wwPDB EM Validation Summary Report (i)

Nov 16, 2022 – 05:09 PM JST

PDB ID	:	7BQX
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:

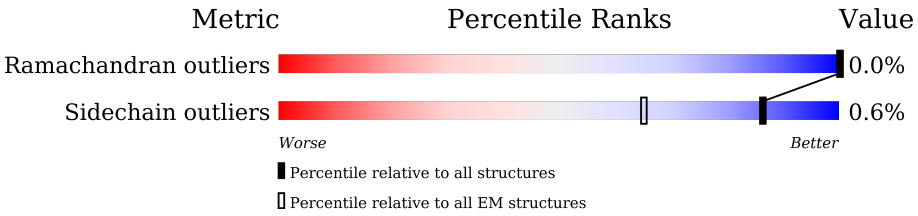
EMDB validation analysis	:	0.0.1.dev43
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
MapQ	:	1.9.9
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.31.2

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)
Ramachandran outliers	154571	4023
Sidechain outliers	154315	3826

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% 77% 22%
1	e	364	6% 87% 12%
2	6	301	35% 98%
2	7	301	33% 94% 5%
2	f	301	14% 96%
2	g	301	16% 95%
3	G	570	8% 14% 85%
3	K	570	6% 12% 88%
4	C	507	36% 67% 32%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
5	B	3149	99%
5	O	3149	99%
6	2	176	38% 40% 58%
6	Y	176	23% 41% 58%
6	Z	176	26% 41% 58%
6	y	176	27% 41% 58%
7	S	1381	16% 93% 6%
7	T	1381	13% 93% 6%
7	W	1381	15% 96%
7	x	1381	15% 92% 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
1	5	284	Total	C	N	O	S	0	0
			2199	1421	374	397	7		
1	e	319	Total	C	N	O	S	0	0
			2505	1608	444	446	7		

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

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

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

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

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

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

- 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:

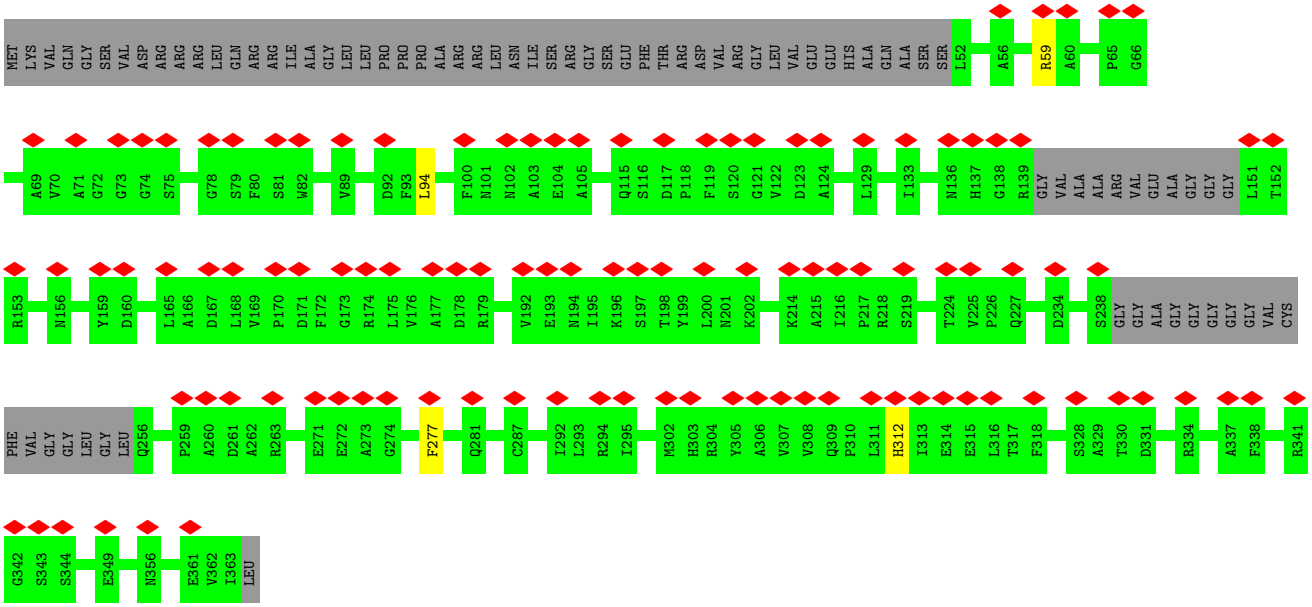
Chain	Residue	Modelled	Actual	Comment	Reference
S	89	ILE	THR	conflict	UNP P03226
T	89	ILE	THR	conflict	UNP P03226
W	89	ILE	THR	conflict	UNP P03226
x	89	ILE	THR	conflict	UNP P03226

3 Residue-property plots (i)

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

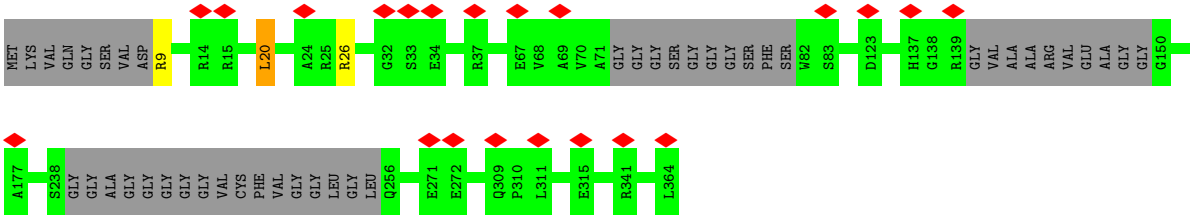
• Molecule 1: Triplex capsid protein 1

Chain 5:



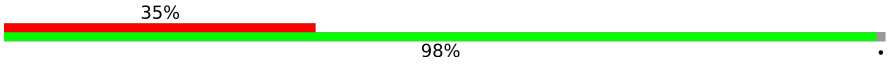
• Molecule 1: Triplex capsid protein 1

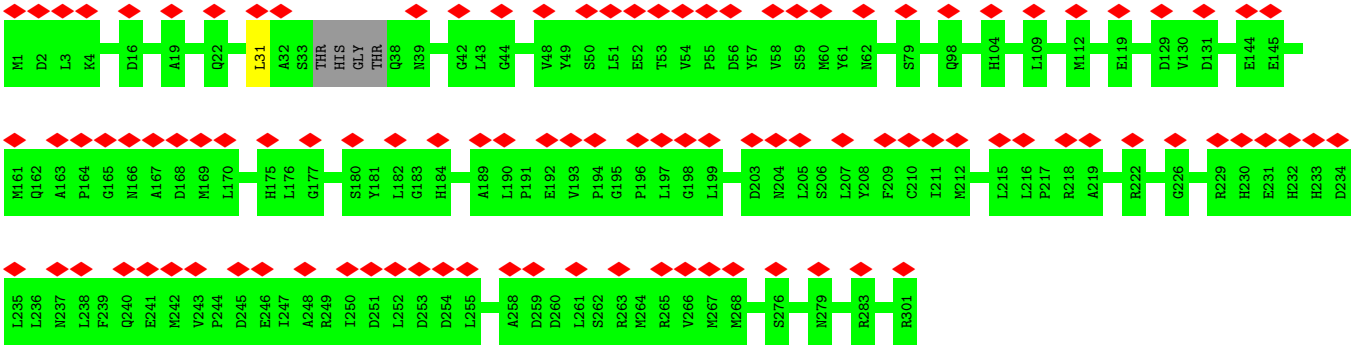
Chain e:



• Molecule 2: Triplex capsid protein 2

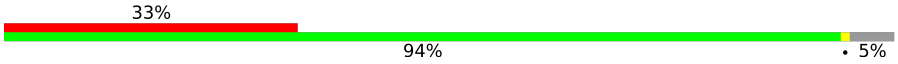
Chain 6:

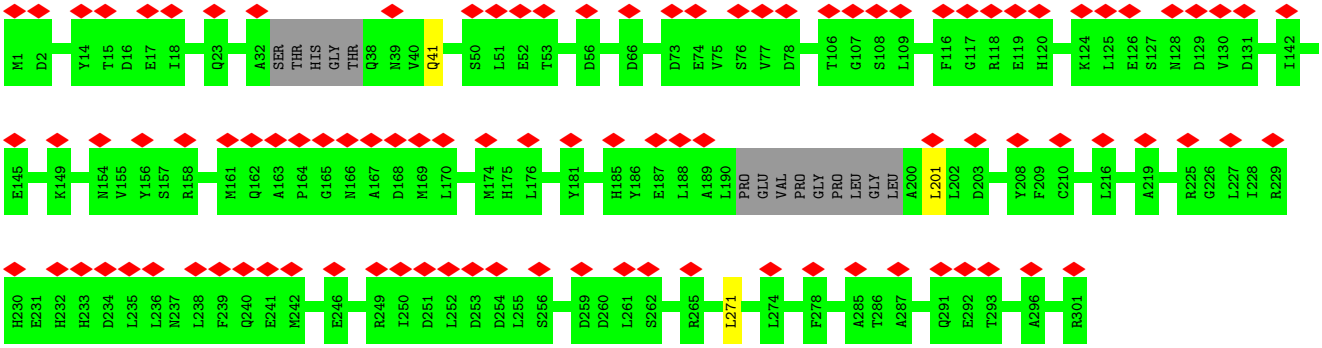




• Molecule 2: Triplex capsid protein 2

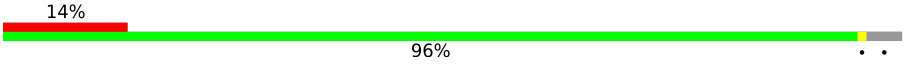
Chain 7:

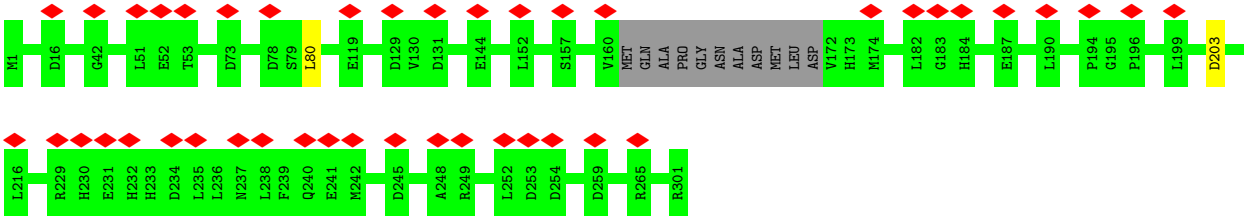




• Molecule 2: Triplex capsid protein 2

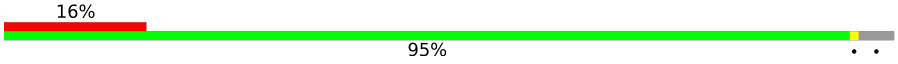
Chain f:

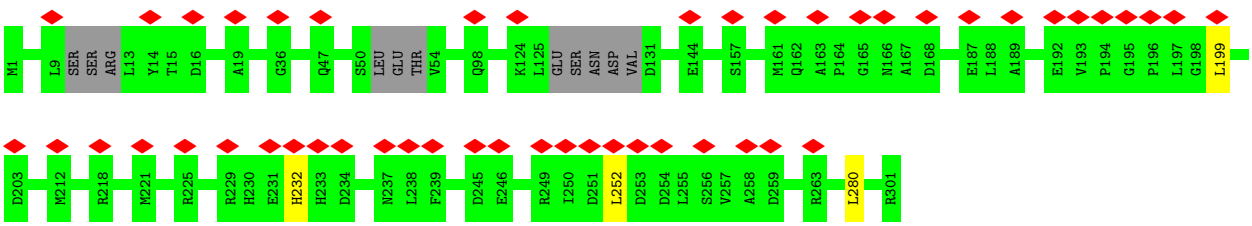




• Molecule 2: Triplex capsid protein 2

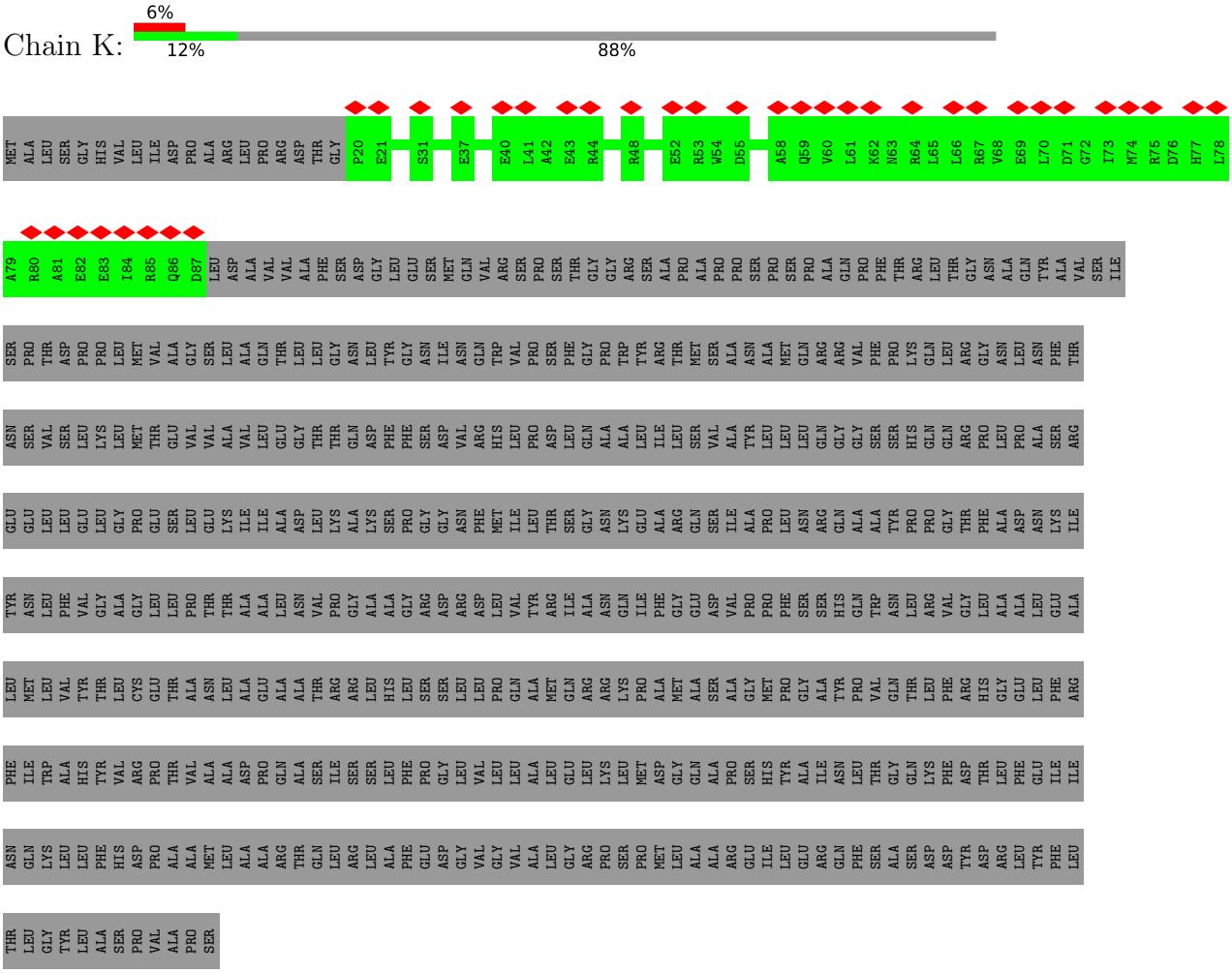
Chain g:





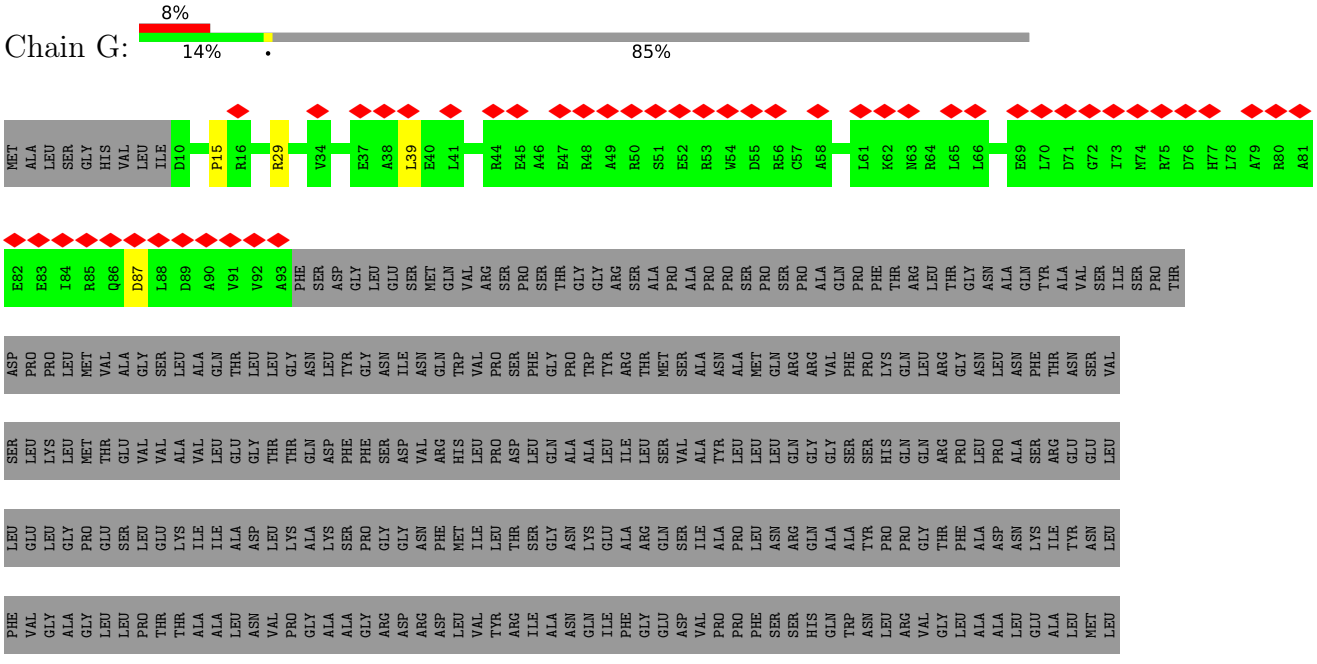
• Molecule 3: Capsid vertex component 2

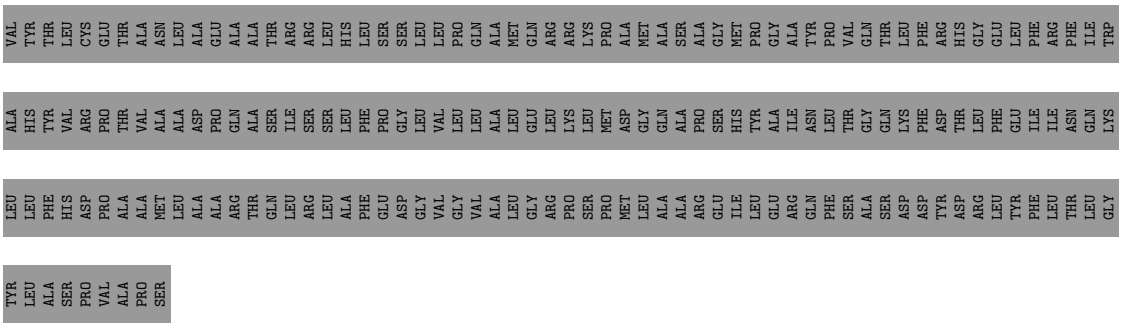
Chain K:



• Molecule 3: Capsid vertex component 2

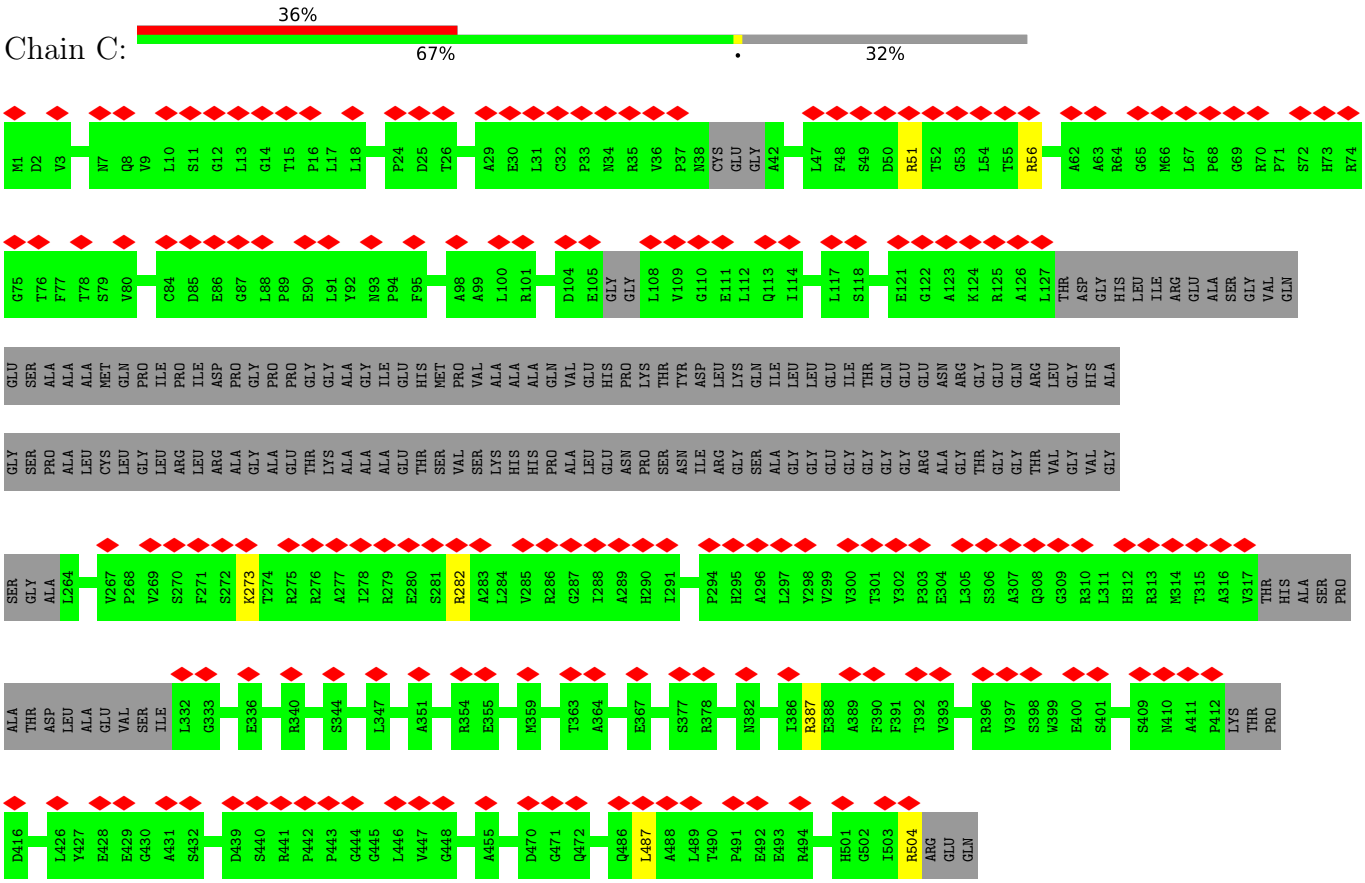
Chain G:





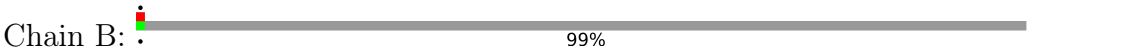
• Molecule 4: Capsid vertex component 1

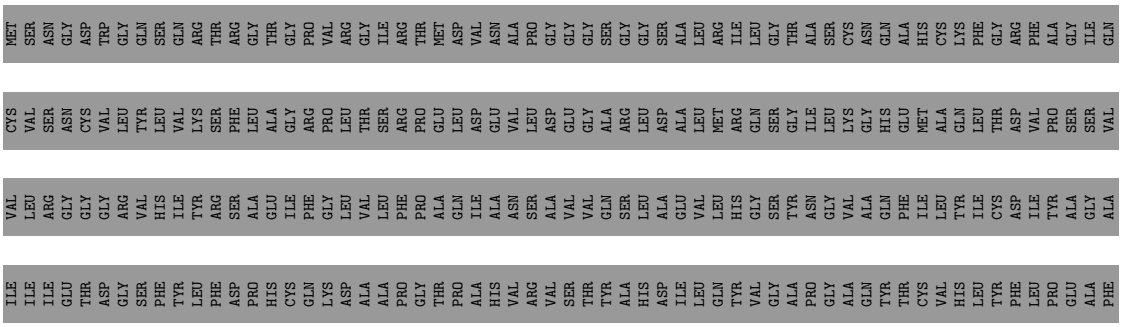
Chain C:

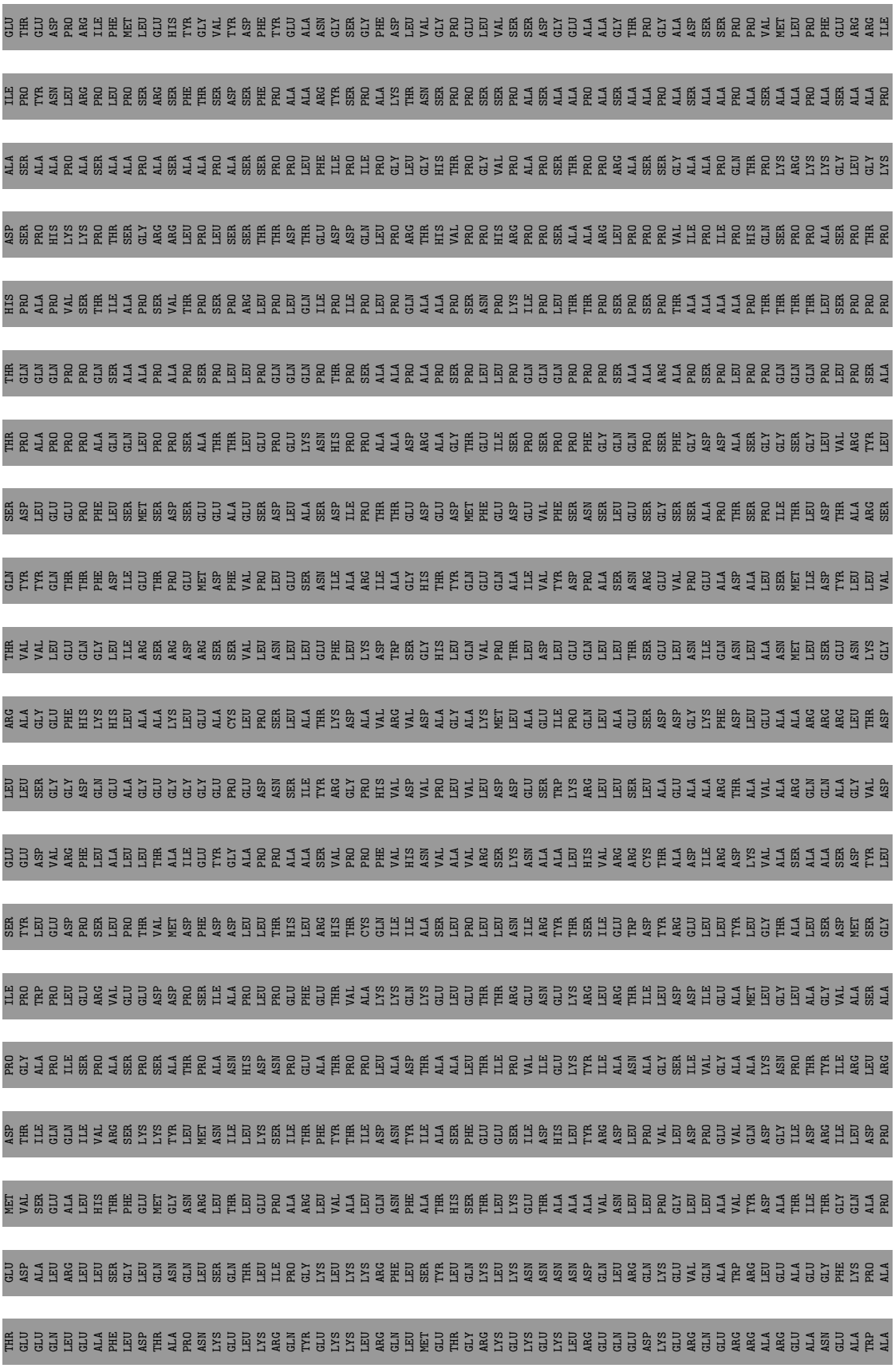


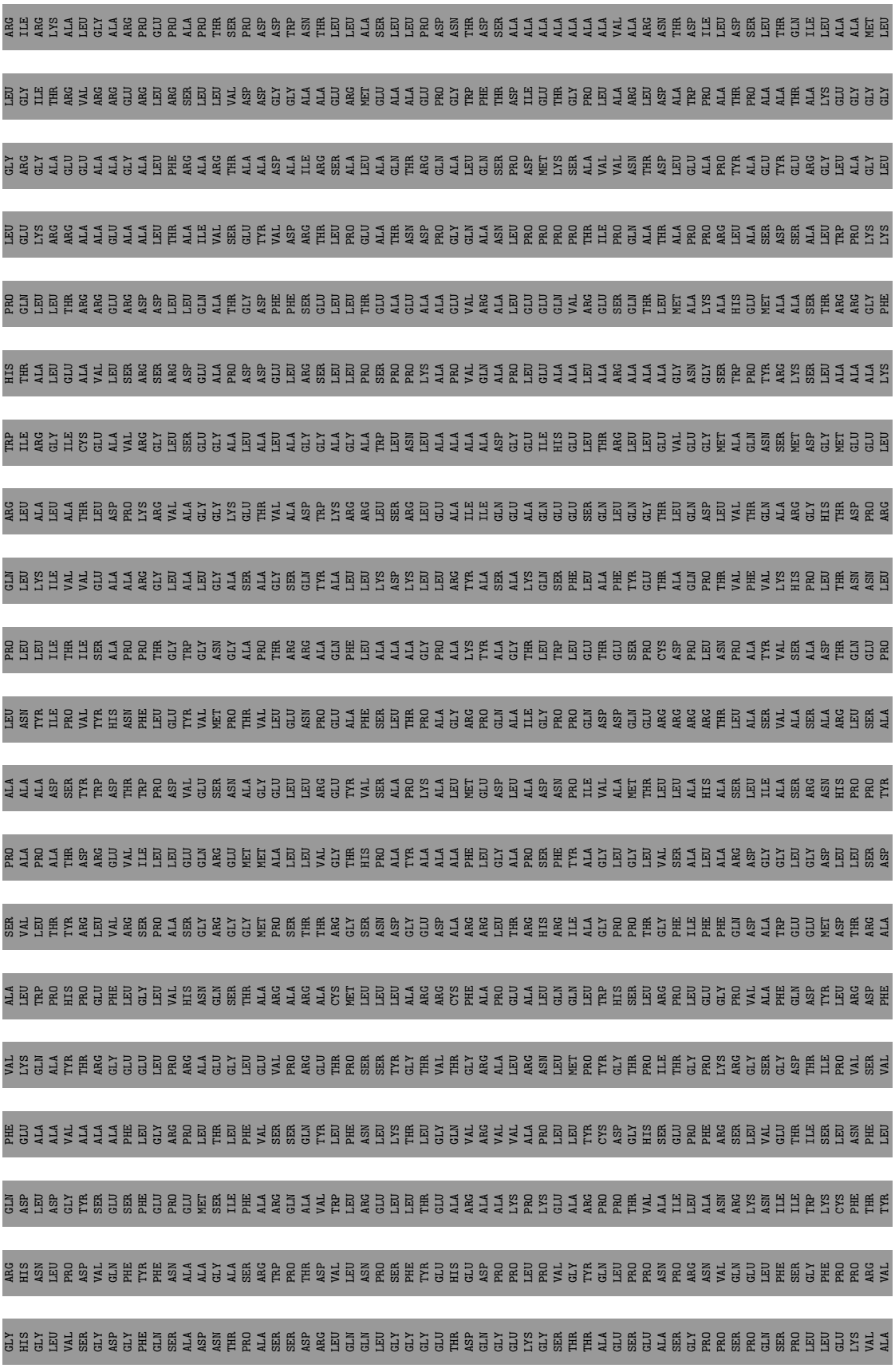
• Molecule 5: Large tegument protein deneddylase

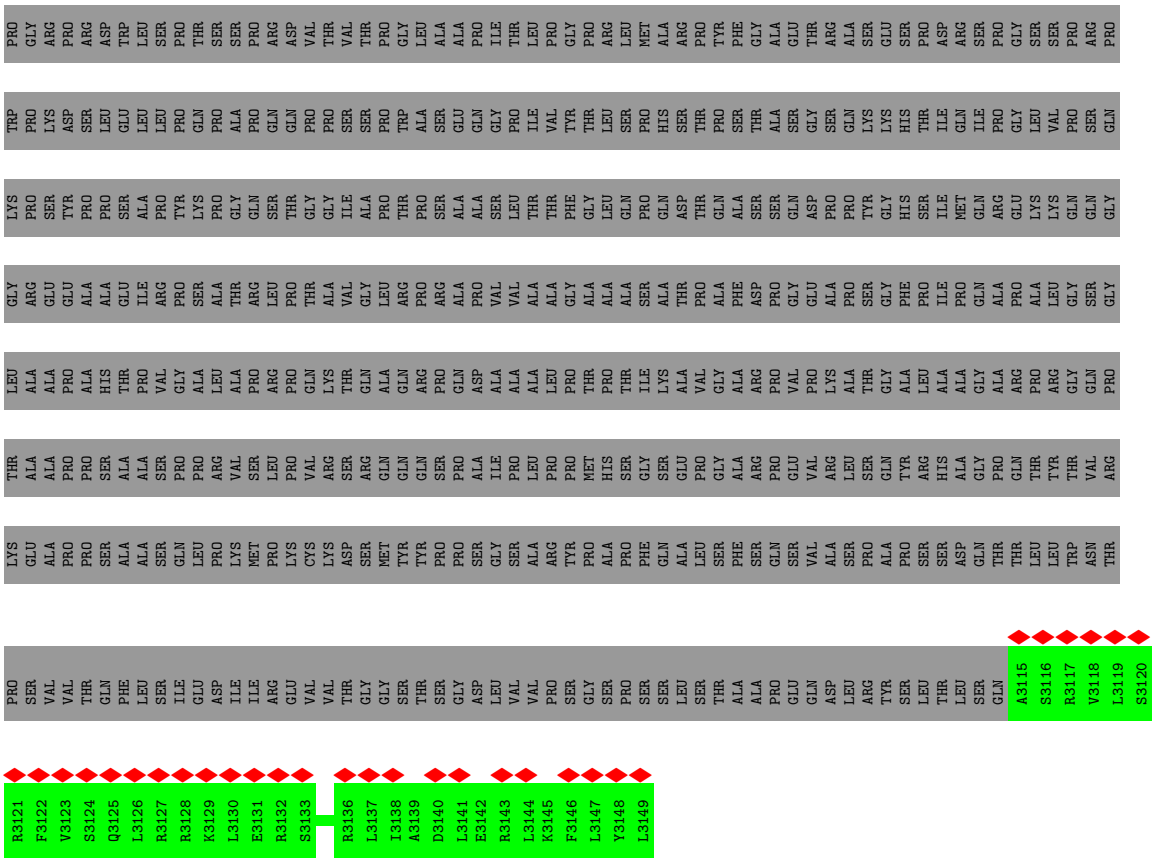
Chain B:





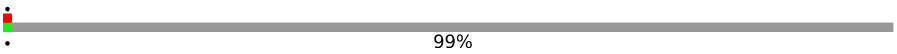

GLU THR GLU ASP PRO ARG ILE PHE MET LEU GLU HIS TYR GLY VAL TYR ASP PHE TYR GLU ALA ASN GLY SER GLY PHE ASP LEU VAL GLY PRO GLU GLN LEU VAL SER SER ASP GLY GLU ALA ALA GLY THR PRO GLY ALA ASP SER SER PRO PRO VAL MET LEU PRO PHE GLU ARG ARG ILE
ILE PRO TYR ASN LEU ARG PRO LEU PRO SER ARG SER PHE THR SER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

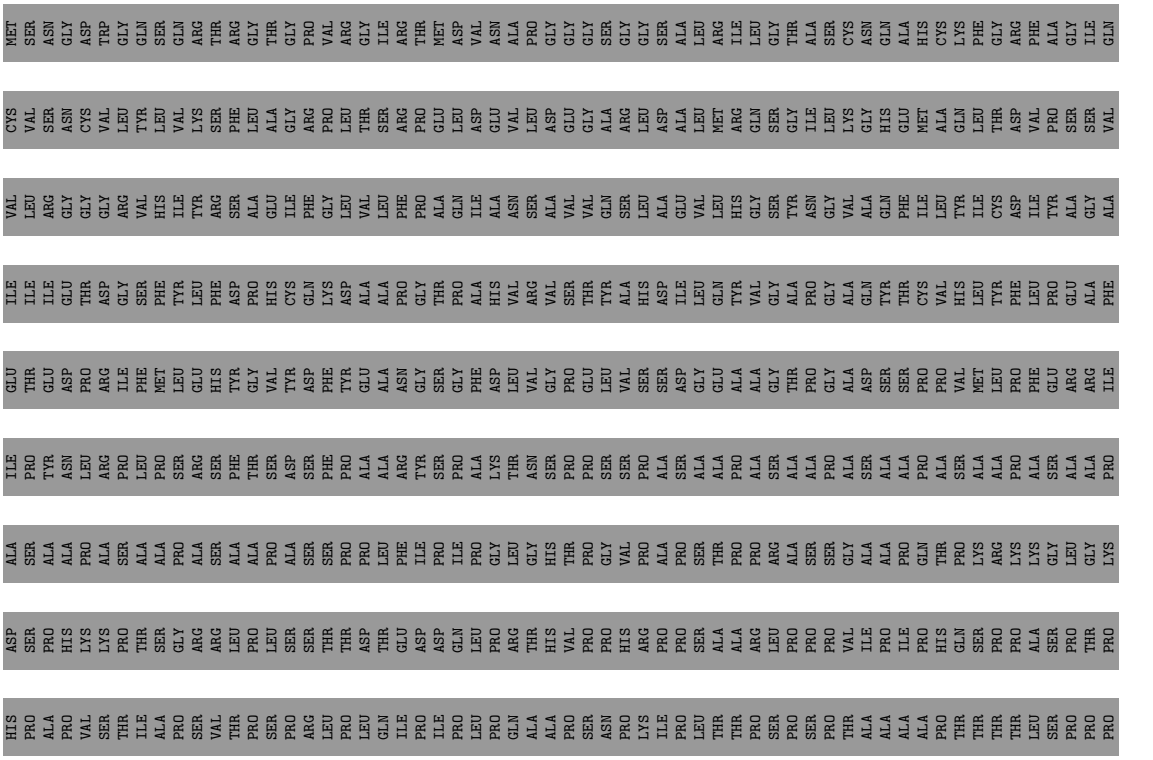

ARG ILE ARG LYS ALA LEU GLY ALA ARG PRO GLU PRO ALA PRO THR SER PRO ASP ASP TRP ASN THR LEU LEU ALA SER LEU LEU PRO ASP ASN THR ASP SER ALA ALA ALA ALA ALA ALA VAL ALA ARG ASN THR ASP ILE LEU ASP SER LEU THR GLN ILE LEU ALA ALA MET LEU
LEU GLY ILE THR ARG VAL ARG ARG GLU ARG LEU ARG SER LEU LEU VAL ASP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

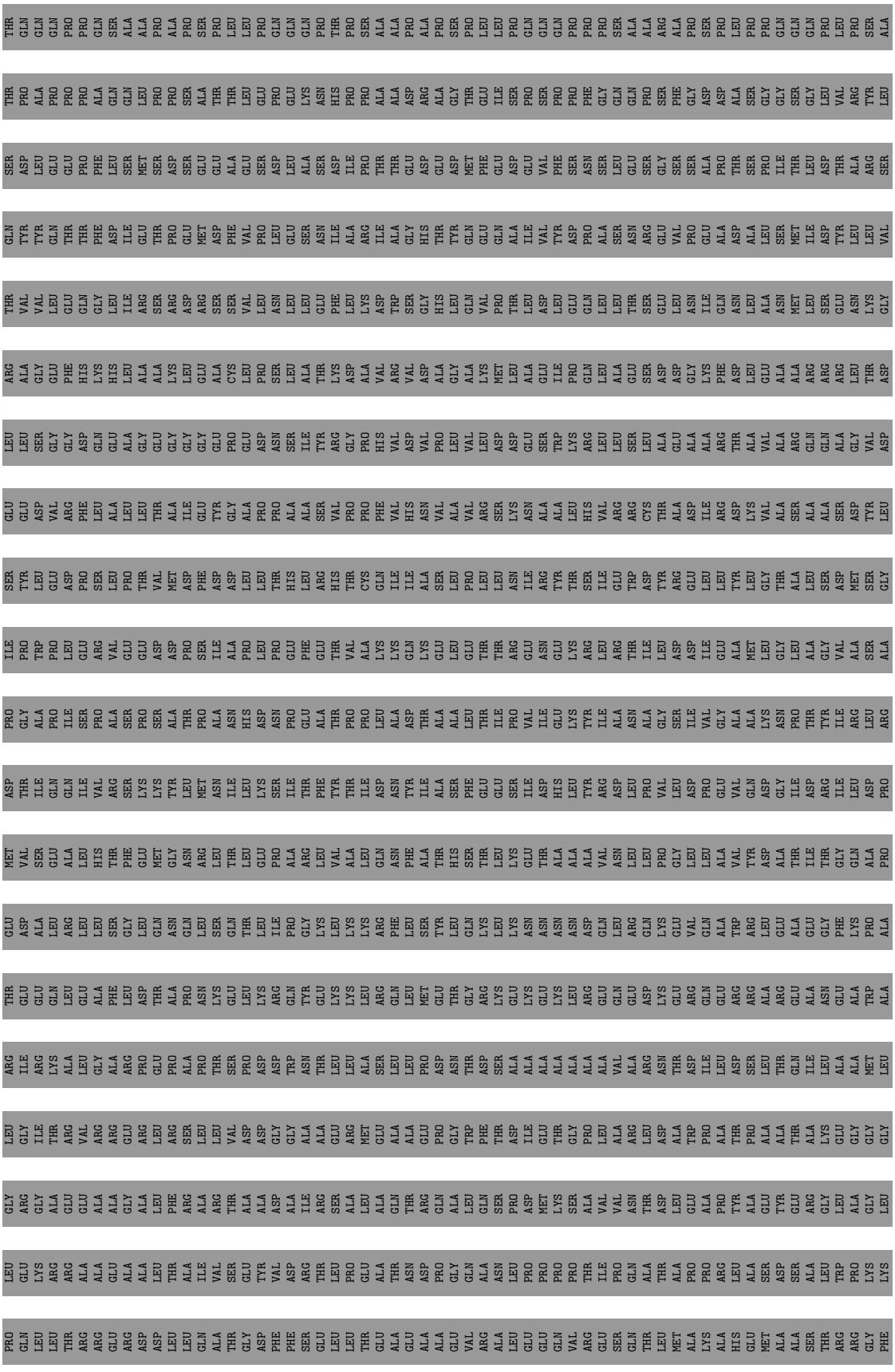


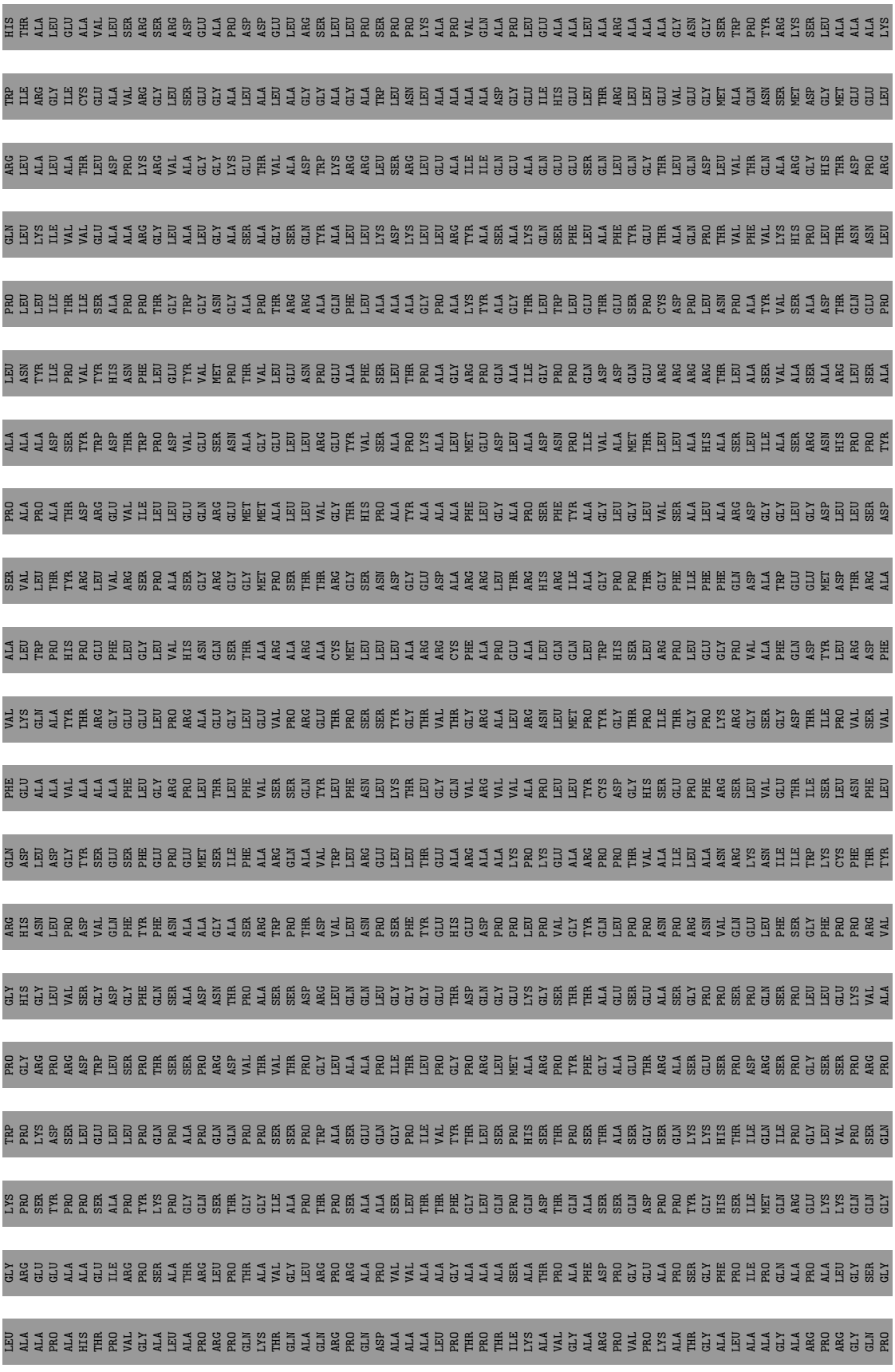
• Molecule 5: Large tegument protein deneddylase

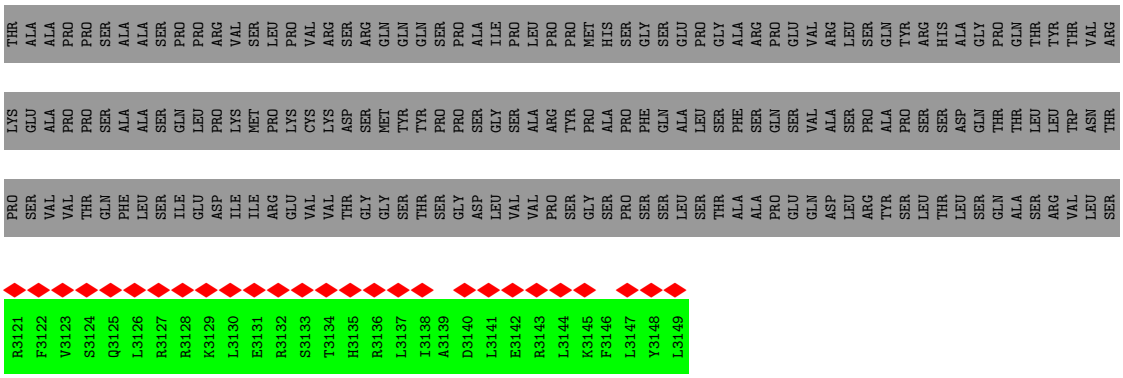
Chain O:





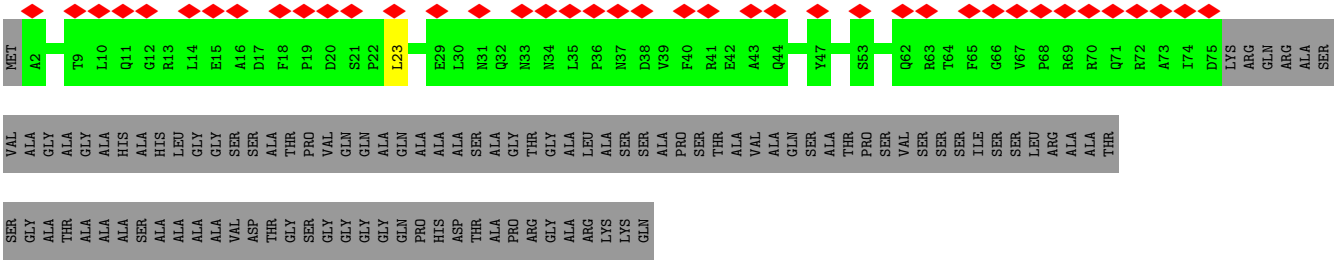

THR GLN GLN GLN PRO PRO GLN SER ALA ALA PRO ALA PRO SER PRO LEU LEU LEU PRO GLN GLN GLN PRO PRO PRO SER ALA ALA ARG ALA PRO SER PRO LEU LEU PRO GLN GLN GLN PRO PRO PRO GLY ALA PRO LEU PRO GLN GLN GLN PRO PRO PRO ALA GLN GLN LEU PRO PRO PRO SER ALA ALA ASP ARG ALA GLY THR GLU ILE SER PRO GLY GLN SER GLY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


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



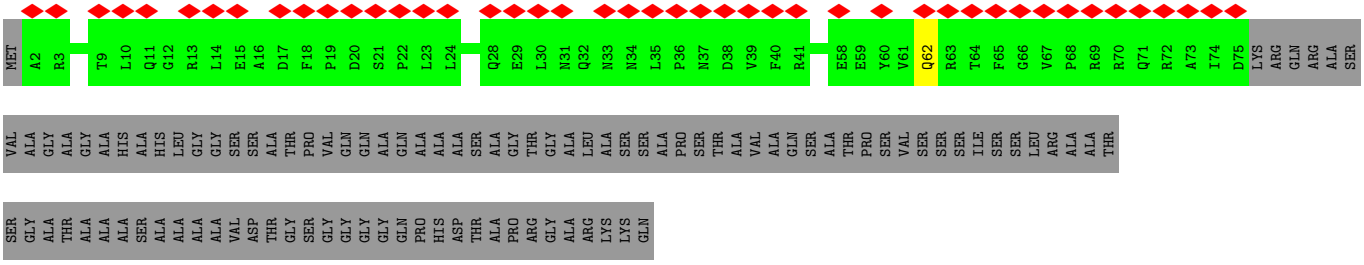
• Molecule 6: Small capsomere-interacting protein

Chain Y:



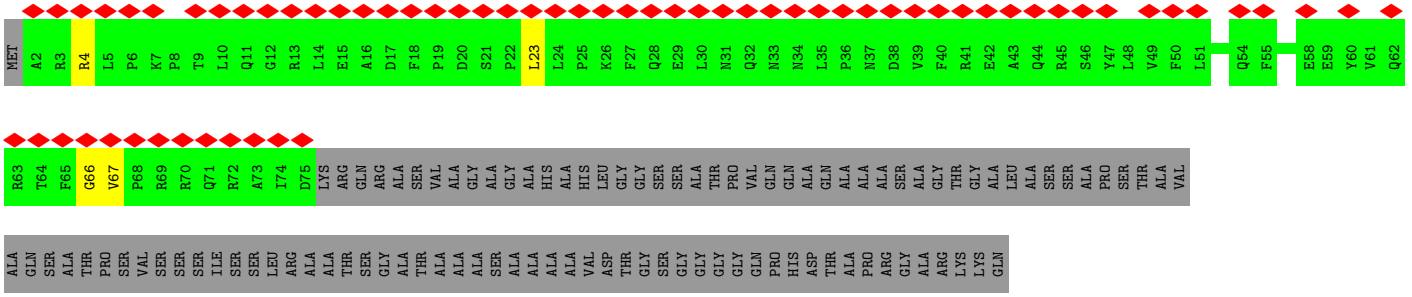
• Molecule 6: Small capsomere-interacting protein

Chain Z:



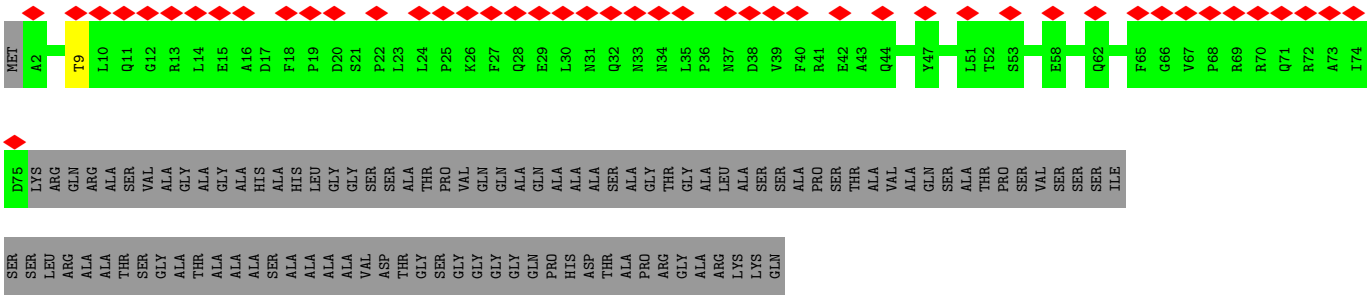
• Molecule 6: Small capsomere-interacting protein

Chain 2:



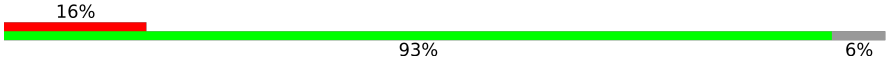
• Molecule 6: Small capsomere-interacting protein

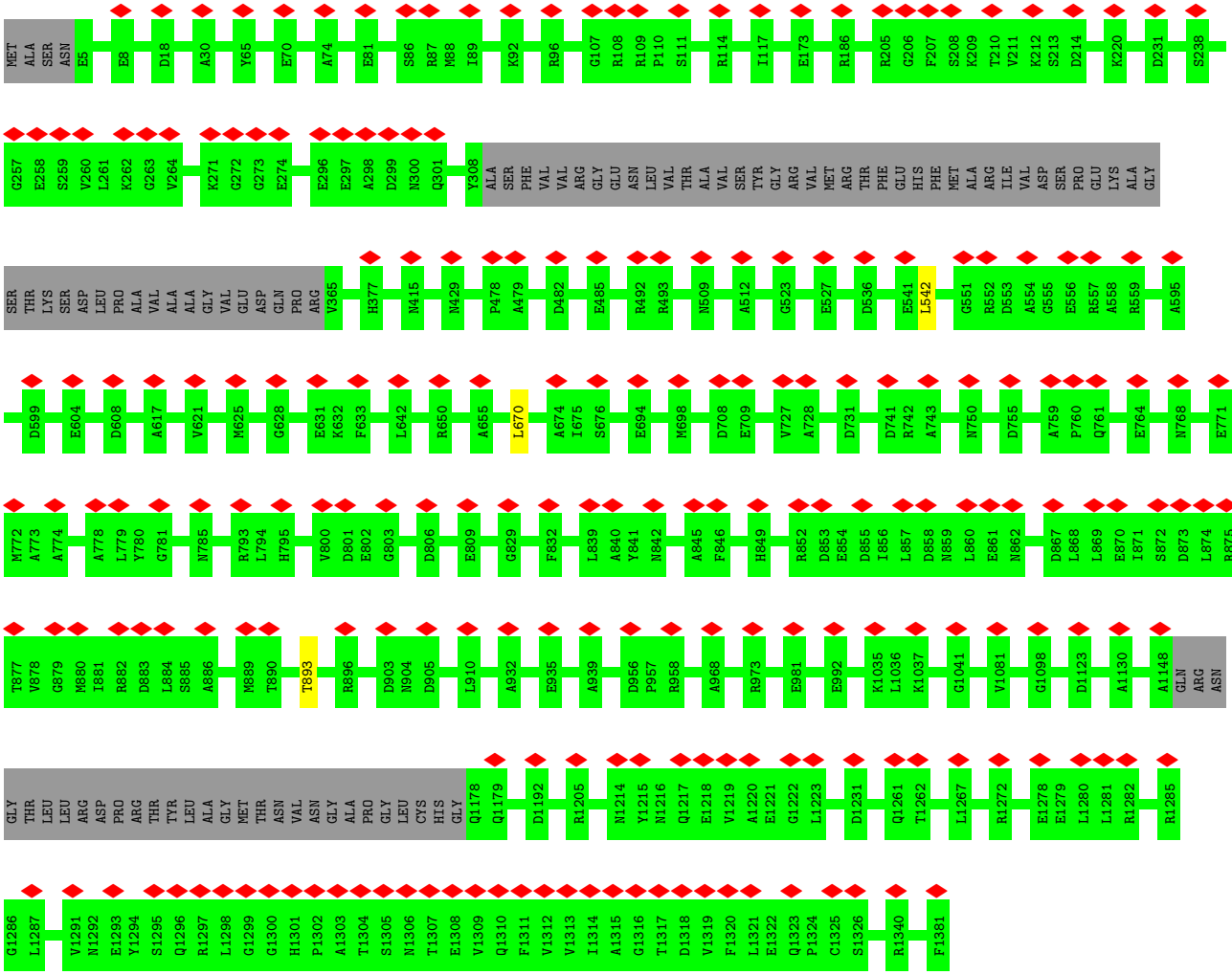
Chain y:



• Molecule 7: Major capsid protein

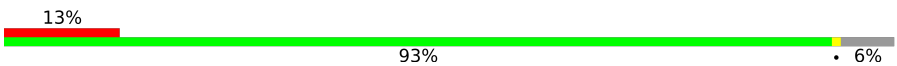
Chain S:

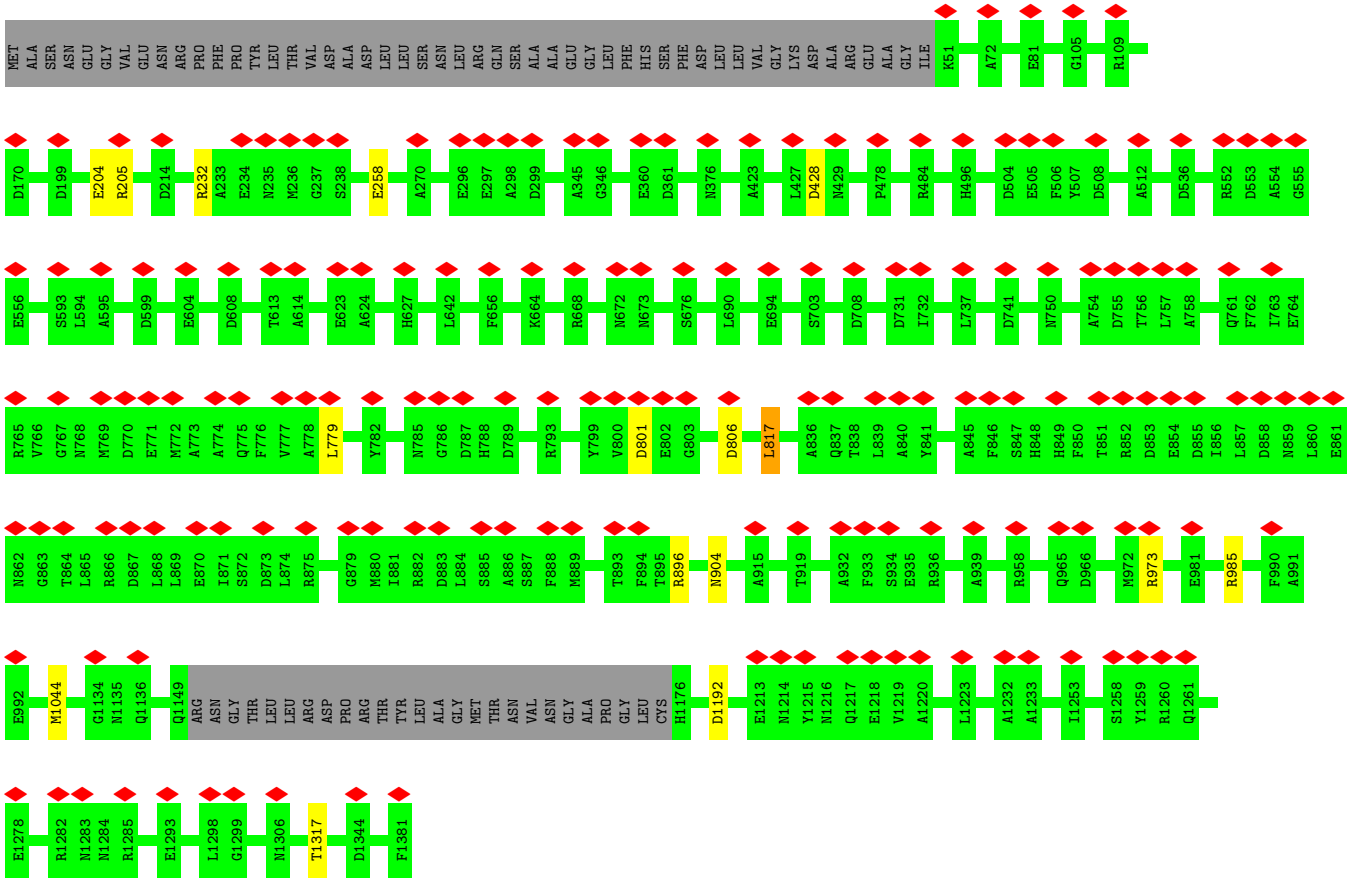




• Molecule 7: Major capsid protein

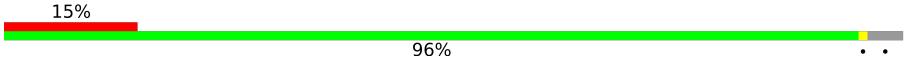
Chain T:

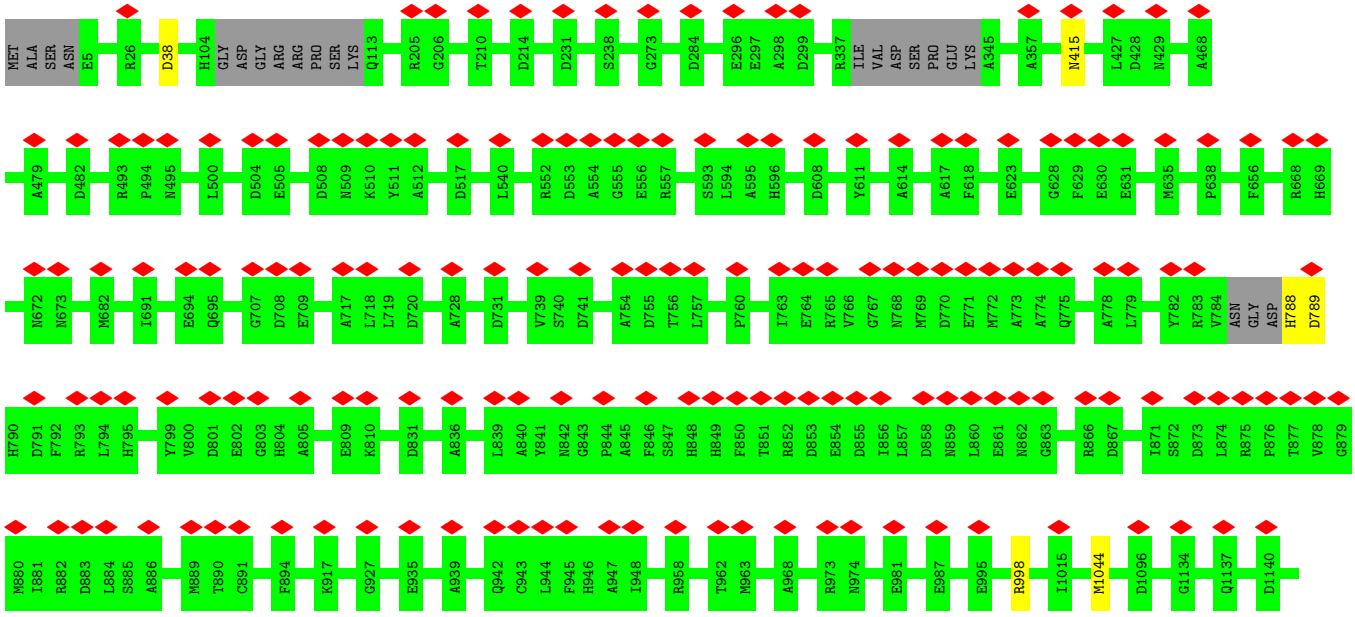


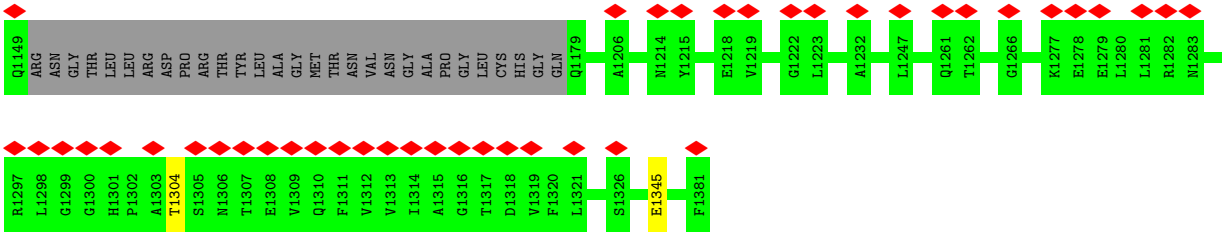


• Molecule 7: Major capsid protein

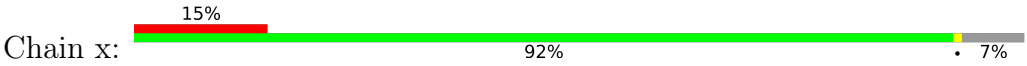
Chain W:

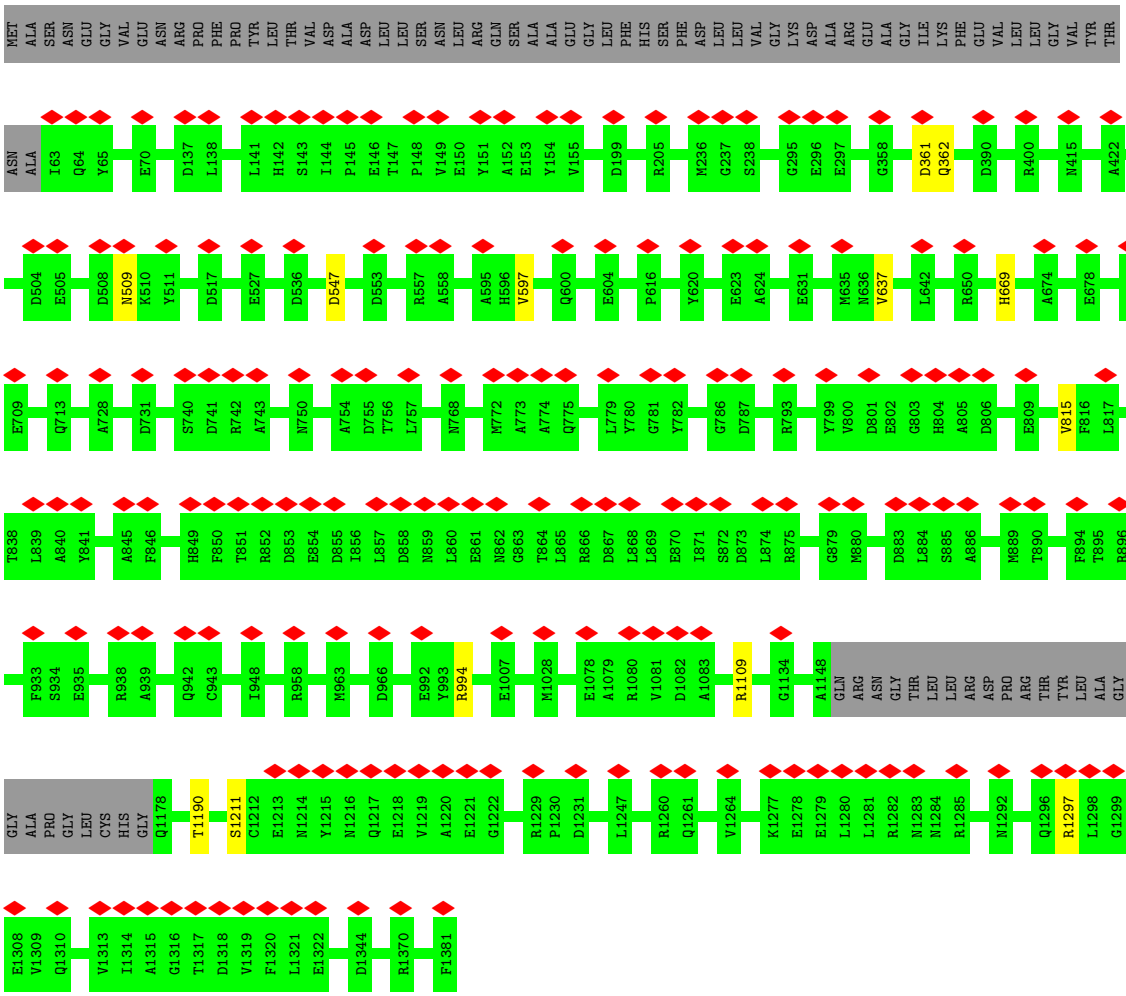






• Molecule 7: Major capsid protein





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 ($e^-/Å^2$)	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

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	5	0.34	0/2263	0.56	1/3088 (0.0%)
1	e	0.42	0/2572	0.58	1/3503 (0.0%)
2	6	0.35	0/2377	0.62	1/3236 (0.0%)
2	7	0.36	0/2306	0.65	2/3135 (0.1%)
2	f	0.36	0/2327	0.57	1/3169 (0.0%)
2	g	0.38	0/2319	0.61	2/3155 (0.1%)
3	G	0.32	0/684	0.66	1/924 (0.1%)
3	K	0.30	0/570	0.56	0/765
4	C	0.35	0/2770	0.62	2/3763 (0.1%)
5	B	0.27	0/304	0.51	0/404
5	O	0.30	0/255	0.69	0/339
6	2	0.34	0/636	0.69	2/861 (0.2%)
6	Y	0.38	0/636	0.59	1/861 (0.1%)
6	Z	0.33	0/636	0.52	0/861
6	y	0.33	0/636	0.56	0/861
7	S	0.43	0/10405	0.58	1/14142 (0.0%)
7	T	0.43	0/10500	0.59	6/14271 (0.0%)
7	W	0.40	0/10699	0.57	1/14539 (0.0%)
7	x	0.43	1/10380 (0.0%)	0.59	0/14109
All	All	0.40	1/63275 (0.0%)	0.59	22/85986 (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 (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
7	x	362	GLN	C-N	8.95	1.51	1.34

The worst 5 of 22 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
7	T	817	LEU	CA-CB-CG	8.12	133.97	115.30
6	2	23	LEU	CA-CB-CG	6.79	130.92	115.30
4	C	487	LEU	CA-CB-CG	6.74	130.81	115.30
2	g	199	LEU	CA-CB-CG	6.58	130.43	115.30
6	Y	23	LEU	CA-CB-CG	6.50	130.25	115.30

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)

Due to software issues we are unable to calculate clashes - this section is therefore empty.

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%)	13	50
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%)	51	85
7	x	1286/1381 (93%)	1196 (93%)	89 (7%)	1 (0%)	51	85
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 (i)

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%)	232 (99%)	2 (1%)	78 87
1	e	268/289 (93%)	266 (99%)	2 (1%)	84 90
2	6	264/267 (99%)	264 (100%)	0	100 100
2	7	256/267 (96%)	255 (100%)	1 (0%)	91 94
2	f	259/267 (97%)	258 (100%)	1 (0%)	91 94
2	g	256/267 (96%)	255 (100%)	1 (0%)	91 94
3	G	70/465 (15%)	68 (97%)	2 (3%)	42 64
3	K	59/465 (13%)	59 (100%)	0	100 100
4	C	288/400 (72%)	283 (98%)	5 (2%)	60 78
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%)	67 (98%)	1 (2%)	65 80
6	Y	68/128 (53%)	68 (100%)	0	100 100
6	Z	68/128 (53%)	67 (98%)	1 (2%)	65 80
6	y	68/128 (53%)	67 (98%)	1 (2%)	65 80
7	S	1101/1171 (94%)	1099 (100%)	2 (0%)	93 96
7	T	1111/1171 (95%)	1101 (99%)	10 (1%)	78 87
7	W	1131/1171 (97%)	1125 (100%)	6 (0%)	88 93
7	x	1099/1171 (94%)	1091 (99%)	8 (1%)	84 90
All	All	6728/13250 (51%)	6685 (99%)	43 (1%)	86 92

5 of 43 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
7	W	1304	THR
7	x	1109	ARG
7	W	1345	GLU
7	x	637	VAL
7	x	1297	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 106 such sidechains are listed below:

Mol	Chain	Res	Type
7	W	433	GLN
7	W	1149	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	e	137	HIS
7	W	534	ASN
7	W	942	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 monosaccharides 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.

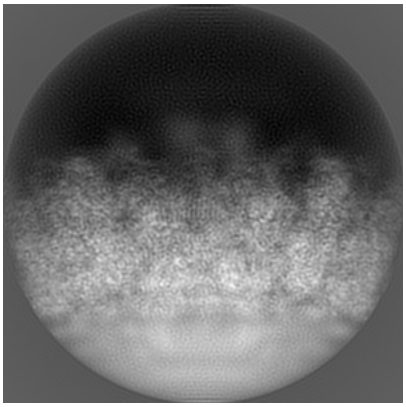
6 Map visualisation

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.

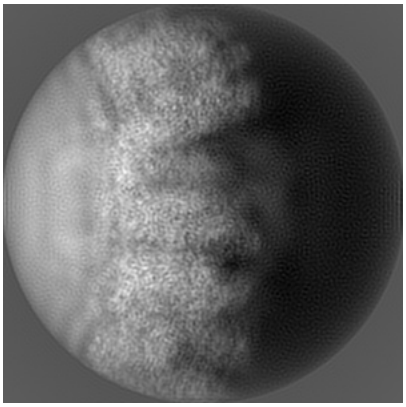
No raw map or half-maps were deposited for this entry and therefore no images, graphs, etc. pertaining to the raw map can be shown.

6.1 Orthogonal projections

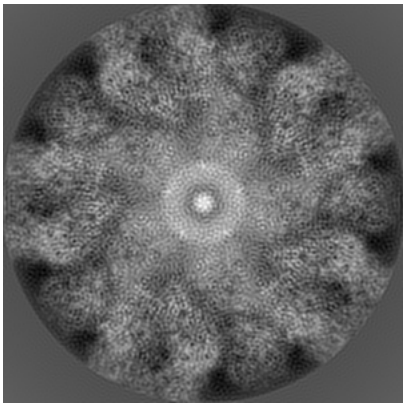
6.1.1 Primary map



X



Y

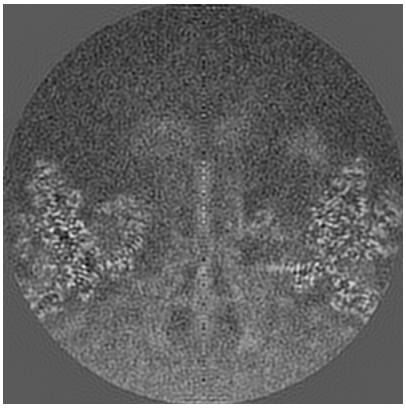


Z

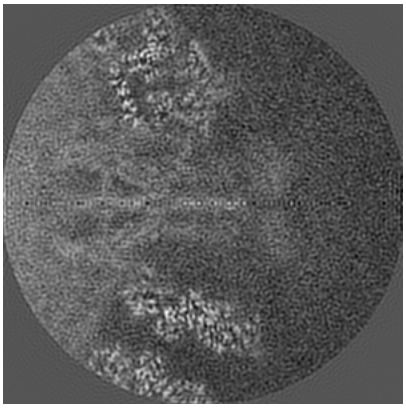
The images above show the map projected in three orthogonal directions.

6.2 Central slices

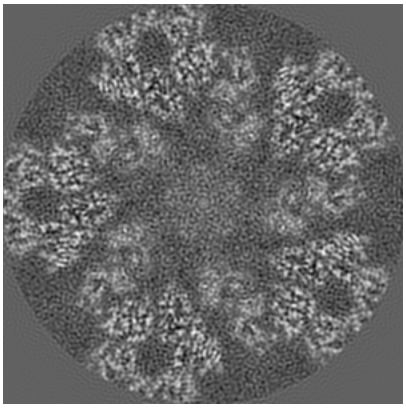
6.2.1 Primary map



X Index: 150



Y Index: 150

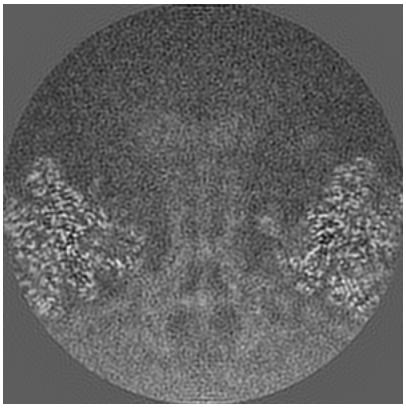


Z Index: 150

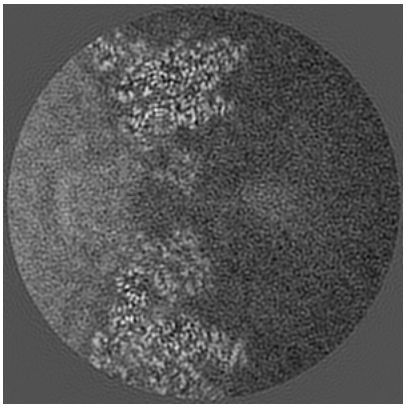
The images above show central slices of the map in three orthogonal directions.

6.3 Largest variance slices (i)

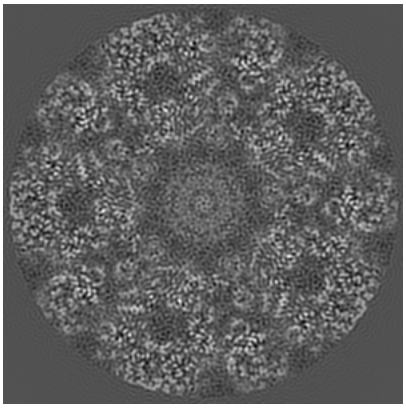
6.3.1 Primary map



X Index: 144



Y Index: 114



Z Index: 109

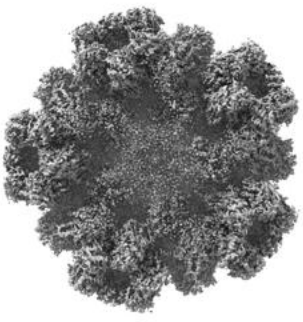
The images above show the largest variance slices of the map in three orthogonal directions.

6.4 Orthogonal surface views (i)

6.4.1 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.

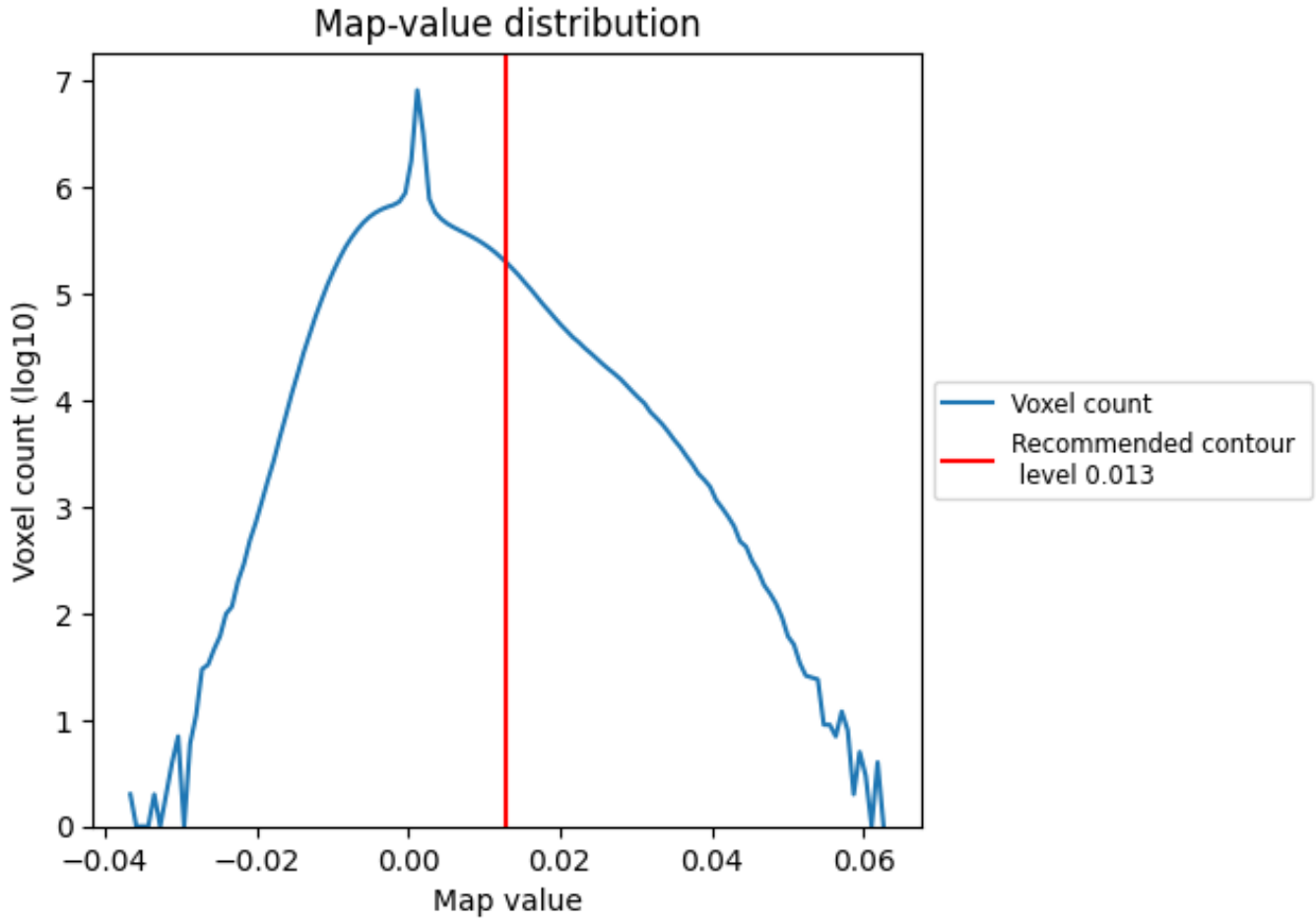
6.5 Mask visualisation (i)

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

7 Map analysis

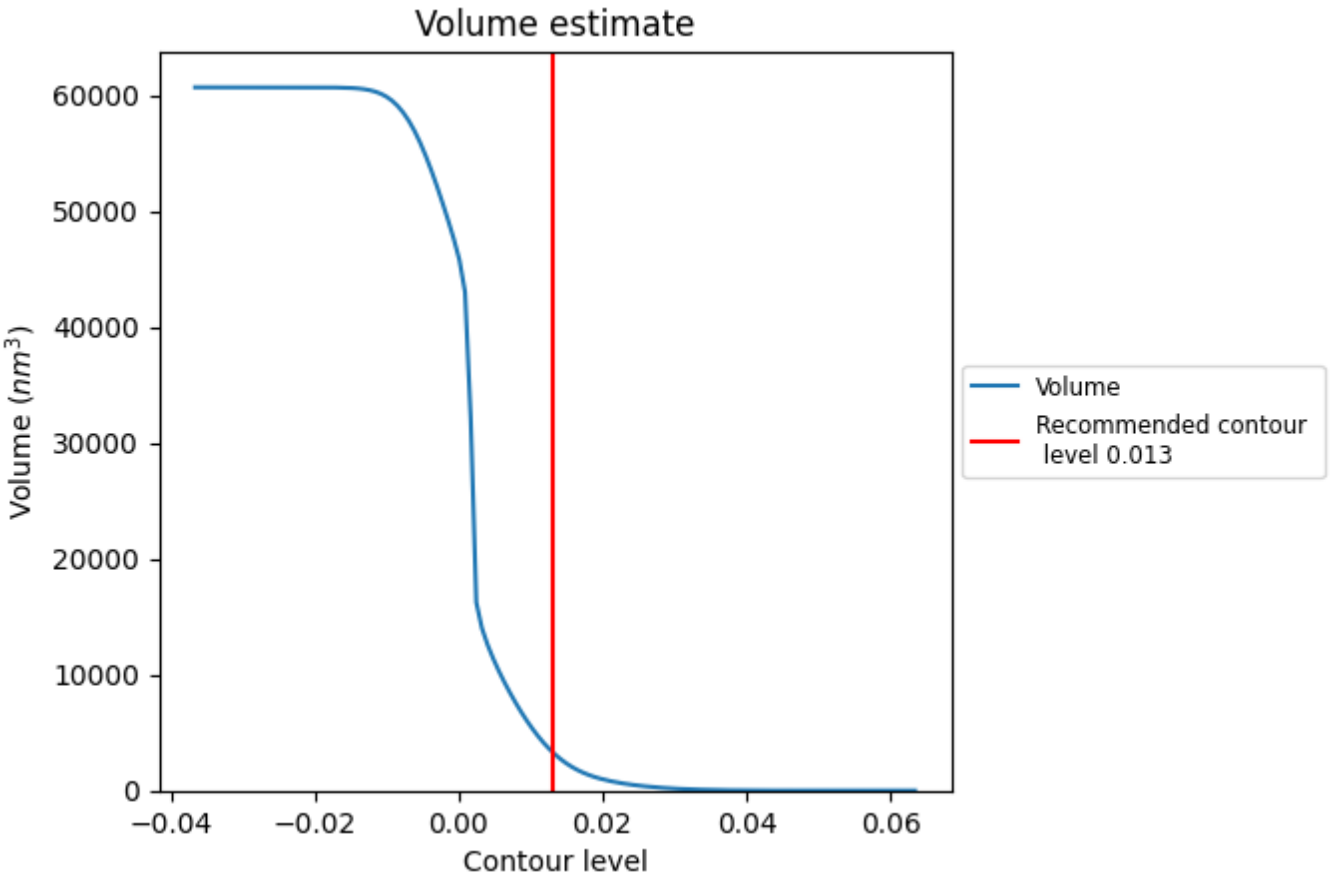
This section contains the results of statistical analysis of the map.

7.1 Map-value distribution



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.

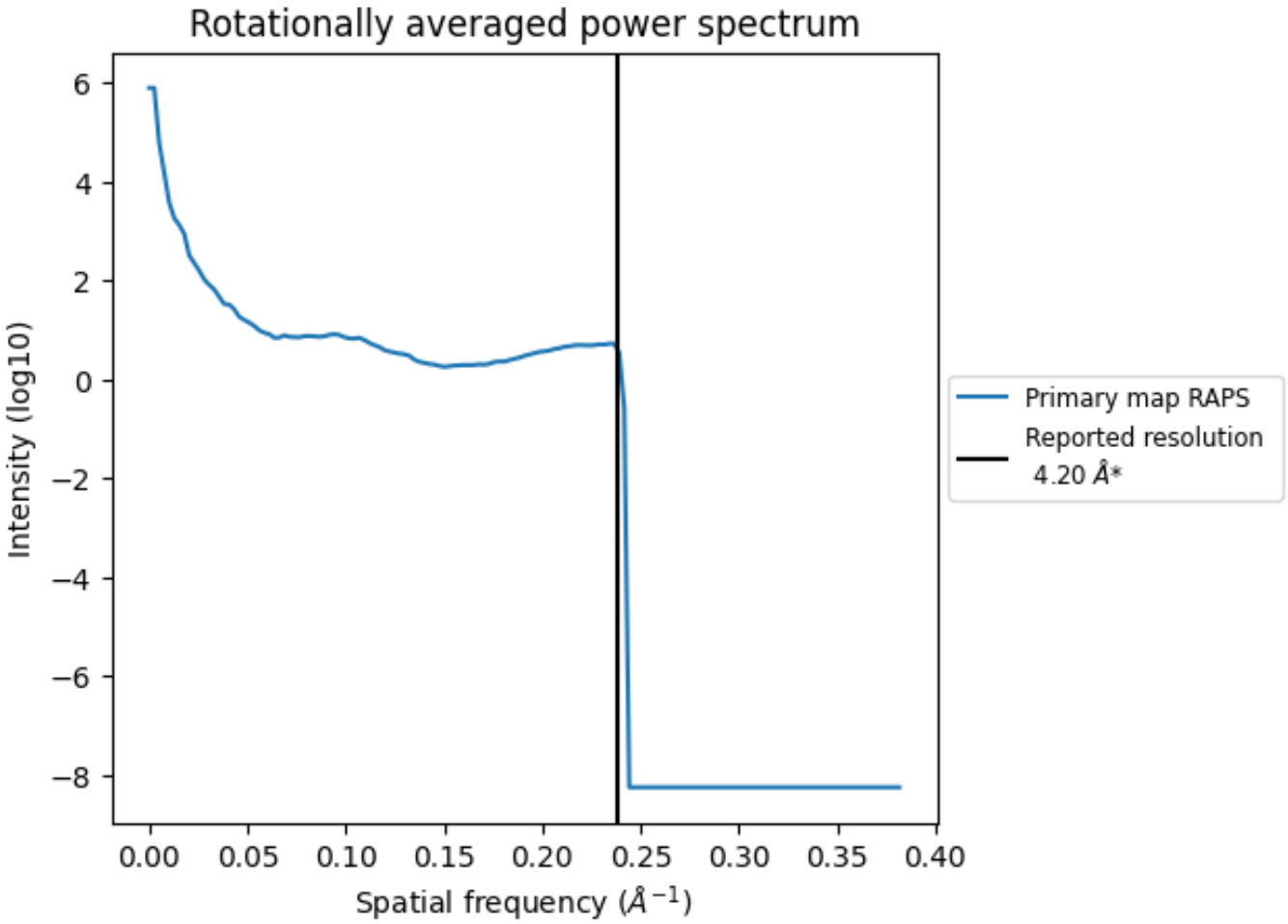
7.2 Volume estimate (i)



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 (i)

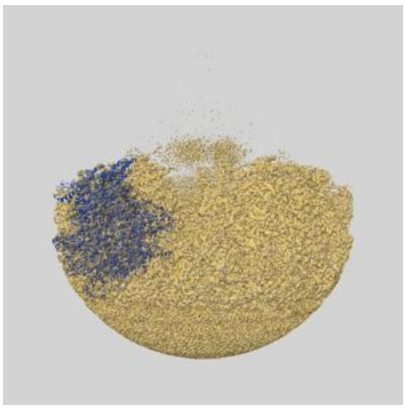
This section was not generated. No FSC curve or half-maps provided.

9 Map-model fit

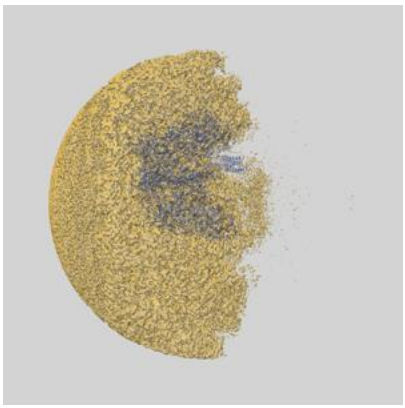
This section contains information regarding the fit between EMDB map EMD-30157 and PDB model 7BQX. Per-residue inclusion information can be found in section 3 on page 6.

9.1 Map-model overlays

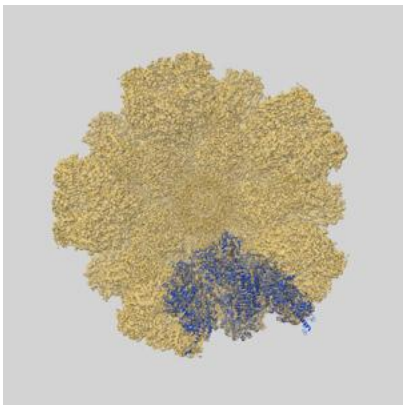
9.1.1 Map-model overlay



X

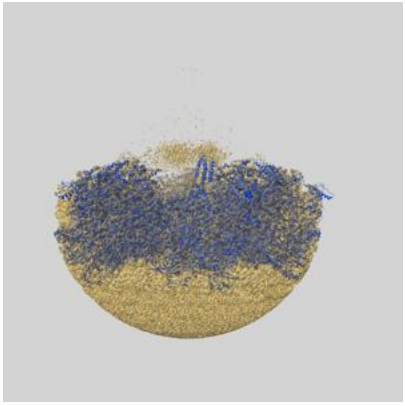


Y

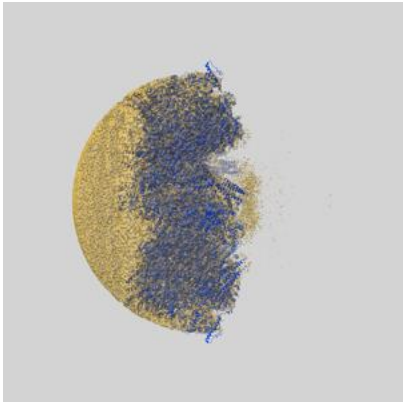


Z

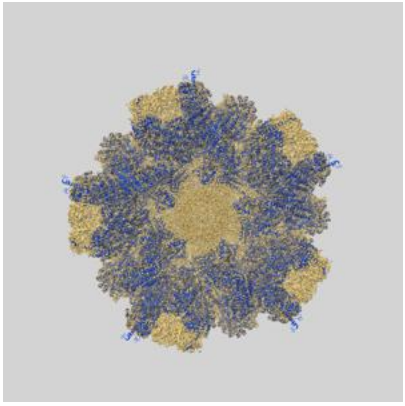
9.1.2 Map-model assembly overlay



X



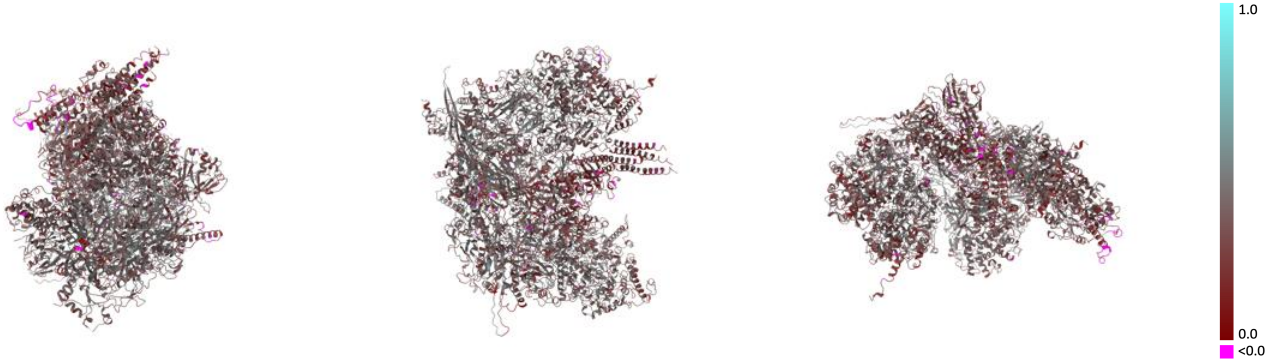
Y



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.

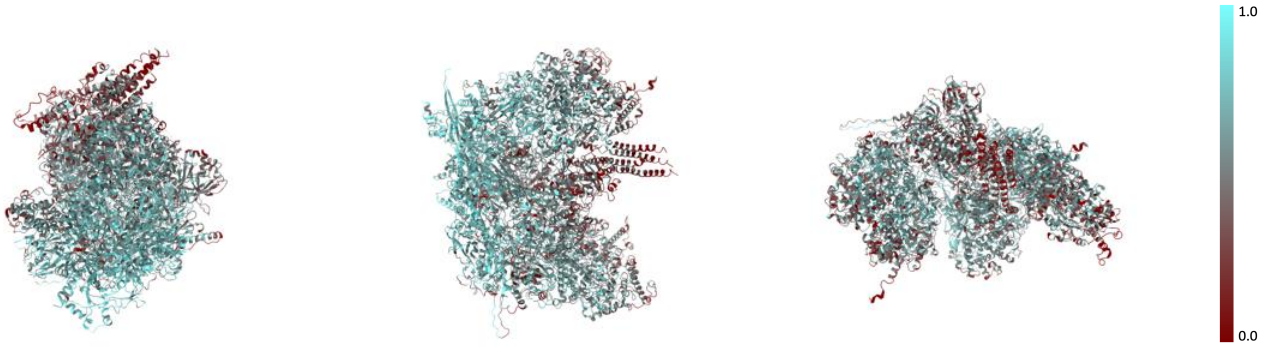
9.2 Q-score mapped to coordinate model (i)



X Y Z

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.

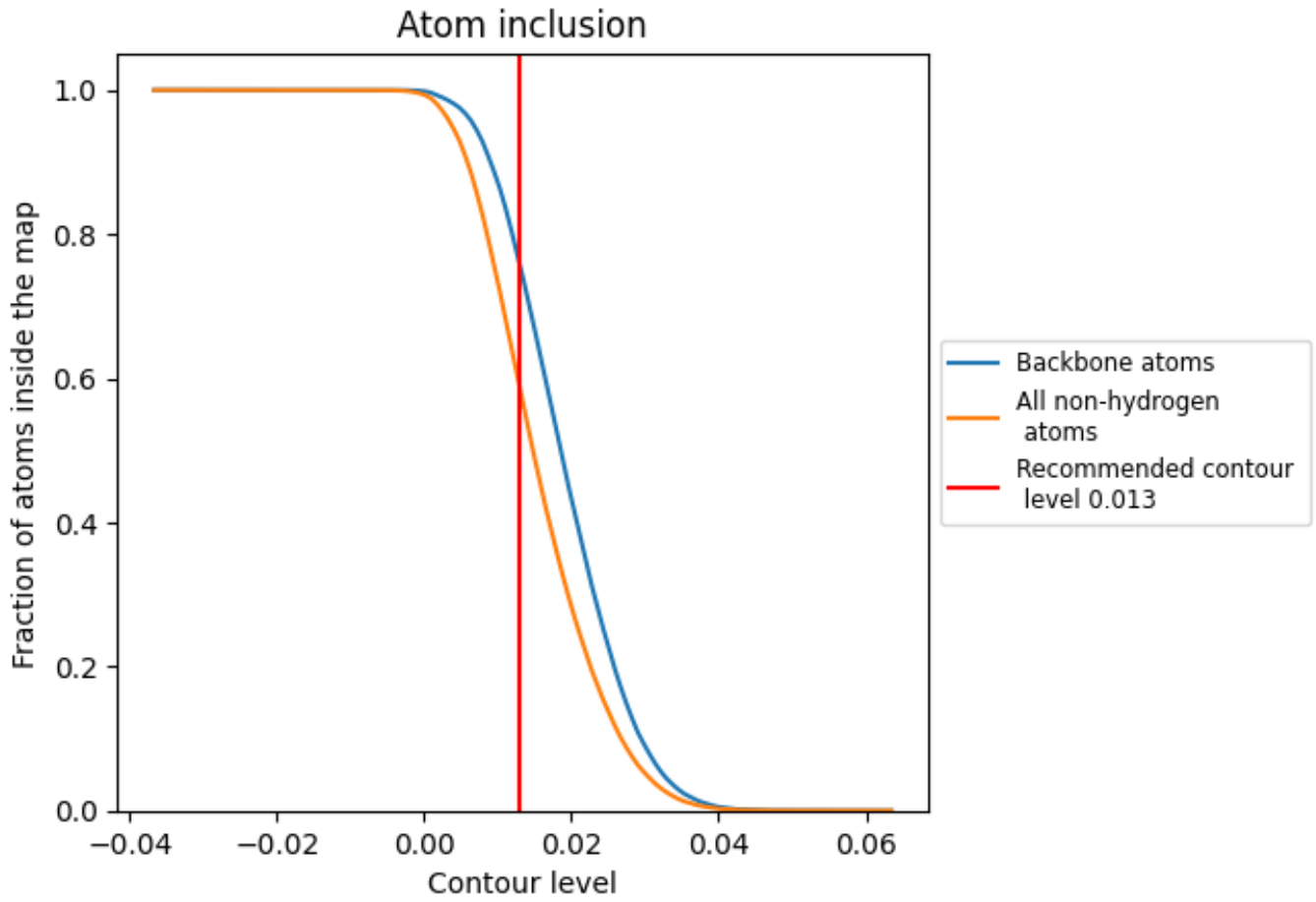
9.3 Atom inclusion mapped to coordinate model (i)



X Y Z

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	0.5866	0.3650
2	0.1333	0.1500
5	0.4498	0.3030
6	0.4873	0.3080
7	0.4885	0.3280
B	0.1818	0.2560
C	0.4062	0.2870
G	0.3643	0.2950
K	0.3985	0.3100
O	0.1369	0.2440
S	0.6199	0.3830
T	0.6582	0.3980
W	0.6295	0.3780
Y	0.3750	0.3180
Z	0.3150	0.2840
e	0.7125	0.4020
f	0.6291	0.3720
g	0.6382	0.3760
x	0.6270	0.3830
y	0.3167	0.2980