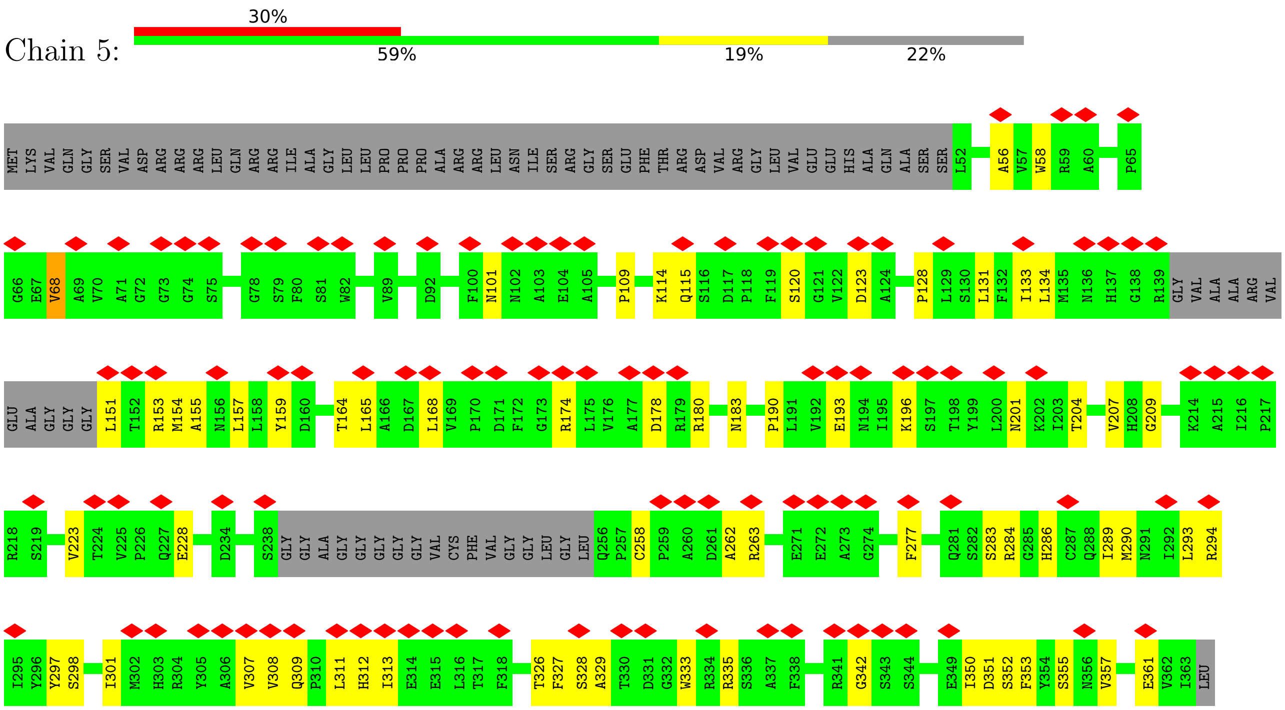
There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
S	89	ILE	THR	conflict	UNP P03226
T	89	ILE	THR	conflict	UNP P03226
W	89	ILE	THR	conflict	UNP P03226
x	89	ILE	THR	conflict	UNP P03226

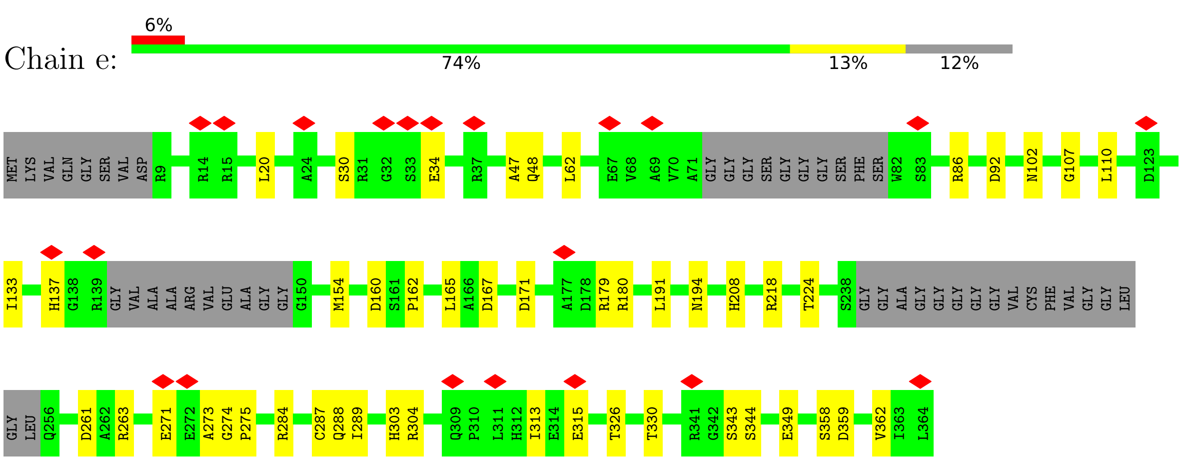
3 Residue-property plots

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 atom inclusion in map 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 diamond above a residue indicates a poor fit to the EM map for this residue (all-atom inclusion < 40%). 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: Triplex capsid protein 1

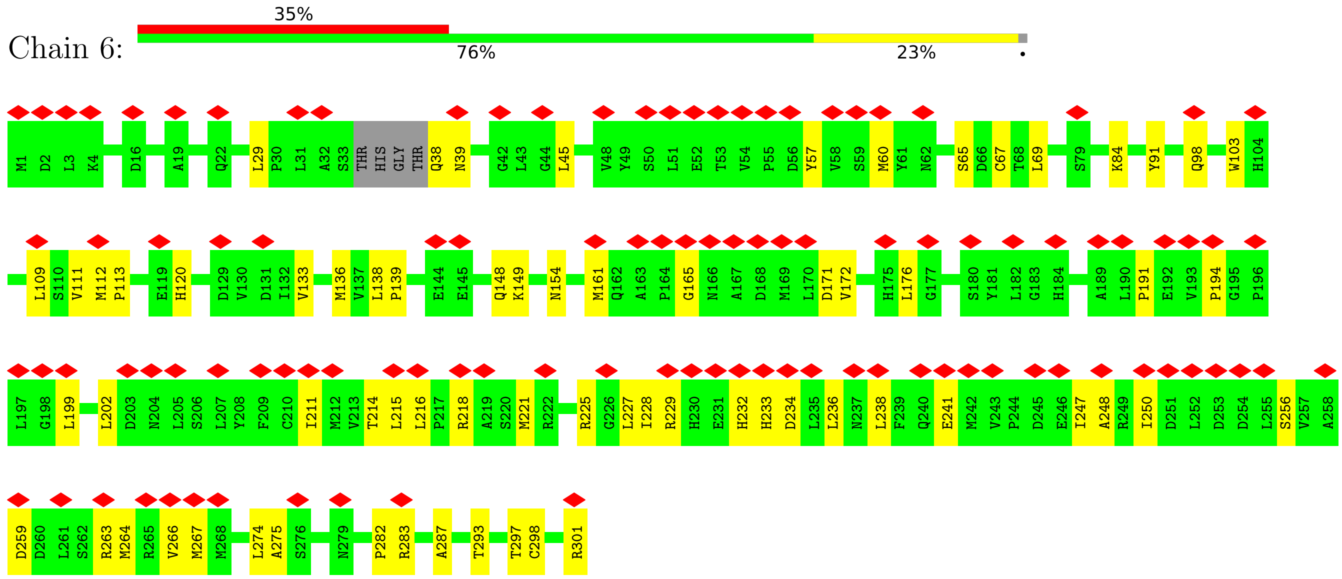


- Molecule 1: Triplex capsid protein 1



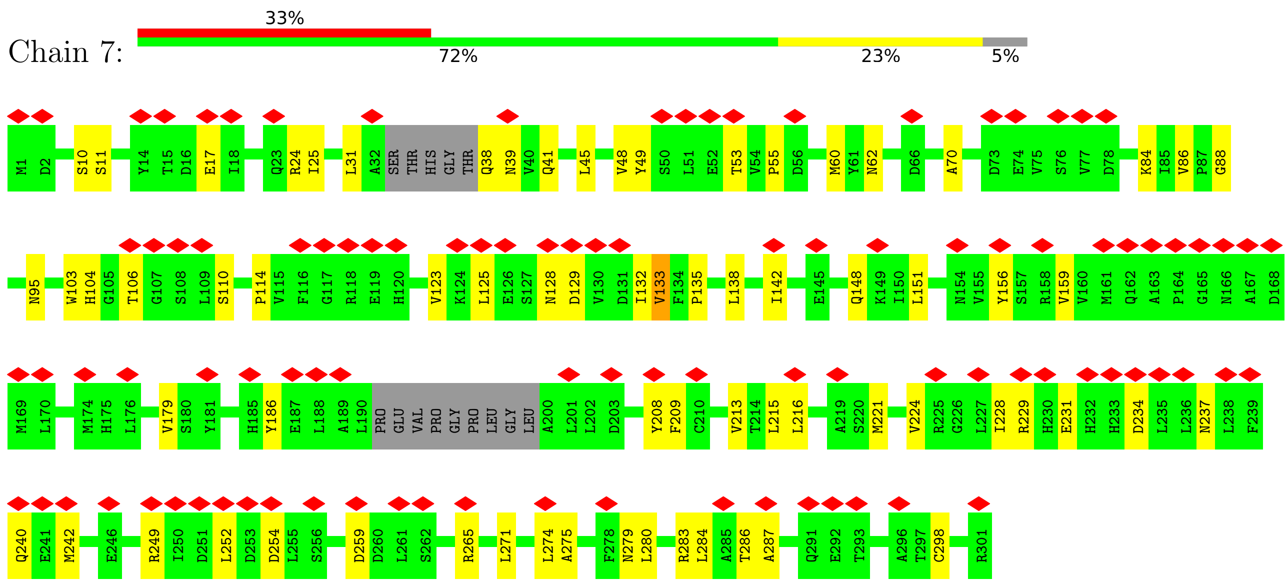
- Molecule 2: Triplex capsid protein 2

Chain 6:



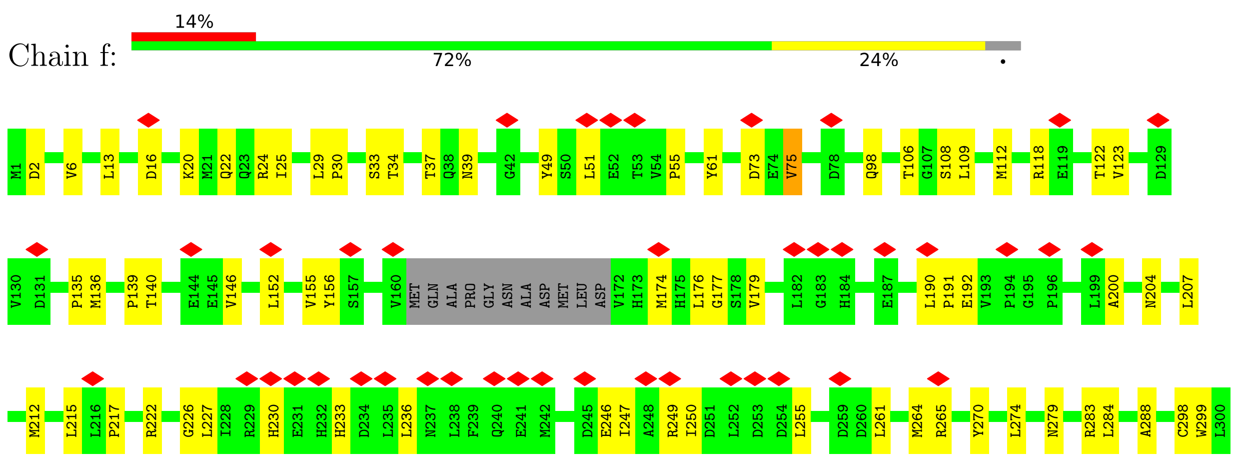
• Molecule 2: Triplex capsid protein 2

Chain 7:



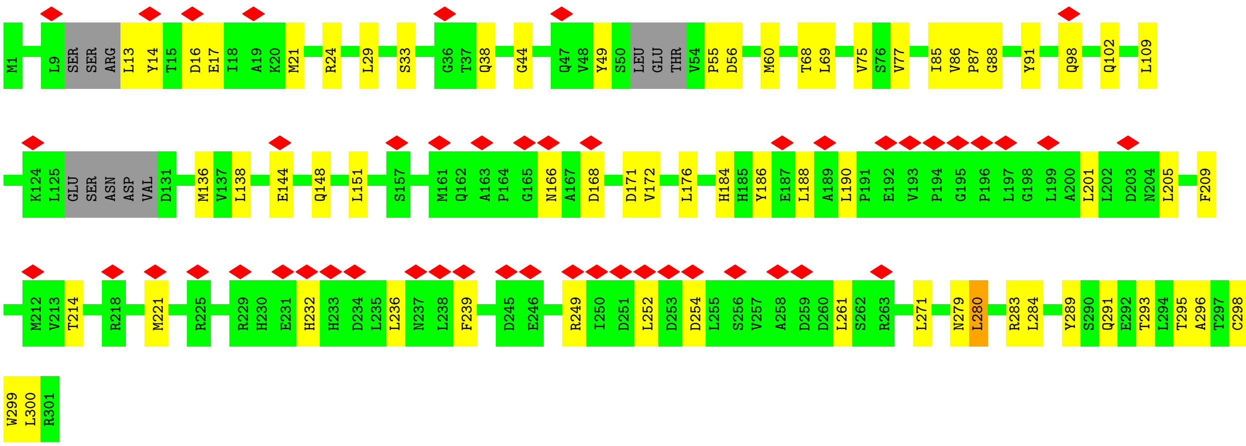
• Molecule 2: Triplex capsid protein 2

Chain f:



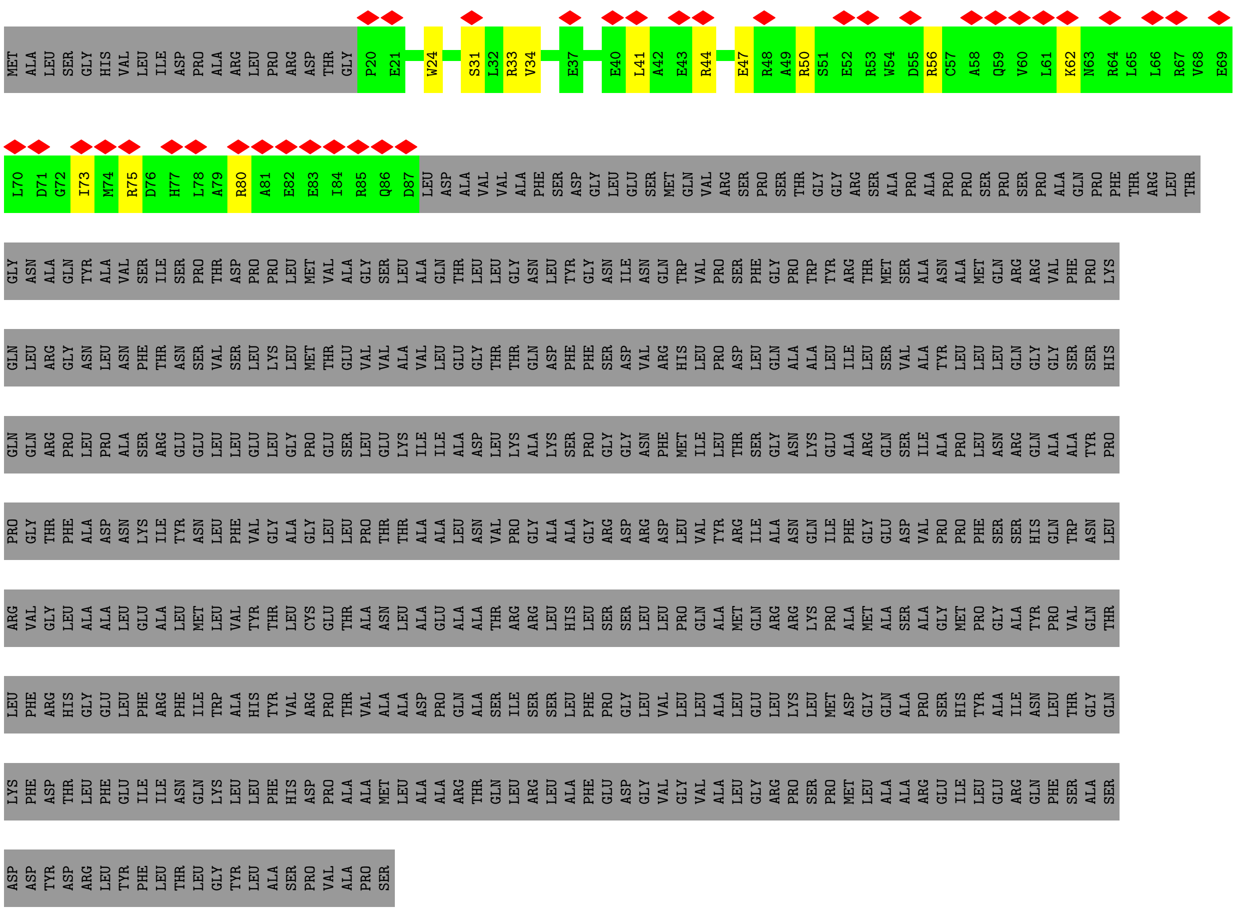
• Molecule 2: Triplex capsid protein 2

Chain g:



• Molecule 3: Capsid vertex component 2

Chain K:



• Molecule 3: Capsid vertex component 2

Chain G: 8% 12% • 85%

MET ALA LEU SER GLY HIS VAL LEU ILE D10 L14 P15 R16 W24 V34 E37 A38 L39 E40 L41 A42 E43 R44 E45 A46 E47 R48 A49 R50 S51 E52 R53 W54 D55 R56 C57 A58 L61 K62 N63 R64 L65 L66 R67 V68 E69 L70 D71 G72 I73 M74 R75 D76 H77 L78

A79 R80 A81 E82 E83 I84 R85 Q86 D87 L88 D89 A90 V91 V92 A93 PHE SER ASP GLY LEU GLU SER MET GLN VAL ARG SER PRO SER THR GLY GLY ARG SER ALA PRO ALA PRO PRO SER PRO SER PRO ALA GLN PRO PHE THR ARG LEU THR GLY ASN ALA GLN TYR ALA VAL SER ILE

SER PRO THR ASP PRO PRO LEU MET VAL ALA GLY SER LEU ALA GLN THR LEU LEU GLY ASN LEU TYR GLY ASN ILE ASN GLN TRP VAL PRO SER PHE GLY PRO TRP TYR ARG THR MET SER ALA ASN ALA MET GLN ARG ARG VAL PHE PRO LYS GLN LEU ARG GLY ASN LEU ASN PHE THR

ASN SER VAL SER LEU LYS LEU MET THR GLU VAL VAL ALA VAL LEU GLU GLY THR THR GLN ASP PHE PHE SER ASP VAL ARG HIS LEU PRO ASP LEU GLN ALA ALA LEU ILE LEU SER VAL ALA TYR LEU LEU LEU GLN GLY GLY SER SER HIS GLN GLN ARG PRO LEU PRO ALA SER ARG

GLU GLU LEU LEU GLU LEU GLY PRO GLU SER LEU GLU LYS ILE ILE ALA ASP LEU LYS ALA LYS SER PRO GLY GLY ASN PHE MET ILE LEU THR SER GLY ASN LYS GLU ALA ARG GLN SER ILE ALA PRO LEU ASN ARG GLN ALA ALA TYR PRO PRO GLY THR PHE ALA ASP ASN LYS ILE

TYR ASN LEU PHE VAL GLY ALA GLY LEU LEU PRO THR THR ALA ALA LEU ASN VAL PRO GLY ALA ALA GLY ARG ASP ARG ASP LEU VAL TYR ARG ILE ALA ASN GLN ILE PHE GLY GLU ASP VAL PRO PRO PHE SER SER HIS GLN TRP ASN LEU ARG VAL GLY LEU ALA ALA LEU GLU ALA

LEU MET LEU VAL TYR THR LEU CYS GLU THR ALA ASN LEU ALA GLU ALA ALA THR ARG ARG LEU HIS LEU SER SER LEU LEU PRO GLN ALA MET GLN ARG ARG LYS PRO ALA MET ALA SER ALA GLY MET PRO GLY ALA TYR PRO VAL GLN THR LEU PHE ARG HIS GLY GLU LEU PHE ARG

PHE ILE TRP ALA HIS TYR VAL ARG PRO THR VAL ALA ALA ASP PRO GLN ALA SER ILE SER SER LEU PHE PRO GLY LEU VAL LEU LEU ALA LEU GLU LEU LYS LEU MET ASP GLY GLN ALA PRO SER HIS TYR ALA ILE ASN LEU THR GLY GLN LYS PHE ASP THR LEU PHE GLU ILE ILE

ASN GLN LYS LEU LEU PHE HIS ASP PRO ALA ALA MET LEU ALA ALA ARG THR GLN LEU ARG LEU ALA PHE GLU ASP GLY VAL GLY VAL ALA LEU GLY ARG PRO SER PRO MET LEU ALA ALA ARG GLU ILE LEU GLU ARG GLN PHE SER ALA SER ASP ASP TYR ASP ARG LEU TYR PHE LEU

THR LEU GLY TYR LEU ALA SER PRO VAL ALA PRO SER

• Molecule 4: Capsid vertex component 1

Chain C: 36% 51% 17% 32%

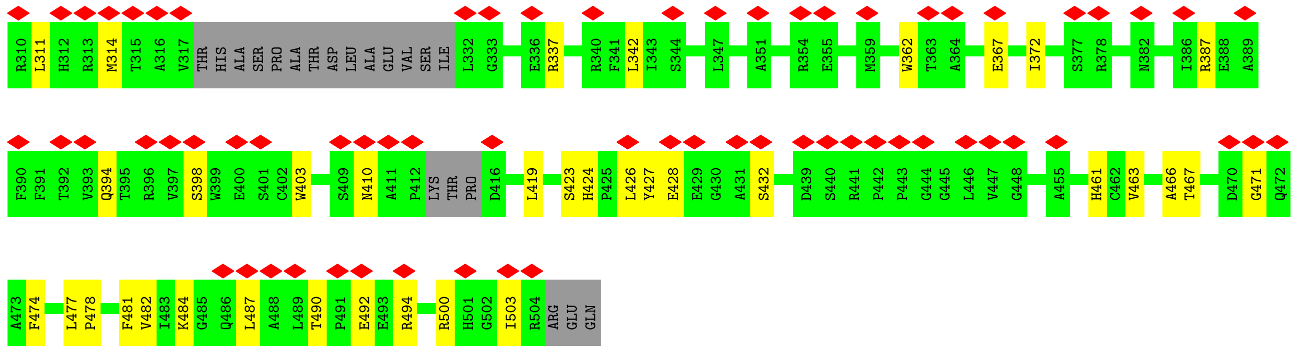
M1 D2 V3 H4 I5 D6 N7 Q8 V9 L10 S11 G12 L13 G14 T15 P16 L17 L18 L21 P24 D25 T26 V27 M28 A29 E30 L31 C32 P33 N34 R35 V36 P37 N38 CYS GLU GLY A42 Q45 T46 L47 F48 S49 D50 R51 T52 G53 L54 T55 R56 V60 F61 A62 A63 R64 G65

M66 L67 P68 G69 R70 P71 S72 H73 R74 G75 T76 F77 T78 S79 V80 P81 V82 Y83 C84 D85 E86 G87 L88 P89 E90 L91 Y92 N93 P94 F95 A98 A99 L100 R101 D104 E105 GLY GLY L108 V109 G110 E111 L112 Q113 I114 Y115 Y116 L117 S118 E121 G122 A123 K124 R125 A126 L127 THR

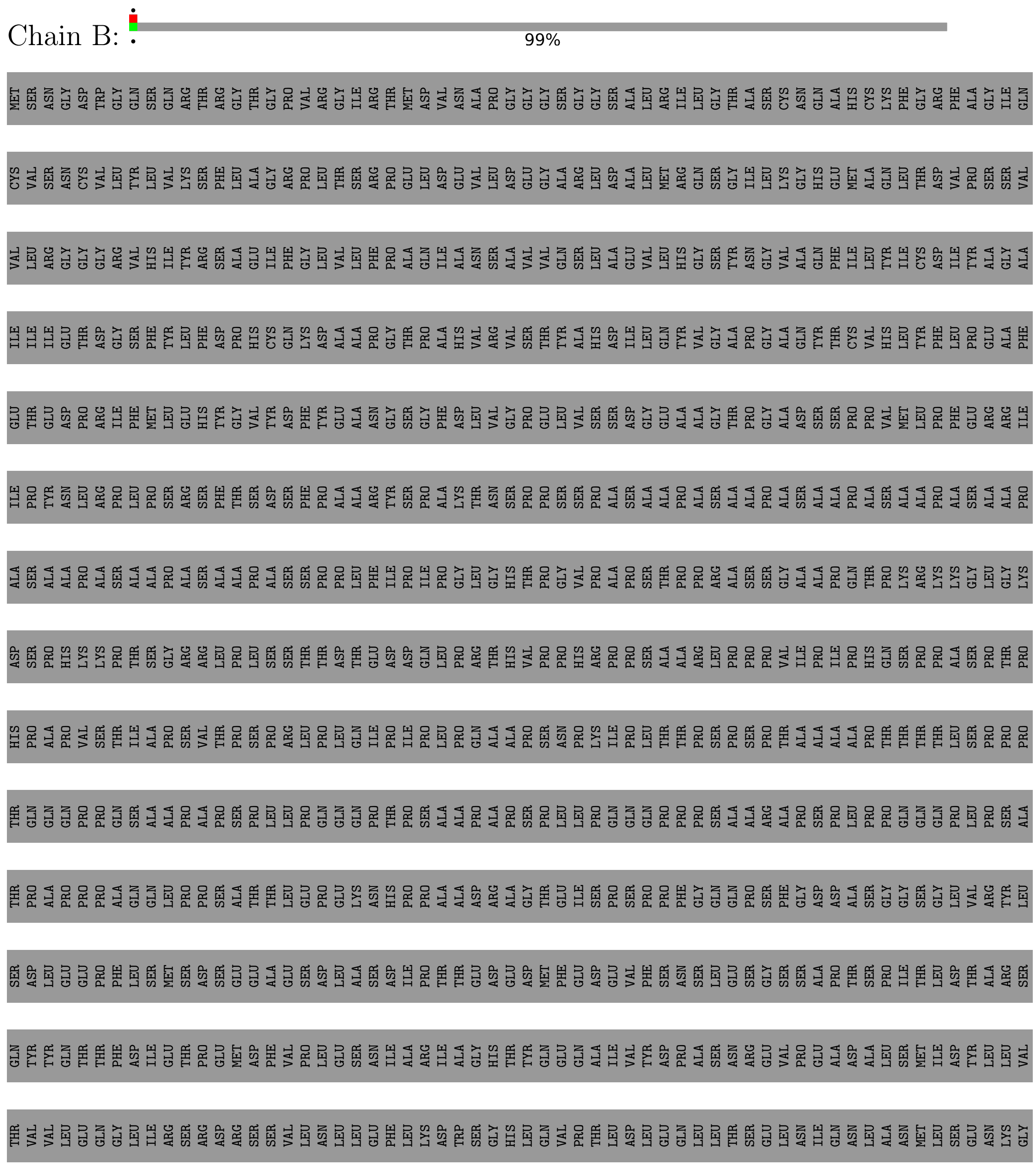
ASP GLY HIS LEU ILE ARG GLU ALA SER GLY VAL GLN GLU SER ALA ALA ALA MET GLN PRO ILE PRO ILE ASP PRO GLY PRO PRO GLY ALA GLY ILE GLU HIS MET PRO VAL ALA ALA ALA GLN VAL GLU HIS PRO LYS THR TYR ASP LEU LYS GLN ILE LEU LEU GLU ILE THR GLN

GLU GLU ASN ARG GLY GLU GLN ARG LEU GLY HIS ALA GLY SER PRO ALA LEU CYS LEU GLY LEU ARG LEU ARG ALA GLY ALA GLU THR LYS ALA ALA ALA GLU THR SER VAL SER LYS HIS HIS PRO ALA LEU GLU ASN PRO SER ASN ILE ARG GLY SER ALA GLY GLY GLU GLY GLY GLY

GLY ARG ALA GLY THR GLY GLY THR VAL GLY VAL GLY SER GLY ALA L264 S265 R266 V267 P268 V269 S270 F271 S272 K273 T274 R275 R276 A277 I278 R279 E280 S281 R282 A283 L284 V285 R286 G287 I288 A289 H290 I291 P294 H295 A296 L297 Y298 V299 V300 T301 Y302 P303 E304 L305 S306 A307 Q308 G309



• Molecule 5: Large tegument protein deneddylase



ARG ALA GLY GLU PHE HIS LYS HIS LEU ALA ALA LYS LEU GLU ALA CYS LEU PRO SER LEU ALA THR LYS ASP ALA VAL ARG VAL ASP ALA GLY ALA LYS MET LEU ALA GLU ILE PRO GLN LEU ALA GLU SER ASP ASP GLY LYS PHE ASP LEU GLU ALA ALA ARG ARG ARG LEU THR ASP

LEU LEU SER GLY GLY ASP GLN GLU ALA GLY GLU GLY GLY GLY GLU PRO GLU ASP ASN SER ILE TYR ARG GLY PRO HIS VAL ASP VAL PRO LEU VAL LEU ASP ASP GLU SER TRP LYS ARG LEU LEU SER LEU ALA GLU ALA ALA ARG THR ALA VAL ALA ARG GLN GLN ALA GLY VAL ASP

GLU GLU ASP VAL ARG PHE LEU ALA LEU LEU THR ALA ILE GLU TYR GLY ALA PRO PRO ALA ALA SER VAL PRO PRO PHE VAL HIS ASN VAL ALA VAL ARG SER LYS ASN ALA ALA LEU HIS VAL ARG ARG CYS THR ALA ASP ILE ARG ASP LYS VAL ALA SER ALA ALA SER ASP TYR LEU

SER TYR LEU GLU ASP PRO SER LEU PRO THR VAL MET ASP PHE ASP ASP LEU LEU THR HIS LEU ARG HIS THR CYS GLN ILE ILE ALA SER LEU PRO LEU LEU ASN ILE ARG TYR THR SER ILE GLU TRP ASP TYR ARG GLU LEU LEU TYR LEU GLY THR ALA LEU SER ASP MET SER GLY

ILE PRO TRP PRO LEU GLU ARG VAL GLU GLU ASP ASP PRO SER ILE ALA PRO LEU PRO GLU PHE GLU THR VAL ALA LYS LYS GLN LYS GLU LEU GLU THR THR ARG GLU ASN GLU LYS ARG LEU ARG THR ILE LEU ASP ASP ILE GLU ALA MET LEU GLY LEU ALA GLY VAL ALA SER ALA

PRO GLY ALA PRO ILE SER PRO ALA SER PRO SER ALA THR PRO ALA ASN HIS ASP ASN PRO GLU ALA THR PRO PRO LEU ALA ASP THR ALA ALA LEU THR ILE PRO VAL ILE GLU LYS TYR ILE ALA ASN ALA GLY SER ILE VAL GLY ALA ALA LYS ASN PRO THR TYR ILE ARG LEU ARG

ASP THR ILE GLN GLN ILE VAL ARG SER LYS LYS TYR LEU MET ASN ILE LEU LYS SER ILE THR PHE TYR THR ILE ASP ASN TYR ILE ALA SER PHE GLU GLU SER ILE ASP HIS LEU TYR ARG ASP LEU PRO VAL LEU ASP PRO GLU VAL GLN ASP GLY ILE ASP ARG ILE LEU ASP PRO

MET VAL SER GLU ALA LEU HIS THR PHE GLU MET GLY ASN ARG LEU THR LEU GLU PRO ALA ARG LEU VAL ALA LEU GLN ASN PHE ALA THR HIS SER THR LEU LYS GLU THR ALA ALA ALA VAL ASN LEU LEU PRO GLY LEU LEU ALA VAL TYR ASP ALA THR ILE THR GLY GLN ALA PRO

GLU ASP ALA LEU ARG LEU LEU SER GLY LEU GLN ASN GLN LEU SER GLN THR LEU ILE PRO GLY LYS LEU LYS LYS ARG PHE LEU SER TYR LEU GLN LYS LEU LYS ASN ASN ASN ASN ASP GLN LEU ARG GLN LYS GLU VAL GLN ALA TRP ARG LEU GLU ALA GLU GLY PHE LYS PRO ALA

THR GLU GLU GLN LEU GLU ALA PHE LEU ASP THR ALA PRO ASN LYS GLU LEU LYS ARG GLN TYR GLU LYS LYS LEU ARG GLN LEU MET GLU THR GLY ARG LYS GLU LYS GLU LYS LEU ARG GLU GLN GLU ASP LYS GLU ARG GLN GLU ARG ARG ALA ARG GLU ALA ASN GLU ALA TRP ALA

ARG ILE ARG LYS ALA LEU GLY ALA ARG PRO GLU PRO ALA PRO THR SER PRO ASP ASP TRP ASN THR LEU LEU ALA SER LEU LEU PRO ASP ASN THR ASP SER ALA ALA ALA ALA ALA ALA ALA VAL ALA ARG ASN THR ASP ILE LEU ASP SER LEU THR GLN ILE LEU ALA ALA MET LEU

LEU GLY ILE THR ARG VAL ARG ARG GLU ARG LEU ARG SER LEU LEU VAL ASP ASP GLY GLY ALA ALA GLU ARG MET GLU ALA ALA GLU PRO GLY TRP PHE THR ASP ILE GLU THR GLY PRO LEU ALA ARG LEU ASP ALA TRP PRO ALA THR PRO ALA ALA THR ALA LYS GLU GLY GLY GLY

GLY ARG GLY ALA GLU GLU ALA ALA GLY ALA LEU PHE ARG ALA ARG THR ALA ALA ASP ALA ILE ARG SER ALA LEU ALA GLN THR ARG GLN ALA LEU GLN SER PRO ASP MET LYS SER ALA VAL VAL ASN THR ASP LEU GLU ALA PRO TYR ALA GLU TYR GLU ARG GLY LEU ALA GLY LEU

LEU GLU LYS ARG ARG ALA ALA GLU ALA ALA LEU THR ALA ILE VAL SER GLU TYR VAL ASP ARG THR LEU PRO GLU ALA THR ASN ASP PRO GLY GLN ALA ASN LEU PRO PRO PRO PRO THR ILE PRO GLN ALA THR ALA PRO PRO ARG LEU ALA SER ASP SER ALA LEU TRP PRO LYS LYS

PRO GLN LEU LEU THR ARG ARG GLU ARG ASP ASP LEU LEU GLN ALA THR GLY ASP PHE PHE SER GLU LEU LEU THR GLU ALA GLU ALA ALA GLU VAL ARG ALA LEU GLU GLU GLN VAL ARG GLU SER GLN THR LEU MET ALA LYS ALA HIS GLU MET ALA ALA SER THR ARG ARG GLY PHE

HIS THR ALA LEU GLU ALA VAL LEU SER ARG SER ARG ASP GLU ALA PRO ASP ASP GLU LEU ARG SER LEU LEU PRO SER PRO PRO LYS ALA PRO VAL GLN ALA PRO LEU GLU ALA ALA LEU ALA ARG ALA ALA ALA GLY ASN GLY SER TRP PRO TYR ARG LYS SER LEU ALA ALA ALA LYS

TRP ILE ARG GLY ILE CYS GLU ALA VAL ARG GLY LEU SER GLU GLY ALA LEU ALA LEU ALA GLY GLY ALA GLY ALA TRP LEU ASN LEU ALA ALA ALA ALA ASP GLY GLU ILE HIS GLU LEU THR ARG LEU LEU GLU VAL GLU GLY MET ALA GLN ASN SER MET ASP GLY MET GLU GLU LEU

ARG LEU ALA LEU ALA THR LEU ASP PRO LYS ARG VAL ALA GLY GLY LYS GLU THR VAL ALA ASP TRP LYS ARG ARG LEU SER ARG LEU GLU ALA ILE ILE GLN GLU ALA GLN GLU GLU SER GLN LEU GLN GLY THR LEU GLN ASP LEU VAL THR GLN ALA ARG GLY HIS THR ASP PRO ARG

GLN LEU LYS ILE VAL VAL GLU ALA ALA ARG GLY LEU ALA LEU GLY ALA SER ALA GLY SER GLN TYR ALA LEU LEU LYS ASP LYS LEU LEU ARG TYR ALA SER ALA LYS GLN SER PHE LEU ALA PHE TYR GLU THR ALA GLN PRO THR VAL PHE VAL LYS HIS PRO LEU THR ASN ASN LEU

PRO LEU LEU ILE THR ILE SER ALA PRO PRO THR GLY TRP GLY ASN GLY ALA PRO THR ARG ARG ALA GLN PHE LEU ALA ALA ALA GLY PRO ALA LYS TYR ALA GLY THR LEU TRP LEU GLU THR GLU SER PRO CYS ASP PRO LEU ASN PRO ALA TYR VAL SER ALA ASP THR GLN GLU PRO

LEU ASN TYR ILE PRO VAL TYR HIS ASN PHE LEU GLU TYR VAL MET PRO THR VAL LEU GLU ASN PRO GLU ALA PHE SER LEU THR PRO ALA GLY ARG PRO GLN ALA ILE GLY PRO PRO GLN ASP ASP GLN GLU ARG ARG ARG ARG THR LEU ALA SER VAL ALA SER ALA ARG LEU SER ALA

ALA ALA ALA ASP SER TYR TRP ASP THR TRP PRO ASP VAL GLU SER ASN ALA GLY GLU LEU LEU ARG GLU TYR VAL SER ALA PRO LYS ALA LEU MET GLU ASP LEU ALA ASP ASN PRO ILE VAL ALA MET THR LEU LEU ALA HIS ALA SER LEU ILE ALA SER ARG ASN HIS PRO PRO TYR

PRO ALA PRO ALA THR ASP ARG GLU VAL ILE LEU LEU GLU GLN ARG GLU MET MET ALA LEU LEU VAL GLY THR HIS PRO ALA TYR ALA ALA ALA PHE LEU GLY ALA PRO SER PHE TYR ALA GLY LEU GLY LEU VAL SER ALA LEU ALA ARG ASP GLY GLY LEU GLY ASP LEU LEU SER ASP

SER VAL LEU THR TYR ARG LEU VAL ARG SER PRO ALA SER GLY ARG GLY GLY MET PRO SER THR THR ARG GLY SER ASN ASP GLY GLU ASP ALA ARG ARG LEU THR ARG HIS ARG ILE ALA GLY PRO PRO THR GLY PHE ILE PHE PHE GLN ASP ALA TRP GLU GLU MET ASP THR ARG ALA

ALA LEU TRP PRO HIS PRO GLU PHE LEU GLY LEU VAL HIS ASN GLN SER THR ALA ARG ALA ARG ALA CYS MET LEU LEU LEU ALA ARG ARG CYS PHE ALA PRO GLU ALA LEU GLN GLN LEU TRP HIS SER LEU ARG PRO LEU GLU GLY PRO VAL ALA PHE GLN ASP TYR LEU ARG ASP PHE

VAL LYS GLN ALA TYR THR ARG GLY GLU GLU LEU PRO ARG ALA GLU GLY LEU GLU VAL PRO ARG GLU THR PRO SER SER TYR GLY THR VAL THR GLY ARG ALA LEU ARG ASN LEU MET PRO TYR GLY THR PRO ILE THR GLY PRO LYS ARG GLY SER GLY ASP THR ILE PRO VAL SER VAL

PHE GLU ALA ALA VAL ALA ALA ALA PHE LEU GLY ARG PRO LEU THR LEU PHE VAL SER SER GLN TYR LEU PHE ASN LEU LYS THR LEU GLY GLN VAL ARG VAL VAL ALA PRO LEU LEU TYR CYS ASP GLY HIS SER GLU PRO PHE ARG SER LEU VAL GLU THR ILE SER LEU ASN PHE LEU

GLN ASP LEU ASP GLY TYR SER GLU SER PHE GLU PRO GLU MET SER ILE PHE ALA ARG GLN ALA VAL TRP LEU ARG GLU LEU LEU THR GLU ALA ARG ALA ALA LYS PRO LYS GLU ALA ARG PRO PRO THR VAL ALA ILE LEU ALA ASN ARG LYS ASN ILE ILE TRP LYS CYS PHE THR TYR

ARG HIS ASN LEU PRO ASP VAL GLN PHE TYR PHE ASN ALA ALA GLY ALA SER ARG TRP PRO THR ASP VAL LEU ASN PRO SER PHE TYR GLU HIS GLU ASP PRO PRO LEU PRO VAL GLY TYR GLN LEU PRO PRO ASN PRO ARG ASN VAL GLN GLU LEU PHE SER GLY PHE PRO PRO ARG VAL

GLY HIS GLY LEU VAL SER GLY ASP GLY PHE GLN SER ALA ASP ASN THR PRO ALA SER SER ASP ARG LEU GLN GLN LEU GLY GLY GLY GLU THR ASP GLN GLY GLU LYS GLY SER THR THR ALA GLU SER GLU ALA SER GLY PRO PRO SER PRO GLN SER PRO LEU LEU GLU LYS VAL ALA

PRO GLY ARG PRO ARG ASP TRP LEU SER PRO THR SER SER PRO ARG ASP VAL THR VAL THR PRO GLY LEU ALA ALA PRO ILE THR LEU PRO GLY PRO ARG LEU MET ALA ARG PRO TYR PHE GLY ALA GLU THR ARG ALA SER GLU SER PRO ASP ARG SER PRO GLY SER SER PRO ARG PRO

TRP PRO LYS ASP SER LEU GLU LEU LEU PRO GLN PRO ALA PRO GLN GLN PRO PRO SER SER PRO TRP ALA SER GLU GLN GLY PRO ILE VAL TYR THR LEU SER PRO HIS SER THR PRO SER THR ALA SER GLY SER GLN LYS LYS HIS THR ILE GLN ILE PRO GLY LEU VAL PRO SER GLN

LYS PRO SER TYR PRO PRO SER ALA PRO TYR LYS PRO GLY GLN SER THR GLY GLY ILE ALA PRO THR PRO SER ALA ALA SER LEU THR THR PHE GLY LEU GLN PRO GLN ASP THR GLN ALA SER SER GLN ASP PRO PRO TYR GLY HIS SER ILE MET GLN ARG GLU LYS LYS GLN GLN GLY

GLY ARG GLU GLU ALA ALA GLU ILE ARG PRO SER ALA THR ARG LEU PRO THR ALA VAL GLY LEU ARG PRO ARG ALA PRO VAL VAL ALA ALA GLY ALA ALA ALA SER ALA THR PRO ALA PHE ASP PRO GLY GLU ALA PRO SER GLY PHE PRO ILE PRO GLN ALA PRO ALA LEU GLY SER GLY

LEU ALA ALA PRO ALA HIS THR PRO VAL GLY ALA LEU ALA PRO ARG PRO GLN LYS THR GLN ALA GLN ARG PRO GLN ASP ALA ALA ALA LEU PRO THR PRO THR ILE LYS ALA VAL GLY ALA ARG PRO VAL PRO LYS ALA THR GLY ALA LEU ALA ALA GLY ALA ARG PRO ARG GLY GLN PRO

THR ALA ALA PRO PRO SER ALA ALA SER PRO PRO ARG VAL SER LEU PRO VAL ARG SER ARG GLN GLN GLN SER PRO ALA ILE PRO LEU PRO PRO MET HIS SER GLY SER GLU PRO GLY ALA ARG PRO GLU VAL ARG LEU SER GLN TYR ARG HIS ALA GLY PRO GLN THR TYR THR VAL ARG

LYS GLU ALA PRO PRO SER ALA ALA SER GLN LEU PRO LYS MET PRO LYS CYS LYS ASP SER MET TYR TYR PRO PRO SER GLY SER ALA ARG TYR PRO ALA PRO PHE GLN ALA LEU SER PHE SER GLN SER VAL ALA SER PRO ALA PRO SER SER ASP GLN THR THR LEU LEU TRP ASN THR

PRO SER VAL VAL THR GLN PHE LEU SER ILE GLU ASP ILE ILE ARG GLU VAL VAL THR GLY GLY SER THR SER GLY ASP LEU VAL VAL PRO SER GLY SER PRO SER SER LEU SER THR ALA ALA PRO GLU GLN ASP LEU ARG TYR SER LEU THR LEU SER GLN A3115 S3116 R3117 V3118 L3119 S3120

R3121 F3122 V3123 S3124 Q3125 L3126 R3127 R3128 K3129 L3130 E3131 R3132 S3133 R3136 L3137 I3138 A3139 D3140 L3141 E3142 R3143 L3144 K3145 F3146 L3147 Y3148 L3149

• Molecule 5: Large tegument protein deneddylase

Chain O:

99%

MET SER ASN GLY ASP TRP GLY GLN SER GLN ARG THR ARG GLY THR GLY PRO VAL ARG GLY ILE ARG THR MET ASP VAL ASN ALA PRO GLY GLY GLY SER GLY GLY SER ALA LEU ARG ILE LEU GLY THR ALA SER CYS ASN GLN ALA HIS CYS LYS PHE GLY ARG PHE ALA GLY ILE GLN

CYS VAL SER ASN CYS VAL LEU TYR LEU VAL LYS SER PHE LEU ALA GLY ARG PRO LEU THR SER ARG PRO GLU LEU ASP GLU VAL LEU ASP GLU GLY ALA ARG LEU ASP ALA LEU MET ARG GLN SER GLY ILE LEU LYS GLY HIS GLU MET ALA GLN LEU THR ASP VAL PRO SER SER VAL

VAL LEU ARG GLY GLY GLY ARG VAL HIS ILE TYR ARG SER ALA GLU ILE PHE GLY LEU VAL LEU PHE PRO ALA GLN ILE ALA ASN SER ALA VAL VAL GLN SER LEU ALA GLU VAL LEU HIS GLY SER TYR ASN GLY VAL ALA GLN PHE ILE LEU TYR ILE CYS ASP ILE TYR ALA GLY ALA

ILE ILE ILE GLU THR ASP GLY SER PHE TYR LEU PHE ASP PRO HIS CYS GLN LYS ASP ALA ALA PRO GLY THR PRO ALA HIS VAL ARG VAL SER THR TYR ALA HIS ASP ILE LEU GLN TYR VAL GLY ALA PRO GLY ALA GLN TYR THR CYS VAL HIS LEU TYR PHE LEU PRO GLU ALA PHE

GLU THR GLU ASP PRO ARG ILE PHE MET LEU GLU HIS TYR GLY VAL TYR ASP PHE TYR GLU ALA ASN GLY SER GLY PHE ASP LEU VAL GLY PRO GLU LEU VAL SER SER ASP GLY GLU ALA ALA GLY THR PRO GLY ALA ASP SER SER PRO PRO VAL MET LEU PRO PHE GLU ARG ARG ILE

ILE PRO TYR ASN LEU ARG PRO LEU PRO SER ARG SER PHE THR SER ASP SER PHE PRO ALA ALA ARG TYR SER PRO ALA LYS THR ASN SER PRO PRO SER SER PRO ALA SER ALA ALA PRO ALA SER ALA ALA PRO ALA SER ALA ALA PRO ALA SER ALA ALA PRO ALA SER ALA ALA PRO

ALA SER ALA ALA PRO ALA SER ALA ALA PRO ALA SER ALA ALA PRO ALA SER SER PRO PRO LEU PHE ILE PRO ILE PRO GLY LEU GLY HIS THR PRO GLY VAL PRO ALA PRO SER THR PRO PRO ARG ALA SER SER GLY ALA ALA PRO GLN THR PRO LYS ARG LYS LYS GLY LEU GLY LYS

ASP SER PRO HIS LYS LYS PRO THR SER GLY ARG ARG LEU PRO LEU SER SER THR THR ASP THR GLU ASP ASP GLN LEU PRO ARG THR HIS VAL PRO PRO HIS ARG PRO PRO SER ALA ALA ARG LEU PRO PRO PRO VAL ILE PRO ILE PRO HIS GLN SER PRO PRO ALA SER PRO THR PRO

HIS PRO ALA PRO VAL SER THR ILE ALA PRO SER VAL THR PRO SER PRO ARG LEU PRO LEU GLN ILE PRO ILE PRO LEU PRO GLN ALA ALA PRO SER ASN PRO LYS ILE PRO LEU THR THR PRO SER PRO SER PRO THR ALA ALA ALA ALA PRO THR THR THR THR LEU SER PRO PRO PRO

THR GLN GLN GLN PRO PRO GLN SER ALA ALA PRO ALA PRO SER PRO LEU LEU PRO GLN GLN GLN PRO THR PRO SER ALA ALA PRO ALA PRO SER PRO LEU LEU PRO GLN GLN GLN PRO PRO PRO SER ALA ALA ARG ALA PRO SER PRO LEU PRO PRO GLN GLN GLN PRO LEU PRO SER ALA

THR PRO ALA PRO PRO PRO ALA GLN GLN LEU PRO PRO SER ALA THR THR LEU GLU PRO GLU LYS ASN HIS PRO PRO ALA ALA ASP ARG ALA GLY THR GLU ILE SER PRO SER PRO PRO PHE GLY GLN GLN PRO SER PHE GLY ASP ASP ALA SER GLY GLY SER GLY LEU VAL ARG TYR LEU

SER ASP LEU GLU GLU PRO PHE LEU SER MET SER ASP SER GLU GLU ALA GLU SER ASP LEU ALA SER ASP ILE PRO THR THR GLU ASP GLU ASP MET PHE GLU ASP GLU VAL PHE SER ASN SER LEU GLU SER GLY SER SER ALA PRO THR SER PRO ILE THR LEU ASP THR ALA ARG SER

GLN TYR TYR GLN THR THR PHE ASP ILE GLU THR PRO GLU MET ASP PHE VAL PRO LEU GLU SER ASN ILE ALA ARG ILE ALA GLY HIS THR TYR GLN GLU GLN ALA ILE VAL TYR ASP PRO ALA SER ASN ARG GLU VAL PRO GLU ALA ASP ALA LEU SER MET ILE ASP TYR LEU LEU VAL

THR VAL VAL LEU GLU GLN GLY LEU ILE ARG SER ARG ASP ARG SER SER VAL LEU ASN LEU LEU GLU PHE LEU LYS ASP TRP SER GLY HIS LEU GLN VAL PRO THR LEU ASP LEU GLU GLN LEU LEU THR SER GLU LEU ASN ILE GLN ASN LEU ALA ASN MET LEU SER GLU ASN LYS GLY

ARG ALA GLY GLU PHE HIS LYS HIS LEU ALA ALA LYS LEU GLU ALA CYS LEU PRO SER LEU ALA THR LYS ASP ALA VAL ARG VAL ASP ALA GLY ALA LYS MET LEU ALA GLU ILE PRO GLN LEU ALA GLU SER ASP ASP GLY LYS PHE ASP LEU GLU ALA ALA ARG ARG ARG LEU THR ASP

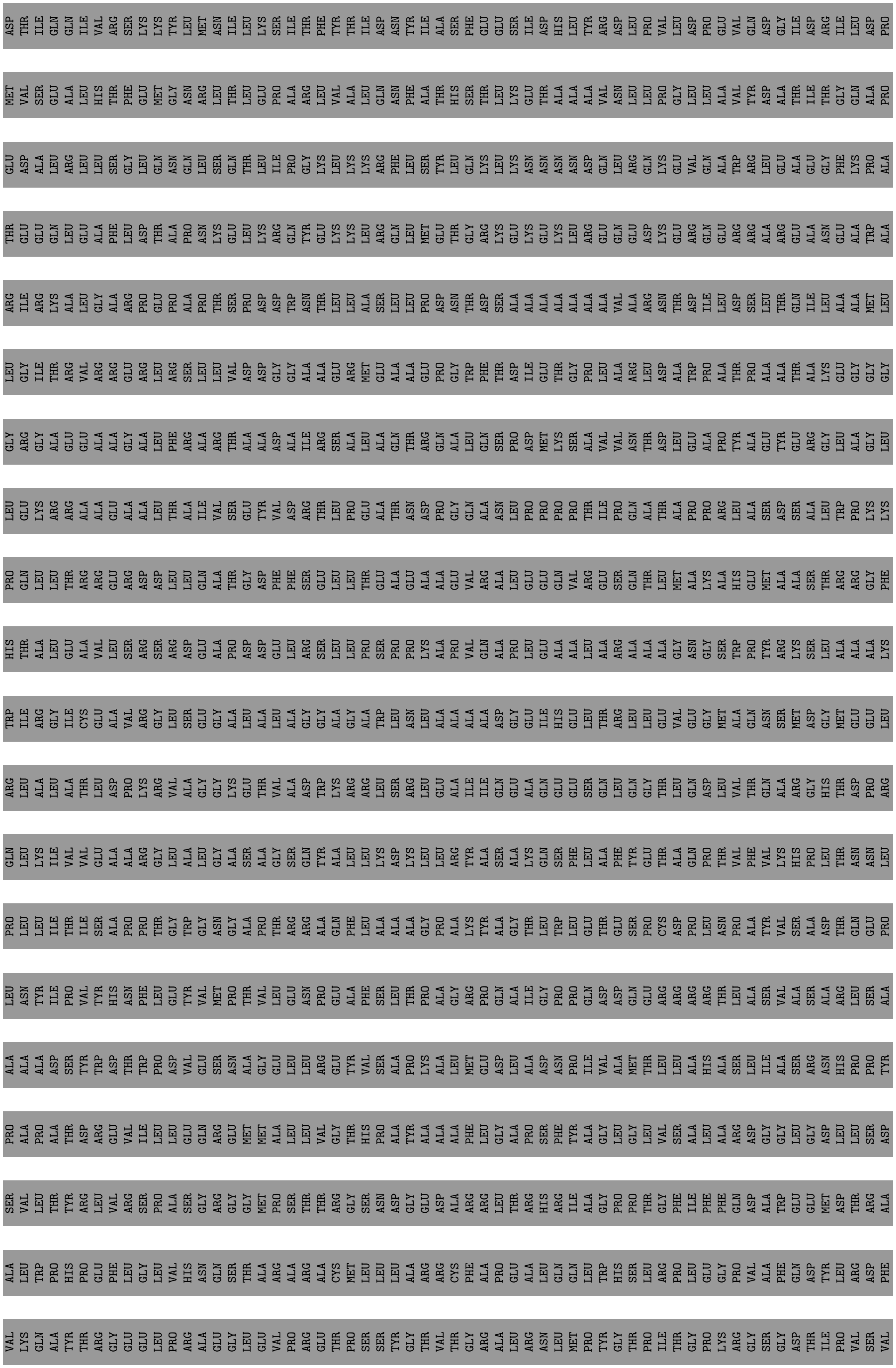
LEU LEU SER GLY GLY ASP GLN GLU ALA GLY GLU GLY GLY GLY GLU PRO GLU ASP ASN SER ILE TYR ARG GLY PRO HIS VAL ASP VAL PRO LEU VAL LEU ASP ASP GLU SER TRP LYS ARG LEU LEU SER LEU ALA GLU ALA ALA ARG THR ALA VAL ALA ARG GLN GLN ALA GLY VAL ASP

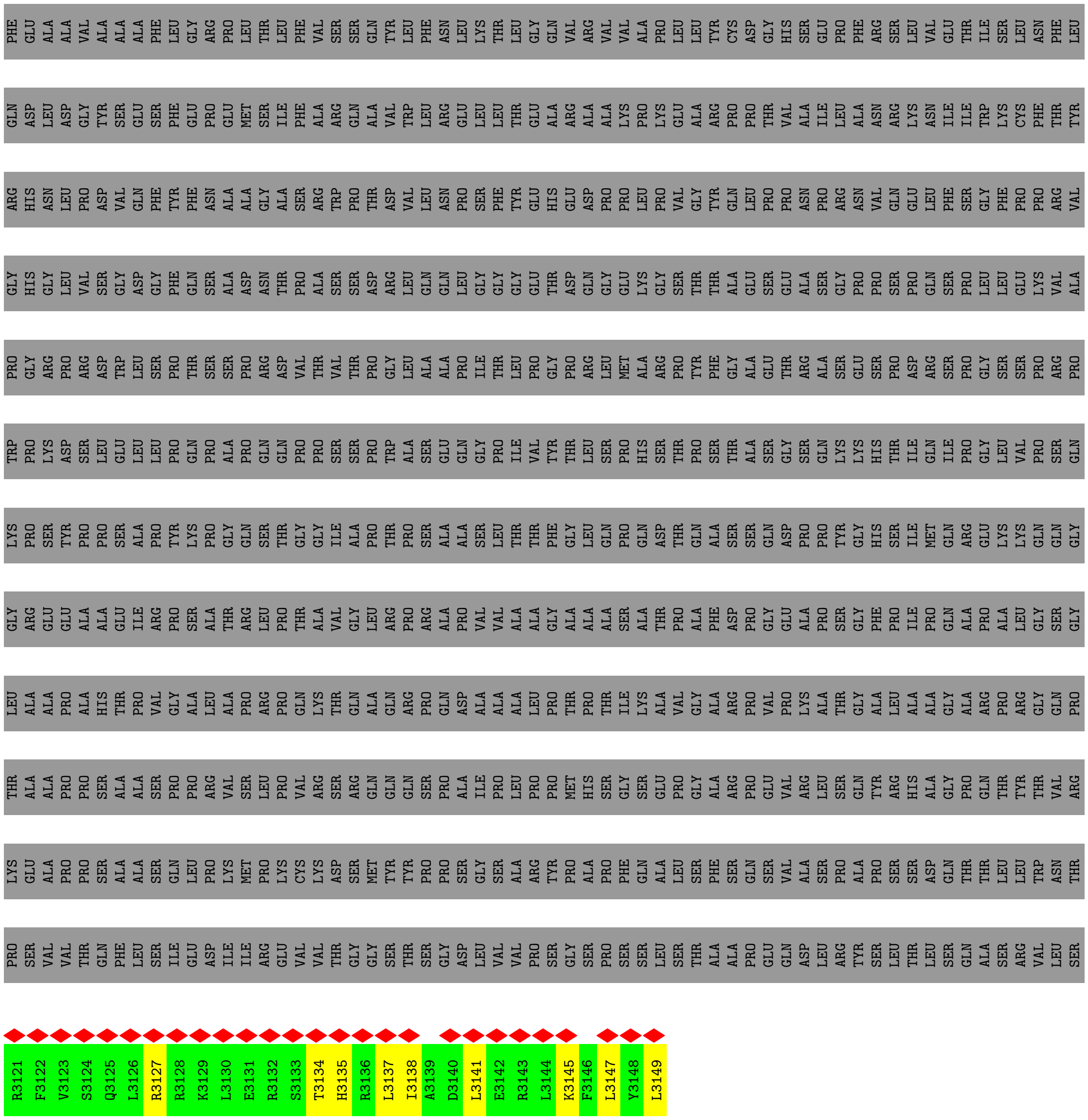
GLU GLU ASP VAL ARG PHE LEU ALA LEU LEU THR ALA ILE GLU TYR GLY ALA PRO PRO ALA ALA SER VAL PRO PRO PHE VAL HIS ASN VAL ALA VAL ARG SER LYS ASN ALA ALA LEU HIS VAL ARG ARG CYS THR ALA ASP ILE ARG ASP LYS VAL ALA SER ALA ALA SER ASP TYR LEU

SER TYR LEU GLU ASP PRO SER LEU PRO THR VAL MET ASP PHE ASP ASP LEU LEU THR HIS LEU ARG HIS THR CYS GLN ILE ILE ALA SER LEU PRO LEU LEU ASN ILE ARG TYR THR SER ILE GLU TRP ASP TYR ARG GLU LEU LEU TYR LEU GLY THR ALA LEU SER ASP MET SER GLY

ILE PRO TRP PRO LEU GLU ARG VAL GLU GLU ASP ASP PRO SER ILE ALA PRO LEU PRO GLU PHE GLU THR VAL ALA LYS LYS GLN LYS GLU LEU GLU THR THR ARG GLU ASN GLU LYS ARG LEU ARG THR ILE LEU ASP ASP ILE GLU ALA MET LEU GLY LEU ALA GLY VAL ALA SER ALA

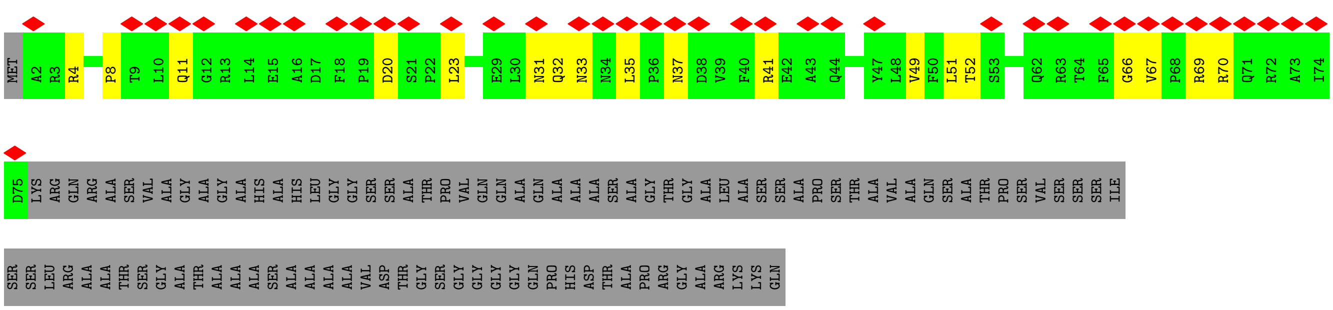
PRO GLY ALA PRO ILE SER PRO ALA SER PRO SER ALA THR PRO ALA ASN HIS ASP ASN PRO GLU ALA THR PRO PRO LEU ALA ASP THR ALA ALA LEU THR ILE PRO VAL ILE GLU LYS TYR ILE ALA ASN ALA GLY SER ILE VAL GLY ALA ALA LYS ASN PRO THR TYR ILE ARG LEU ARG


ASP THR ILE GLN GLN ILE VAL ARG SER LYS LYS TYR LEU MET ASN ILE LEU LYS SER ILE THR PHE TYR THR ILE ASP ASN TYR ILE ALA SER PHE GLU GLU SER ILE ASP HIS LEU TYR ARG ASP LEU PRO VAL LEU ASP PRO GLU VAL GLN ASP GLY ILE ASP ARG ILE LEU ASP PRO
MET VAL SER GLU ALA LEU HIS THR PHE GLU MET GLY ASN ARG LEU THR LEU GLU PRO ALA ARG LEU VAL ALA LEU GLN ASN PHE ALA THR HIS SER THR LEU LYS GLU THR ALA ALA ALA VAL ASN LEU LEU PRO GLY LEU LEU ALA VAL TYR ASP ALA THR ILE THR GLY GLN ALA PRO
GLU ASP ALA LEU ARG LEU LEU SER GLY LEU GLN ASN GLN LEU SER GLN THR LEU ILE PRO GLY LYS LEU LYS LYS ARG PHE LEU SER TYR LEU GLN LYS LEU LYS ASN ASN ASN ASN ASP GLN LEU ARG GLN LYS GLU VAL GLN ALA TRP ARG LEU GLU ALA GLU GLY PHE LYS PRO ALA
THR GLU GLU GLN LEU GLU ALA PHE LEU ASP THR ALA PRO ASN LYS GLU LEU LYS ARG GLN TYR GLU LYS LYS LEU ARG GLN LEU MET GLU THR GLY ARG LYS GLU LYS GLU LYS LEU ARG GLU GLN GLU ASP LYS GLU ARG GLN GLU ARG ARG ALA ARG GLU ALA ASN GLU ALA TRP ALA
ARG ILE ARG LYS ALA LEU GLY ALA ARG PRO GLU PRO ALA PRO THR SER PRO ASP ASP TRP ASN THR LEU LEU ALA SER LEU LEU PRO ASP ASN THR ASP SER ALA ALA ALA ALA ALA ALA VAL ALA ARG ASN THR ASP ILE LEU ASP SER LEU THR GLN ILE LEU ALA ALA MET LEU
LEU GLY ILE THR ARG VAL ARG ARG GLU ARG LEU ARG SER LEU LEU VAL ASP ASP GLY GLY ALA ALA GLU ARG MET GLU ALA ALA GLU PRO GLY TRP PHE THR ASP ILE GLU THR GLY PRO LEU ALA ARG LEU ASP ALA TRP PRO ALA THR PRO ALA ALA THR ALA LYS GLU GLY GLY GLY
GLY ARG GLY ALA GLU GLU ALA ALA GLY ALA LEU PHE ARG ALA ARG THR ALA ALA ASP ALA ILE ARG SER ALA LEU ALA GLN THR ARG GLN ALA LEU GLN SER PRO ASP MET LYS SER ALA VAL VAL ASN THR ASP LEU GLU ALA PRO TYR ALA GLU TYR GLU ARG GLY LEU ALA GLY LEU
LEU GLU LYS ARG ARG ALA ALA GLU ALA ALA LEU THR ALA ILE VAL SER GLU TYR VAL ASP ARG THR LEU PRO GLU ALA THR ASN ASP PRO GLY GLN ALA ASN LEU PRO PRO PRO THR ILE PRO GLN ALA THR ALA PRO PRO ARG LEU ALA SER ASP SER ALA LEU TRP PRO LYS LYS
PRO GLN LEU LEU THR ARG ARG GLU ARG ASP ASP LEU LEU GLN ALA THR GLY ASP PHE PHE SER GLU LEU LEU THR GLU ALA GLU ALA ALA GLU VAL ARG ALA LEU GLU GLU GLN VAL ARG GLU SER GLN THR LEU MET ALA LYS ALA HIS GLU MET ALA ALA SER THR ARG GLY PHE
HIS THR ALA LEU GLU ALA VAL LEU SER ARG SER ARG ASP GLU ALA PRO ASP ASP GLU LEU ARG SER LEU LEU PRO SER PRO PRO LYS ALA PRO VAL GLN ALA PRO LEU GLU ALA ALA LEU ALA ARG ALA ALA ALA GLY ASN GLY SER TRP PRO TYR ARG LYS SER LEU ALA ALA ALA LYS
TRP ILE ARG GLY ILE CYS GLU ALA VAL ARG GLY LEU SER GLU GLY ALA LEU ALA LEU ALA GLY GLY ALA GLY ALA TRP LEU ASN LEU ALA ALA ALA ALA ASP GLY GLU ILE HIS GLU LEU THR ARG LEU LEU GLU VAL GLU GLY MET ALA GLN ASN SER MET ASP GLY MET GLU GLU LEU
ARG LEU ALA LEU ALA THR LEU ASP PRO LYS ARG VAL ALA GLY GLY LYS GLU THR VAL ALA ASP TRP LYS ARG ARG LEU SER ARG LEU GLU ALA ILE ILE GLN GLU ALA GLN GLU GLU SER GLN LEU GLN GLY THR LEU GLN ASP LEU VAL THR GLN ALA ARG GLY HIS THR ASP PRO ARG
GLN LEU LYS ILE VAL VAL GLU ALA ALA ARG GLY LEU ALA LEU GLY ALA SER ALA GLY SER GLN TYR ALA LEU LEU LYS ASP LYS LEU LEU ARG TYR ALA SER ALA LYS GLN SER PHE LEU ALA PHE TYR GLU THR ALA GLN PRO THR VAL PHE VAL LYS HIS PRO LEU THR ASN ASN LEU
PRO LEU LEU ILE THR ILE SER ALA PRO PRO THR GLY TRP GLY ASN GLY ALA PRO THR ARG ARG ALA GLN PHE LEU ALA ALA ALA GLY PRO ALA LYS TYR ALA GLY THR LEU TRP LEU GLU THR GLU SER PRO CYS ASP PRO LEU ASN PRO ALA TYR VAL SER ALA ASP THR GLN GLU PRO
LEU ASN TYR ILE PRO VAL TYR HIS ASN PHE LEU GLU TYR VAL MET PRO THR VAL LEU GLU ASN PRO GLU ALA PHE SER LEU THR PRO ALA GLY ARG PRO GLN ALA ILE GLY PRO PRO GLN ASP ASP GLN GLU ARG ARG ARG THR LEU ALA SER VAL ALA SER ALA ARG LEU SER ALA
ALA ALA ALA ASP SER TYR TRP ASP THR TRP PRO ASP VAL GLU SER ASN ALA GLY GLU LEU LEU ARG GLU TYR VAL SER ALA PRO LYS ALA LEU MET GLU ASP LEU ALA ASP ASN PRO ILE VAL ALA MET THR LEU LEU ALA HIS ALA SER LEU ILE ALA SER ARG ASN HIS PRO PRO TYR
PRO ALA PRO ALA THR ASP ARG GLU VAL ILE LEU LEU GLU GLN ARG GLU MET MET ALA LEU LEU LEU VAL GLY THR HIS PRO ALA TYR ALA ALA ALA PHE LEU GLY ALA PRO SER PHE TYR ALA GLY LEU GLY LEU VAL SER ALA LEU ALA ARG ASP GLY GLY LEU GLY ASP LEU LEU SER ASP
SER VAL LEU THR TYR ARG LEU VAL ARG SER PRO ALA SER GLY ARG GLY GLY MET PRO SER THR THR ARG GLY SER ASN ASP GLY GLU ASP ALA ARG LEU THR ARG HIS ARG ILE ALA GLY PRO PRO THR GLY PHE ILE PHE PHE GLN ASP ALA TRP GLU GLU MET ASP THR ARG ALA
ALA LEU TRP PRO HIS PRO GLU PHE LEU GLY LEU VAL HIS ASN GLN SER THR ALA ARG ALA ARG ALA CYS MET LEU LEU LEU ALA ARG ARG CYS PHE ALA PRO GLU ALA LEU GLN GLN LEU TRP HIS SER LEU ARG PRO LEU GLU GLY PRO VAL ALA PHE GLN ASP TYR LEU ARG ASP PHE
VAL LYS GLN ALA TYR THR ARG GLY GLU GLU LEU PRO ARG ALA GLU GLY LEU GLU VAL PRO ARG GLU THR PRO SER SER TYR GLY THR VAL THR GLY ARG ALA LEU ARG ASN LEU MET PRO TYR GLY THR PRO ILE THR GLY PRO LYS ARG GLY SER GLY ASP THR ILE PRO VAL SER VAL



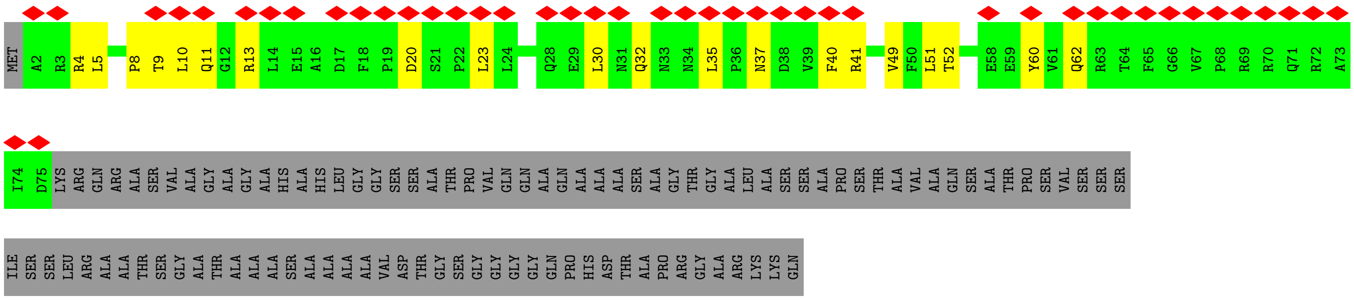
• Molecule 6: Small capsomere-interacting protein

Chain Y:



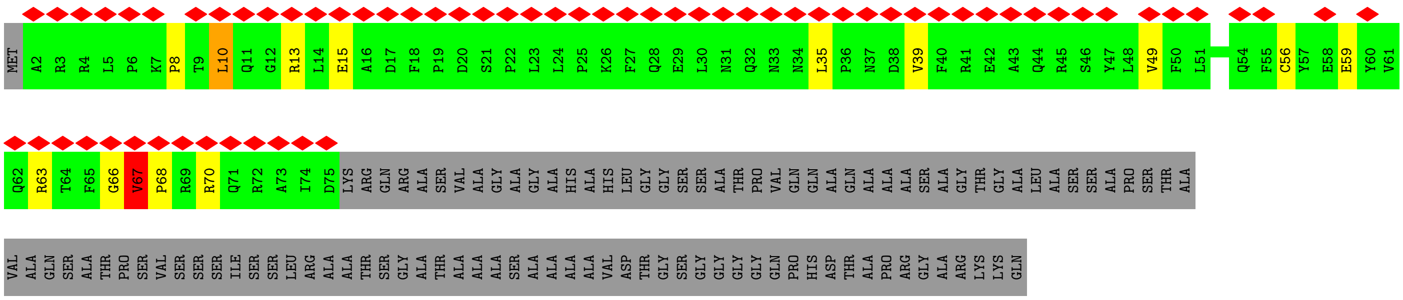
• Molecule 6: Small capsomere-interacting protein

Chain Z:



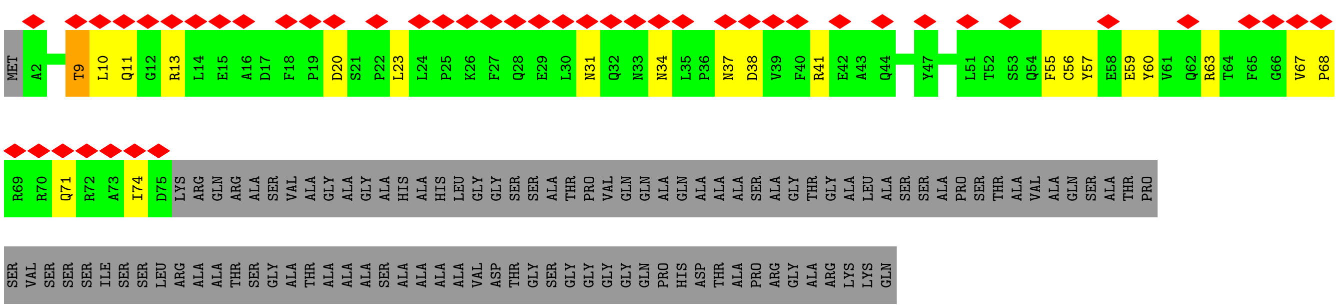
- Molecule 6: Small capsomere-interacting protein

Chain 2:



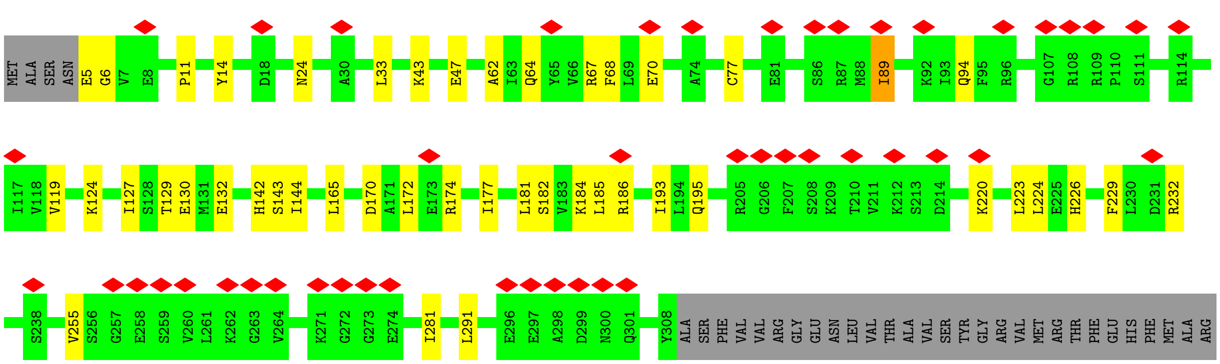
- Molecule 6: Small capsomere-interacting protein

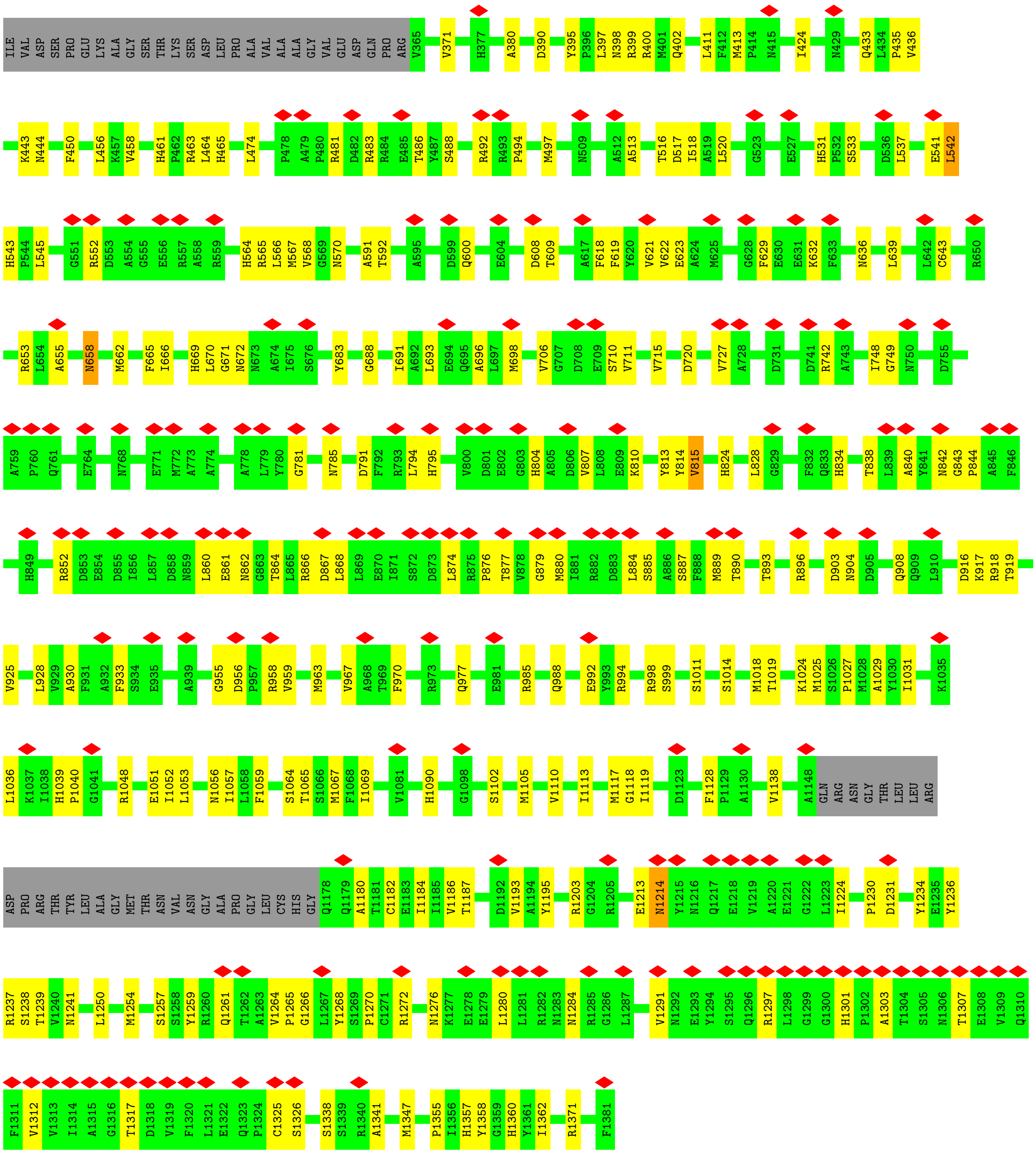
Chain y:



- Molecule 7: Major capsid protein

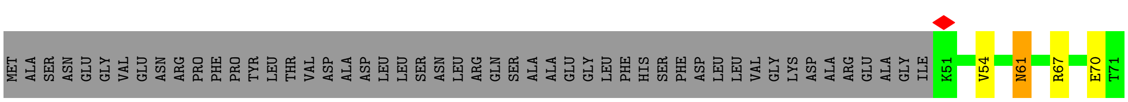
Chain S:

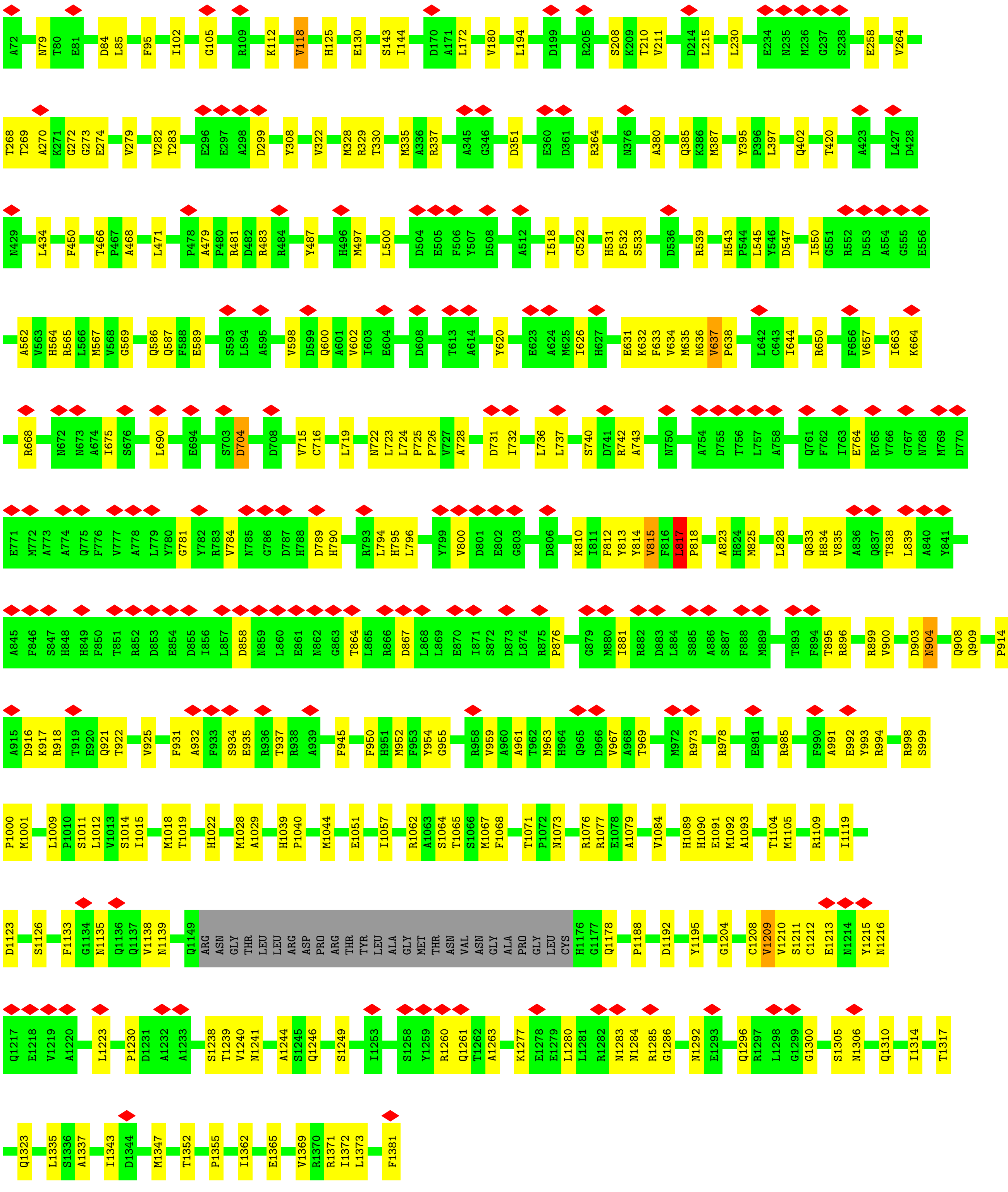




• Molecule 7: Major capsid protein

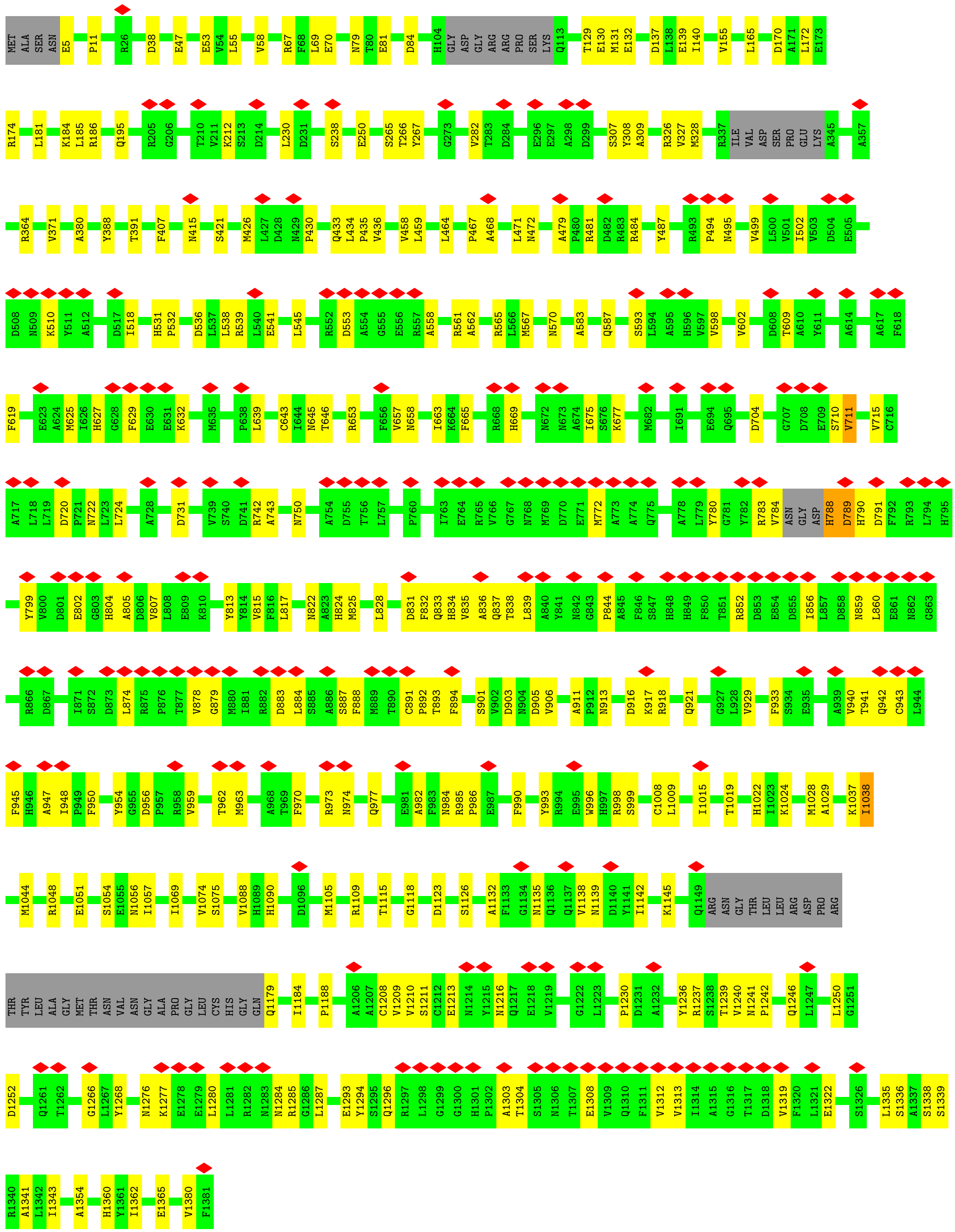
Chain T:



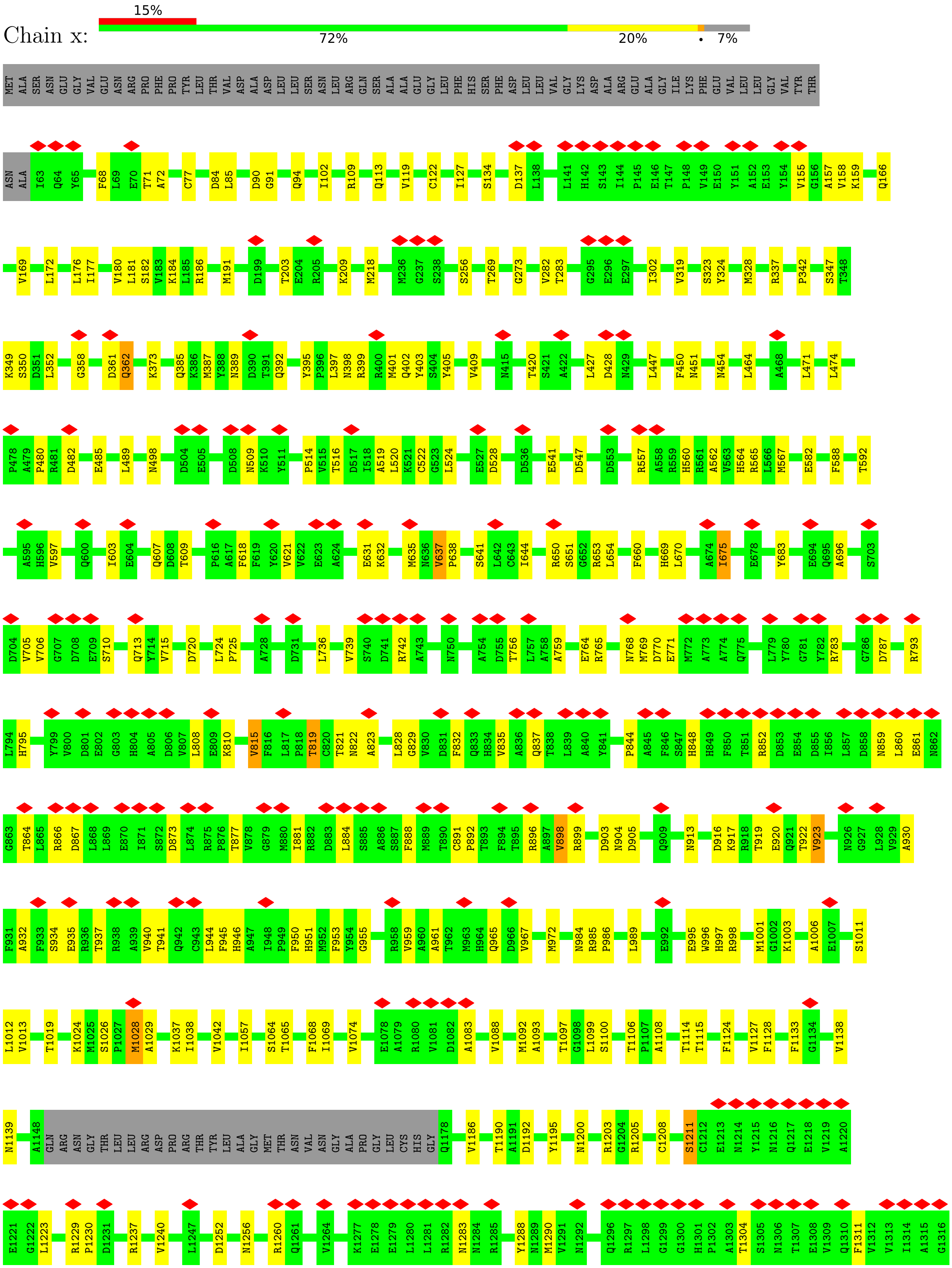


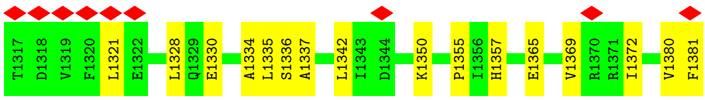
- Molecule 7: Major capsid protein

Chain W:



• Molecule 7: Major capsid protein


Chain x:
15%
72%
20%
•
7%
MET ALA SER ASN GLU GLY VAL GLU ASN ARG PRO PHE PRO TYR LEU THR VAL ASP ALA ASP LEU LEU SER ASN LEU ARG GLN SER ALA ALA GLU GLY LEU PHE HIS SER PHE ASP LEU LEU VAL GLY LYS ASP ALA ARG GLU ALA GLY ILE LYS PHE GLU VAL LEU LEU GLY VAL TYR THR
ASN ALA I63 Q64 Y65 F68 L69 E70 T71 A72 C77 D84 L85 D90 G91 Q94 I102 R109 Q113 V119 C122 I127 S134 D137 L138 L141 H142 S143 I144 P145 E146 T147 P148 V149 E150 Y151 A152 E153 Y154 V155 G156 A157 V158 K159 Q166
V169 L172 L176 I177 V180 L181 S182 V183 K184 L185 R186 M191 D199 T203 E204 R205 K209 M218 M236 G237 S238 S256 T269 G273 V282 T283 G295 E296 E297 I302 V319 S323 Y324 M328 R337 P342 S347 T348
K349 S350 D351 L352 G358 D361 Q362 K373 Q385 K386 M387 Y388 N389 D390 T391 Q392 Y395 P396 L397 N398 R399 R400 M401 Q402 Y403 S404 Y405 V409 N415 T420 S421 A422 L427 D428 N429 L447 F450 N451 N454 L464 A468 L471 L474
P478 A479 P480 R481 D482 E485 L489 N498 D504 E505 D508 N509 K510 Y511 P514 V515 T516 D517 I518 A519 L520 K521 C522 G523 L524 E527 D528 D536 E541 D547 D553 R557 A558 R559 H560 R561 A562 V563 H564 R565 L566 M567 E582 F588 T592
A595 H596 V597 Q600 I603 E604 Q607 D608 T609 P616 A617 F618 F619 Y620 V621 V622 E623 A624 E631 K632 M635 N636 V637 P638 S641 L642 C643 I644 R650 S651 G652 R653 L654 F660 H669 L670 A674 I675 E678 Y683 E694 Q695 A696 S703
D704 V705 V706 G707 D708 E709 S710 Q713 Y714 V715 D720 L724 P725 A728 D731 L736 V739 S740 D741 R742 A743 N750 A754 D755 T756 L757 A758 A759 E764 R765 N768 M769 D770 E771 M772 A773 A774 Q775 L779 Y780 G781 Y782 R783 G786 D787 R793
L794 H795 Y799 V800 D801 E802 G803 H804 A805 D806 V807 L808 E809 K810 V815 F816 L817 P818 T819 C820 T821 N822 A823 L828 G829 V830 D831 F832 Q833 H834 V835 A836 Q837 T838 L839 A840 Y841 P844 A845 F846 S847 H848 H849 F850 T851 R852 D853 E854 D855 I856 L857 D858 N859 L860 E861 N862
G863 T864 L865 R866 D867 L868 L869 E870 I871 S872 D873 L874 R875 P876 T877 V878 G879 M880 I881 R882 D883 L884 S885 A886 S887 F888 M889 T890 C891 P892 T893 F894 T895 R896 A897 V898 R899 D903 N904 D905 Q909 N913 D916 K917 R918 T919 E920 Q921 T922 V923 N926 G927 L928 V929 A930
F931 A932 F933 S934 E935 R936 T937 R938 A939 V940 T941 Q942 C943 L944 F945 H946 A947 I948 P949 F950 H951 M952 F953 Y954 G955 R958 V959 A960 A961 T962 M963 H964 Q965 D966 V967 M972 N984 R985 P986 L989 E992 E995 W996 H997 R998 M1001 G1002 K1003 A1006 E1007 S1011
L1012 V1013 T1019 K1024 M1025 S1026 P1027 M1028 A1029 K1037 I1038 V1042 I1057 S1064 T1065 F1068 I1069 V1074 E1078 A1079 R1080 V1081 D1082 A1083 V1088 M1092 A1093 T1097 G1098 L1099 S1100 T1106 P1107 A1108 T1114 T1115 F1124 V1127 F1128 F1133 G1134 V1138
N1139 A1148 GLN ARG ASN GLY THR LEU LEU ARG ASP PRO ARG THR TYR LEU ALA GLY MET THR ASN VAL ASN GLY ALA PRO GLY LEU CYS HIS GLY Q1178 V1186 T1190 A1191 D1192 Y1195 N1200 R1203 G1204 R1205 C1208 S1211 C1212 E1213 N1214 Y1215 N1216 Q1217 E1218 V1219 A1220
E1221 G1222 L1223 R1229 P1230 D1231 R1237 V1240 L1247 D1252 N1256 R1260 Q1261 V1264 K1277 E1278 E1279 L1280 L1281 R1282 N1283 N1284 R1285 Y1288 N1289 M1290 V1291 N1292 Q1296 R1297 L1298 G1299 G1300 H1301 P1302 A1303 T1304 S1305 N1306 T1307 E1308 V1309 Q1310 F1311 V1312 V1313 I1314 A1315 G1316


T1317
D1318
V1319
F1320
L1321
E1322
L1328
Q1329
E1330
A1334
L1335
S1336
A1337
L1342
I1343
D1344
K1350
P1355
I1356
H1357
E1365
V1369
R1370
R1371
I1372
V1380
F1381

4 Experimental information (i)

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, C5	Depositor
Number of particles used	28639	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	PHASE FLIPPING AND AMPLITUDE CORRECTION	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ($e^-/\text{Å}^2$)	48	Depositor
Minimum defocus (nm)	Not provided	
Maximum defocus (nm)	Not provided	
Magnification	Not provided	
Image detector	GATAN K2 QUANTUM (4k x 4k)	Depositor
Maximum map value	0.063	Depositor
Minimum map value	-0.037	Depositor
Average map value	0.002	Depositor
Map value standard deviation	0.006	Depositor
Recommended contour level	0.013	Depositor
Map size (Å)	392.99997, 392.99997, 392.99997	wwPDB
Map dimensions	300, 300, 300	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	1.31, 1.31, 1.31	Depositor

5 Model quality (i)

5.1 Standard geometry (i)

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	5	0.27	0/2263	0.61	0/3088
1	e	0.36	0/2572	0.59	0/3503
2	6	0.29	0/2377	0.61	0/3236
2	7	0.32	0/2306	0.66	0/3135
2	f	0.30	0/2327	0.55	2/3169 (0.1%)
2	g	0.34	0/2319	0.59	0/3155
3	G	0.31	0/684	0.72	1/924 (0.1%)
3	K	0.26	0/570	0.59	0/765
4	C	0.30	0/2770	0.65	2/3763 (0.1%)
5	B	0.24	0/304	0.55	0/404
5	O	0.28	0/255	0.64	0/339
6	2	0.28	0/636	0.68	1/861 (0.1%)
6	Y	0.31	0/636	0.59	0/861
6	Z	0.27	0/636	0.58	0/861
6	y	0.27	0/636	0.60	0/861
7	S	0.38	0/10405	0.61	3/14142 (0.0%)
7	T	0.38	0/10500	0.62	5/14271 (0.0%)
7	W	0.36	0/10699	0.59	4/14539 (0.0%)
7	x	0.41	2/10380 (0.0%)	0.64	2/14109 (0.0%)
All	All	0.36	2/63275 (0.0%)	0.62	20/85986 (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	5	0	1
1	e	0	1
2	g	0	1
6	2	0	1
7	T	0	1
7	x	0	3

Continued on next page...

Continued from previous page...

Mol	Chain	#Chirality outliers	#Planarity outliers
All	All	0	8

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
7	x	362	GLN	C-N	14.74	1.51	1.33
7	x	392	GLN	C-N	-5.56	1.24	1.33

The worst 5 of 20 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
7	x	362	GLN	CA-C-N	13.07	133.94	119.83
7	x	362	GLN	C-N-CA	13.07	133.94	119.83
7	S	890	THR	CA-C-N	7.83	131.29	120.65
7	S	890	THR	C-N-CA	7.83	131.29	120.65
7	W	436	VAL	N-CA-C	-6.90	107.15	113.71

There are no chirality outliers.

5 of 8 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
6	2	66	GLY	Peptide
1	5	312	HIS	Peptide
7	T	817	LEU	Peptide
7	x	361	ASP	Mainchain
7	x	669	HIS	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	5	2199	0	2144	40	0
1	e	2505	0	2480	34	0
2	6	2329	0	2363	54	0
2	7	2262	0	2293	53	0
2	f	2279	0	2316	51	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	g	2272	0	2308	43	0
3	G	675	0	678	15	0
3	K	563	0	566	12	0
4	C	2705	0	2712	70	0
5	B	301	0	328	4	0
5	O	252	0	269	7	0
6	2	621	0	611	10	0
6	Y	621	0	611	11	0
6	Z	621	0	611	14	0
6	y	621	0	611	15	0
7	S	10166	0	9995	193	0
7	T	10258	0	10093	211	0
7	W	10456	0	10286	192	0
7	x	10141	0	9977	218	0
All	All	61847	0	61252	1150	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 9.

The worst 5 of 1150 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)
7:x:395:TYR:CE1	7:x:397:LEU:HB2	1.14	1.63
7:S:395:TYR:HE1	7:S:397:LEU:CB	1.13	1.55
7:x:395:TYR:HE1	7:x:397:LEU:CB	1.21	1.50
7:S:395:TYR:CE1	7:S:397:LEU:CB	2.01	1.44
7:S:395:TYR:CE1	7:S:397:LEU:HB2	1.53	1.41

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 PDB entries followed by that with respect to all EM entries.

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	5	278/364 (76%)	260 (94%)	18 (6%)	0	100	100
1	e	311/364 (85%)	291 (94%)	20 (6%)	0	100	100
2	6	293/301 (97%)	276 (94%)	17 (6%)	0	100	100
2	7	281/301 (93%)	260 (92%)	21 (8%)	0	100	100
2	f	286/301 (95%)	270 (94%)	16 (6%)	0	100	100
2	g	282/301 (94%)	258 (92%)	24 (8%)	0	100	100
3	G	82/570 (14%)	76 (93%)	5 (6%)	1 (1%)	11	44
3	K	66/570 (12%)	65 (98%)	1 (2%)	0	100	100
4	C	334/507 (66%)	301 (90%)	33 (10%)	0	100	100
5	B	33/3149 (1%)	32 (97%)	1 (3%)	0	100	100
5	O	27/3149 (1%)	27 (100%)	0	0	100	100
6	2	72/176 (41%)	60 (83%)	12 (17%)	0	100	100
6	Y	72/176 (41%)	66 (92%)	6 (8%)	0	100	100
6	Z	72/176 (41%)	66 (92%)	6 (8%)	0	100	100
6	y	72/176 (41%)	64 (89%)	8 (11%)	0	100	100
7	S	1286/1381 (93%)	1205 (94%)	81 (6%)	0	100	100
7	T	1301/1381 (94%)	1196 (92%)	105 (8%)	0	100	100
7	W	1320/1381 (96%)	1224 (93%)	95 (7%)	1 (0%)	48	82
7	x	1286/1381 (93%)	1196 (93%)	89 (7%)	1 (0%)	48	82
All	All	7754/16105 (48%)	7193 (93%)	558 (7%)	3 (0%)	100	100

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
7	W	789	ASP
7	x	509	ASN
3	G	15	PRO

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 PDB entries followed by that with respect to all EM entries.

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	5	234/289 (81%)	231 (99%)	3 (1%)	65	76
1	e	268/289 (93%)	267 (100%)	1 (0%)	89	91
2	6	264/267 (99%)	264 (100%)	0	100	100
2	7	256/267 (96%)	254 (99%)	2 (1%)	79	84
2	f	259/267 (97%)	256 (99%)	3 (1%)	67	78
2	g	256/267 (96%)	254 (99%)	2 (1%)	79	84
3	G	70/465 (15%)	69 (99%)	1 (1%)	62	75
3	K	59/465 (13%)	59 (100%)	0	100	100
4	C	288/400 (72%)	287 (100%)	1 (0%)	91	92
5	B	33/2539 (1%)	33 (100%)	0	100	100
5	O	27/2539 (1%)	27 (100%)	0	100	100
6	2	68/128 (53%)	66 (97%)	2 (3%)	37	58
6	Y	68/128 (53%)	67 (98%)	1 (2%)	60	74
6	Z	68/128 (53%)	67 (98%)	1 (2%)	60	74
6	y	68/128 (53%)	67 (98%)	1 (2%)	60	74
7	S	1101/1171 (94%)	1092 (99%)	9 (1%)	79	84
7	T	1111/1171 (95%)	1100 (99%)	11 (1%)	73	81
7	W	1131/1171 (97%)	1122 (99%)	9 (1%)	79	84
7	x	1099/1171 (94%)	1085 (99%)	14 (1%)	65	76
All	All	6728/13250 (51%)	6667 (99%)	61 (1%)	74	83

5 of 61 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
7	T	1209	VAL
1	e	362	VAL
7	W	929	VAL
6	y	9	THR
2	g	138	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 141 such sidechains are listed below:

Mol	Chain	Res	Type
7	x	1120	HIS
7	x	1217	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	e	281	GLN
7	S	1353	HIS
7	S	1306	ASN

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.

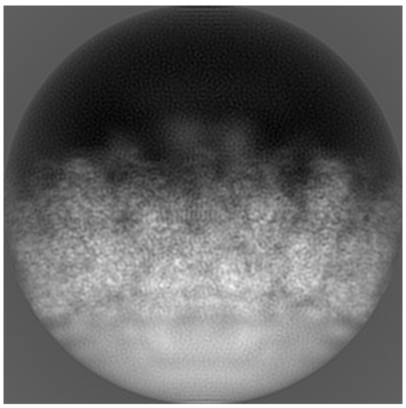
6 Map visualisation (i)

This section contains visualisations of the EMDB entry EMD-30157. These allow visual inspection of the internal detail of the map and identification of artifacts.

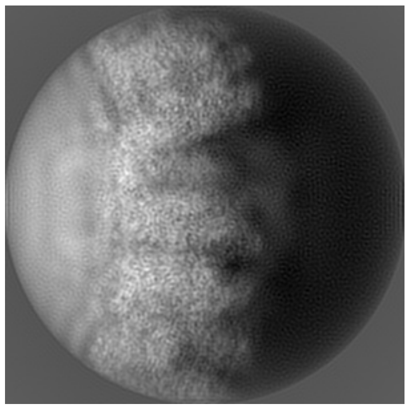
No raw map or half-maps were deposited for this entry and therefore no images, graphs, etc. pertaining to the raw map can be shown.

6.1 Orthogonal projections (i)

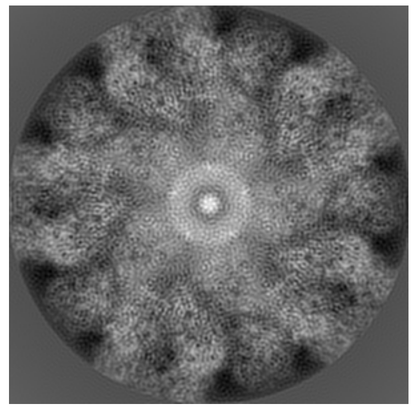
6.1.1 Primary map



X



Y

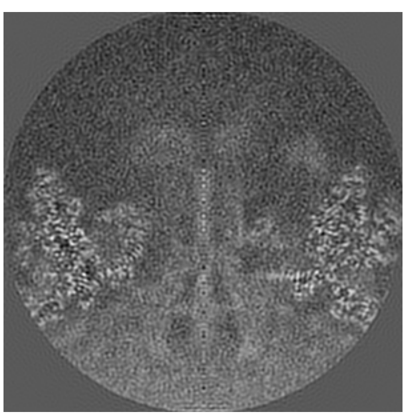


Z

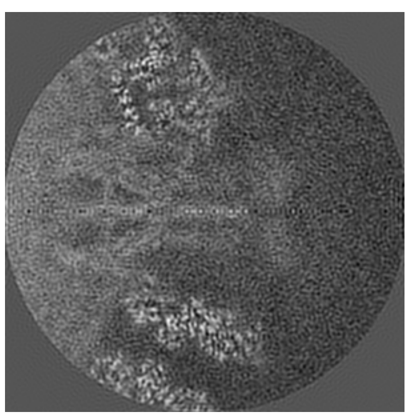
The images above show the map projected in three orthogonal directions.

6.2 Central slices (i)

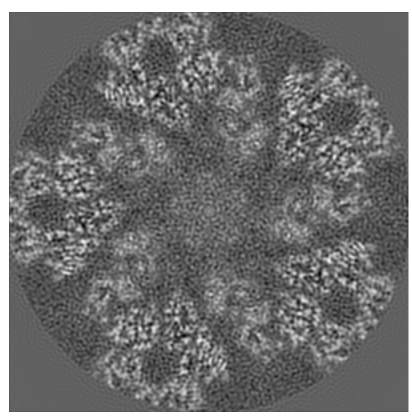
6.2.1 Primary map



X Index: 150



Y Index: 150

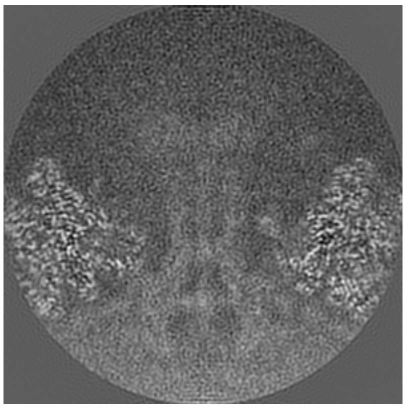


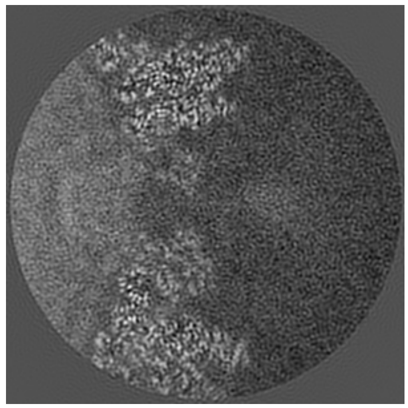
Z Index: 150

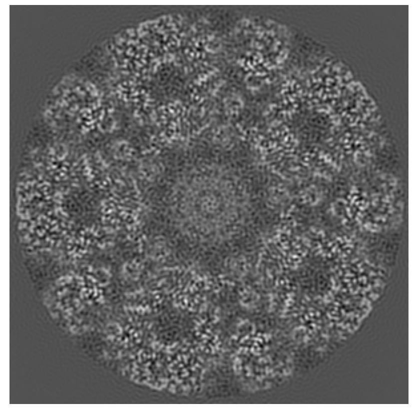
The images above show central slices of the map in three orthogonal directions.

6.3 Largest variance slices

6.3.1 Primary map


X Index: 144

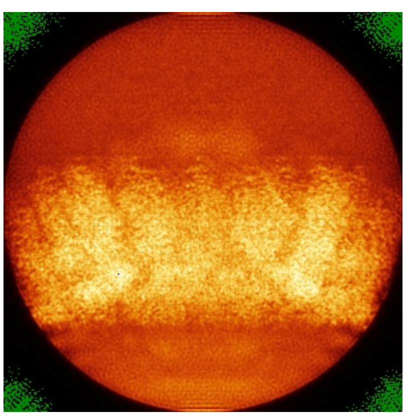

Y Index: 114

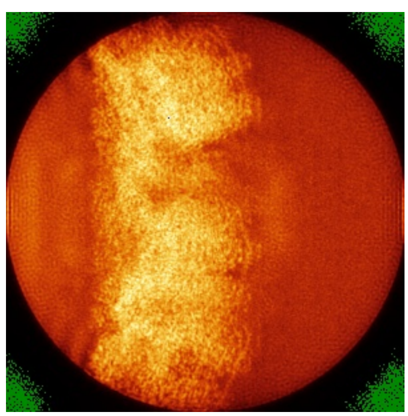

Z Index: 109

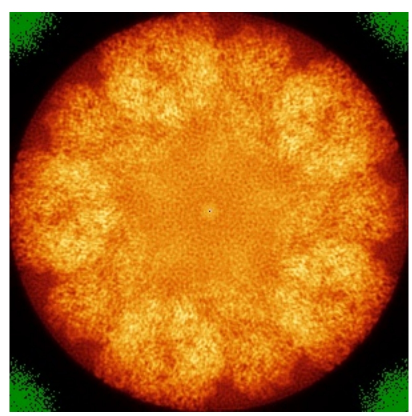
The images above show the largest variance slices of the map in three orthogonal directions.

6.4 Orthogonal standard-deviation projections (False-color)

6.4.1 Primary map


X


Y


Z

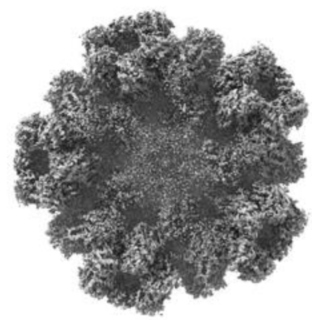
The images above show the map standard deviation projections with false color in three orthogonal directions. Minimum values are shown in green, max in blue, and dark to light orange shades represent small to large values respectively.

6.5 Orthogonal surface views (i)

6.5.1 Primary map

X

Y



Z

The images above show the 3D surface view of the map at the recommended contour level 0.013. These images, in conjunction with the slice images, may facilitate assessment of whether an appropriate contour level has been provided.

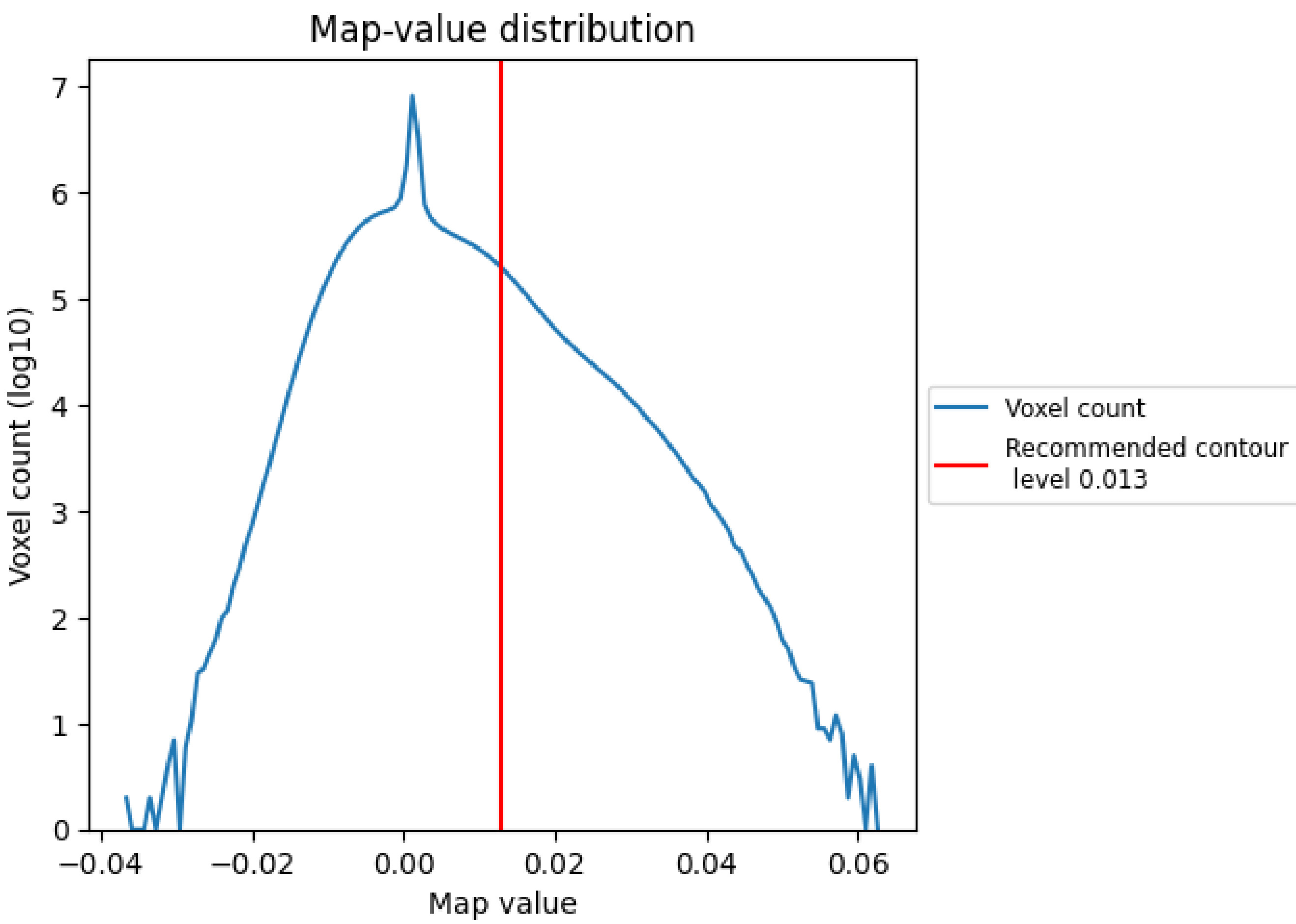
6.6 Mask visualisation (i)

This section was not generated. No masks/segmentation were deposited.

7 Map analysis (i)

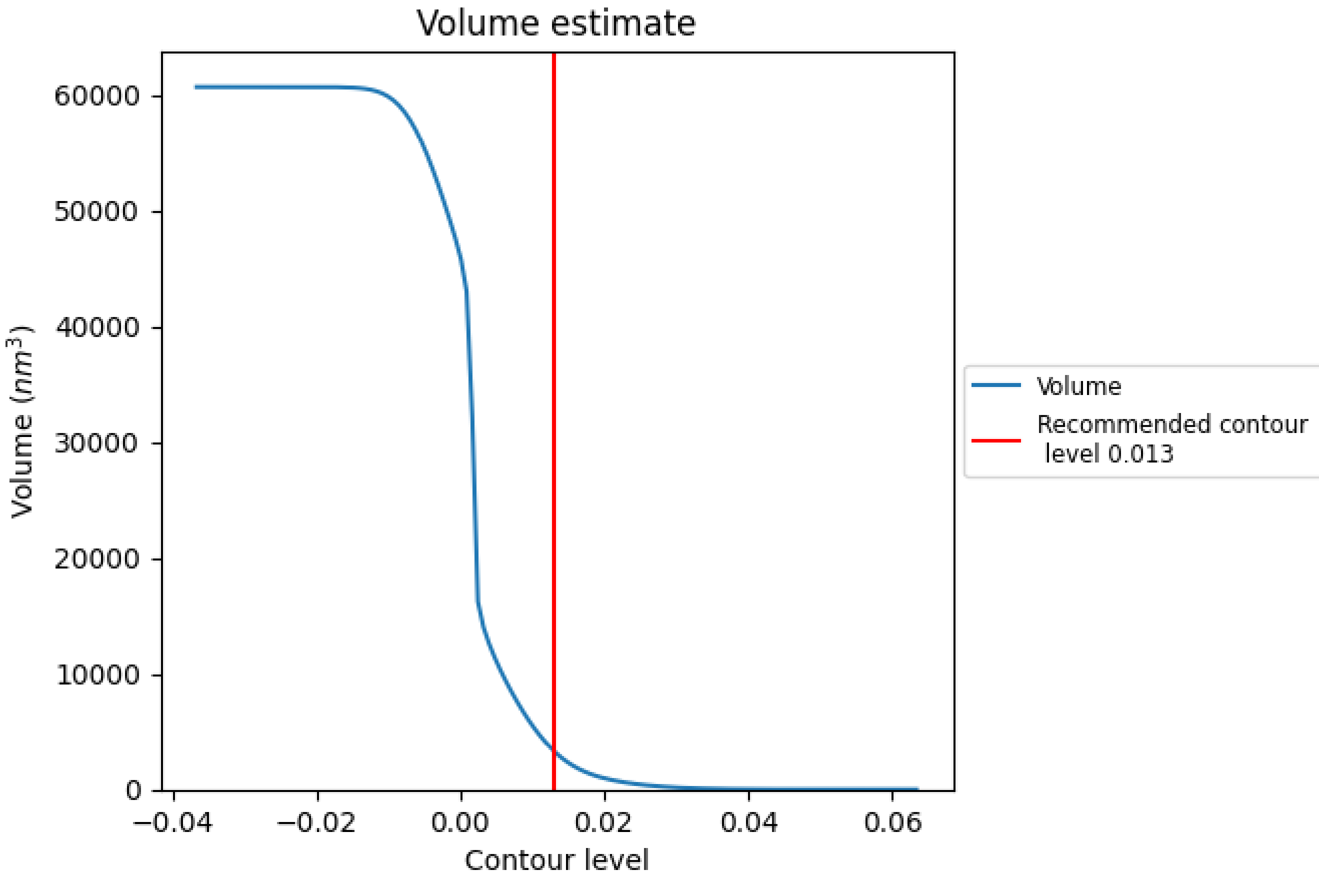
This section contains the results of statistical analysis of the map.

7.1 Map-value distribution (i)



The map-value distribution is plotted in 128 intervals along the x-axis. The y-axis is logarithmic. A spike in this graph at zero usually indicates that the volume has been masked.

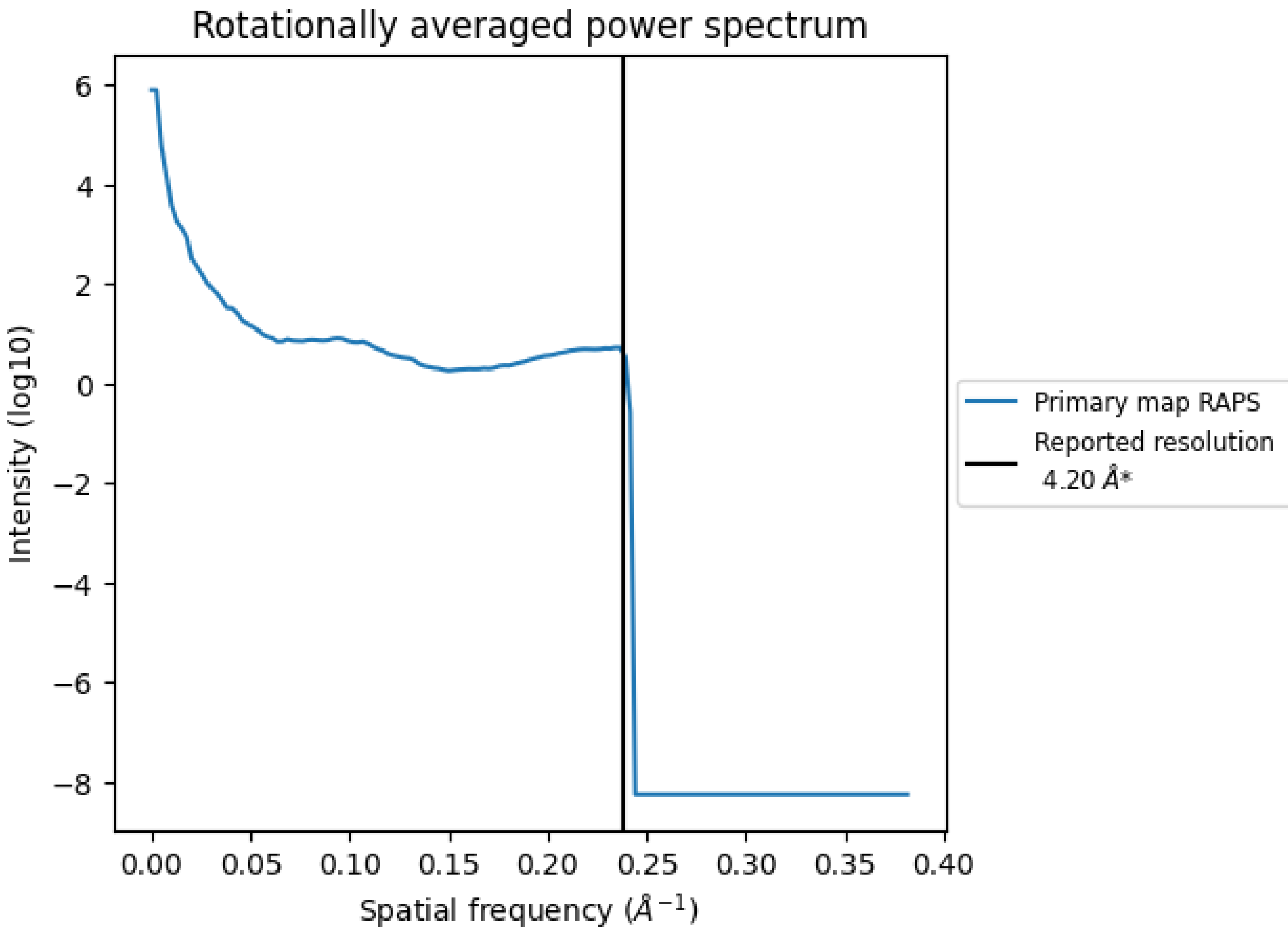
7.2 Volume estimate (i)



The volume at the recommended contour level is 3365 nm^3; this corresponds to an approximate mass of 3040 kDa.

The volume estimate graph shows how the enclosed volume varies with the contour level. The recommended contour level is shown as a vertical line and the intersection between the line and the curve gives the volume of the enclosed surface at the given level.

7.3 Rotationally averaged power spectrum (i)



*Reported resolution corresponds to spatial frequency of 0.238 Å^{-1}

8 Fourier-Shell correlation

This section was not generated. No FSC curve or half-maps provided.

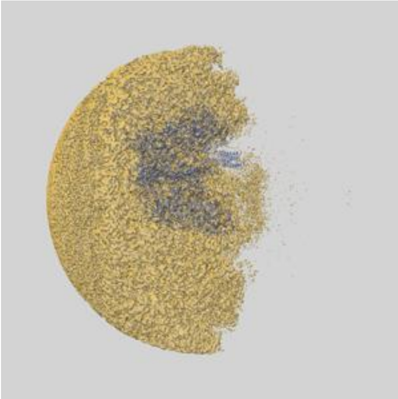
9 Map-model fit (i)

This section contains information regarding the fit between EMDB map EMD-30157 and PDB model 7BQX. Per-residue inclusion information can be found in section 3 on page 6.

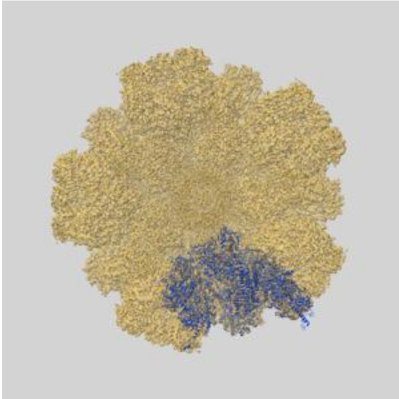
9.1 Map-model overlays

9.1.1 Map-model overlay (i)

X

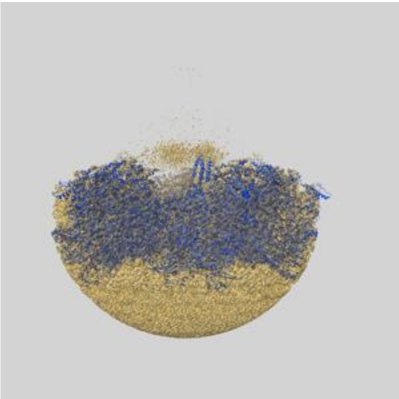


Y

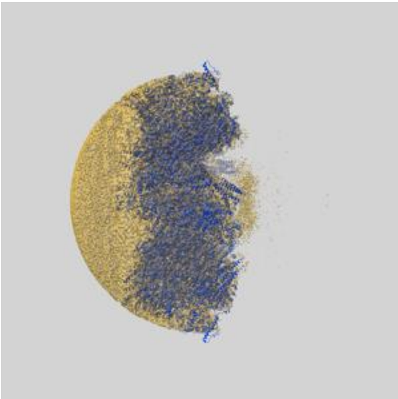


Z

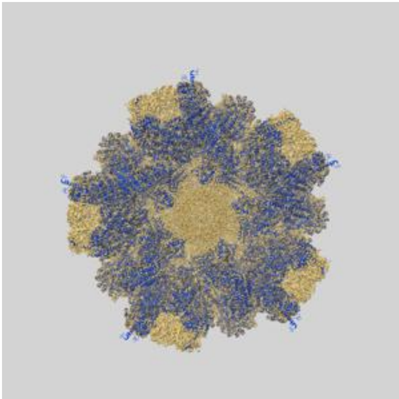
9.1.2 Map-model assembly overlay (i)



X



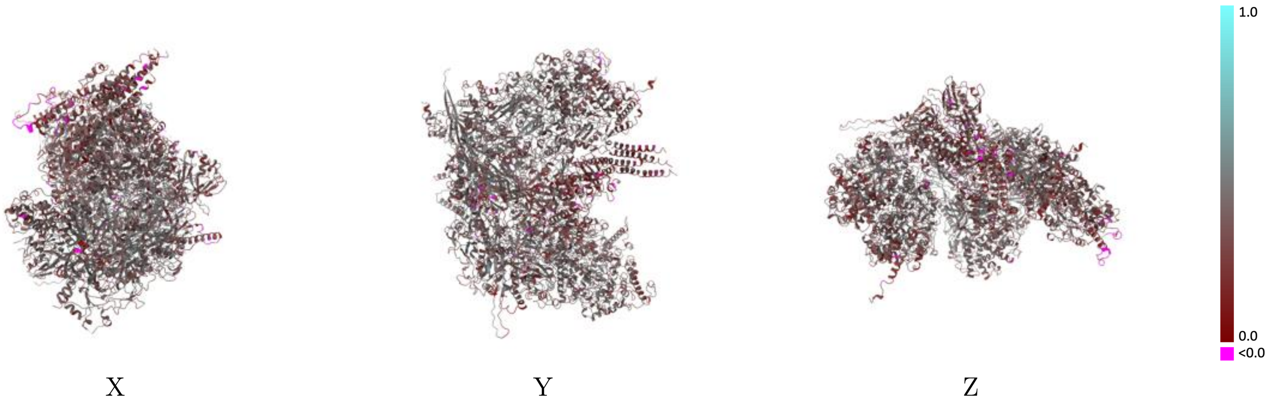
Y



Z

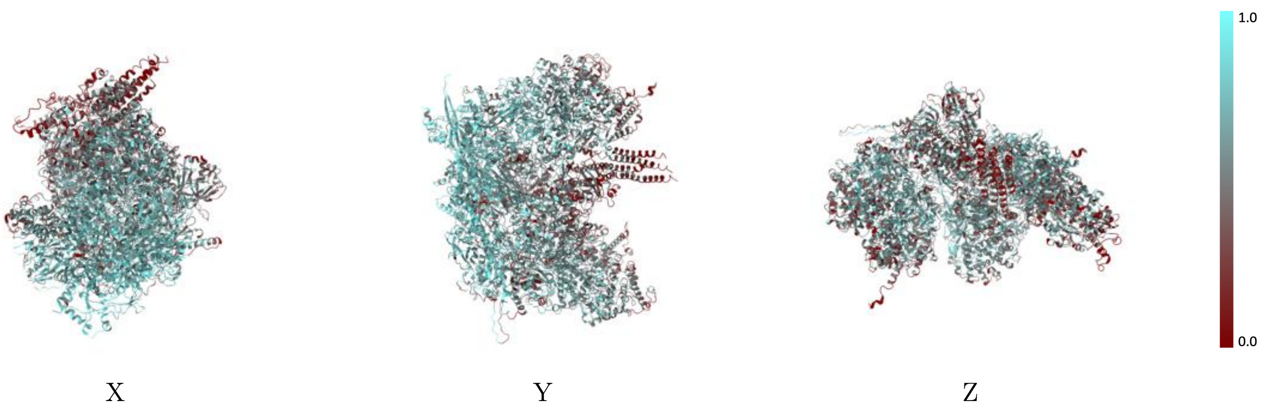
The images above show the 3D surface view of the map at the recommended contour level 0.013 at 50% transparency in yellow overlaid with a ribbon representation of the model coloured in blue. These images allow for the visual assessment of the quality of fit between the atomic model and the map.

9.2 Q-score mapped to coordinate model (i)



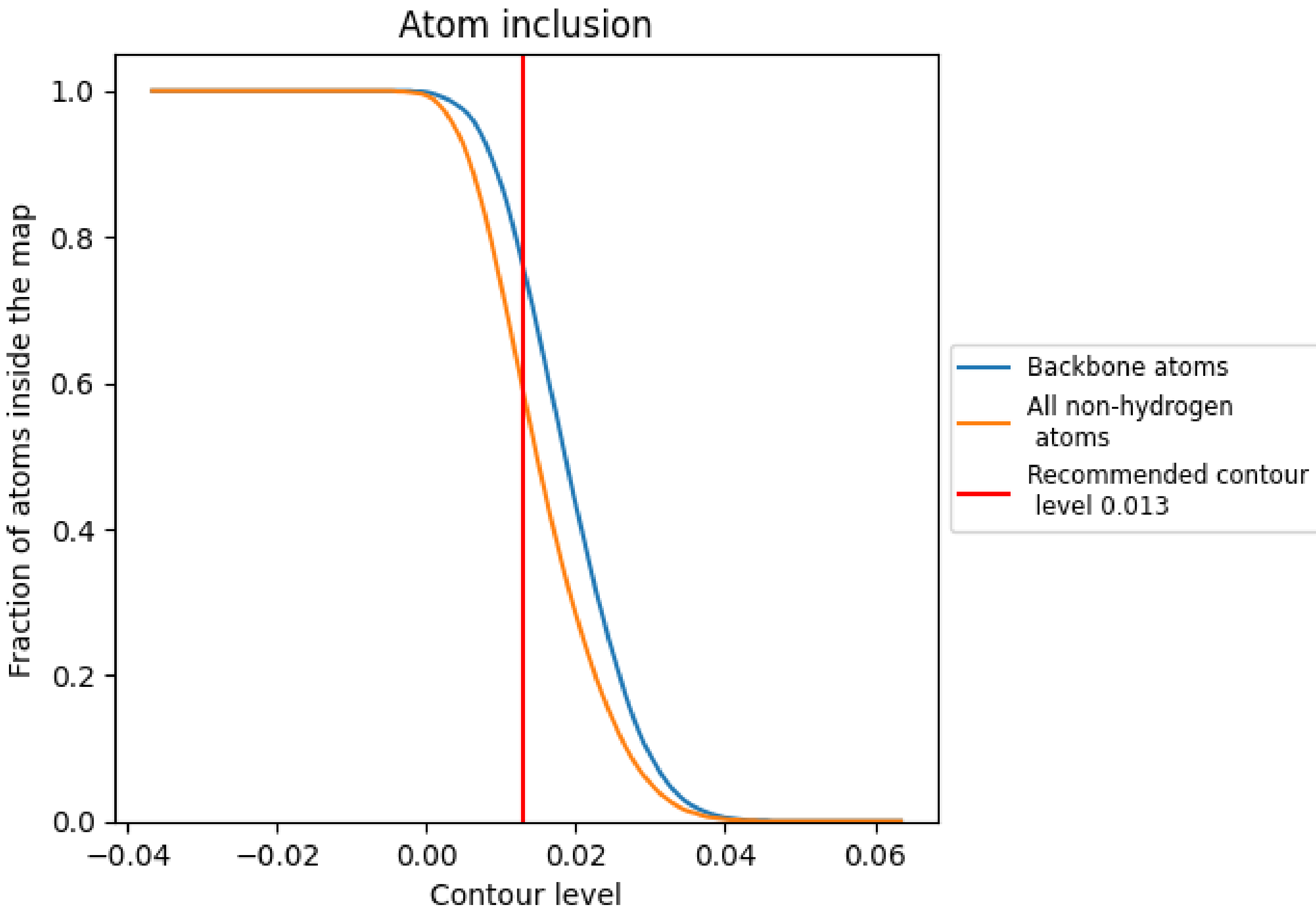
The images above show the model with each residue coloured according its Q-score. This shows their resolvability in the map with higher Q-score values reflecting better resolvability. Please note: Q-score is calculating the resolvability of atoms, and thus high values are only expected at resolutions at which atoms can be resolved. Low Q-score values may therefore be expected for many entries.

9.3 Atom inclusion mapped to coordinate model (i)



The images above show the model with each residue coloured according to its atom inclusion. This shows to what extent they are inside the map at the recommended contour level (0.013).

9.4 Atom inclusion (i)



At the recommended contour level, 76% of all backbone atoms, 59% of all non-hydrogen atoms, are inside the map.

9.5 Map-model fit summary

The table lists the average atom inclusion at the recommended contour level (0.013) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	0.5870	0.3650
2	0.1330	0.1500
5	0.4500	0.3030
6	0.4870	0.3080
7	0.4880	0.3280
B	0.1820	0.2560
C	0.4060	0.2870
G	0.3640	0.2950
K	0.3990	0.3100
O	0.1370	0.2440
S	0.6200	0.3830
T	0.6580	0.3980
W	0.6290	0.3780
Y	0.3750	0.3180
Z	0.3150	0.2840
e	0.7130	0.4020
f	0.6290	0.3720
g	0.6380	0.3760
x	0.6270	0.3830
y	0.3170	0.2980