wwPDB EM Validation Summary Report (i)

Feb 25, 2024 – 08:06 AM EST

PDB ID	:	6X6K
EMDB ID	:	EMD-22076
Title	:	Cryo-EM Structure of the Helicobacter pylori dCag3 OMC
Authors	:	Sheedlo, M.J.; Chung, J.M.; Sawhney, N.; Durie, C.L.; Cover, T.L.; Ohi, M.D.; Lacy, D.B.
Deposited on	:	2020-05-28
Resolution	:	3.10 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:

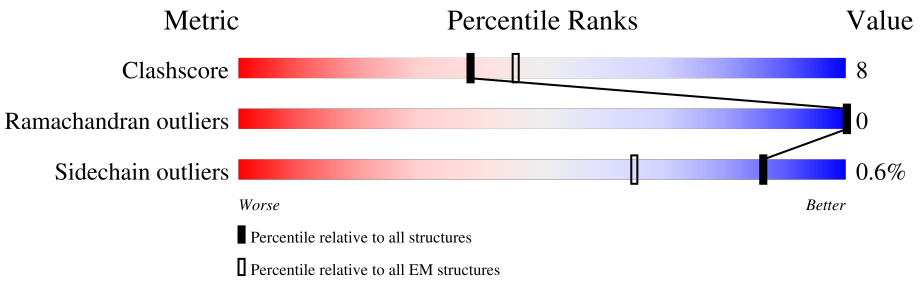
EMDB validation analysis	:	0.0.1.dev70
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
MapQ	:	1.9.13
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

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 3.10 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	158937	4297
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	AT	278	38% 53% 11% • 36%
1	BT	278	36% 52% 12% • 36%
1	CT	278	36% 51% 12% • 36%
1	DT	278	37% 52% 12% • 36%
1	ET	278	36% 53% 11% • 36%
1	FT	278	37% 51% 13% • 36%
1	GT	278	38% 53% 10% • 36%
1	HT	278	37% 52% 12% • 36%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	IT	278	37% 52% 12% • 36%
1	JT	278	37% 52% 12% • 36%
1	KT	278	37% 51% 13% • 36%
1	LT	278	36% 51% 13% • 36%
1	MT	278	37% 51% 12% • 36%
1	NT	278	37% 53% 10% • 36%
2	AX	522	7% 24% 6% 70%
2	BX	522	7% 24% 6% 70%
2	CX	522	7% 24% 5% 70%
2	DX	522	7% 24% 6% 70%
2	EX	522	7% 24% 6% 70%
2	FX	522	6% 23% 7% 70%
2	GX	522	7% 23% 7% 70%
2	HX	522	7% 23% 7% 70%
2	IX	522	7% 25% 5% 70%
2	JX	522	7% 24% 6% 70%
2	KX	522	6% 24% 6% 70%
2	LX	522	7% 24% 6% 70%
2	MX	522	6% 24% 6% 70%
2	NX	522	7% 23% 6% 70%
3	AY	1927	• 8% • 89%
3	BY	1927	• 8% • 89%
3	CY	1927	• 8% • 89%
3	DY	1927	• 8% • 89%
3	EY	1927	• 8% • 89%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	FY	1927	8% 89%
3	GY	1927	8% 89%
3	HY	1927	8% 89%
3	IY	1927	8% 89%
3	JY	1927	8% 89%
3	KY	1927	8% 89%
3	LY	1927	8% 89%
3	MY	1927	8% 89%
3	NY	1927	8% 89%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 60970 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 Cag pathogenicity island protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
1	AT	179	1502	957	261	282	2	0	0
1	BT	179	1502	957	261	282	2	0	0
1	CT	179	1502	957	261	282	2	0	0
1	DT	179	1502	957	261	282	2	0	0
1	ET	179	1502	957	261	282	2	0	0
1	FT	179	1502	957	261	282	2	0	0
1	GT	179	1502	957	261	282	2	0	0
1	HT	179	1502	957	261	282	2	0	0
1	IT	179	1502	957	261	282	2	0	0
1	JT	179	1502	957	261	282	2	0	0
1	KT	179	1502	957	261	282	2	0	0
1	LT	179	1502	957	261	282	2	0	0
1	MT	179	1502	957	261	282	2	0	0
1	NT	179	1502	957	261	282	2	0	0

- Molecule 2 is a protein called Type IV secretion system apparatus protein CagX.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
2	AX	155	1275	819	222	231	3	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
2	BX	155	1275	819	222	231	3	0	0
2	CX	155	1275	819	222	231	3	0	0
2	DX	155	1275	819	222	231	3	0	0
2	EX	155	1275	819	222	231	3	0	0
2	FX	155	1275	819	222	231	3	0	0
2	GX	155	1275	819	222	231	3	0	0
2	HX	155	1275	819	222	231	3	0	0
2	IX	155	1275	819	222	231	3	0	0
2	JX	155	1275	819	222	231	3	0	0
2	KX	155	1275	819	222	231	3	0	0
2	LX	155	1275	819	222	231	3	0	0
2	MX	155	1275	819	222	231	3	0	0
2	NX	155	1275	819	222	231	3	0	0

- Molecule 3 is a protein called Cag pathogenicity island protein (Cag7).

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
3	AY	209	1578	1014	255	300	9	0	0
3	BY	209	1578	1014	255	300	9	0	0
3	CY	209	1578	1014	255	300	9	0	0
3	DY	209	1578	1014	255	300	9	0	0
3	EY	209	1578	1014	255	300	9	0	0
3	FY	209	1578	1014	255	300	9	0	0

Continued on next page...

Continued from previous page...

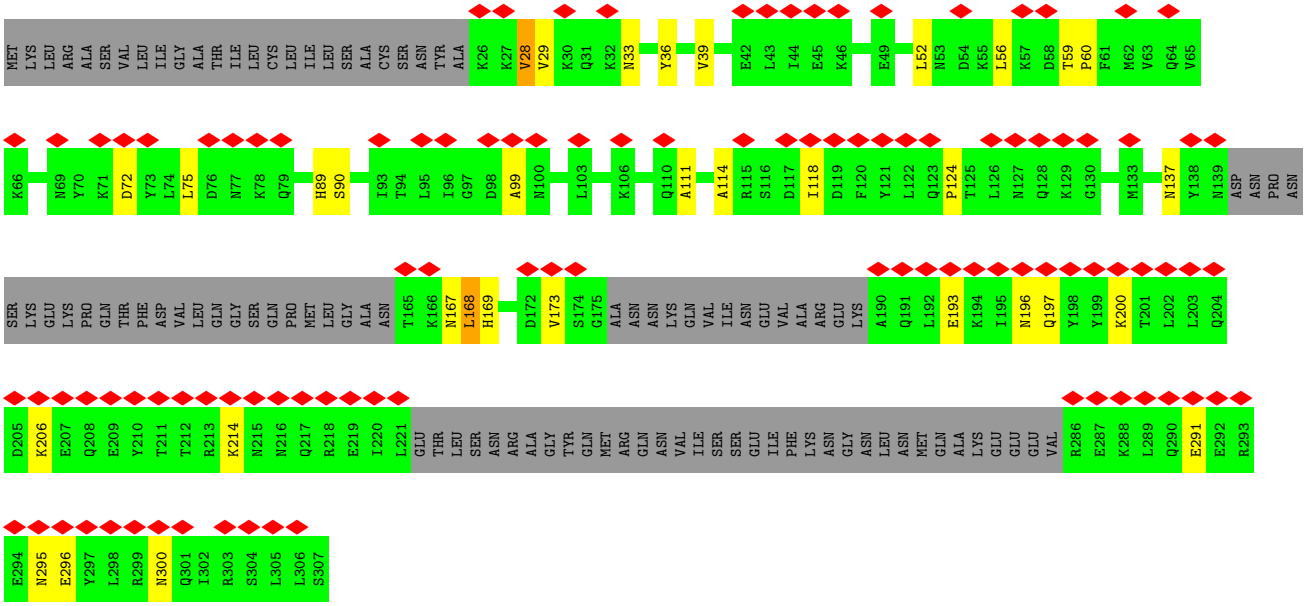
Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
3	GY	209	1578	1014	255	300	9	0	0
			Total	C	N	O	S		
3	HY	209	1578	1014	255	300	9	0	0
			Total	C	N	O	S		
3	IY	209	1578	1014	255	300	9	0	0
			Total	C	N	O	S		
3	JY	209	1578	1014	255	300	9	0	0
			Total	C	N	O	S		
3	KY	209	1578	1014	255	300	9	0	0
			Total	C	N	O	S		
3	LY	209	1578	1014	255	300	9	0	0
			Total	C	N	O	S		
3	MY	209	1578	1014	255	300	9	0	0
			Total	C	N	O	S		
3	NY	209	1578	1014	255	300	9	0	0

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.

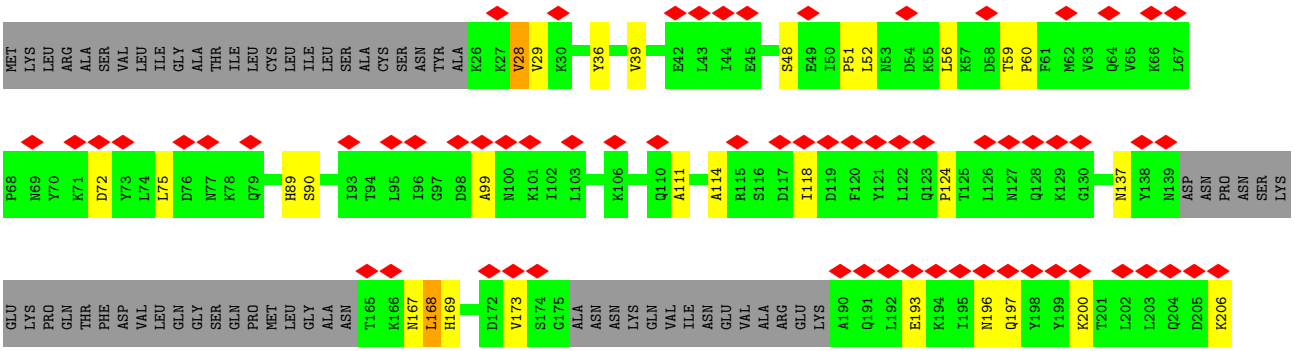
- Molecule 1: Cag pathogenicity island protein

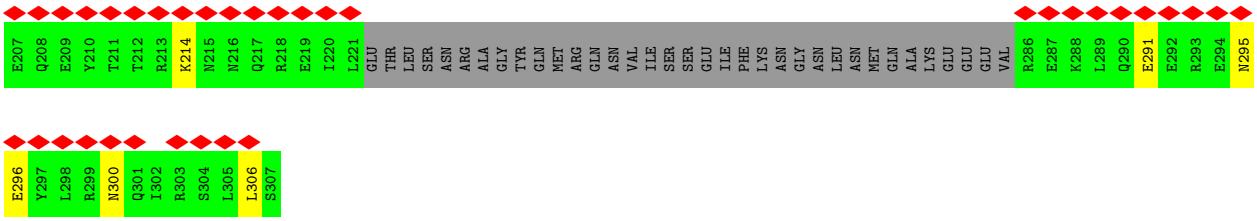
Chain AT:



- Molecule 1: Cag pathogenicity island protein

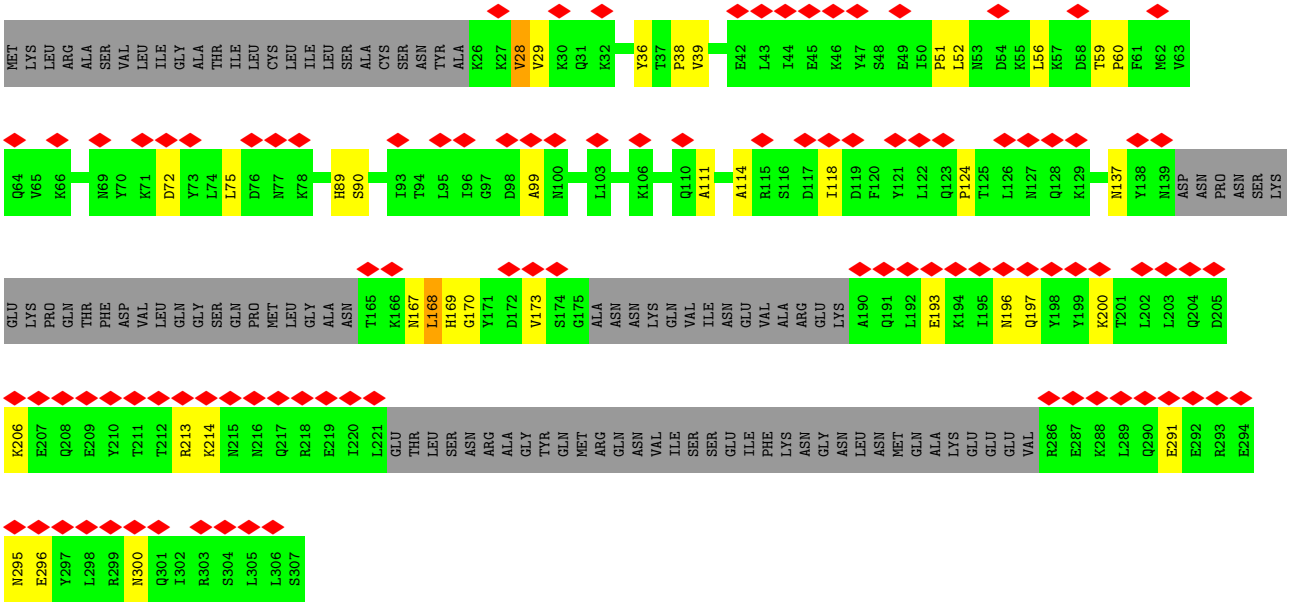
Chain BT:





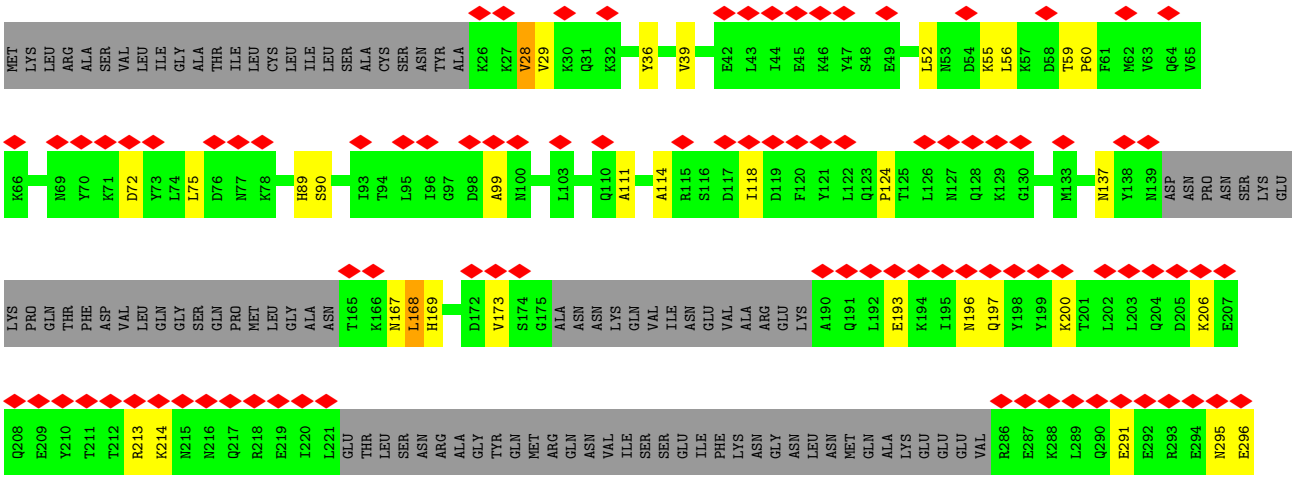
• Molecule 1: Cag pathogenicity island protein

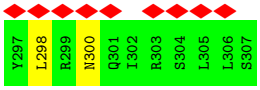
Chain CT:



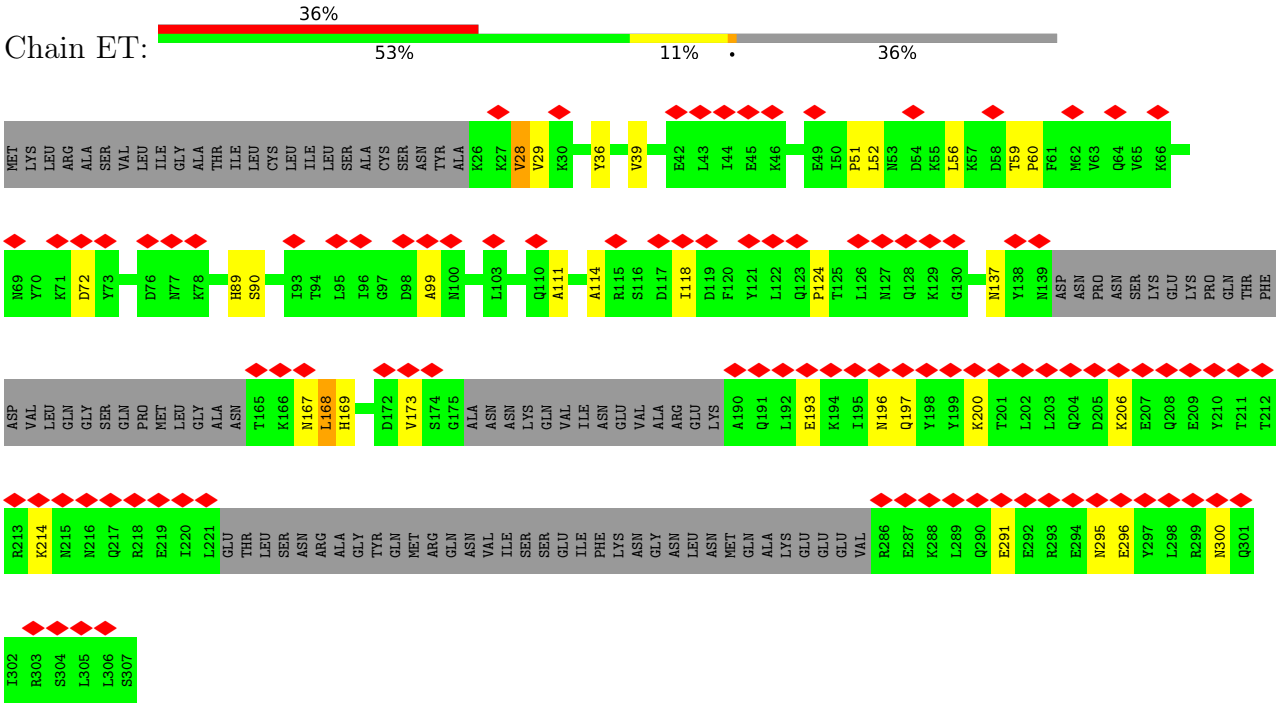
• Molecule 1: Cag pathogenicity island protein

Chain DT:

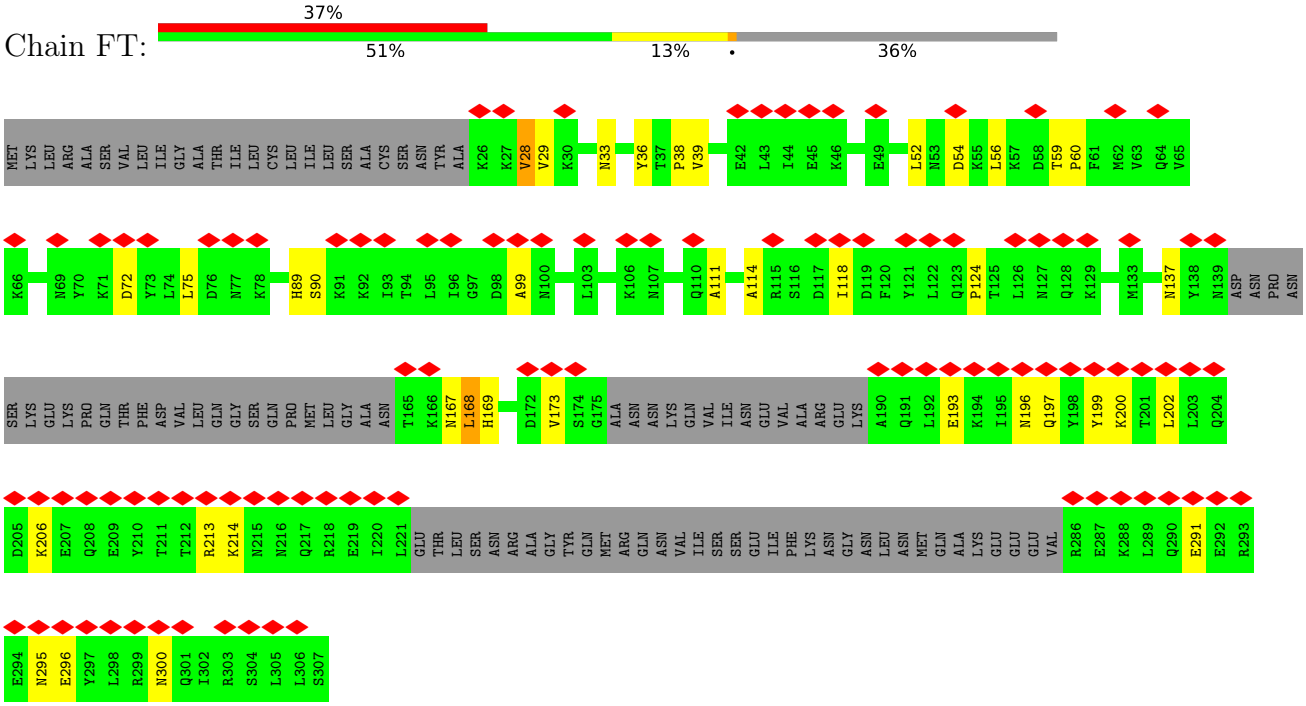




• Molecule 1: Cag pathogenicity island protein

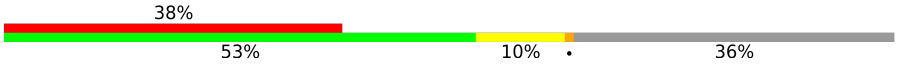


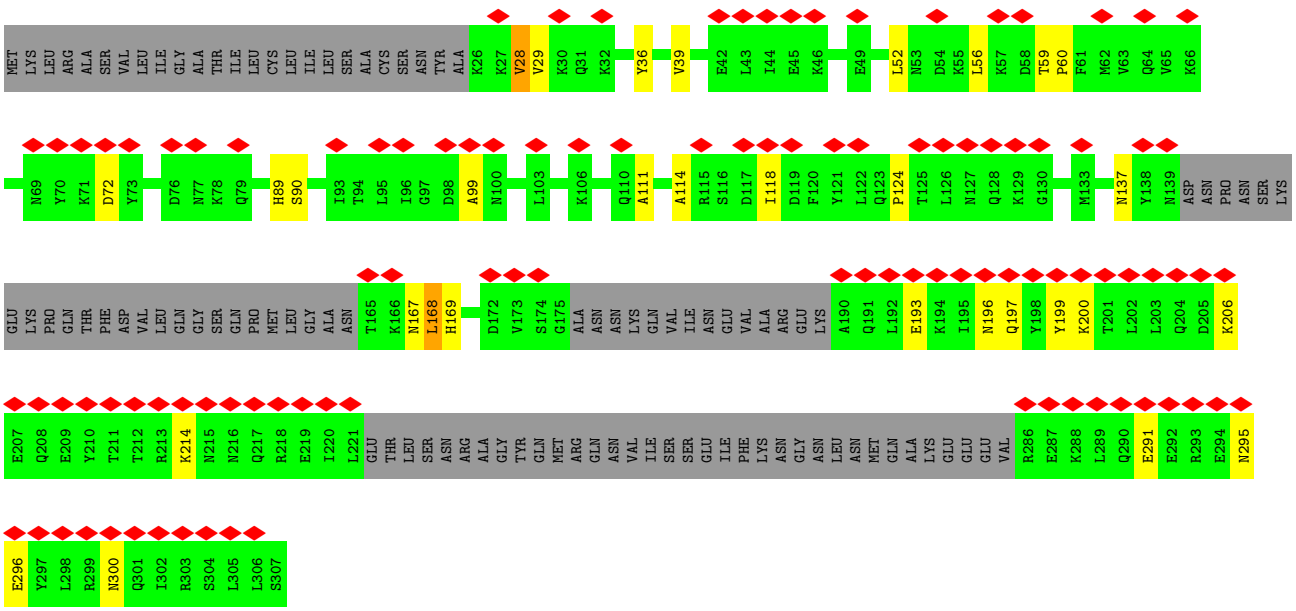
• Molecule 1: Cag pathogenicity island protein



• Molecule 1: Cag pathogenicity island protein

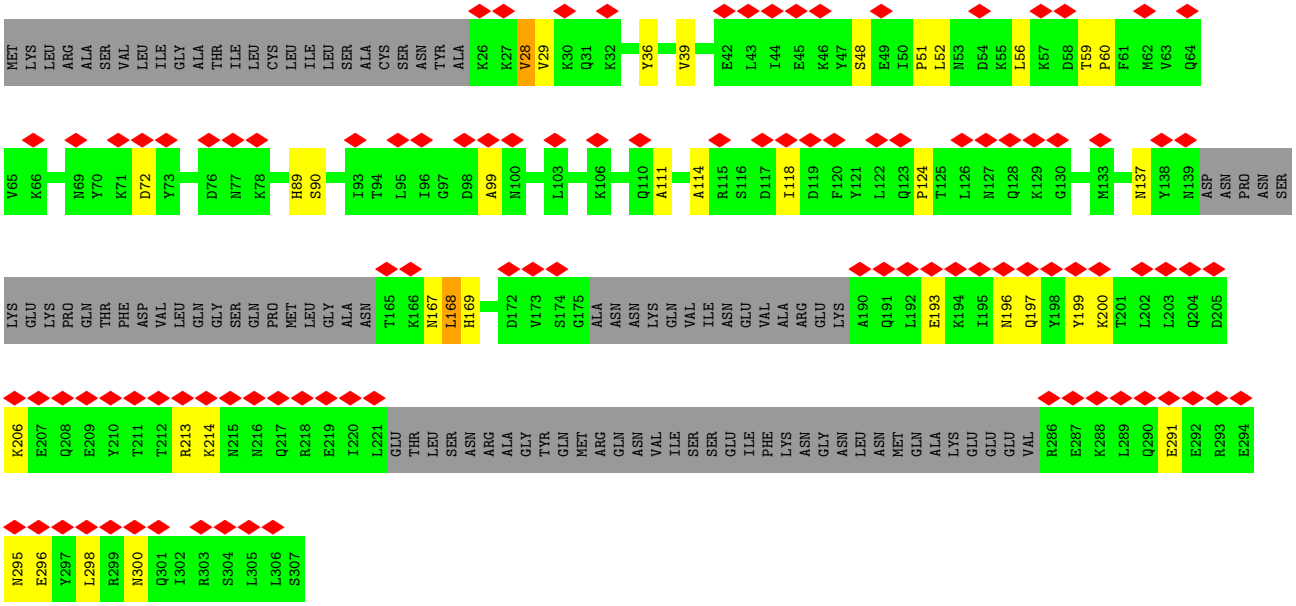
Chain GT:





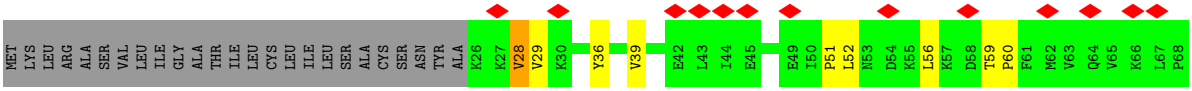
- Molecule 1: Cag pathogenicity island protein

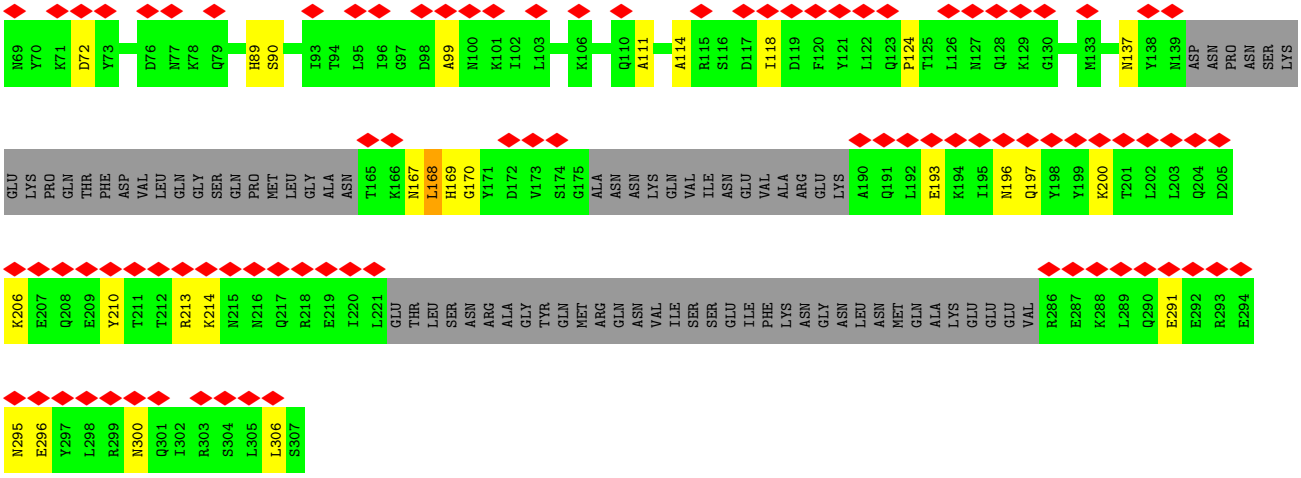
Chain HT:



- Molecule 1: Cag pathogenicity island protein

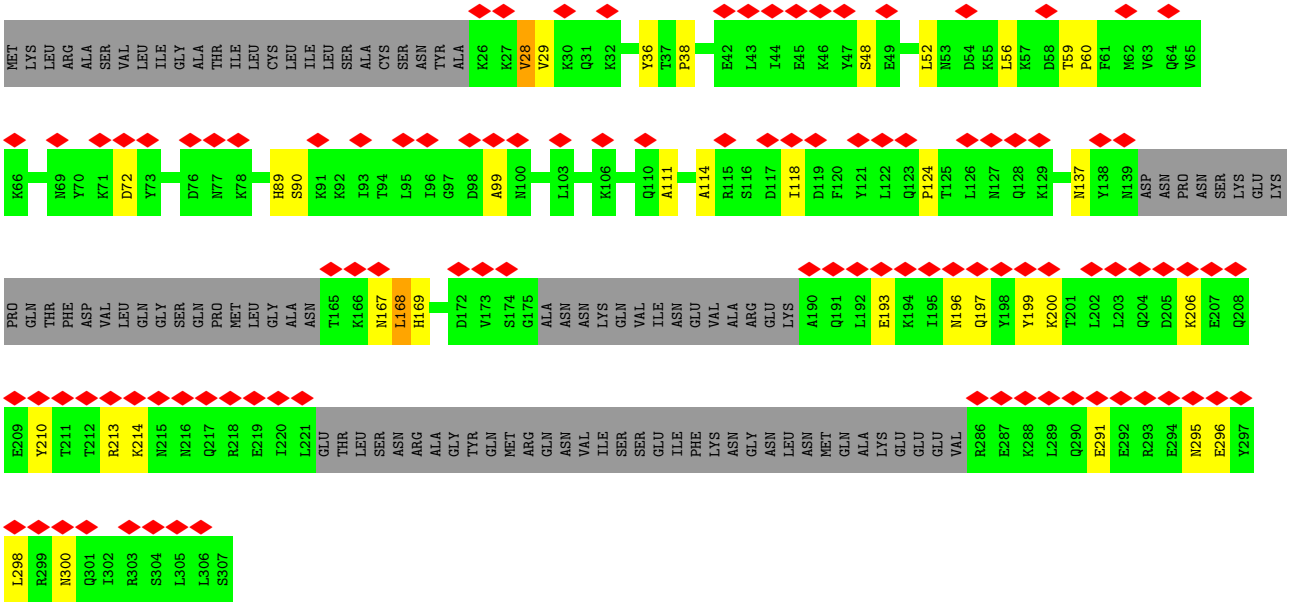
Chain IT:





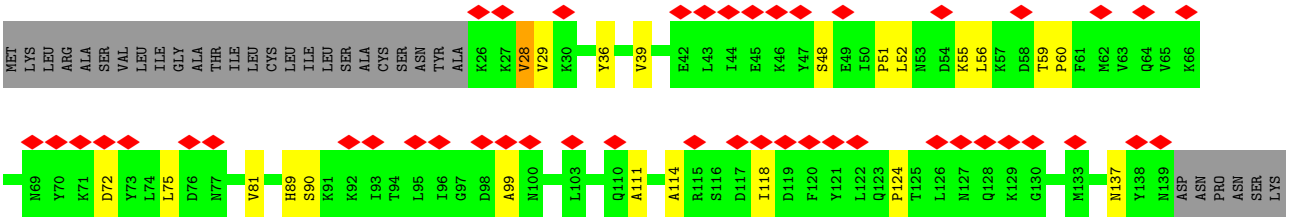
- Molecule 1: Cag pathogenicity island protein

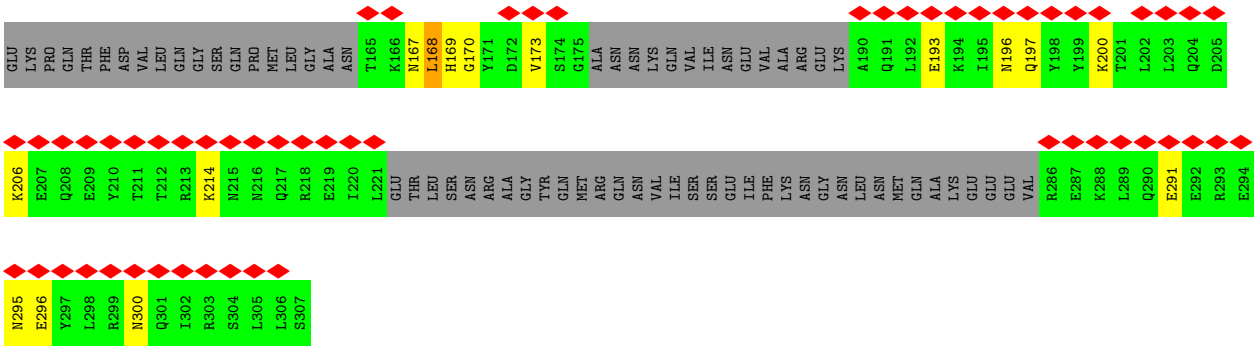
Chain JT:



- Molecule 1: Cag pathogenicity island protein

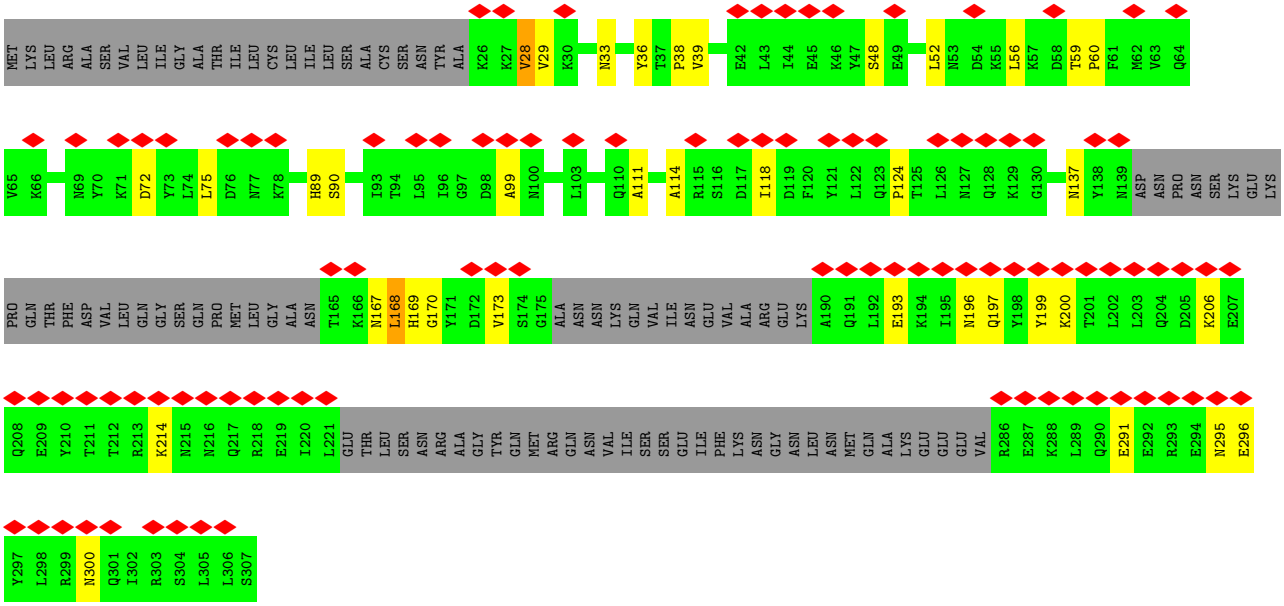
Chain KT:





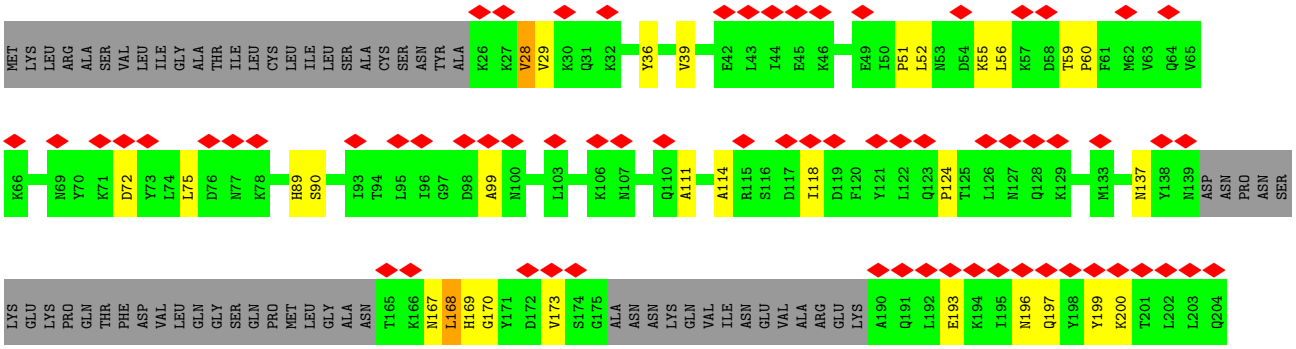
• Molecule 1: Cag pathogenicity island protein

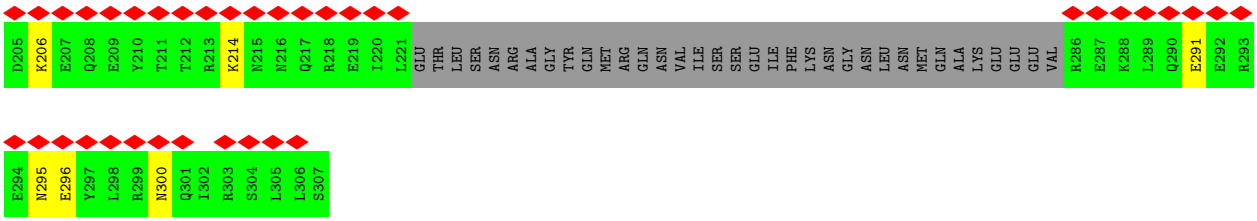
Chain LT:



• Molecule 1: Cag pathogenicity island protein

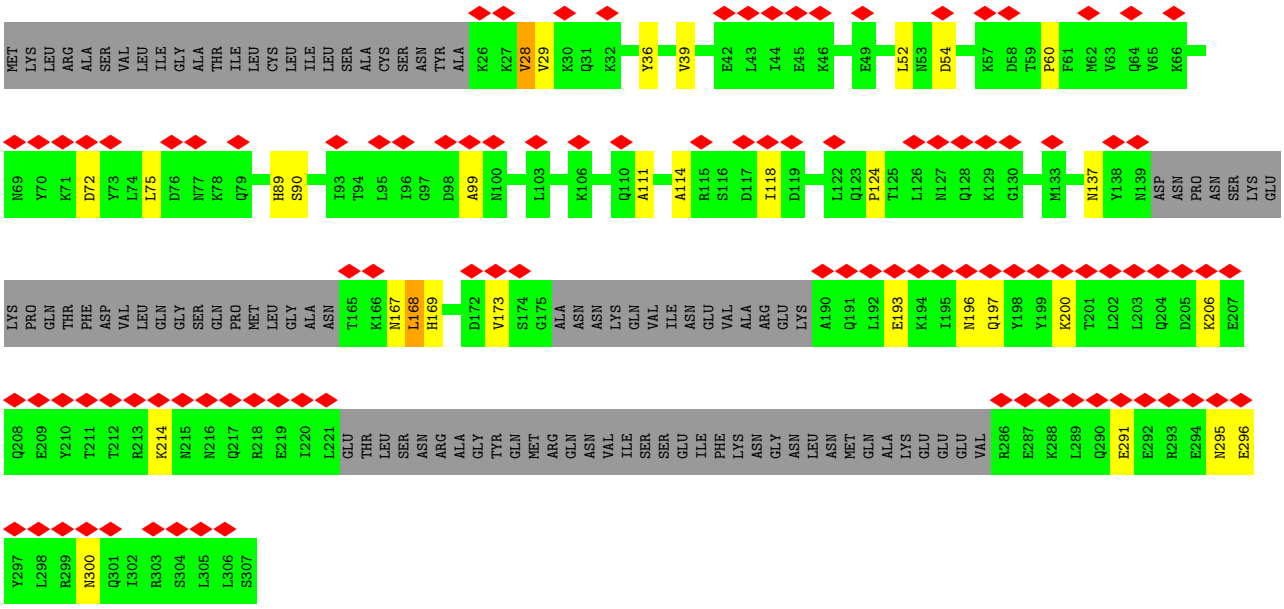
Chain MT:





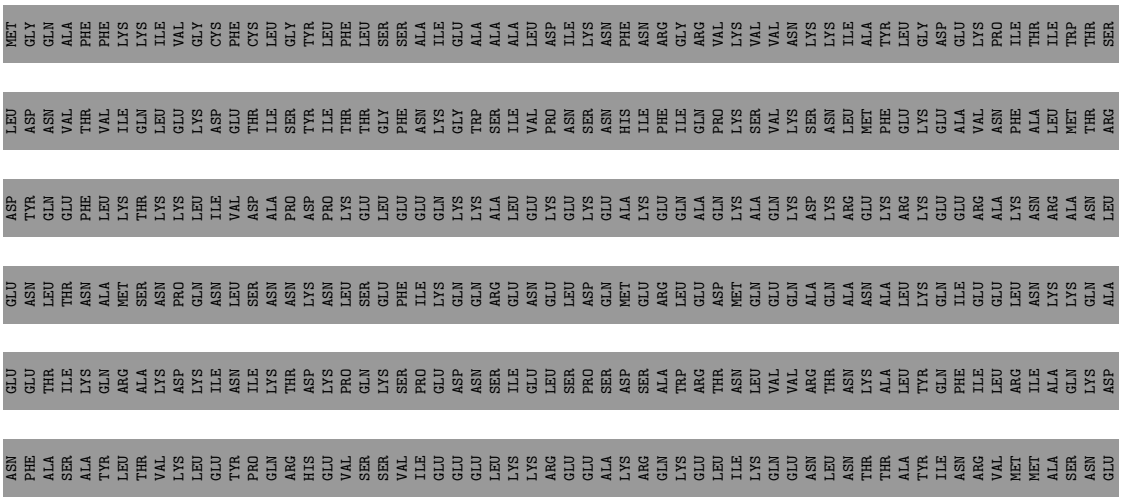
• Molecule 1: Cag pathogenicity island protein

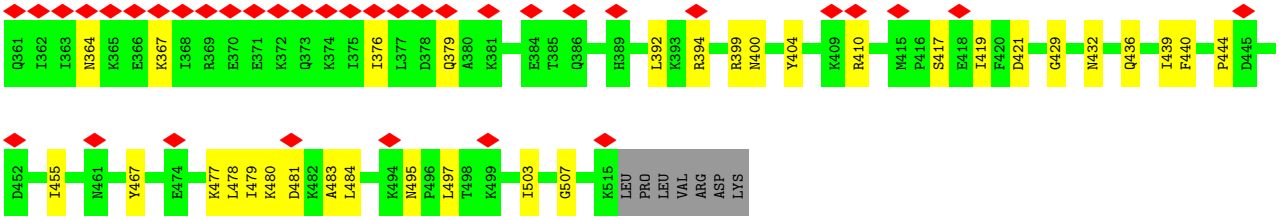
Chain NT:



• Molecule 2: Type IV secretion system apparatus protein CagX

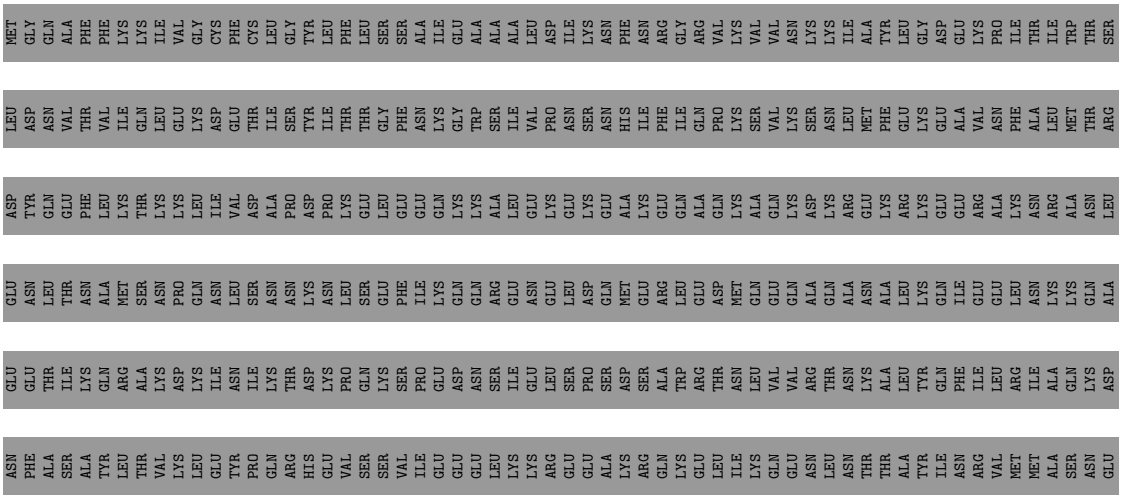
Chain AX:

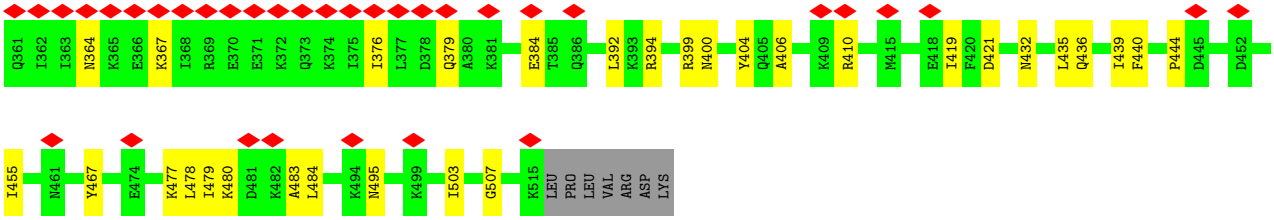




• Molecule 2: Type IV secretion system apparatus protein CagX

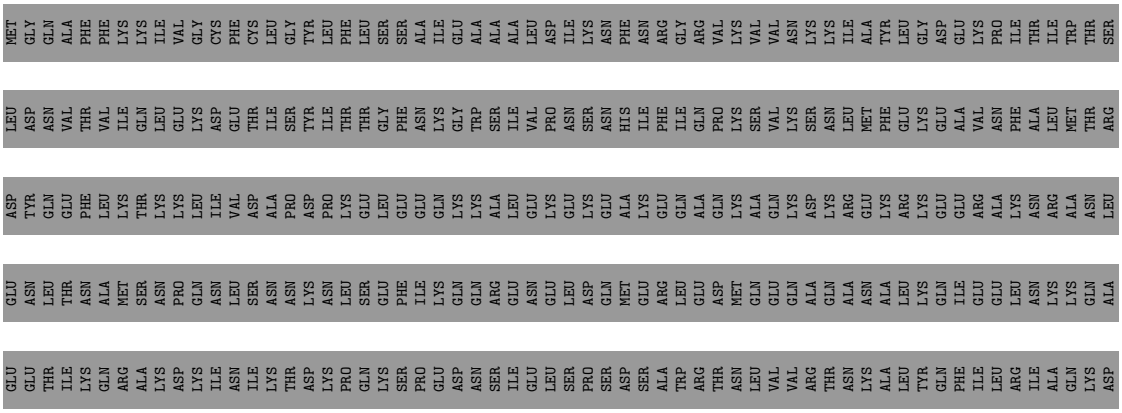
Chain BX:

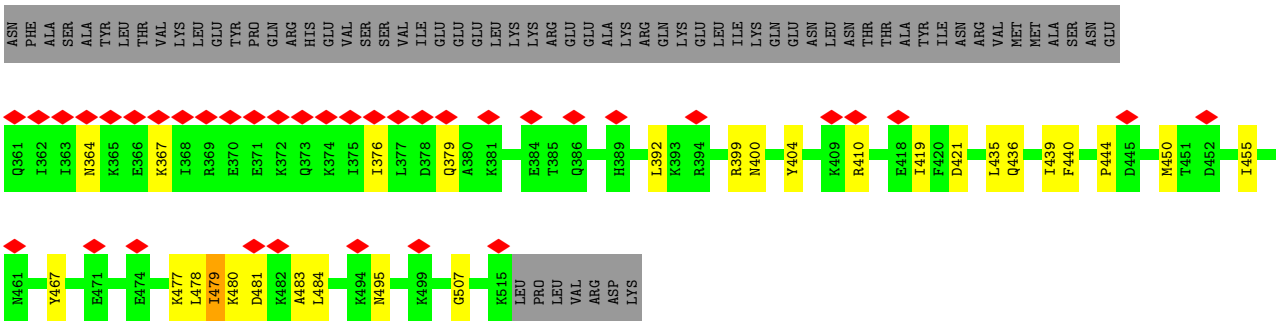




• Molecule 2: Type IV secretion system apparatus protein CagX

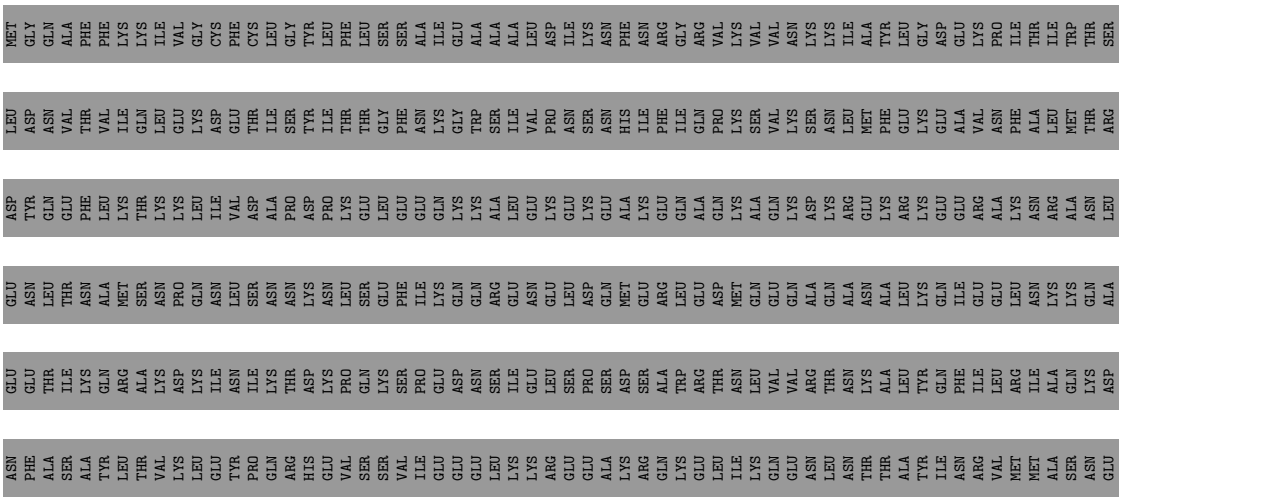
Chain CX:

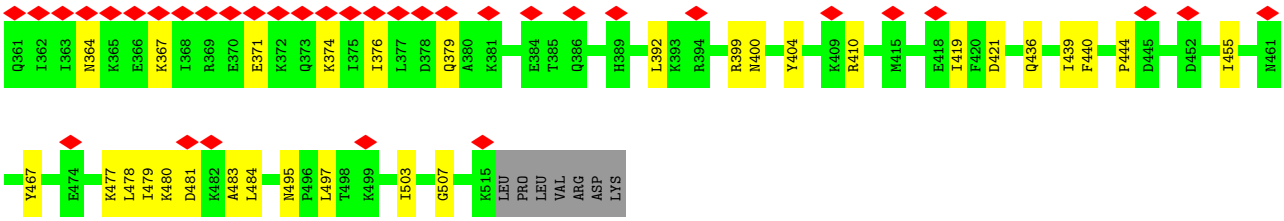




• Molecule 2: Type IV secretion system apparatus protein CagX

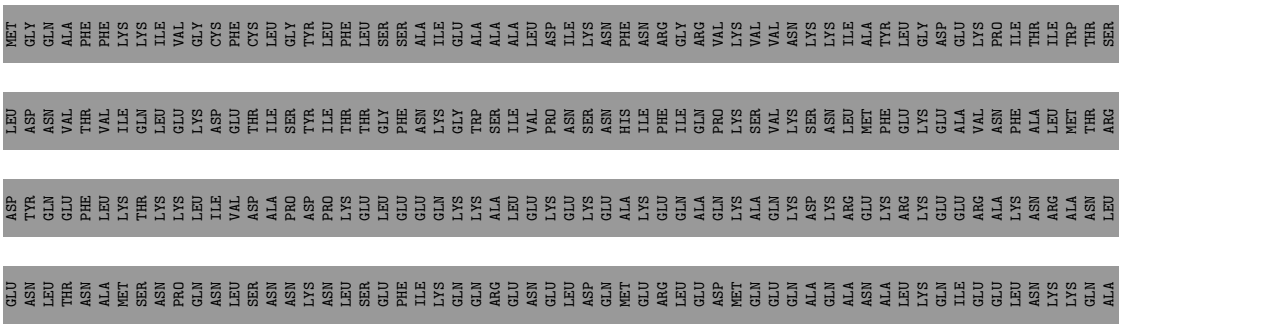
Chain DX:

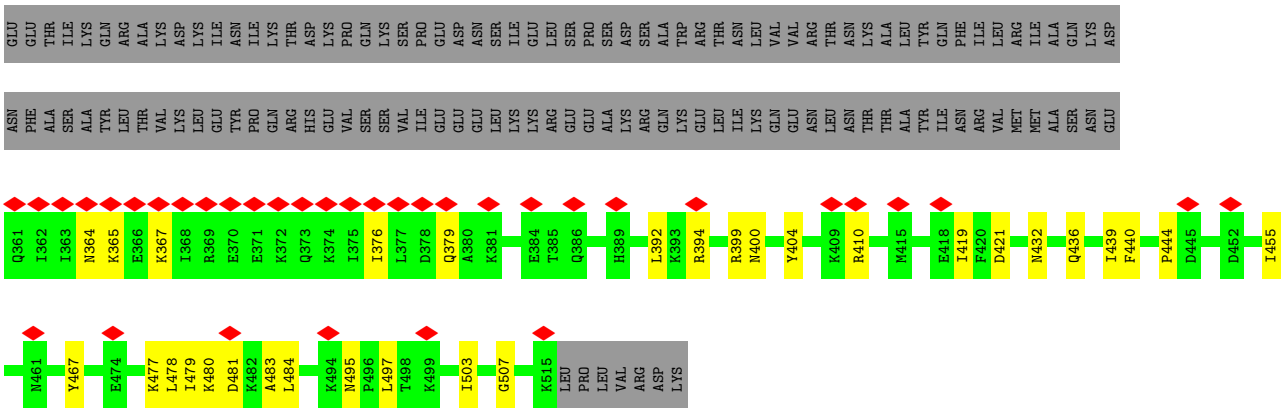




• Molecule 2: Type IV secretion system apparatus protein CagX

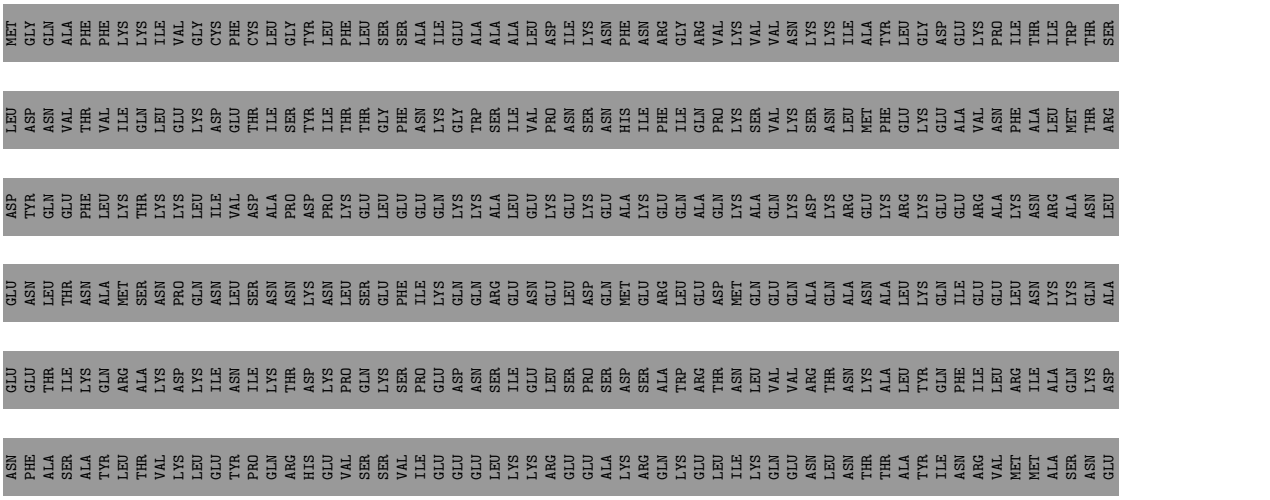
Chain EX:





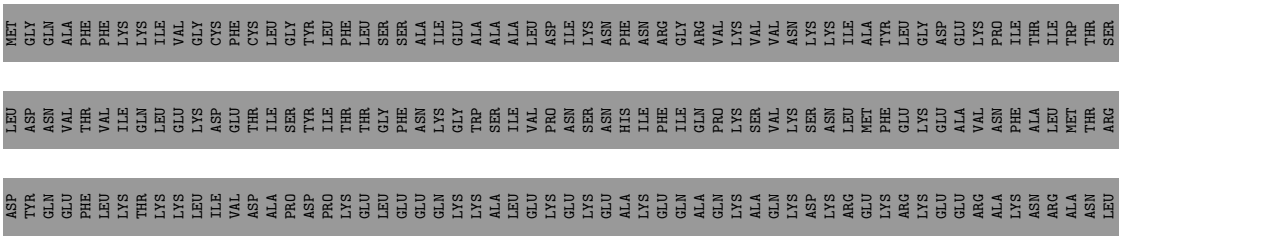
• Molecule 2: Type IV secretion system apparatus protein CagX

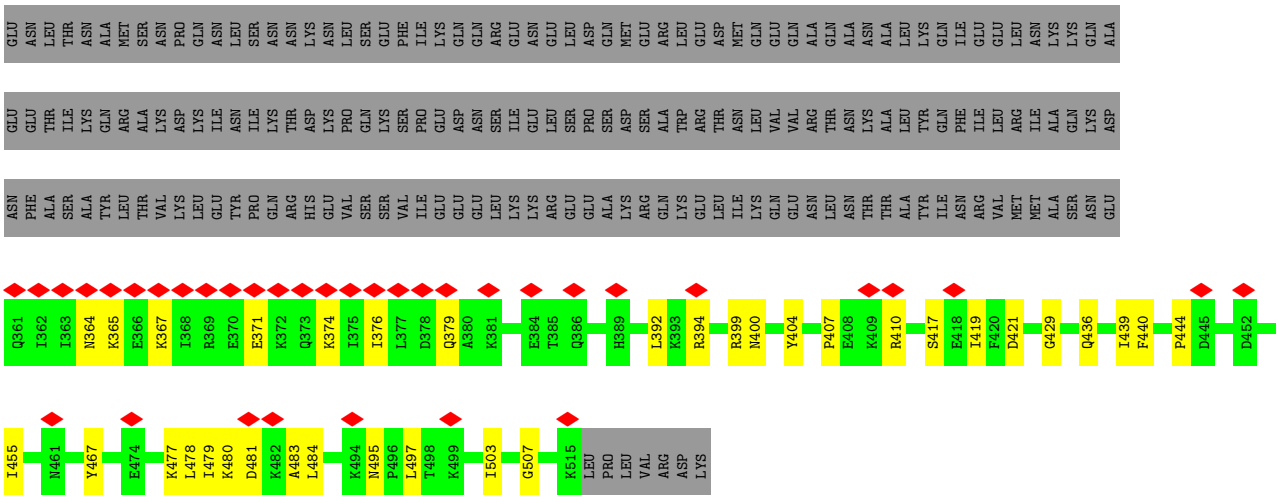
Chain FX:



• Molecule 2: Type IV secretion system apparatus protein CagX

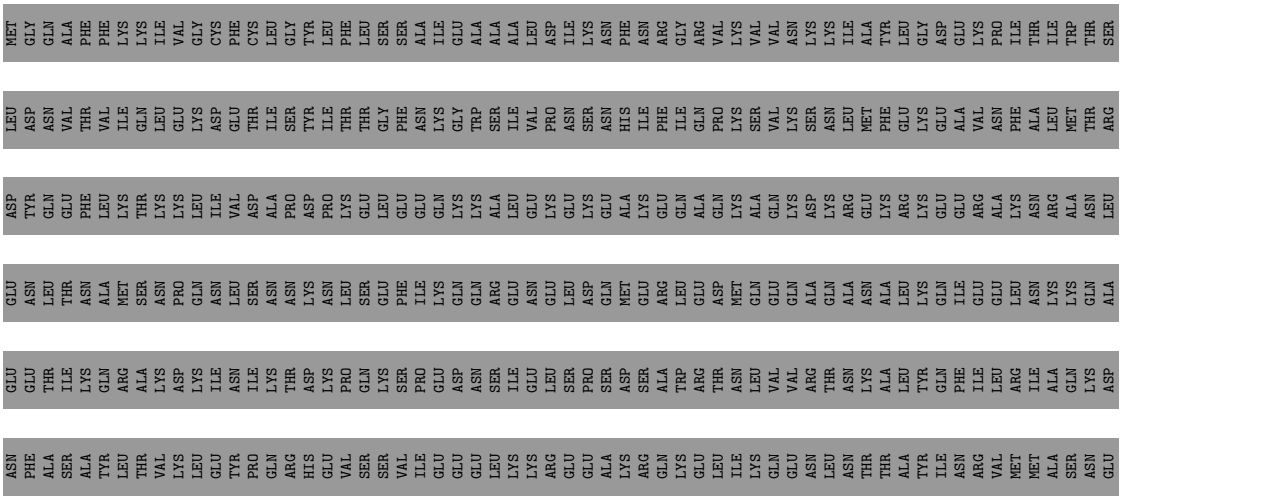
Chain GX:





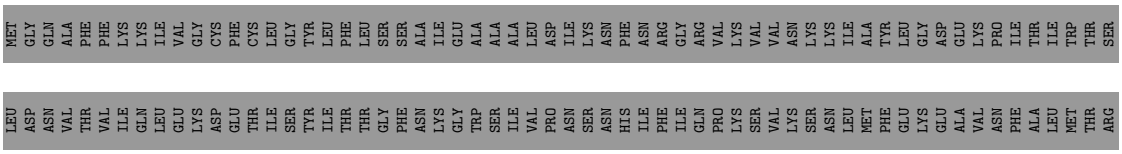
• Molecule 2: Type IV secretion system apparatus protein CagX

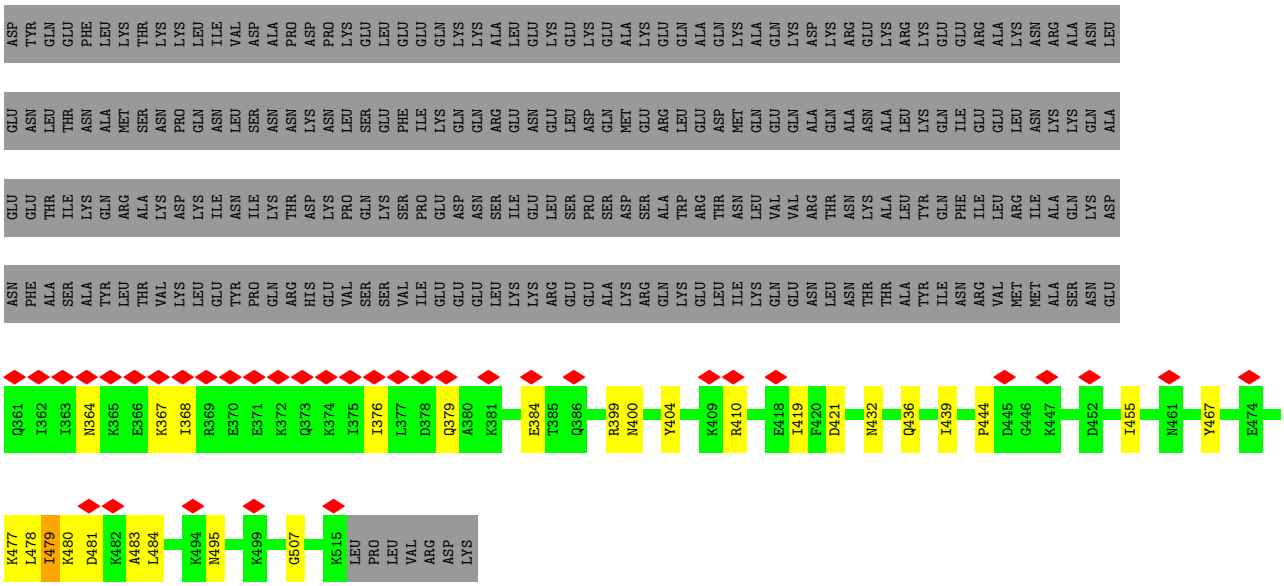
Chain HX:



• Molecule 2: Type IV secretion system apparatus protein CagX

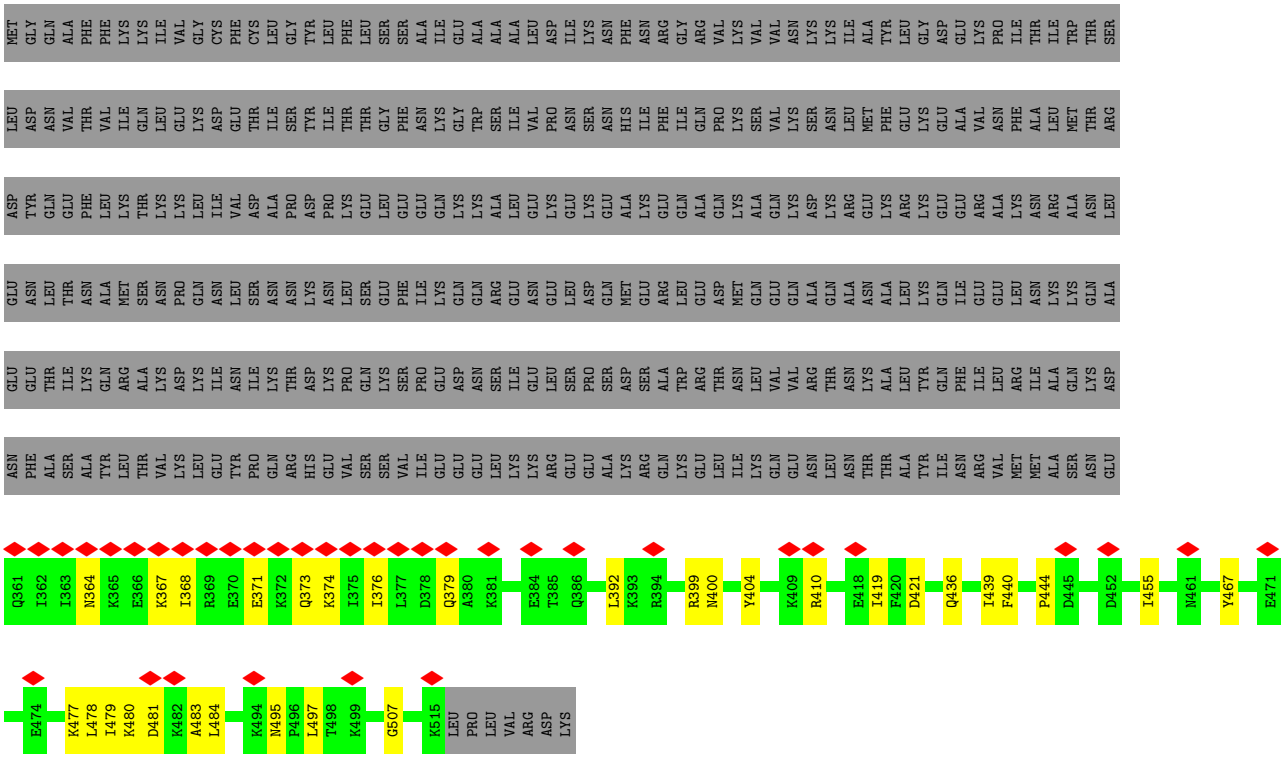
Chain IX:





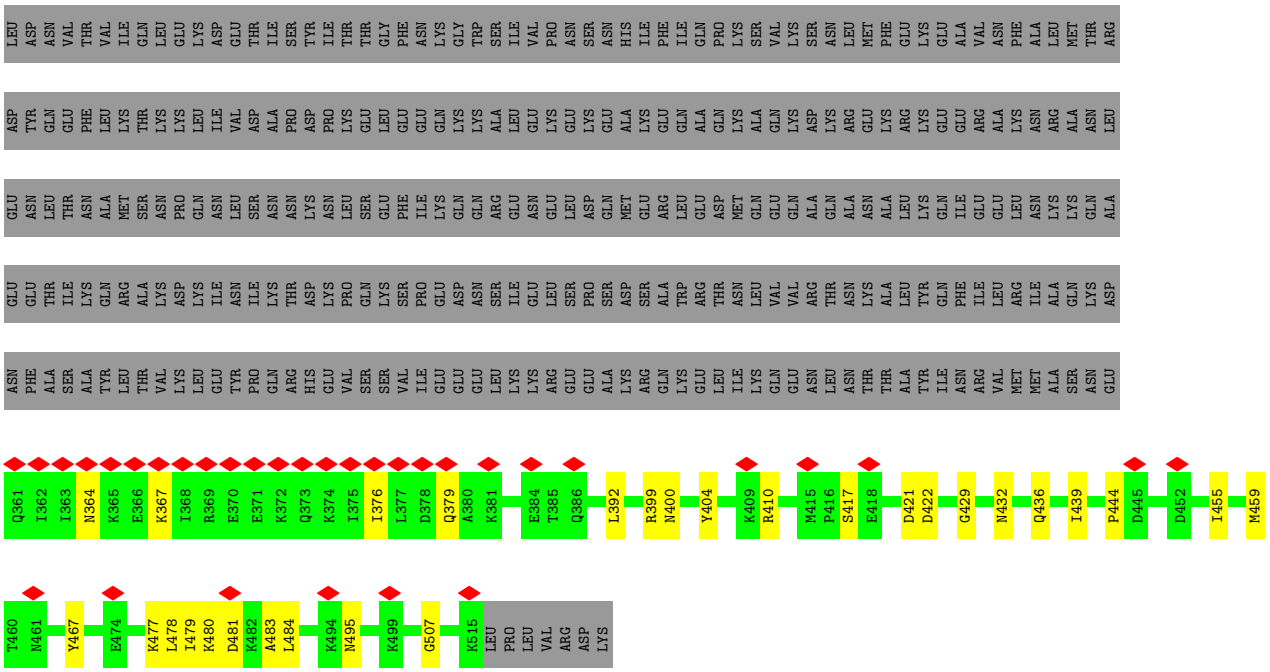
• Molecule 2: Type IV secretion system apparatus protein CagX

Chain JX:



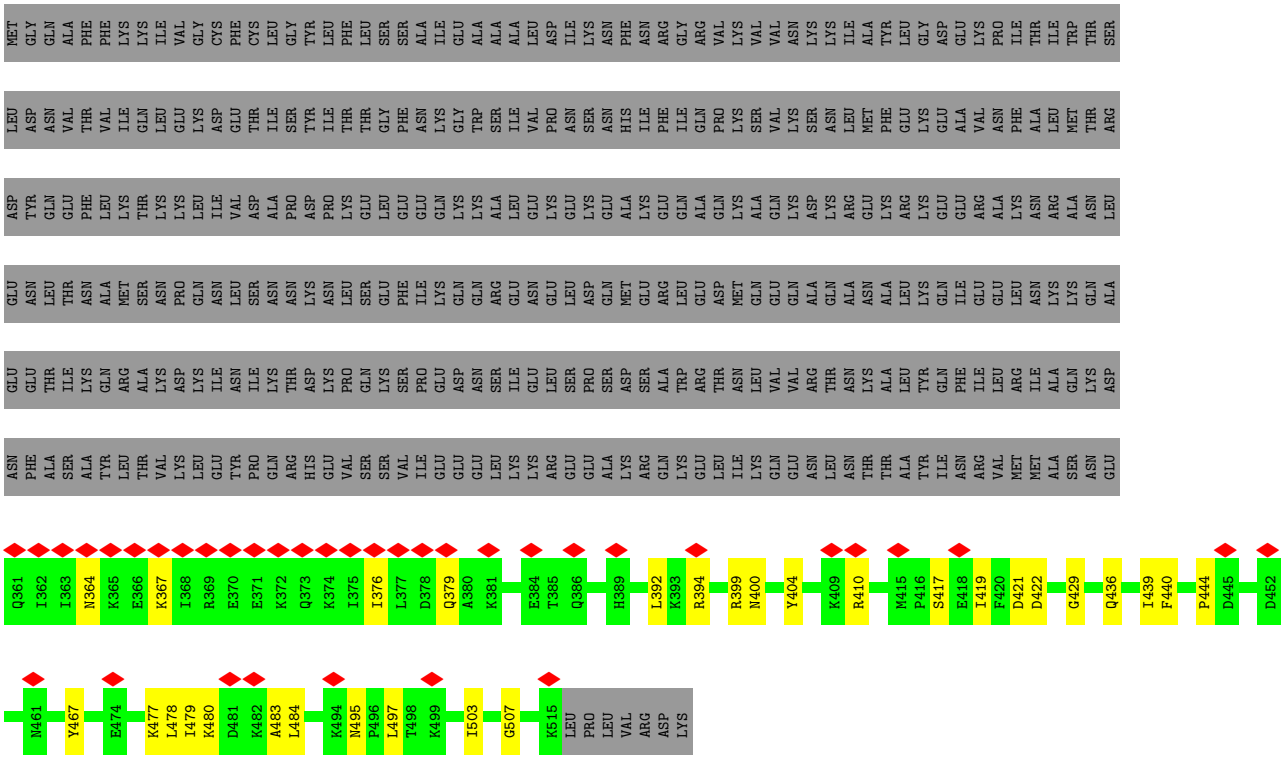
• Molecule 2: Type IV secretion system apparatus protein CagX

Chain KX:



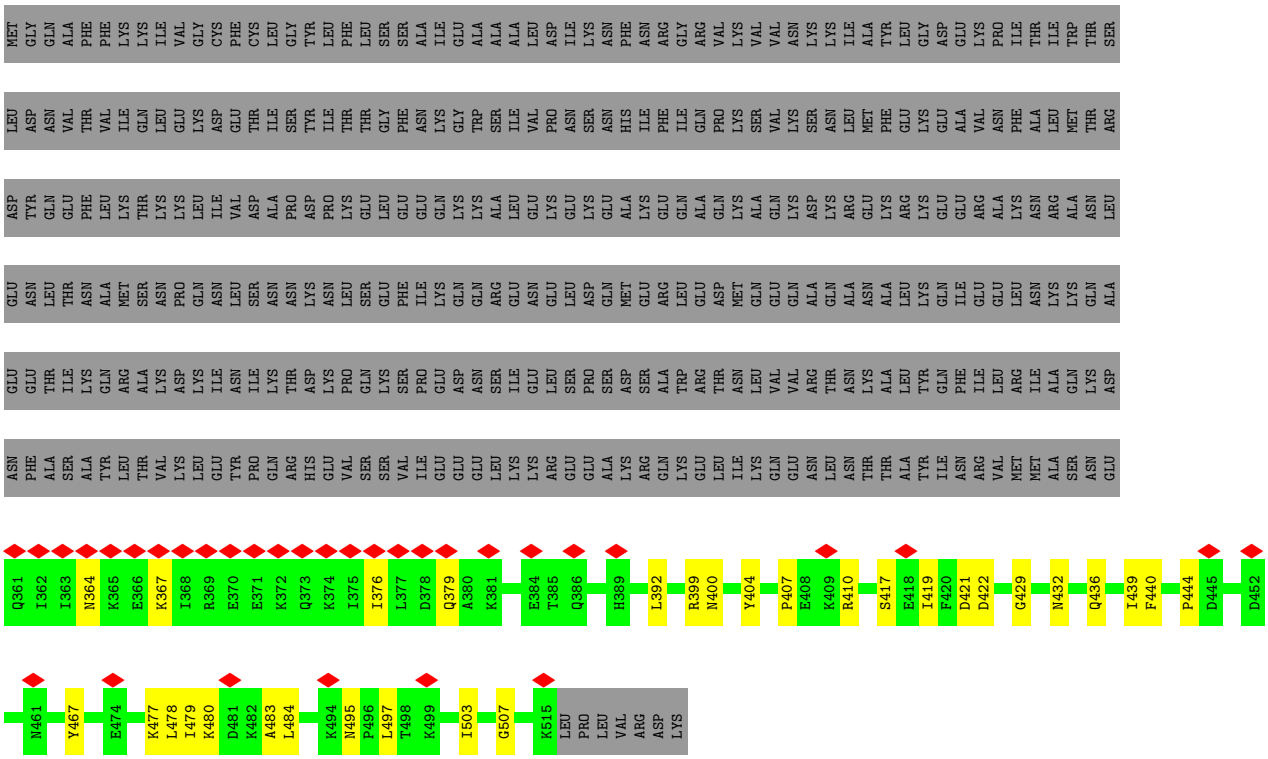
• Molecule 2: Type IV secretion system apparatus protein CagX

Chain LX:



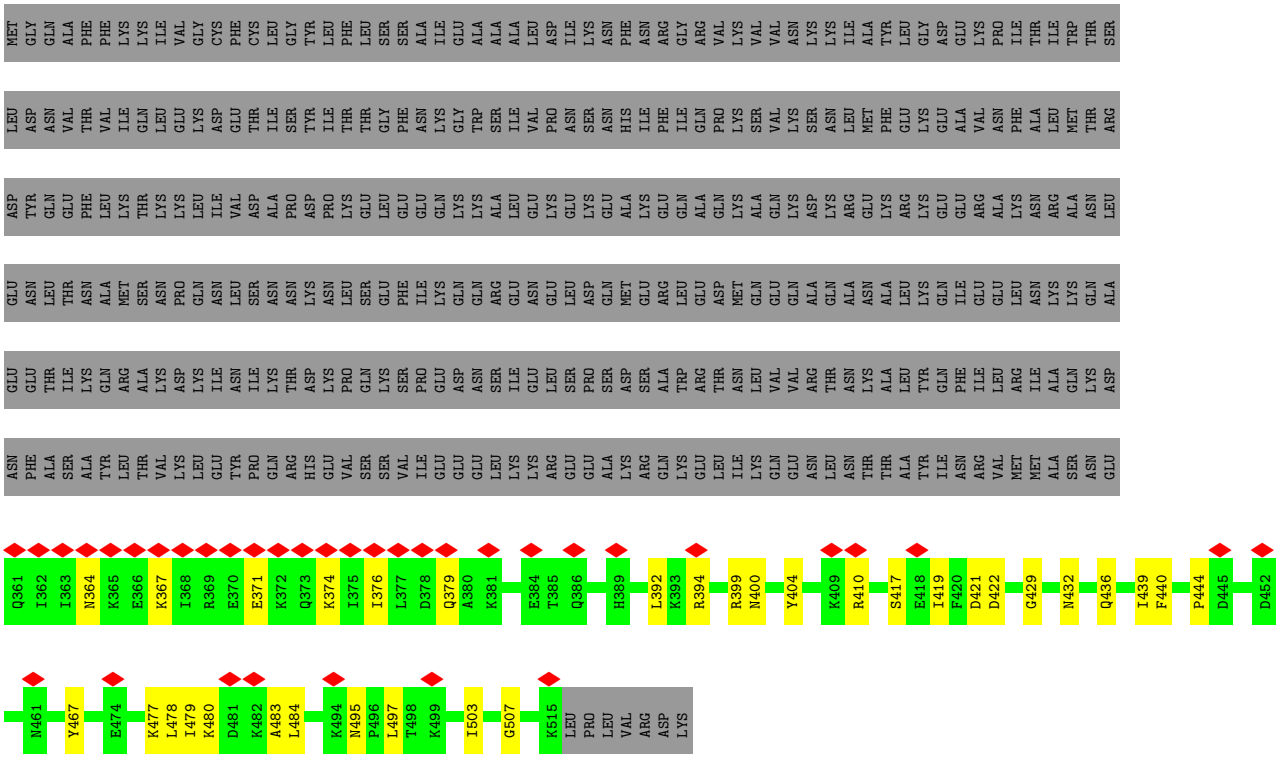
• Molecule 2: Type IV secretion system apparatus protein CagX

Chain MX:



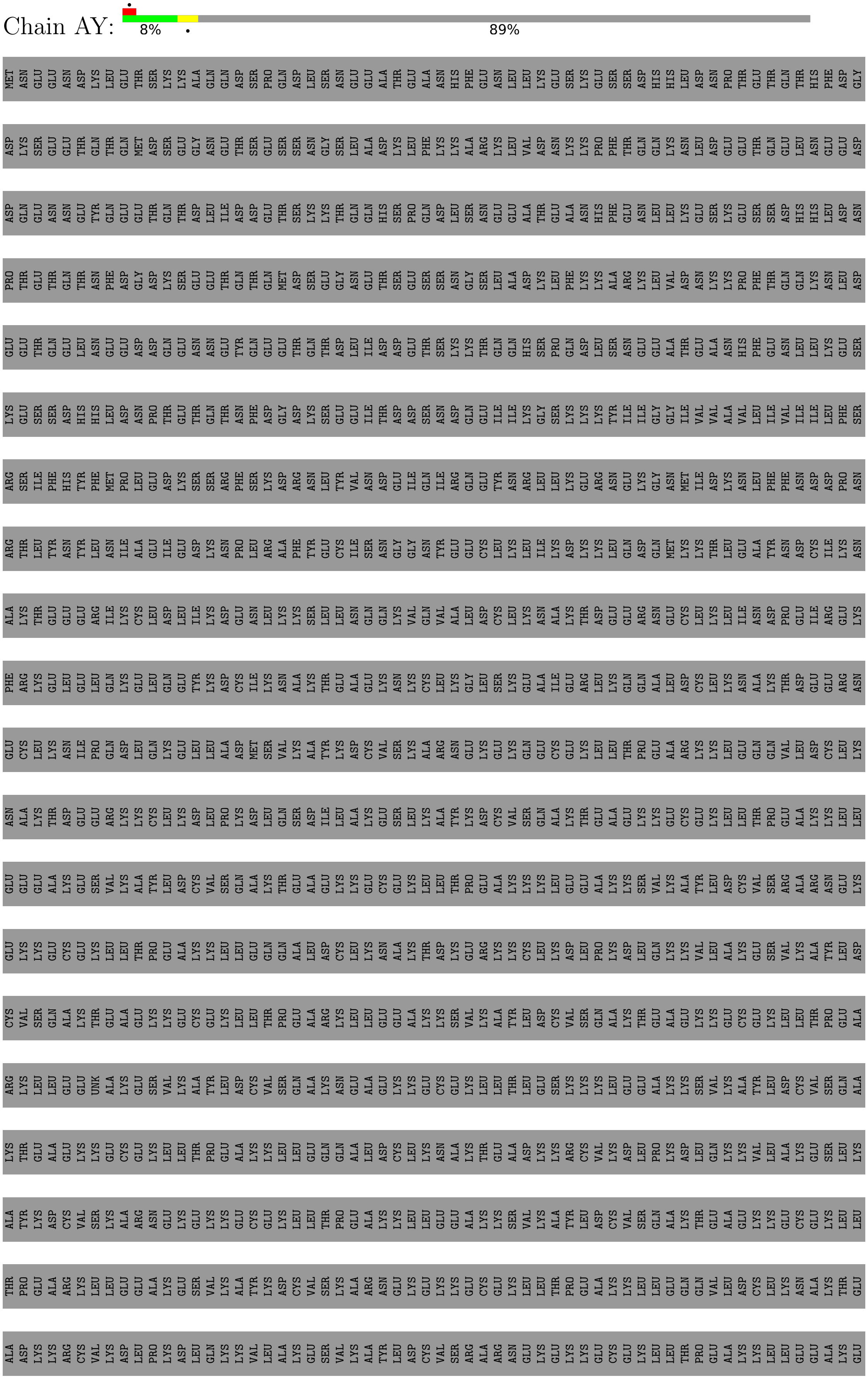
- Molecule 2: Type IV secretion system apparatus protein CagX

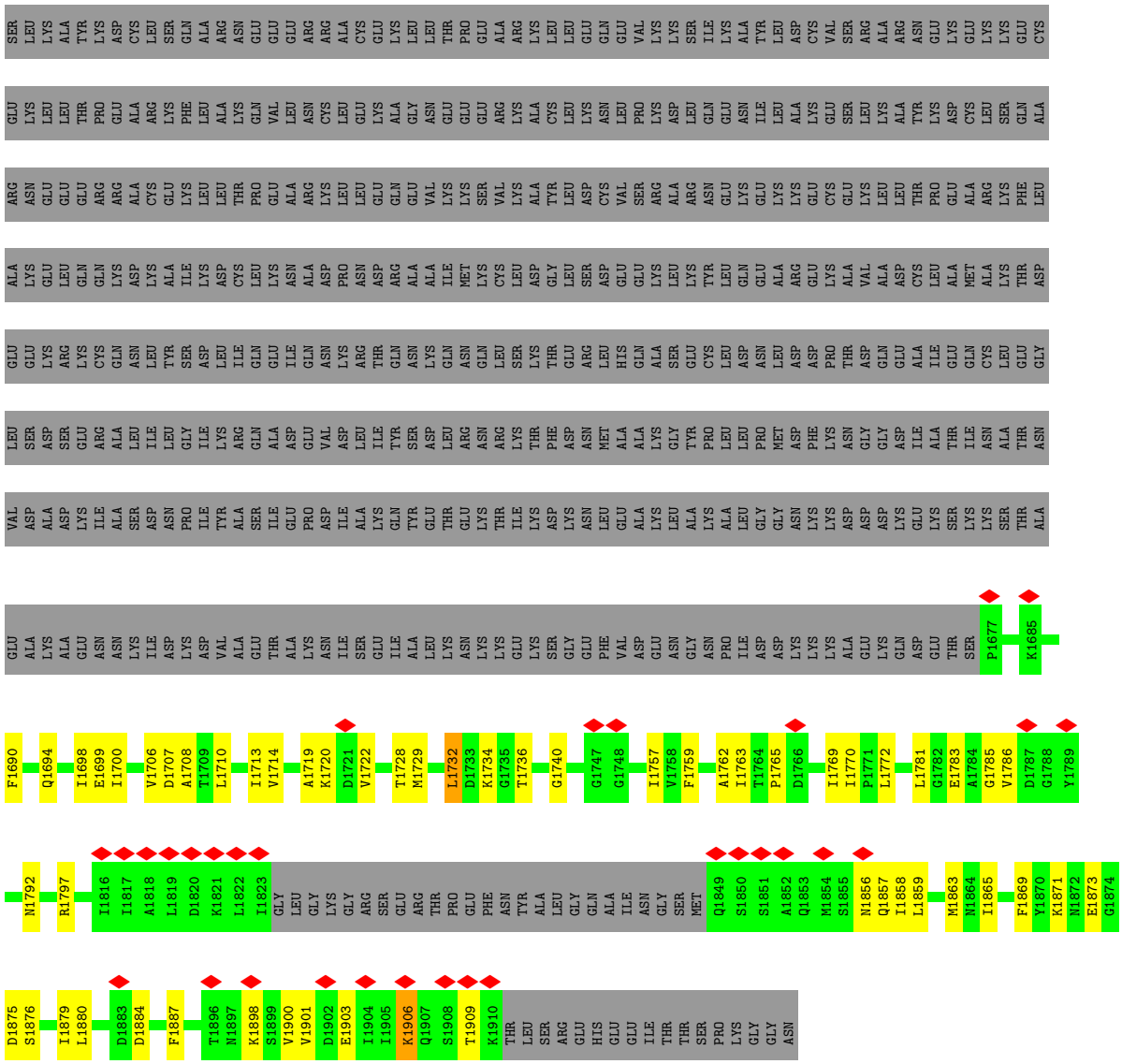
Chain NX:



- Molecule 3: Cag pathogenicity island protein (Cag7)

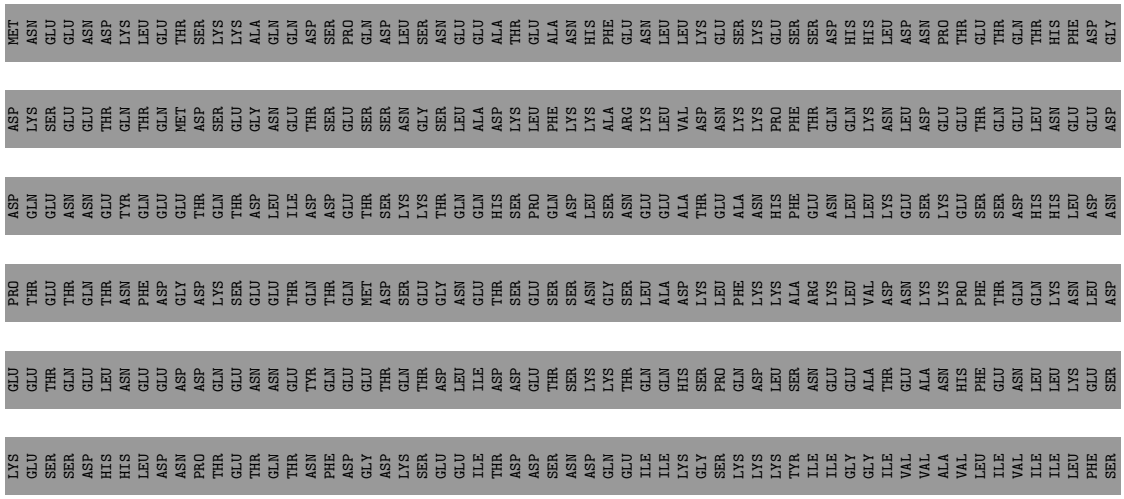
Chain AY:





- Molecule 3: Cag pathogenicity island protein (Cag7)

Chain BY:



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

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

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

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

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

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

GLU GLU GLU ALA LYS GLU SER VAL LYS ALA TYR LEU ASP CYS VAL SER GLN ALA LYS THR GLU ALA GLU LYS LYS GLU CYS GLU LYS LEU LEU THR PRO GLU ALA LYS LYS LYS LEU GLU GLU ALA LYS LYS SER VAL LYS ALA TYR LEU ASP CYS VAL SER ARG ALA ARG ASN GLU LYS

GLU LYS LYS GLU CYS GLU LYS LEU LEU THR PRO GLU ALA LYS LYS LEU LEU GLU GLN GLN ALA LEU ASP CYS LEU LYS ASN ALA LYS THR ASP LYS GLU ARG LYS LYS CYS LEU LYS ASP LEU PRO LYS ASP LEU GLN LYS LYS VAL LEU ALA LYS GLU SER VAL LYS ALA TYR LEU ASP

CYS VAL SER GLN ALA LYS THR GLU ALA GLU LYS LYS GLU CYS GLU LYS LEU LEU THR PRO GLU ALA ARG LYS LEU LEU GLU GLU ALA LYS LYS SER VAL LYS ALA TYR LEU ASP CYS VAL SER GLN ALA LYS THR GLU ALA GLU LYS LYS GLU CYS GLU LYS LEU LEU THR PRO GLU ALA

ARG LYS LEU LEU GLU GLU UNK ALA LYS GLU SER VAL LYS ALA TYR LEU ASP CYS VAL SER GLN ALA LYS ASN GLU ALA GLU LYS LYS GLU CYS GLU LYS LEU LEU THR LEU GLU SER LYS LYS LEU GLU GLU ALA LYS LYS SER VAL LYS ALA TYR LEU ASP CYS VAL SER GLN ALA

LYS THR GLU ALA GLU LYS LYS GLU CYS GLU LYS LEU LEU THR PRO GLU ALA LYS LYS LEU LEU GLU GLN GLN ALA LEU ASP CYS LEU LYS ASN ALA LYS THR GLU ALA ASP LYS LYS ARG CYS VAL LYS ASP LEU PRO LYS ASP LEU GLN LYS LYS VAL LEU ALA LYS GLU SER LEU LYS

ALA TYR LYS ASP CYS VAL SER LYS ALA ARG ASN GLU LYS GLU LYS LYS GLU CYS GLU LYS LEU LEU THR PRO GLU ALA LYS LYS LEU LEU GLU GLU ALA LYS LYS SER VAL LYS ALA TYR LEU ASP CYS VAL SER GLN ALA LYS THR GLU ALA GLU LYS LYS GLU CYS GLU LYS LEU LEU

THR PRO GLU ALA ARG LYS LEU LEU GLU GLU ALA LYS GLU SER VAL LYS ALA TYR LYS ASP CYS VAL SER LYS ALA ARG ASN GLU LYS GLU LYS LYS GLU CYS GLU LYS LEU LEU THR PRO GLU ALA LYS LYS LEU LEU GLU GLN GLN VAL LEU ASP CYS LEU LYS ASN ALA LYS THR GLU

ALA ASP LYS LYS ARG CYS VAL LYS ASP LEU PRO LYS ASP LEU GLN LYS LYS VAL LEU ALA LYS GLU SER VAL LYS ALA TYR LEU ASP CYS VAL SER ARG ALA ARG ASN GLU LYS GLU LYS LYS GLU CYS GLU LYS LEU LEU THR PRO GLU ALA LYS LYS LEU LEU GLU GLU ALA LYS GLU

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

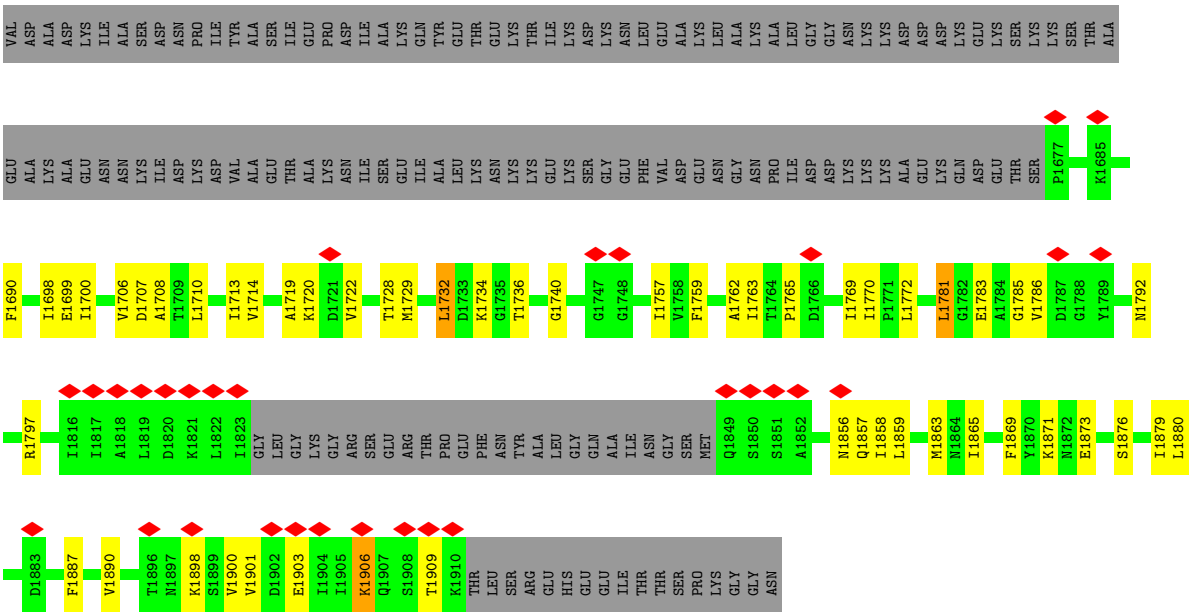
GLU LYS LEU LEU THR PRO GLU ALA ARG LYS PHE LEU ALA LYS GLN VAL LEU ASN CYS LEU GLU LYS ALA GLY ASN GLU GLU ARG LYS ALA CYS LEU LYS ASN LEU PRO LYS ASP LEU GLN GLU ASN ILE LEU ALA LYS GLU SER LEU LYS ALA TYR LYS ASP CYS LEU SER GLN ALA

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

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

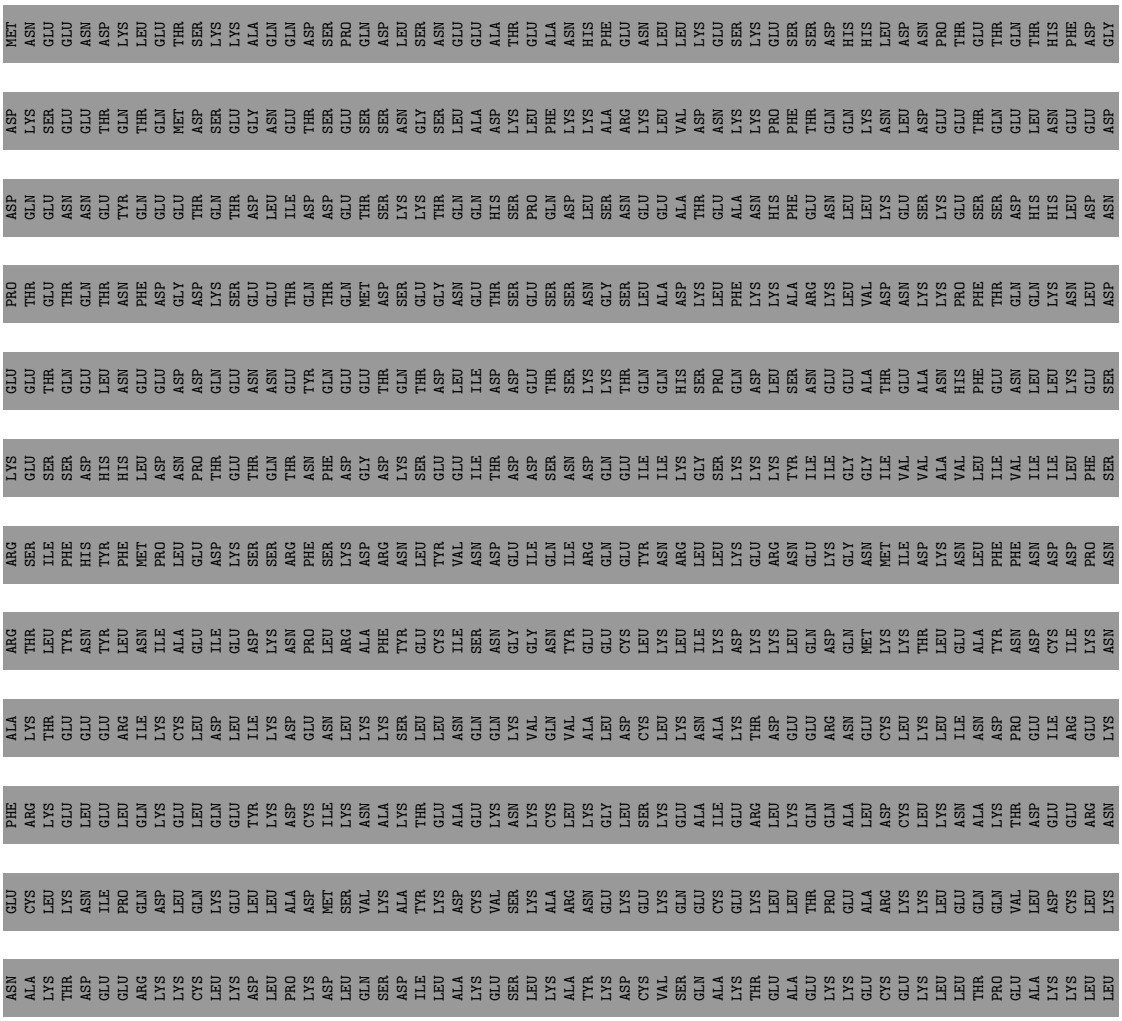
GLU GLU LYS ARG LYS CYS GLN ASN LEU TYR SER ASP LEU ILE GLN GLU ILE GLN ASN LYS ARG THR GLN ASN LYS GLN ASN GLN LEU SER LYS THR GLU ARG LEU HIS GLN ALA SER GLU CYS LEU ASP ASN LEU ASP ASP PRO THR ASP GLN GLU ALA ILE GLU GLN CYS LEU GLN GLY

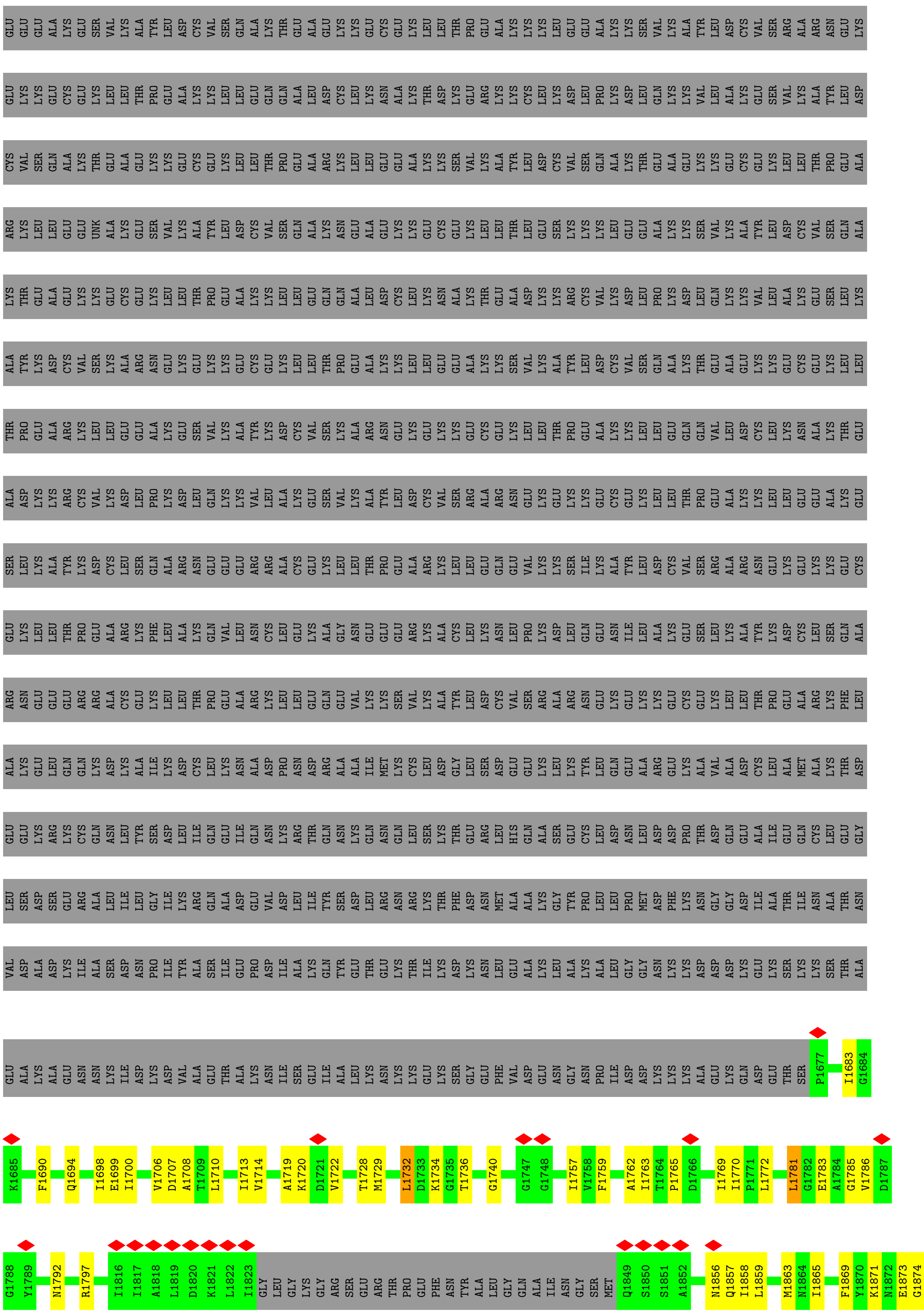
LEU SER ASP SER GLU ARG ALA LEU ILE LEU GLY ILE LYS ARG GLN ALA ASP GLU VAL ASP LEU ILE TYR SER ASP LEU ARG ASN ARG LYS THR PHE ASP ASN MET ALA ALA LYS GLY TYR PRO LEU LEU PRO MET ASP PHE LYS ASN GLY GLY ASP ILE ALA THR ILE ASN ALA THR ASN

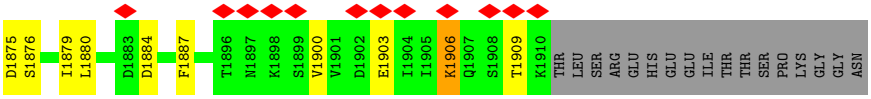


• Molecule 3: Cag pathogenicity island protein (Cag7)

Chain CY: 8% 89%

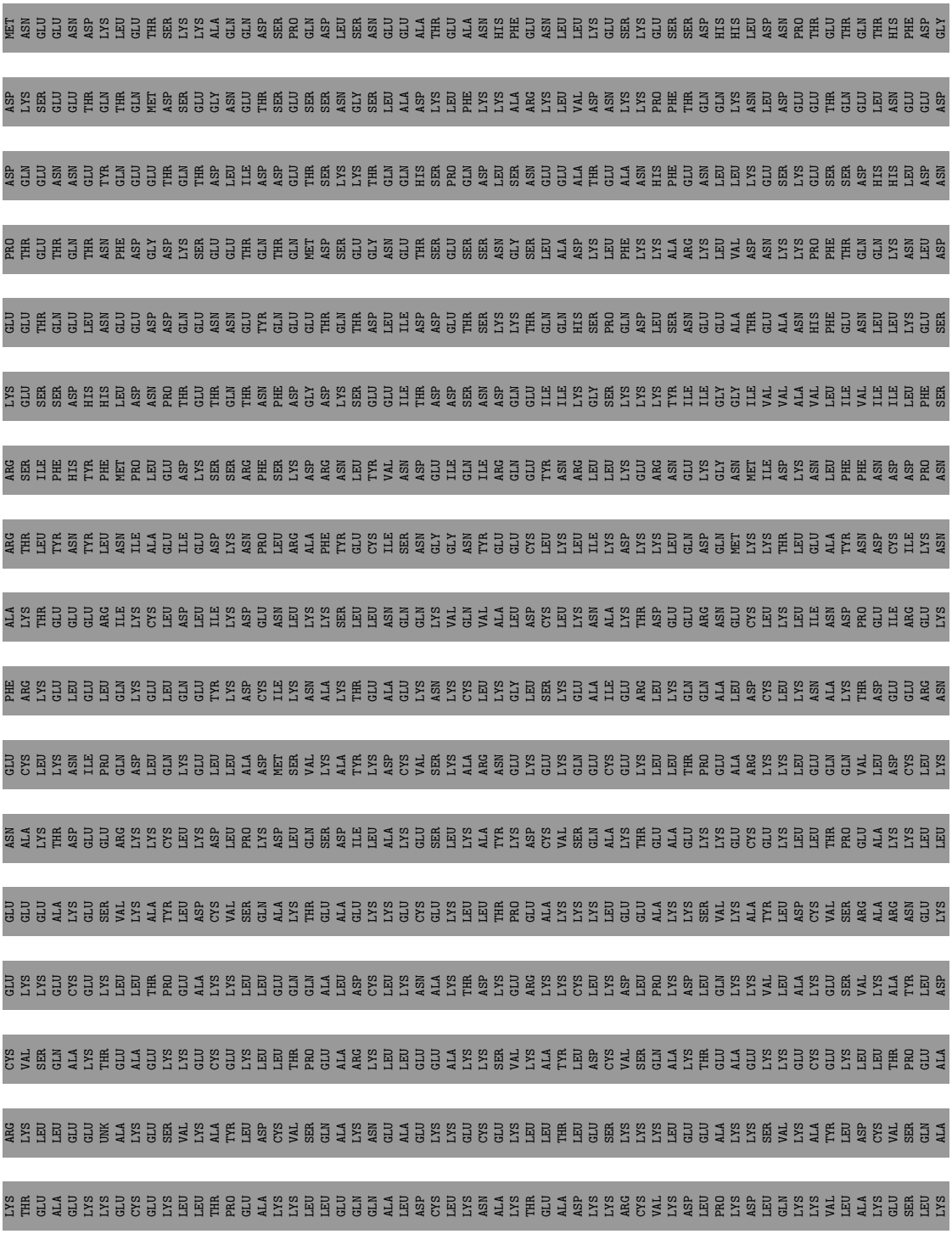


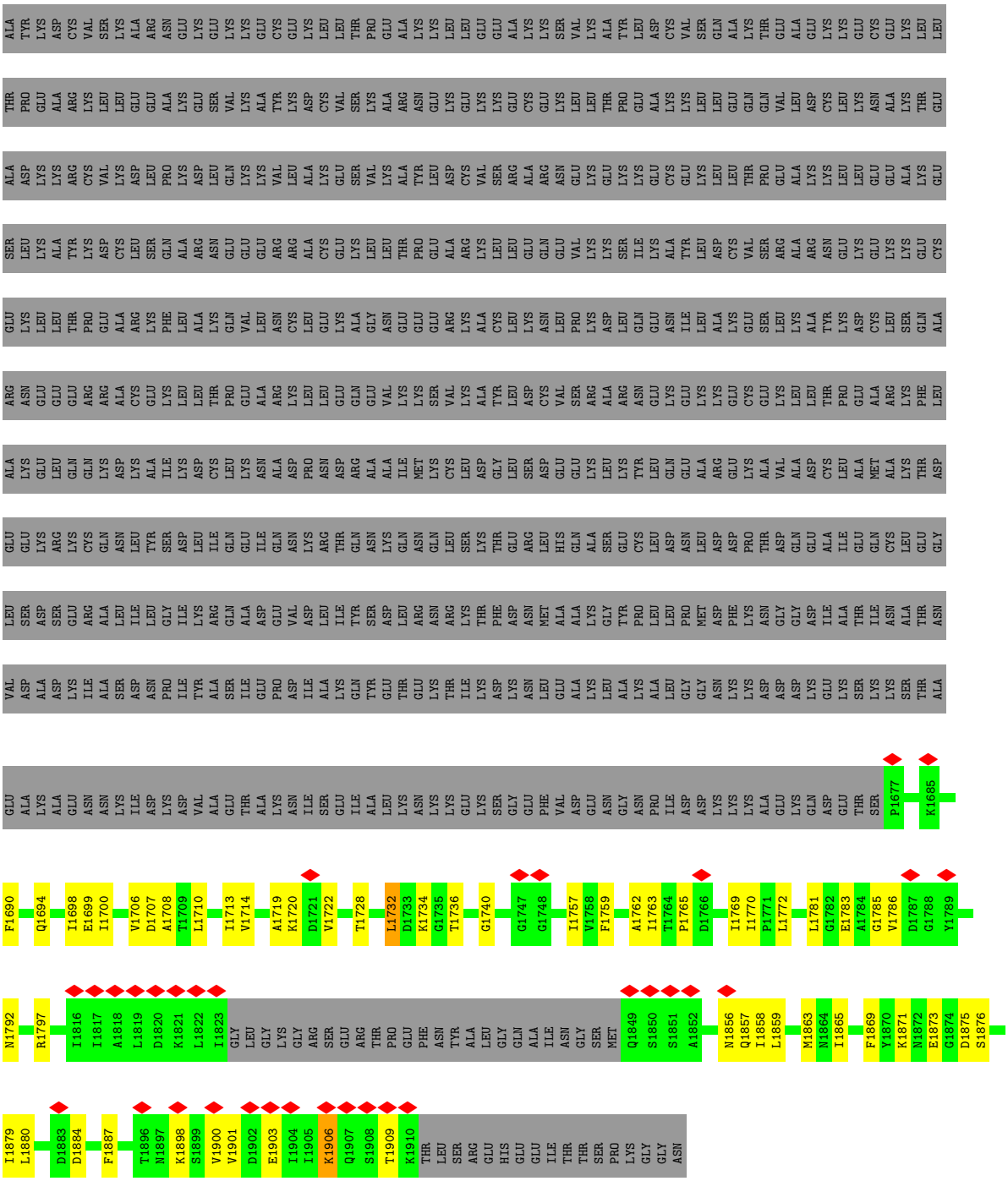

P1677
I1683
G1684
K1685 F1690 Q1694 I1698 E1699 I1700 V1706 D1707 A1708 T1709 L1710 I1713 V1714 A1719 K1720 D1721 V1722 T1728 M1729 L1732 D1733 K1734 G1735 T1736 G1740 G1747 G1748 I1757 V1758 F1759 A1762 I1763 T1764 P1765 D1766 I1769 I1770 P1771 L1772 L1781 G1782 E1783 A1784 G1785 V1786 D1787
G1788 Y1789 N1792 R1797 I1816 I1817 A1818 L1819 D1820 K1821 L1822 I1823 Q1849 S1850 S1851 A1852 N1856 Q1857 I1858 L1859 M1863 N1864 I1865 F1869 Y1870 K1871 N1872 E1873 G1874



- Molecule 3: Cag pathogenicity island protein (Cag7)

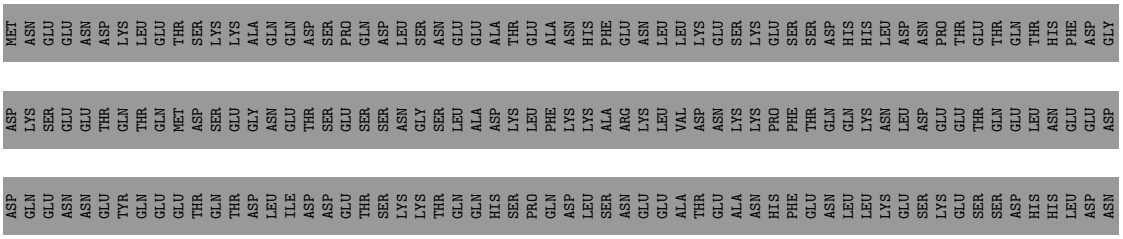
Chain DY: 8% 89%





- Molecule 3: Cag pathogenicity island protein (Cag7)

Chain EY:



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

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

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

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

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

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

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

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

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

GLU GLU GLU ALA LYS GLU SER VAL LYS ALA TYR LEU ASP CYS VAL SER GLN ALA LYS THR GLU ALA GLU LYS LYS GLU CYS GLU LYS LEU LEU THR PRO GLU ALA LYS LYS LYS LEU GLU GLU ALA LYS LYS SER VAL LYS ALA TYR LEU ASP CYS VAL SER ARG ALA ARG ASN GLU LYS

GLU LYS LYS GLU CYS GLU LYS LEU LEU THR PRO GLU ALA LYS LYS LEU LEU GLU GLN GLN ALA LEU ASP CYS LEU LYS ASN ALA LYS THR ASP LYS GLU ARG LYS LYS CYS LEU LYS ASP LEU PRO LYS ASP LEU GLN LYS LYS VAL LEU ALA LYS GLU SER VAL LYS ALA TYR LEU ASP

CYS VAL SER GLN ALA LYS THR GLU ALA GLU LYS LYS GLU CYS GLU LYS LEU LEU THR PRO GLU ALA ARG LYS LEU LEU GLU GLU ALA LYS LYS SER VAL LYS ALA TYR LEU ASP CYS VAL SER GLN ALA LYS THR GLU ALA GLU LYS LYS GLU CYS GLU LYS LEU LEU THR PRO GLU ALA

ARG LYS LEU LEU GLU GLU UNK ALA LYS GLU SER VAL LYS ALA TYR LEU ASP CYS VAL SER GLN ALA LYS ASN GLU ALA GLU LYS LYS GLU CYS GLU LYS LEU LEU THR LEU GLU SER LYS LYS LYS LEU GLU GLU ALA LYS LYS SER VAL LYS ALA TYR LEU ASP CYS VAL SER GLN ALA

LYS THR GLU ALA GLU LYS LYS GLU CYS GLU LYS LEU LEU THR PRO GLU ALA LYS LYS LEU LEU GLU GLN GLN ALA LEU ASP CYS LEU LYS ASN ALA LYS THR GLU ALA ASP LYS LYS ARG CYS VAL LYS ASP LEU PRO LYS ASP LEU GLN LYS LYS VAL LEU ALA LYS GLU SER LEU LYS

ALA TYR LYS ASP CYS VAL SER LYS ALA ARG ASN GLU LYS GLU LYS LYS GLU CYS GLU LYS LEU LEU THR PRO GLU ALA LYS LYS LEU LEU GLU GLU ALA LYS LYS SER VAL LYS ALA TYR LEU ASP CYS VAL SER GLN ALA LYS THR GLU ALA GLU LYS LYS GLU CYS GLU LYS LEU LEU

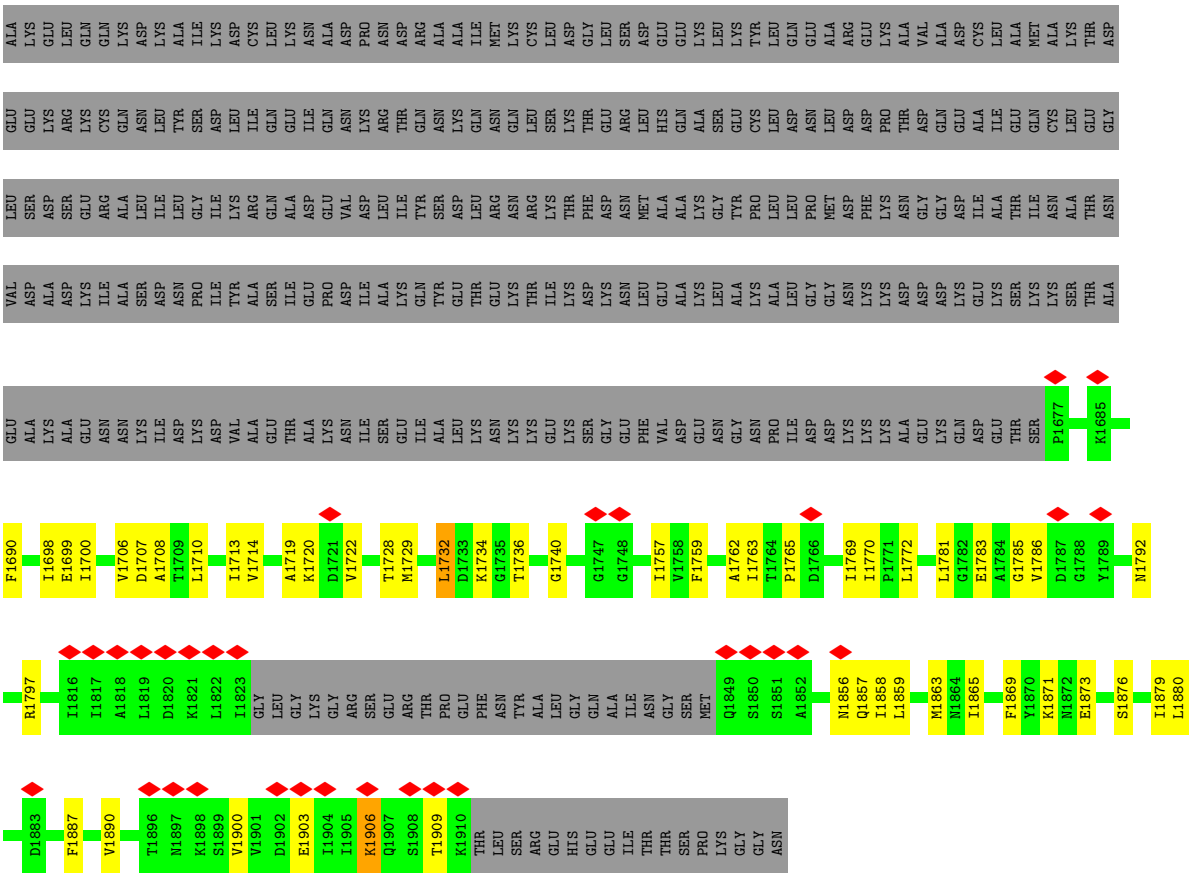
THR PRO GLU ALA ARG LYS LEU LEU GLU GLU ALA LYS GLU SER VAL LYS ALA TYR LYS ASP CYS VAL SER LYS ALA ARG ASN GLU LYS GLU LYS LYS GLU CYS GLU LYS LEU LEU THR PRO GLU ALA LYS LYS LEU LEU GLU GLN GLN VAL LEU ASP CYS LEU LYS ASN ALA LYS THR GLU

ALA ASP LYS LYS ARG CYS VAL LYS ASP LEU PRO LYS ASP LEU GLN LYS LYS VAL LEU ALA LYS GLU SER VAL LYS ALA TYR LEU ASP CYS VAL SER ARG ALA ARG ASN GLU LYS GLU LYS LYS GLU CYS GLU LYS LEU LEU THR PRO GLU ALA LYS LYS LEU LEU GLU GLU ALA LYS GLU

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

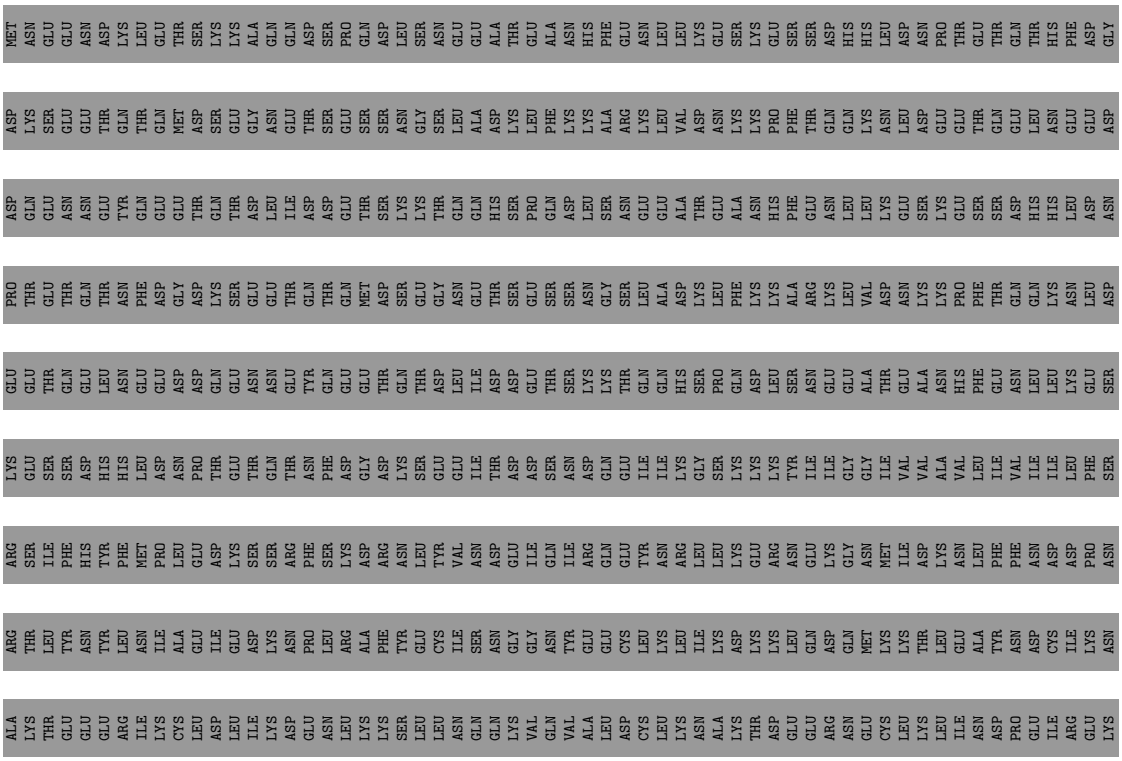
GLU LYS LEU LEU THR PRO GLU ALA ARG LYS PHE LEU ALA LYS GLN VAL LEU ASN CYS LEU GLU LYS ALA GLY ASN GLU GLU ARG LYS ALA CYS LEU LYS ASN LEU PRO LYS ASP LEU GLN GLU ASN ILE LEU ALA LYS GLU SER LEU LYS ALA TYR LYS ASP CYS LEU SER GLN ALA

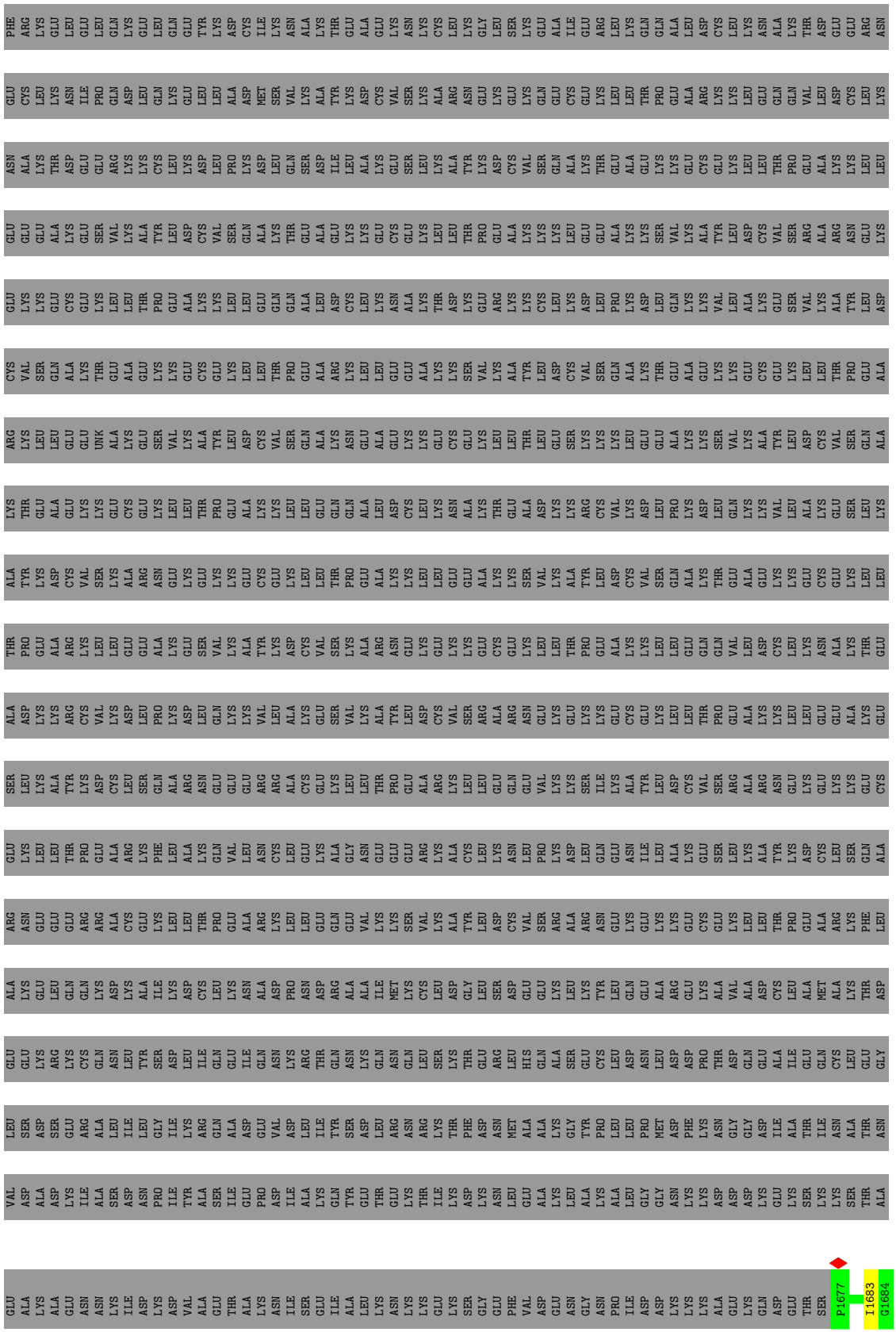
ARG ASN GLU GLU GLU ARG ARG ALA CYS GLU LYS LEU LEU THR PRO GLU ALA ARG LYS LEU LEU GLU GLN GLU VAL LYS LYS SER VAL LYS ALA TYR LEU ASP CYS VAL SER ARG ALA ARG ASN GLU LYS GLU LYS LYS GLU CYS GLU LYS LEU LEU THR PRO GLU ALA ARG LYS PHE LEU



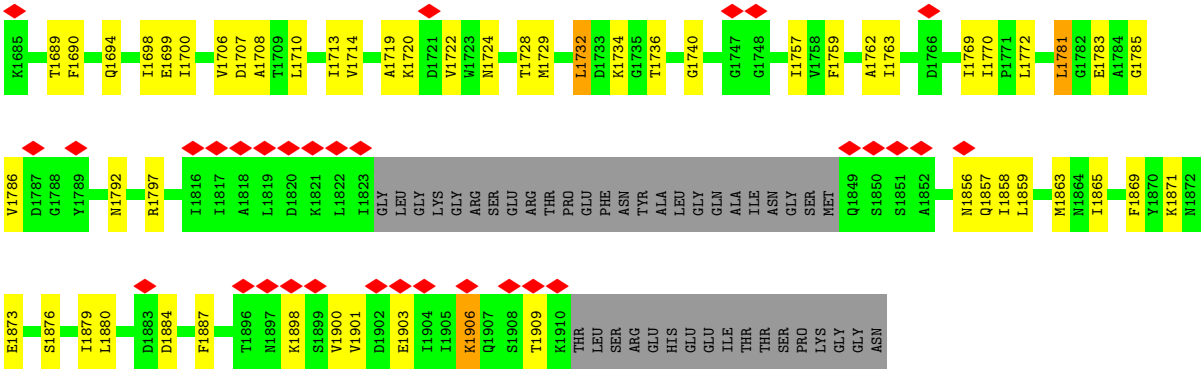
• Molecule 3: Cag pathogenicity island protein (Cag7)

Chain FY: 8% 89%



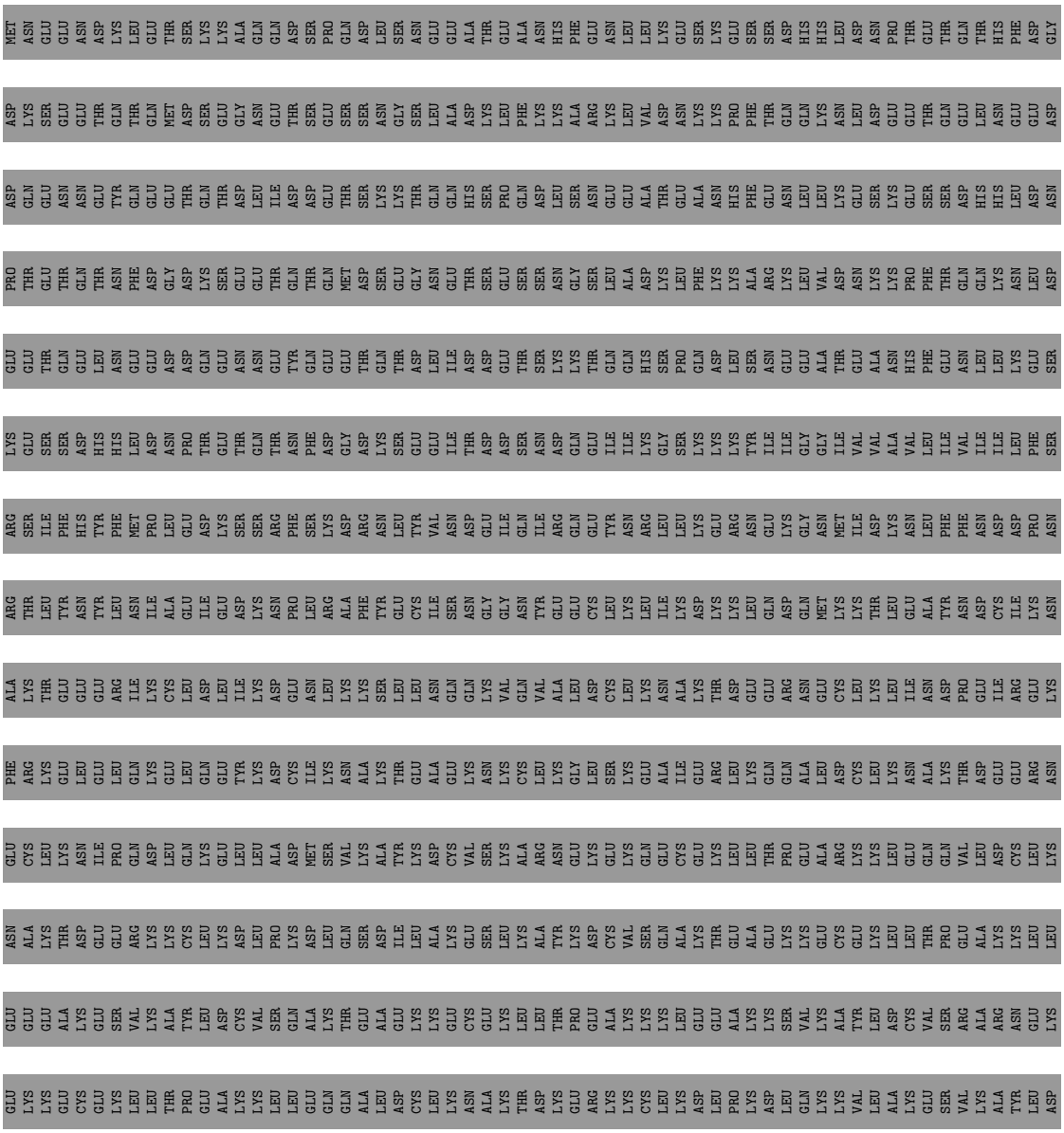


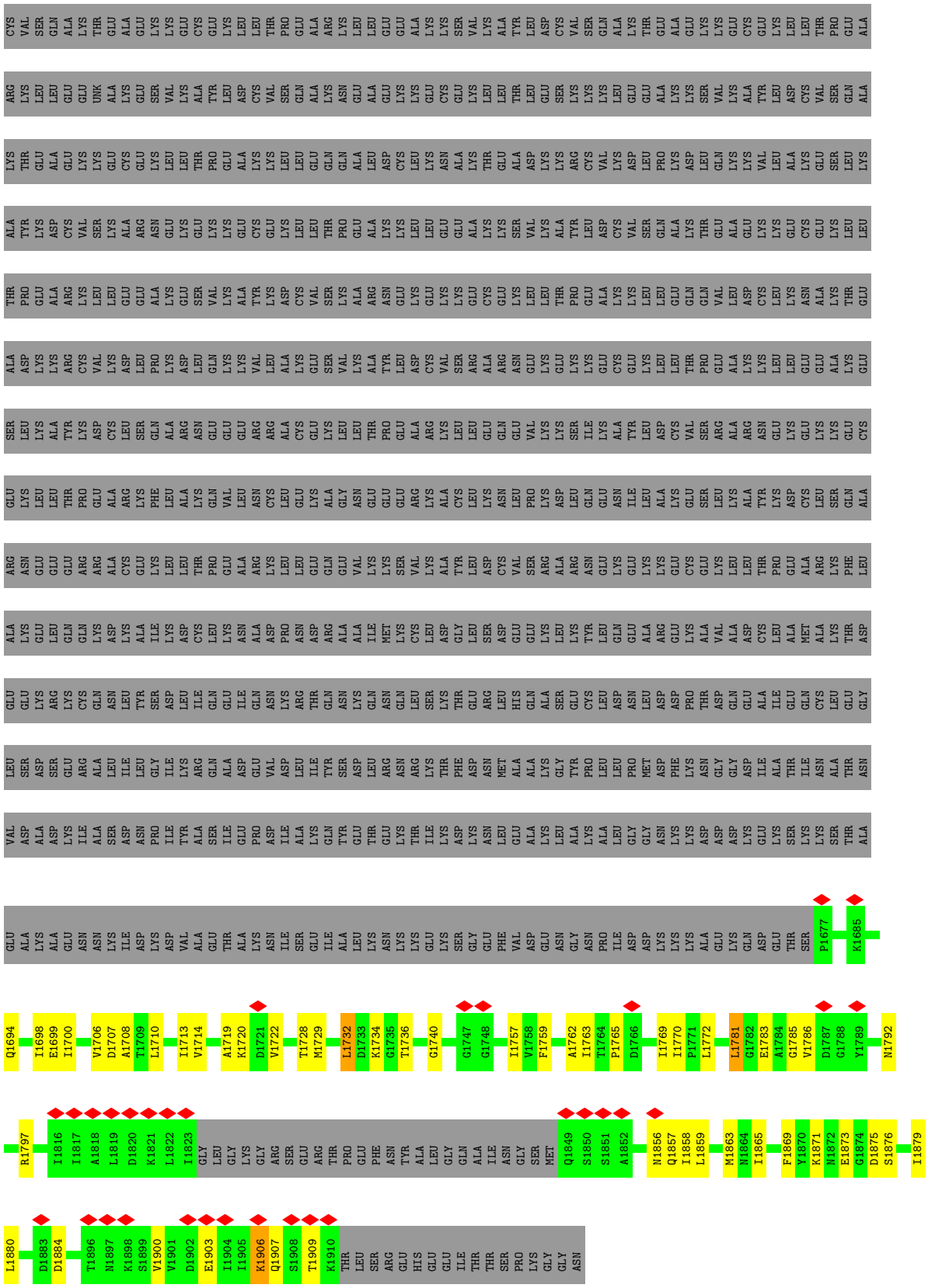
WORLDWIDE PDB PROTEIN DATA BANK



• Molecule 3: Cag pathogenicity island protein (Cag7)

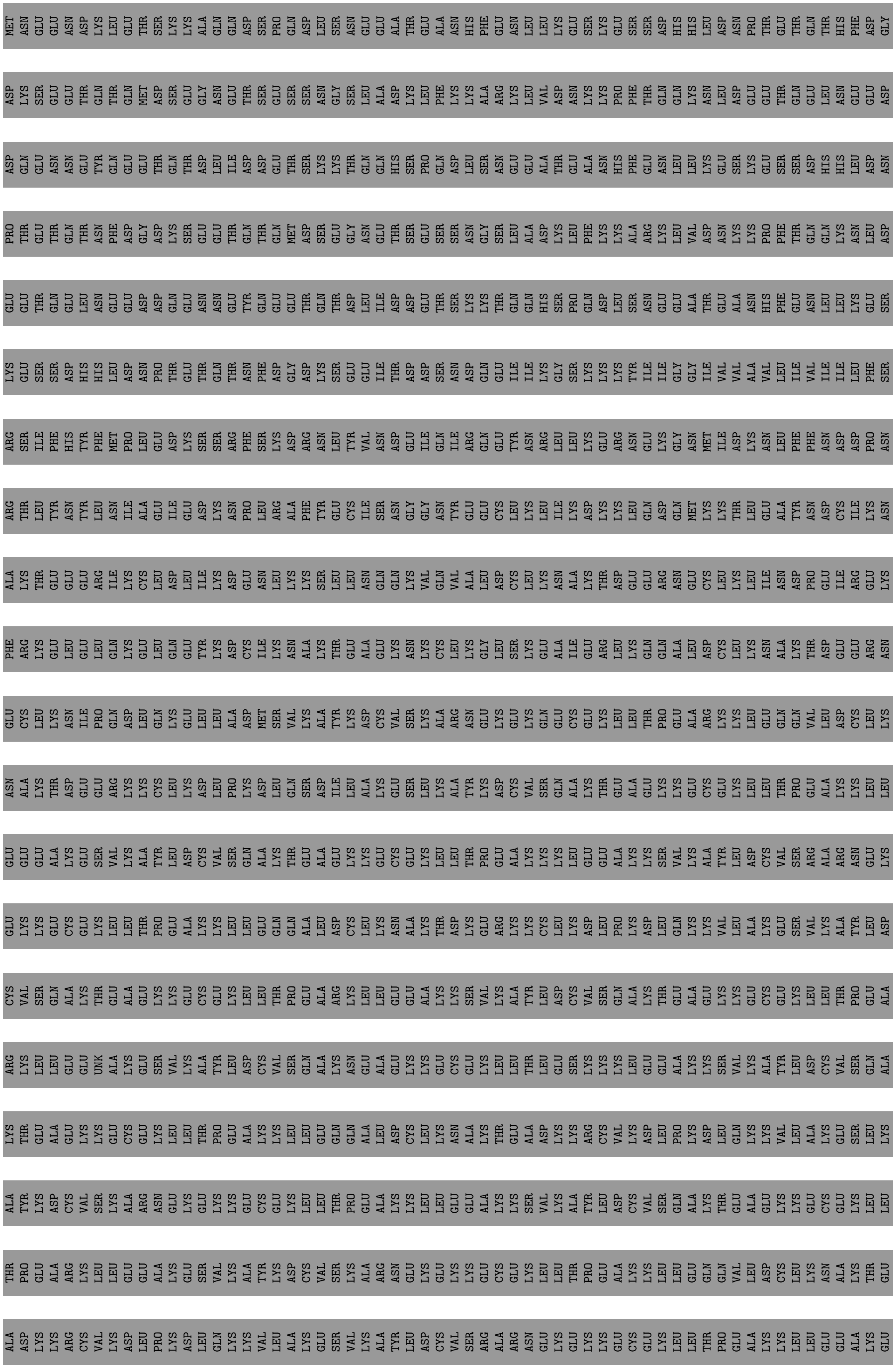
Chain GY: 8% 89%

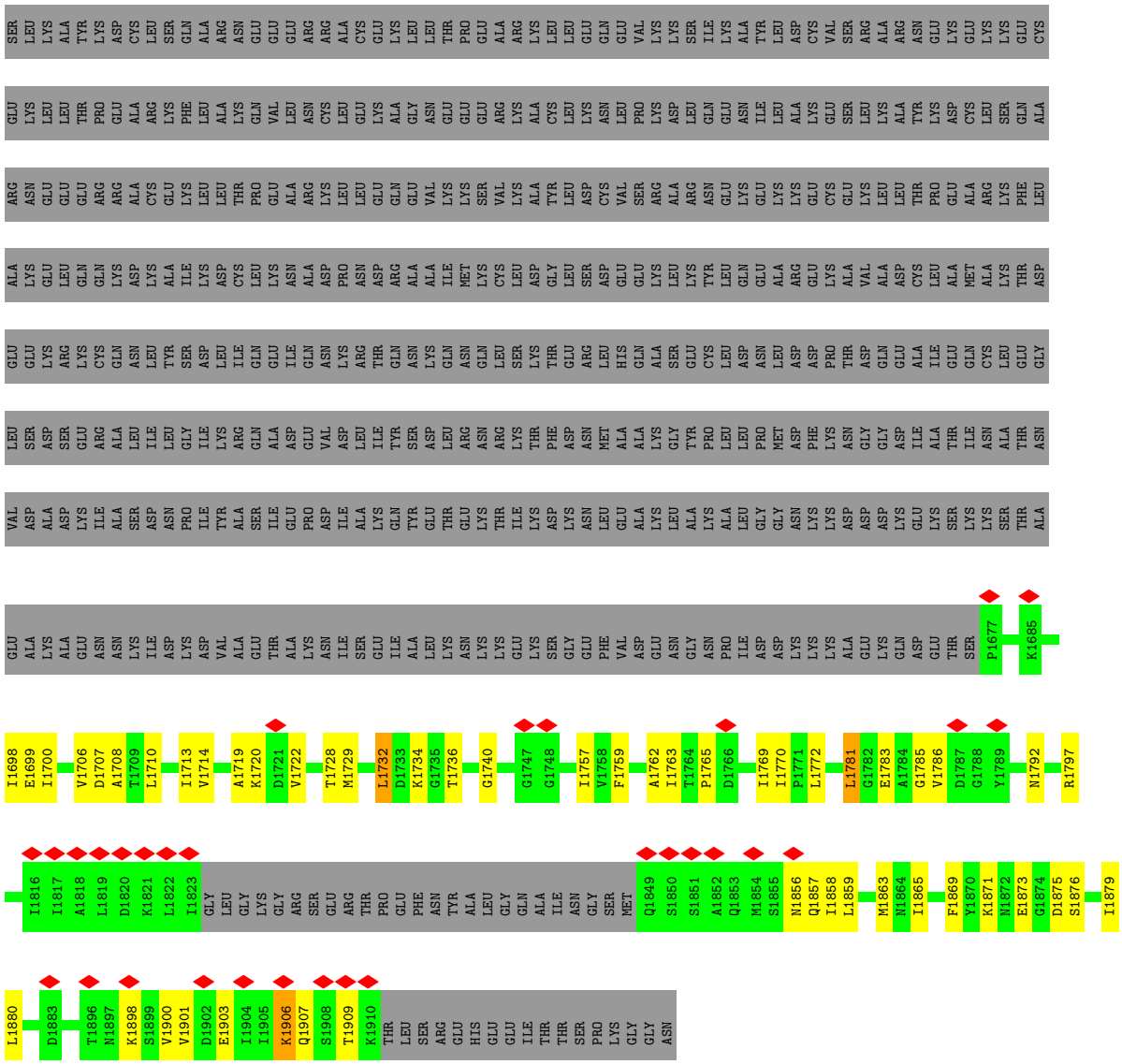




• Molecule 3: Cag pathogenicity island protein (Cag7)

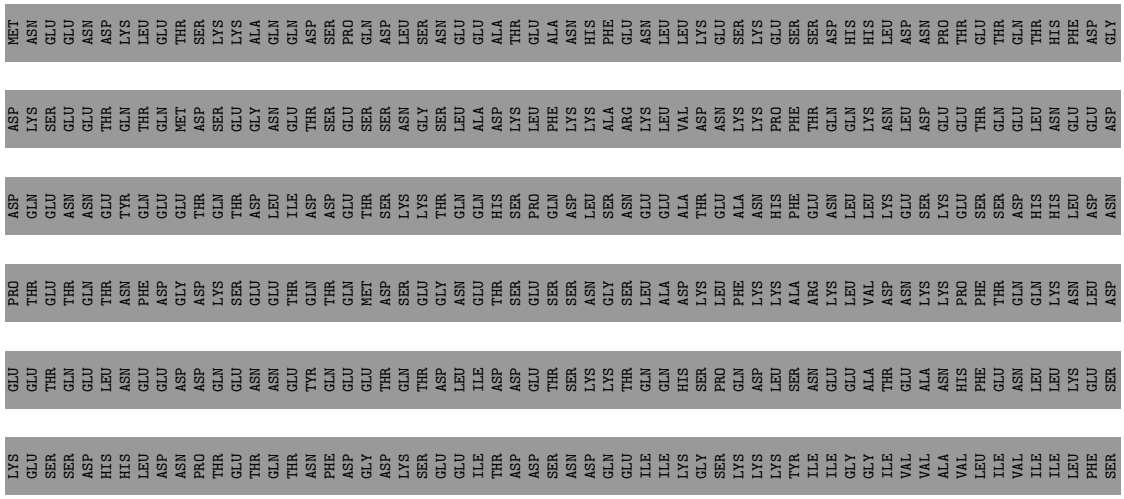
Chain HY: 8% 89%


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



- Molecule 3: Cag pathogenicity island protein (Cag7)

Chain IY:



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

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

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

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

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

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

GLU GLU GLU ALA LYS GLU SER VAL LYS ALA TYR LEU ASP CYS VAL SER GLN ALA LYS THR GLU ALA GLU LYS LYS GLU CYS GLU LYS LEU LEU THR PRO GLU ALA LYS LYS LYS LEU GLU GLU ALA LYS LYS SER VAL LYS ALA TYR LEU ASP CYS VAL SER ARG ALA ARG ASN GLU LYS

GLU LYS LYS GLU CYS GLU LYS LEU LEU THR PRO GLU ALA LYS LYS LEU LEU GLU GLN GLN ALA LEU ASP CYS LEU LYS ASN ALA LYS THR ASP LYS GLU ARG LYS LYS CYS LEU LYS ASP LEU PRO LYS ASP LEU GLN LYS LYS VAL LEU ALA LYS GLU SER VAL LYS ALA TYR LEU ASP

CYS VAL SER GLN ALA LYS THR GLU ALA GLU LYS LYS GLU CYS GLU LYS LEU LEU THR PRO GLU ALA ARG LYS LEU LEU GLU GLU ALA LYS LYS SER VAL LYS ALA TYR LEU ASP CYS VAL SER GLN ALA LYS THR GLU ALA GLU LYS LYS GLU CYS GLU LYS LEU LEU THR PRO GLU ALA

ARG LYS LEU LEU GLU GLU UNK ALA LYS GLU SER VAL LYS ALA TYR LEU ASP CYS VAL SER GLN ALA LYS ASN GLU ALA GLU LYS LYS GLU CYS GLU LYS LEU LEU THR LEU GLU SER LYS LYS LYS LEU GLU GLU ALA LYS LYS SER VAL LYS ALA TYR LEU ASP CYS VAL SER GLN ALA

LYS THR GLU ALA GLU LYS LYS GLU CYS GLU LYS LEU LEU THR PRO GLU ALA LYS LYS LEU LEU GLU GLN GLN ALA LEU ASP CYS LEU LYS ASN ALA LYS THR GLU ALA ASP LYS LYS ARG CYS VAL LYS ASP LEU PRO LYS ASP LEU GLN LYS LYS VAL LEU ALA LYS GLU SER LEU LYS

ALA TYR LYS ASP CYS VAL SER LYS ALA ARG ASN GLU LYS GLU LYS LYS GLU CYS GLU LYS LEU LEU THR PRO GLU ALA LYS LYS LEU LEU GLU GLU ALA LYS LYS SER VAL LYS ALA TYR LEU ASP CYS VAL SER GLN ALA LYS THR GLU ALA GLU LYS LYS GLU CYS GLU LYS LEU LEU

THR PRO GLU ALA ARG LYS LEU LEU GLU GLU ALA LYS GLU SER VAL LYS ALA TYR LYS ASP CYS VAL SER LYS ALA ARG ASN GLU LYS GLU LYS LYS GLU CYS GLU LYS LEU LEU THR PRO GLU ALA LYS LYS LEU LEU GLU GLN GLN VAL LEU ASP CYS LEU LYS ASN ALA LYS THR GLU

ALA ASP LYS LYS ARG CYS VAL LYS ASP LEU PRO LYS ASP LEU GLN LYS LYS VAL LEU ALA LYS GLU SER VAL LYS ALA TYR LEU ASP CYS VAL SER ARG ALA ARG ASN GLU LYS GLU LYS LYS GLU CYS GLU LYS LEU LEU THR PRO GLU ALA LYS LYS LEU LEU GLU GLU ALA LYS GLU

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

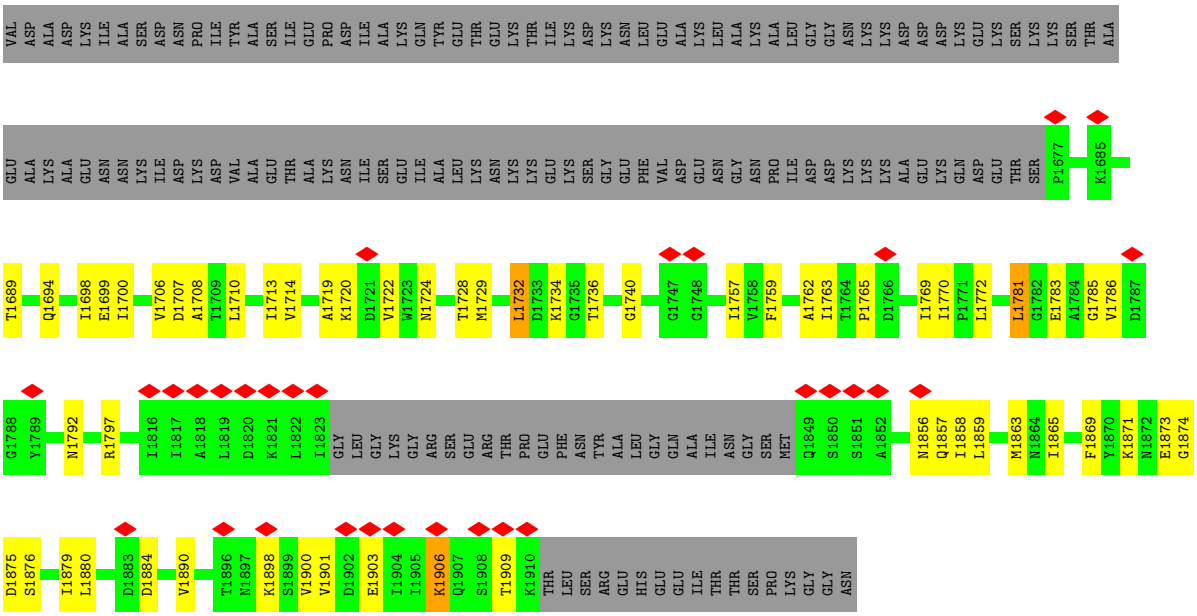
GLU LYS LEU LEU THR PRO GLU ALA ARG LYS PHE LEU ALA LYS GLN VAL LEU ASN CYS LEU GLU LYS ALA GLY ASN GLU GLU GLU ARG LYS ALA CYS LEU LYS ASN LEU PRO LYS ASP LEU GLN GLU ASN ILE LEU ALA LYS GLU SER LEU LYS ALA TYR LYS ASP CYS LEU SER GLN ALA

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

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

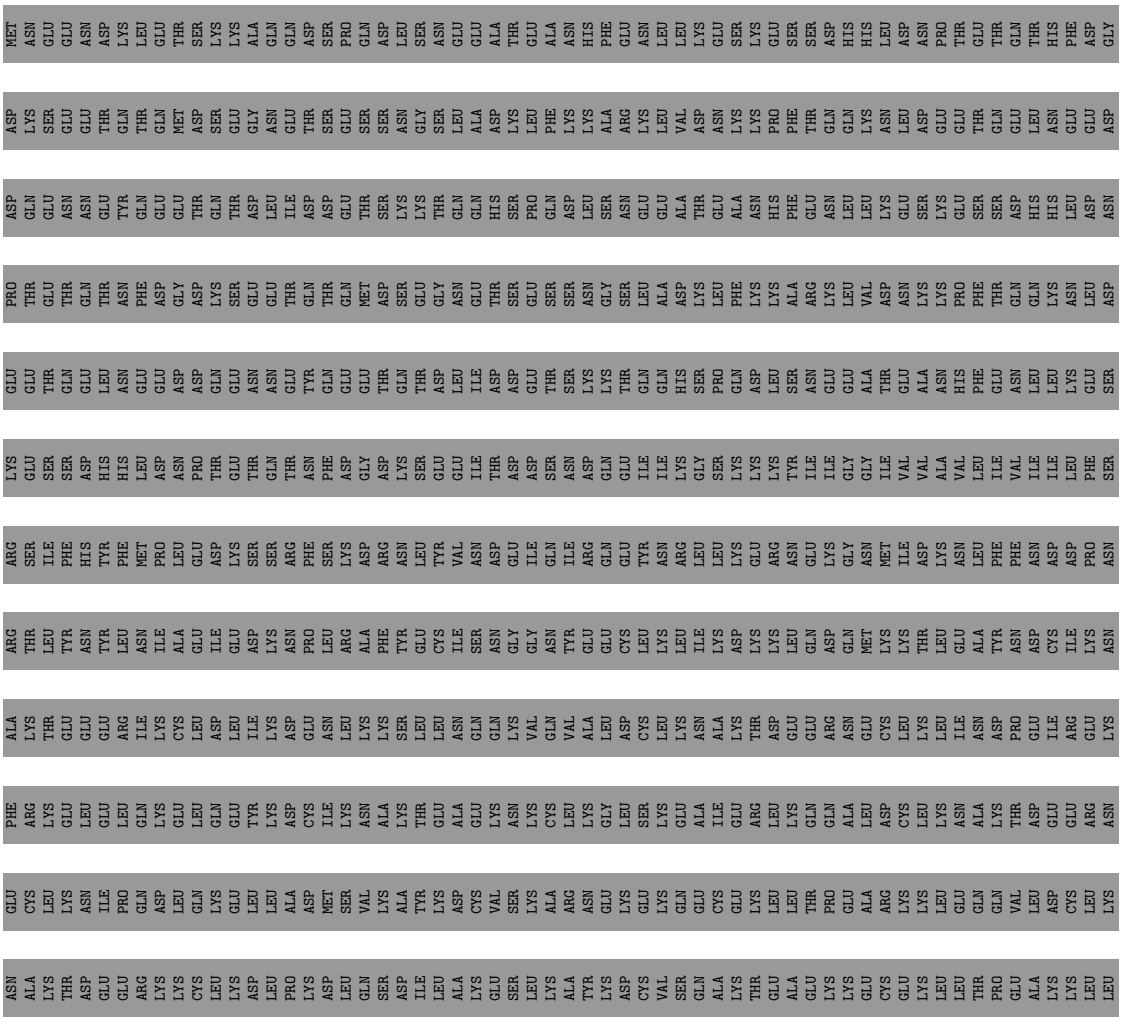
GLU GLU LYS ARG LYS CYS GLN ASN LEU TYR SER ASP LEU ILE GLN GLU ILE GLN ASN LYS ARG THR GLN ASN LYS GLN ASN GLN LEU SER LYS THR GLU ARG LEU HIS GLN ALA SER GLU CYS LEU ASP ASN LEU ASP ASP PRO THR ASP GLN GLU ALA ILE GLU GLN CYS LEU GLU GLY

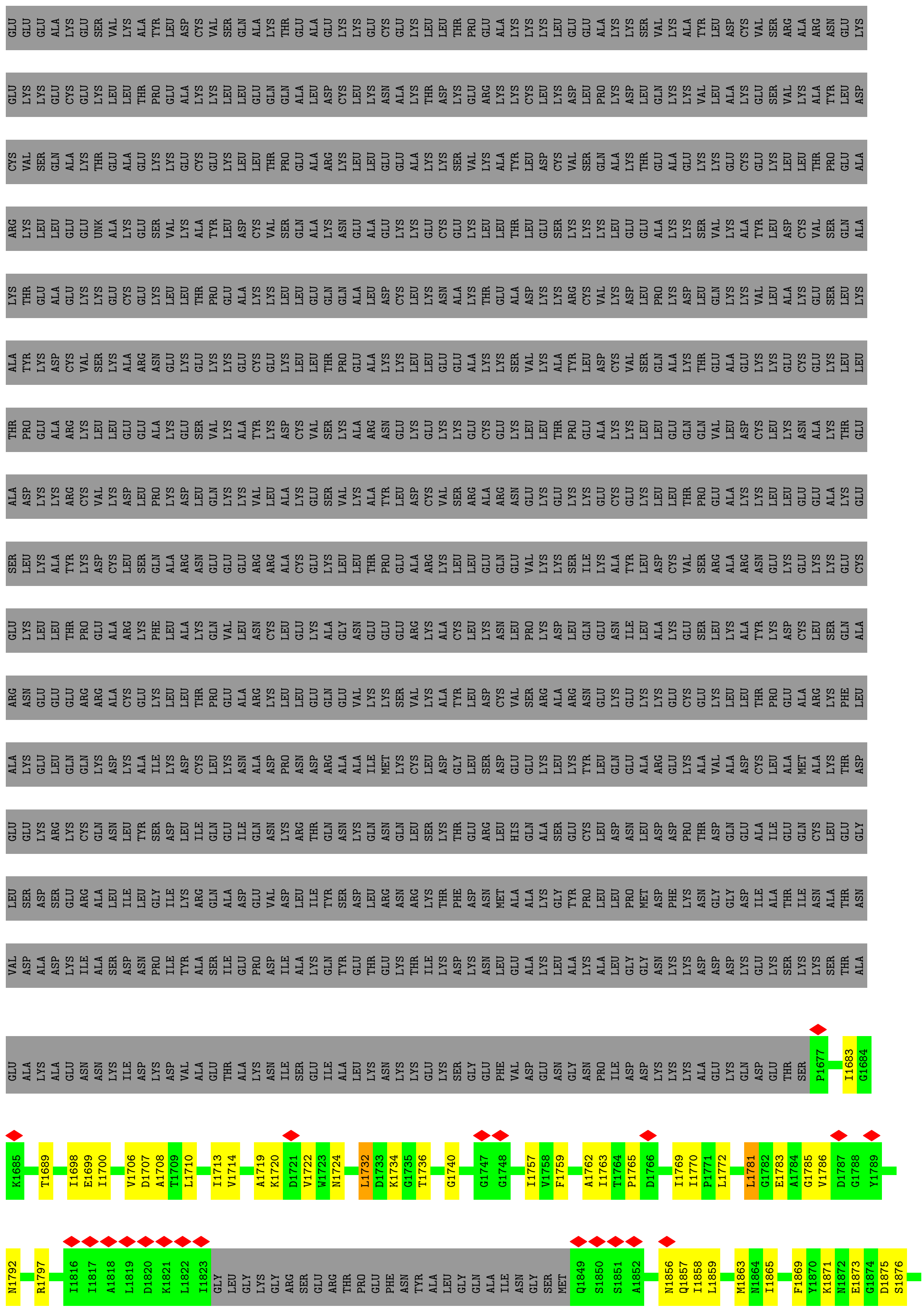
LEU SER ASP SER GLU ARG ALA LEU ILE LEU GLY ILE LYS ARG GLN ALA ASP GLU VAL ASP LEU ILE TYR SER ASP LEU ARG ASN ARG LYS THR PHE ASP ASN MET ALA ALA LYS GLY TYR PRO LEU LEU PRO MET ASP PHE LYS ASN GLY GLY ASP ILE ALA THR ILE ASN ALA THR ASN

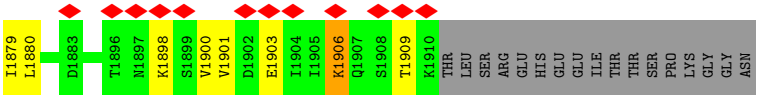


- Molecule 3: Cag pathogenicity island protein (Cag7)

Chain JY:

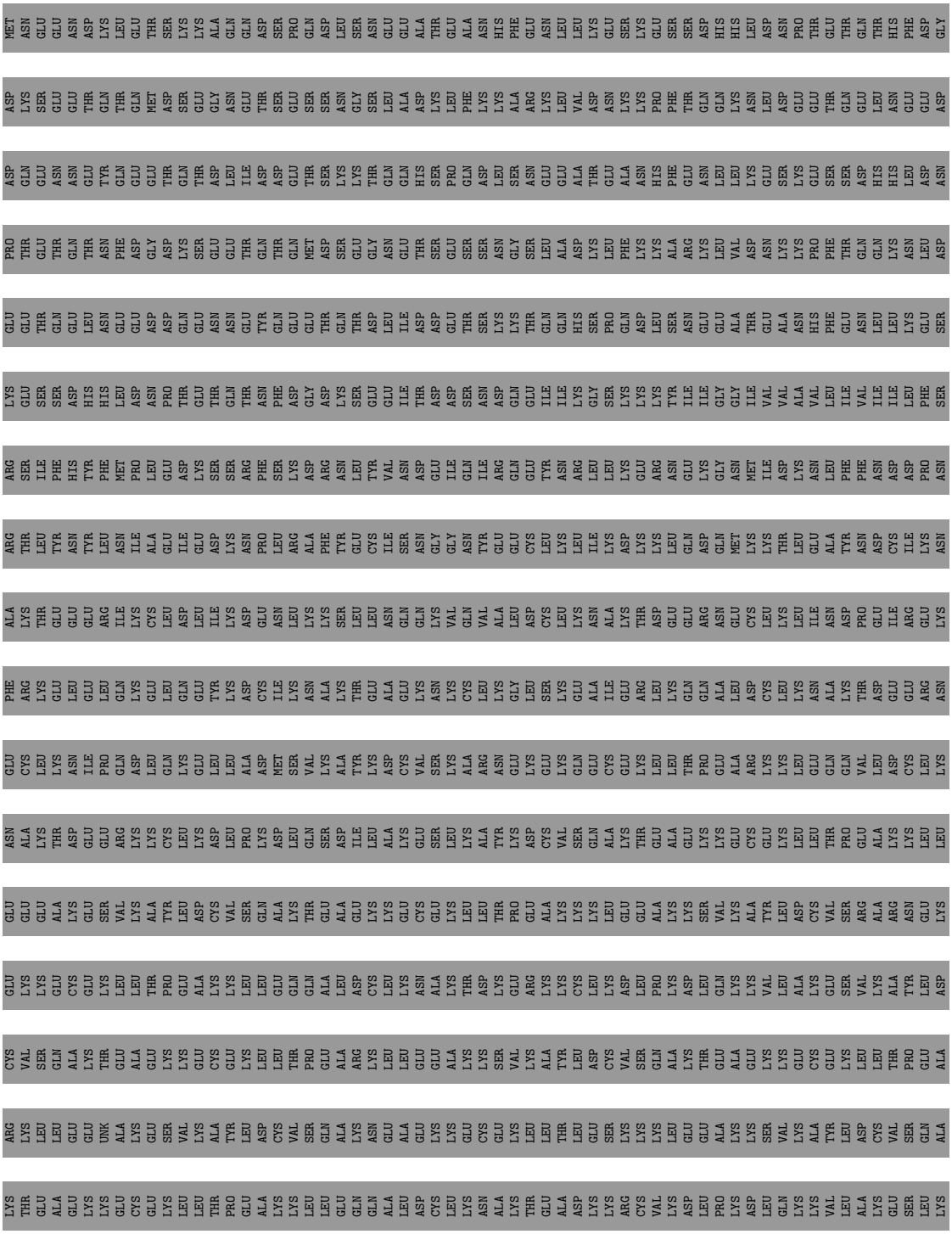


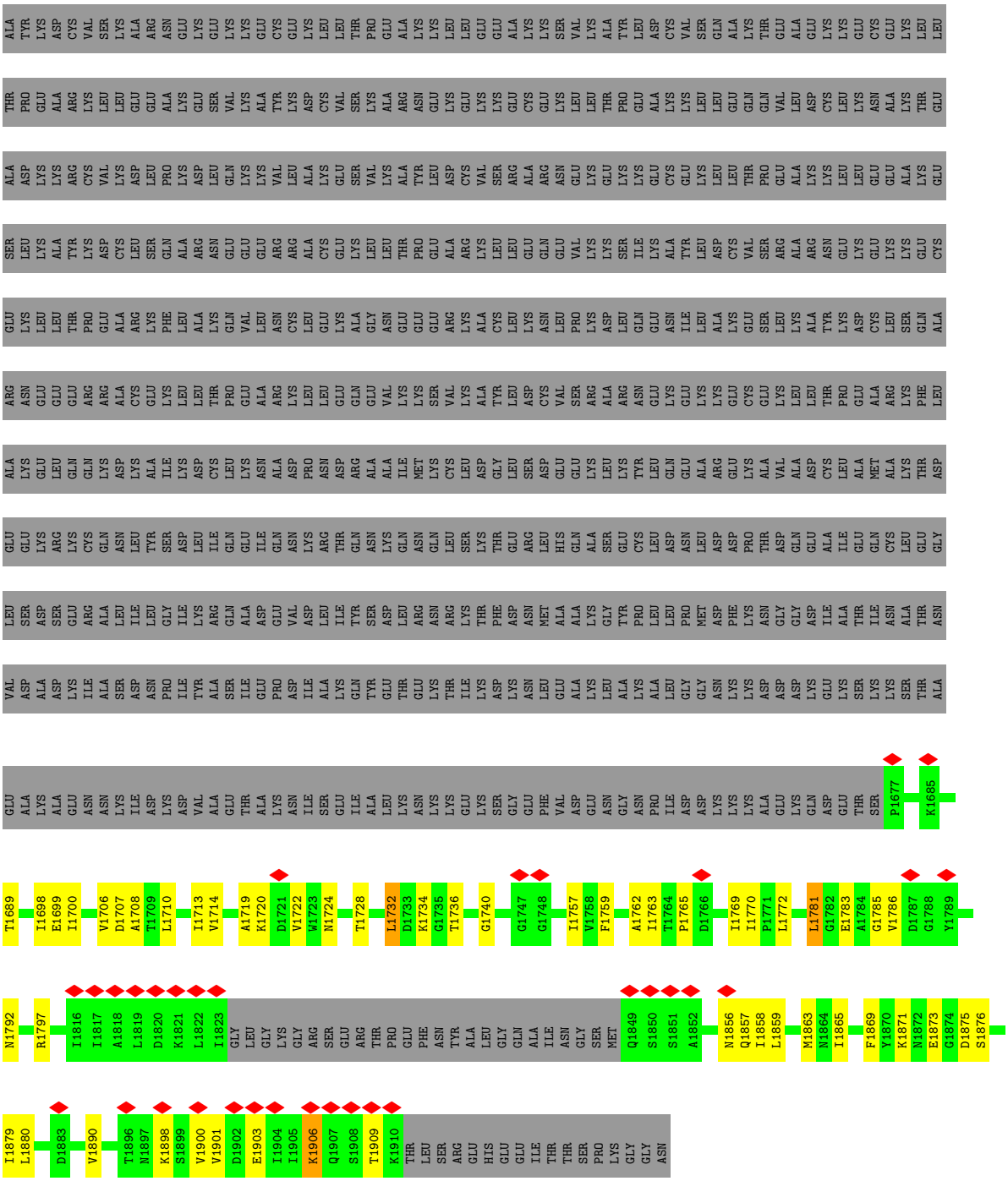

GLU GLU GLU ALA LYS GLU SER VAL LYS ALA TYR LEU ASP CYS VAL SER GLN ALA LYS THR GLU ALA GLU LYS LYS GLU CYS GLU LYS LEU LEU THR PRO GLU ALA LYS LYS LYS LEU GLU GLU ALA LYS LYS SER VAL LYS ALA TYR LEU ASP CYS VAL SER ARG ALA ARG ASN GLU LYS
GLU LYS LYS GLU CYS GLU LYS LEU LEU THR PRO GLU ALA LYS LYS LEU LEU GLU GLN GLN ALA LEU ASP CYS LEU LYS ASN ALA LYS THR ASP LYS GLU ARG LYS LYS CYS LEU LYS ASP LEU PRO LYS ASP LEU GLN LYS LYS VAL LEU ALA LYS GLU SER VAL LYS ALA TYR LEU ASP
CYS VAL SER GLN ALA LYS THR GLU ALA GLU LYS LYS GLU CYS GLU LYS LEU LEU THR PRO GLU ALA ARG LYS LEU LEU GLU GLU ALA LYS LYS SER VAL LYS ALA TYR LEU ASP CYS VAL SER GLN ALA LYS THR GLU ALA GLU LYS LYS GLU CYS GLU LYS LEU LEU THR PRO GLU ALA
ARG LYS LEU LEU GLU GLU UNK ALA LYS GLU SER VAL LYS ALA TYR LEU ASP CYS VAL SER GLN ALA LYS ASN GLU ALA GLU LYS LYS GLU CYS GLU LYS LEU LEU THR LEU GLU SER LYS LYS LEU GLU GLU ALA LYS LYS SER VAL LYS ALA TYR LEU ASP CYS VAL SER GLN ALA
LYS THR GLU ALA GLU LYS LYS GLU CYS GLU LYS LEU LEU THR PRO GLU ALA LYS LYS LEU LEU GLU GLN GLN ALA LEU ASP CYS LEU LYS ASN ALA LYS THR GLU ALA ASP LYS LYS ARG CYS VAL LYS ASP LEU PRO LYS ASP LEU GLN LYS LYS VAL LEU ALA LYS GLU SER LEU LYS
ALA TYR LYS ASP CYS VAL SER LYS ALA ARG ASN GLU LYS GLU LYS LYS GLU CYS GLU LYS LEU LEU THR PRO GLU ALA LYS LYS LEU GLU GLU ALA LYS LYS SER VAL LYS ALA TYR LEU ASP CYS VAL SER GLN ALA LYS THR GLU ALA GLU LYS LYS GLU CYS GLU LYS LEU LEU
THR PRO GLU ALA ARG LYS LEU LEU GLU GLU ALA LYS GLU SER VAL LYS ALA TYR LYS ASP CYS VAL SER LYS ALA ARG ASN GLU LYS GLU LYS LYS GLU CYS GLU LYS LEU LEU THR PRO GLU ALA LYS LYS LEU LEU GLU GLN GLN VAL LEU ASP CYS LEU LYS ASN ALA LYS THR GLU
ALA ASP LYS LYS ARG CYS VAL LYS ASP LEU PRO LYS ASP LEU GLN LYS LYS VAL LEU ALA LYS GLU SER VAL LYS ALA TYR LEU ASP CYS VAL SER ARG ALA ARG ASN GLU LYS GLU LYS LYS GLU CYS GLU LYS LEU LEU THR PRO GLU ALA LYS LYS LEU GLU GLU ALA LYS GLU
SER LEU LYS ALA TYR LYS ASP CYS LEU SER GLN ALA ARG ASN GLU GLU GLU ARG ARG ALA CYS GLU LYS LEU LEU THR PRO GLU ALA ARG LYS LEU LEU GLU GLN GLU VAL LYS LYS SER ILE LYS ALA TYR LEU ASP CYS VAL SER ARG ALA ARG ASN GLU LYS GLU LYS LYS GLU CYS
GLU LYS LEU LEU THR PRO GLU ALA ARG LYS PHE LEU ALA LYS GLN VAL LEU ASN CYS LEU GLU LYS ALA GLY ASN GLU GLU ARG LYS ALA CYS LEU LYS ASN LEU PRO LYS ASP LEU GLN GLU ASN ILE LEU ALA LYS GLU SER LEU LYS ALA TYR LYS ASP CYS LEU SER GLN ALA
ARG ASN GLU GLU GLU ARG ARG ALA CYS GLU LYS LEU LEU THR PRO GLU ALA ARG LYS LEU LEU GLU GLN GLU VAL LYS LYS SER VAL LYS ALA TYR LEU ASP CYS VAL SER ARG ALA ARG ASN GLU LYS GLU LYS LYS GLU CYS GLU LYS LEU LEU THR PRO GLU ALA ARG LYS PHE LEU
ALA LYS GLU LEU GLN GLN LYS ASP LYS ALA ILE LYS ASP CYS LEU LYS ASN ALA ASP PRO ASN ASP ARG ALA ALA ILE MET LYS CYS LEU ASP GLY LEU SER ASP GLU GLU LYS LEU LYS TYR LEU GLN GLU ALA ARG GLU LYS ALA VAL ALA ASP CYS LEU ALA MET ALA LYS THR ASP
GLU GLU LYS ARG LYS CYS GLN ASN LEU TYR SER ASP LEU ILE GLN GLU ILE GLN ASN LYS ARG THR GLN ASN LYS GLN ASN GLN LEU SER LYS THR GLU ARG LEU HIS GLN ALA SER GLU CYS LEU ASP ASN LEU ASP ASP PRO THR ASP GLN GLU ALA ILE GLU GLN CYS LEU GLU GLY
LEU SER ASP SER GLU ARG ALA LEU ILE LEU GLY ILE LYS ARG GLN ALA ASP GLU VAL ASP LEU ILE TYR SER ASP LEU ARG ASN ARG LYS THR PHE ASP ASN MET ALA ALA LYS GLY TYR PRO LEU LEU PRO MET ASP PHE LYS ASN GLY GLY ASP ILE ALA THR ILE ASN ALA THR ASN
VAL ASP ALA ASP LYS ILE ALA SER ASP ASN PRO ILE TYR ALA SER ILE GLU PRO ASP ILE ALA LYS GLN TYR GLU THR GLU LYS THR ILE LYS ASP LYS ASN LEU GLU ALA LYS LEU ALA LYS ALA LEU GLY GLY ASN LYS LYS ASP ASP ASP LYS GLU LYS SER LYS LYS SER THR ALA
GLU ALA LYS ALA GLU ASN ASN LYS ILE ASP LYS ASP VAL ALA GLU THR ALA LYS ASN ILE SER GLU ILE ALA LEU LYS ASN LYS LYS GLU LYS SER GLY GLU PHE VAL ASP GLU ASN GLY ASN PRO ILE ASP ASP LYS LYS LYS ALA GLU LYS GLN ASP GLU THR SER P1677 I1683 G1684
K1685 T1689 I1698 E1699 I1700 V1706 D1707 A1708 T1709 L1710 I1713 V1714 A1719 K1720 D1721 V1722 W1723 N1724 L1732 D1733 K1734 G1735 T1736 G1740 G1747 G1748 I1757 V1758 F1759 A1762 I1763 T1764 P1765 D1766 I1769 I1770 P1771 L1772 L1781 G1782 E1783 A1784 G1785 V1786 D1787 G1788 Y1789
N1792 R1797 I1816 I1817 A1818 L1819 D1820 K1821 L1822 I1823 GLY LEU GLY LYS GLY ARG SER GLU ARG THR PRO GLU PHE ASN TYR ALA LEU GLY GLN ALA ILE ASN GLY SER MET Q1849 S1850 S1851 A1852 N1856 Q1857 I1858 L1859 M1863 N1864 I1865 F1869 Y1870 K1871 N1872 E1873 G1874 D1875 S1876



• Molecule 3: Cag pathogenicity island protein (Cag7)

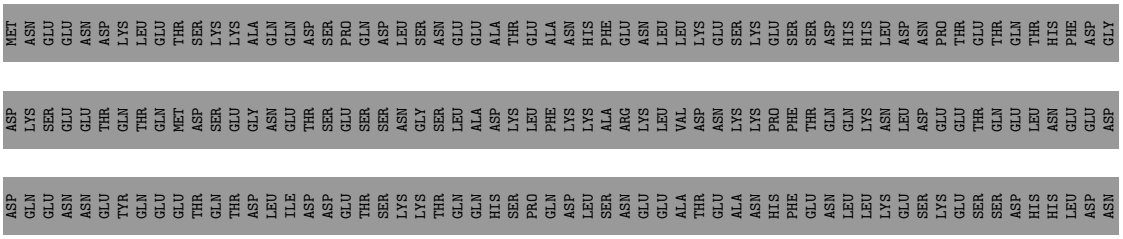
Chain KY:





• Molecule 3: Cag pathogenicity island protein (Cag7)

Chain LY:



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

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

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

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

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

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

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

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

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

GLU GLU GLU ALA LYS GLU SER VAL LYS ALA TYR LEU ASP CYS VAL SER GLN ALA LYS THR GLU ALA GLU LYS LYS GLU CYS GLU LYS LEU LEU THR PRO GLU ALA LYS LYS LYS LEU GLU GLU ALA LYS LYS SER VAL LYS ALA TYR LEU ASP CYS VAL SER ARG ALA ARG ASN GLU LYS

GLU LYS LYS GLU CYS GLU LYS LEU LEU THR PRO GLU ALA LYS LYS LEU LEU GLU GLN GLN ALA LEU ASP CYS LEU LYS ASN ALA LYS THR ASP LYS GLU ARG LYS LYS CYS LEU LYS ASP LEU PRO LYS ASP LEU GLN LYS LYS VAL LEU ALA LYS GLU SER VAL LYS ALA TYR LEU ASP

CYS VAL SER GLN ALA LYS THR GLU ALA GLU LYS LYS GLU CYS GLU LYS LEU LEU THR PRO GLU ALA ARG LYS LEU LEU GLU GLU ALA LYS LYS SER VAL LYS ALA TYR LEU ASP CYS VAL SER GLN ALA LYS THR GLU ALA GLU LYS LYS GLU CYS GLU LYS LEU LEU THR PRO GLU ALA

ARG LYS LEU LEU GLU GLU UNK ALA LYS GLU SER VAL LYS ALA TYR LEU ASP CYS VAL SER GLN ALA LYS ASN GLU ALA GLU LYS LYS GLU CYS GLU LYS LEU LEU THR LEU GLU SER LYS LYS LYS LEU GLU GLU ALA LYS LYS SER VAL LYS ALA TYR LEU ASP CYS VAL SER GLN ALA

LYS THR GLU ALA GLU LYS LYS GLU CYS GLU LYS LEU LEU THR PRO GLU ALA LYS LYS LEU LEU GLU GLN GLN ALA LEU ASP CYS LEU LYS ASN ALA LYS THR GLU ALA ASP LYS LYS ARG CYS VAL LYS ASP LEU PRO LYS ASP LEU GLN LYS LYS VAL LEU ALA LYS GLU SER LEU LYS

ALA TYR LYS ASP CYS VAL SER LYS ALA ARG ASN GLU LYS GLU LYS LYS GLU CYS GLU LYS LEU LEU THR PRO GLU ALA LYS LYS LEU LEU GLU GLU ALA LYS LYS SER VAL LYS ALA TYR LEU ASP CYS VAL SER GLN ALA LYS THR GLU ALA GLU LYS LYS GLU CYS GLU LYS LEU LEU

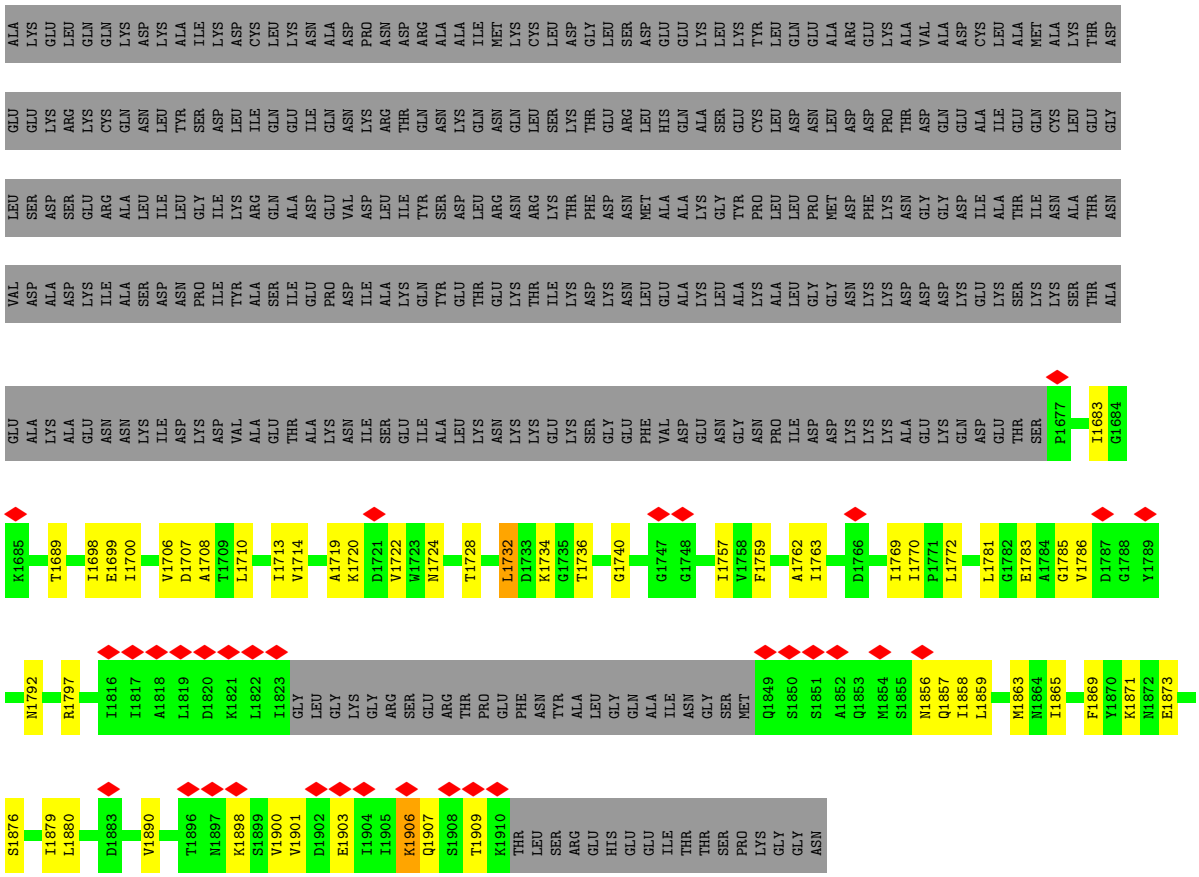
THR PRO GLU ALA ARG LYS LEU LEU GLU GLU ALA LYS GLU SER VAL LYS ALA TYR LYS ASP CYS VAL SER LYS ALA ARG ASN GLU LYS GLU LYS LYS GLU CYS GLU LYS LEU LEU THR PRO GLU ALA LYS LYS LEU LEU GLU GLN GLN VAL LEU ASP CYS LEU LYS ASN ALA LYS THR GLU

ALA ASP LYS LYS ARG CYS VAL LYS ASP LEU PRO LYS ASP LEU GLN LYS LYS VAL LEU ALA LYS GLU SER VAL LYS ALA TYR LEU ASP CYS VAL SER ARG ALA ARG ASN GLU LYS GLU LYS LYS GLU CYS GLU LYS LEU LEU THR PRO GLU ALA LYS LYS LEU LEU GLU GLU ALA LYS GLU

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

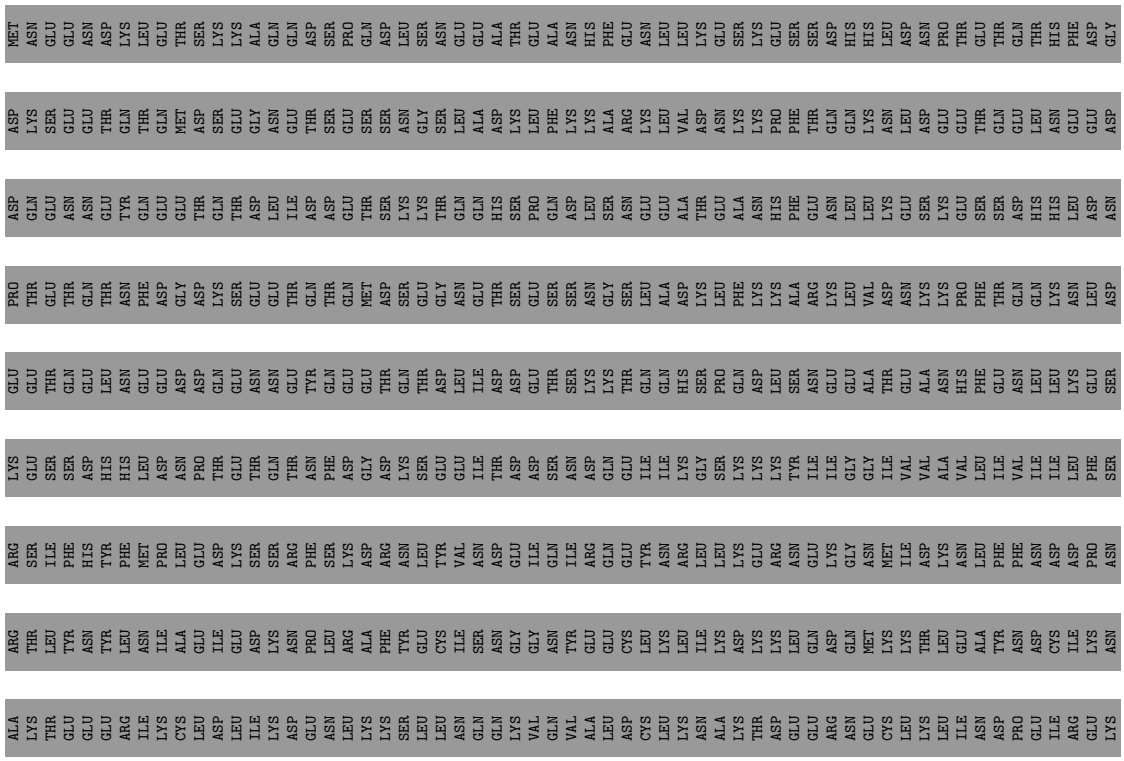
GLU LYS LEU LEU THR PRO GLU ALA ARG LYS PHE LEU ALA LYS GLN VAL LEU ASN CYS LEU GLU LYS ALA GLY ASN GLU GLU GLU ARG LYS ALA CYS LEU LYS ASN LEU PRO LYS ASP LEU GLN GLU ASN ILE LEU ALA LYS GLU SER LEU LYS ALA TYR LYS ASP CYS LEU SER GLN ALA

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

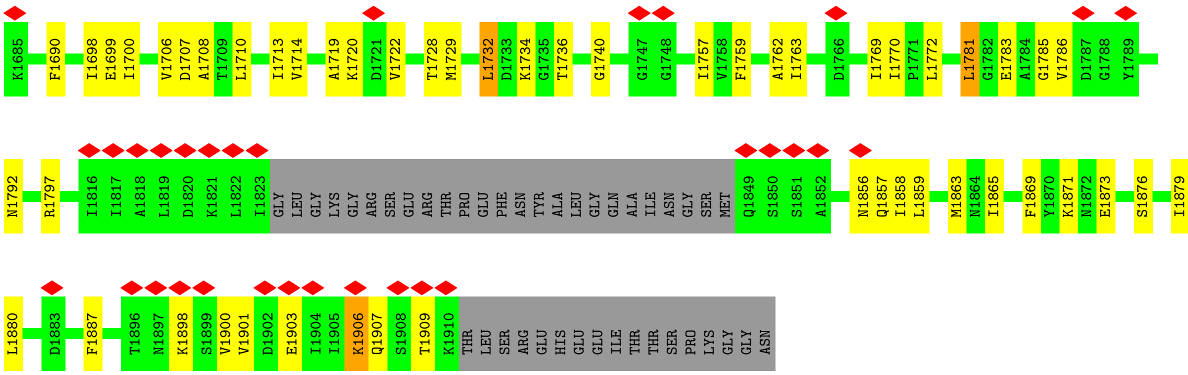


- Molecule 3: Cag pathogenicity island protein (Cag7)

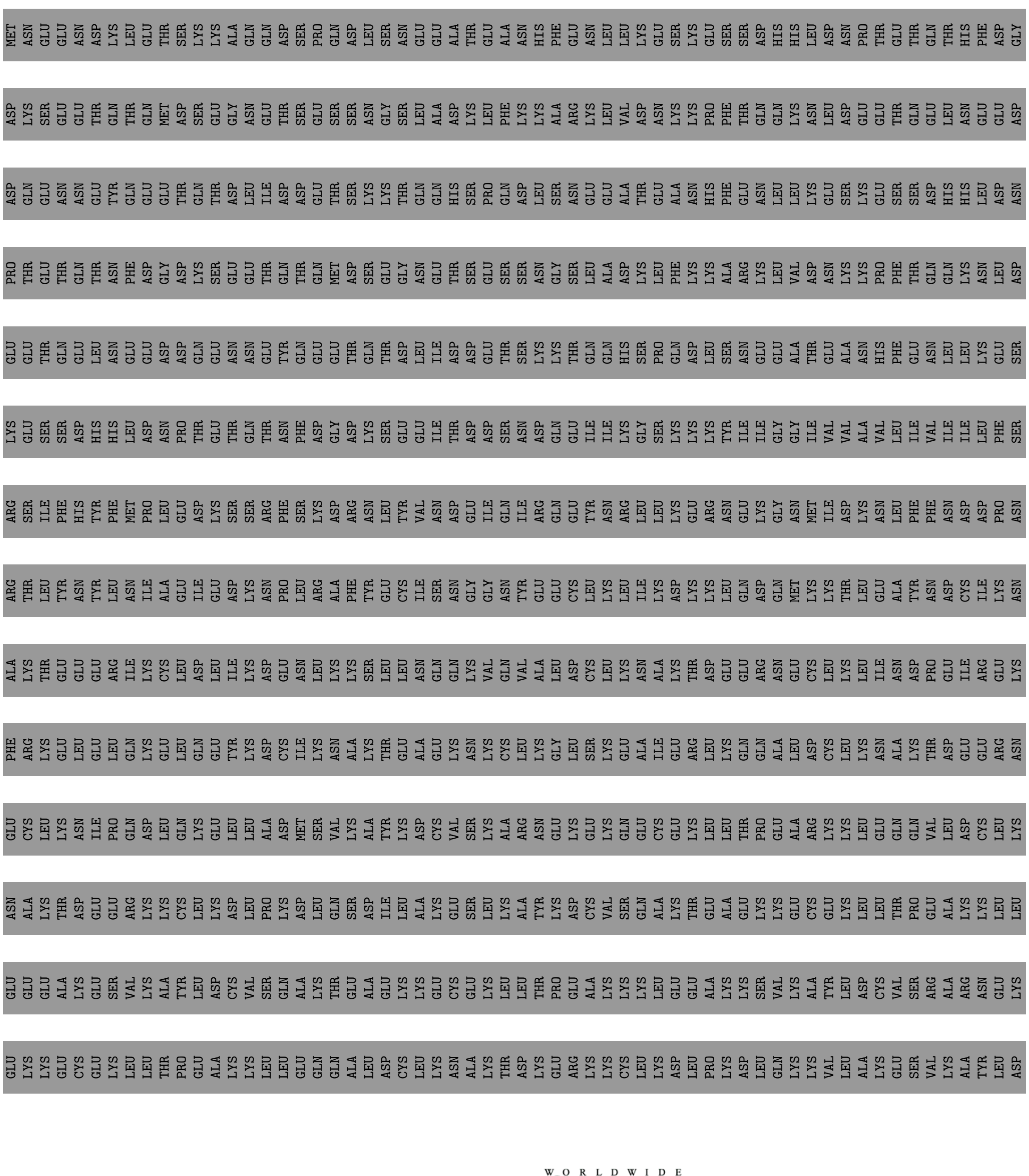
Chain MY:

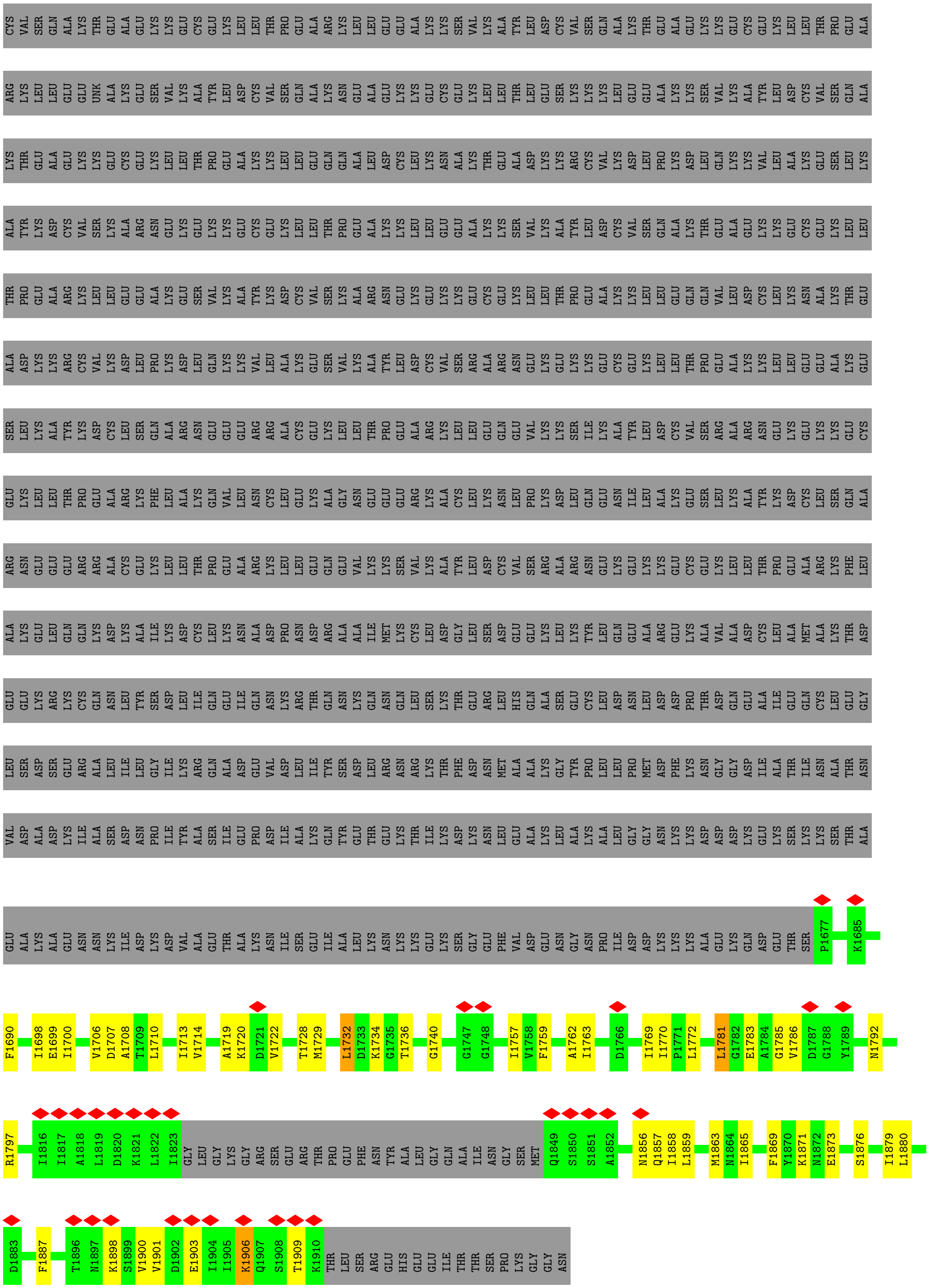


GLU	VAL	LEU	GLU	ALA	ARG	GLU	SER	ALA	THR	ALA	LYS	ARG	CYS	GLU	GLU	ASN	GLU	PHE
ALA	ASP	SER	GLU	LYS	ASN	LYS	LEU	ASP	PRO	TYR	THR	LYS	VAL	LYS	GLU	ALA	CYS	ARG
LYS	ALA	ASP	LYS	GLU	GLU	LEU	LYS	LYS	GLU	LYS	GLU	LEU	SER	LYS	GLU	LYS	LEU	LYS
ALA	ASP	SER	ARG	LEU	GLU	LEU	ALA	LYS	ALA	ASP	ALA	LEU	GLN	GLU	ALA	THR	LYS	GLU
GLU	LYS	GLU	LYS	GLN	GLU	THR	TYR	ARG	ARG	CYS	GLU	GLU	ALA	CYS	LYS	ASP	ASN	LEU
ASN	ILE	ARG	CYS	GLN	ARG	PRO	LYS	CYS	LYS	VAL	LYS	GLU	LYS	GLU	GLU	GLU	ILE	GLU
ASN	ALA	ALA	GLN	LYS	ARG	GLU	ASP	VAL	LEU	SER	LYS	UNK	THR	LYS	SER	GLU	PRO	LEU
LYS	SER	LEU	ASN	ASP	ALA	ALA	CYS	LYS	LEU	LYS	GLU	ALA	GLU	LEU	VAL	ARG	GLN	GLN
ILE	ASP	ILE	LEU	LYS	CYS	ARG	LEU	ASP	GLU	ALA	CYS	LYS	ALA	LEU	LYS	LYS	ASP	LYS
ASP	ASN	LEU	TYR	ALA	GLU	LYS	SER	LEU	GLU	ARG	GLU	GLU	GLU	THR	ALA	LYS	LEU	GLU
LYS	PRO	GLY	SER	ILE	LYS	PHE	GLN	PRO	ALA	ASN	LYS	SER	LYS	PRO	TYR	CYS	GLN	LEU
ASP	ILE	ILE	ASP	LYS	LEU	LEU	ALA	LYS	LYS	GLU	LEU	VAL	LYS	GLU	LEU	LEU	LYS	GLN
VAL	TYR	LYS	LEU	ASP	LEU	ALA	ARG	ASP	GLU	LYS	LEU	LYS	GLU	ALA	ASP	LYS	GLU	GLU
ALA	ALA	ARG	ILE	CYS	THR	LYS	ASN	LEU	SER	GLU	THR	ALA	CYS	LYS	CYS	ASP	LEU	TYR
GLU	SER	GLN	GLN	LEU	PRO	GLN	GLU	GLN	VAL	LYS	PRO	TYR	GLU	LYS	VAL	LEU	LEU	LYS
THR	ILE	ALA	GLU	LYS	GLU	VAL	GLU	LYS	LYS	LYS	GLU	LEU	LYS	LEU	SER	PRO	ALA	ASP
ALA	GLU	ASP	ILE	ASN	ALA	LEU	GLU	LYS	ALA	GLU	ALA	ASP	LEU	LEU	GLN	LYS	ASP	CYS
LYS	PRO	GLU	GLN	ALA	ARG	ASN	ARG	VAL	TYR	CYS	LYS	CYS	LEU	GLU	ALA	ASP	MET	ILE
ASN	ASP	VAL	ASN	ASP	LYS	CYS	ARG	LEU	LYS	GLU	LYS	VAL	THR	GLN	LYS	LEU	SER	LYS
ILE	ILE	ASP	LYS	PRO	LEU	LEU	ALA	ALA	ASP	LYS	LEU	SER	PRO	GLN	THR	GLN	VAL	ASN
SER	ALA	LEU	ARG	ASN	LEU	GLU	CYS	LYS	CYS	LEU	LEU	GLN	GLU	ALA	GLU	SER	LYS	ALA
GLU	LYS	ILE	THR	ASP	GLU	LYS	GLU	GLU	VAL	LEU	GLU	ALA	ALA	LEU	ALA	ASP	ALA	LYS
ILE	GLN	TYR	GLN	ARG	GLN	ALA	LYS	SER	SER	THR	GLN	LYS	ARG	ASP	GLU	ILE	TYR	THR
ALA	TYR	SER	ASN	ALA	GLU	GLY	LEU	VAL	LYS	PRO	GLN	ASN	LYS	CYS	LYS	LEU	LYS	GLU
LEU	GLU	ASP	LYS	ALA	VAL	ASN	LEU	LYS	ALA	GLU	ALA	GLU	LEU	LEU	LYS	ALA	ASP	ALA
LYS	THR	LEU	GLN	ILE	LYS	GLU	THR	ALA	ARG	ALA	LEU	ALA	LEU	LYS	GLU	LYS	CYS	GLU
ASN	GLU	ARG	ASN	MET	LYS	GLU	PRO	TYR	ASN	LYS	ASP	GLU	GLU	ASN	CYS	GLU	VAL	LYS
LYS	LYS	ASN	GLN	LYS	SER	GLU	GLU	LEU	GLU	LYS	CYS	LYS	GLU	ALA	GLU	SER	SER	ASN
LYS	THR	ARG	LEU	CYS	VAL	ARG	ALA	ASP	LYS	LEU	LEU	LYS	ALA	LYS	LYS	LEU	LYS	LYS
GLU	ILE	LYS	SER	LEU	LYS	LYS	ARG	CYS	GLU	LEU	LYS	GLU	LYS	THR	LEU	LYS	ALA	CYS
LYS	LYS	THR	LYS	ASP	ALA	ALA	LYS	VAL	LYS	GLU	ASN	CYS	LYS	ASP	LEU	ALA	ARG	LEU
SER	ASP	PHE	THR	GLY	TYR	CYS	LEU	SER	LYS	GLU	ALA	GLU	SER	LYS	THR	TYR	ASN	LYS
GLY	LYS	ASP	GLU	LEU	LEU	LEU	LEU	ARG	GLU	ALA	LYS	LYS	VAL	GLU	PRO	LYS	GLU	GLY
GLU	ASN	ASN	ARG	SER	ASP	LYS	GLU	ALA	CYS	LYS	THR	LEU	LYS	ARG	GLU	ASP	LYS	LEU
PHE	LEU	MET	LEU	ASP	CYS	ASN	GLN	ARG	GLU	LYS	GLU	LEU	ALA	LYS	ALA	CYS	GLU	SER
VAL	GLU	ALA	HIS	GLU	VAL	LEU	GLU	ASN	LYS	SER	ALA	THR	TYR	LYS	LYS	VAL	LYS	LYS
ASP	ALA	ALA	GLN	GLU	SER	PRO	VAL	GLU	LEU	VAL	ASP	LEU	LEU	CYS	LYS	SER	GLN	GLU
GLU	LYS	LYS	ALA	LYS	ARG	LYS	LYS	LYS	LEU	LYS	LYS	GLU	ASP	LEU	LYS	GLN	GLU	ALA
ASN	LEU	GLY	SER	LEU	ALA	ASP	LYS	GLU	THR	ALA	LYS	SER	CYS	LYS	LEU	ALA	CYS	ILE
GLY	ALA	TYR	GLU	LYS	ARG	LEU	SER	LYS	PRO	TYR	ARG	LYS	VAL	ASP	GLU	LYS	GLU	GLU
ASN	LYS	PRO	CYS	TYR	ASN	GLN	ILE	LYS	GLU	LEU	CYS	LYS	SER	LEU	GLU	THR	LYS	ARG
PRO	ALA	LEU	LEU	LEU	GLU	GLU	LYS	GLU	ALA	ASP	VAL	LYS	GLN	PRO	ALA	GLU	LEU	LEU
ILE	LEU	LEU	ASP	GLN	LYS	ASN	ALA	CYS	LYS	CYS	LYS	LEU	ALA	LYS	LYS	ALA	LEU	LYS
ASP	GLY	PRO	ASN	GLU	GLU	ILE	TYR	GLU	LYS	VAL	ASP	GLU	LYS	ASP	LYS	GLU	THR	GLN
ASP	GLY	MET	LEU	ALA	LYS	LEU	LEU	LYS	LEU	SER	LEU	GLU	THR	LEU	SER	LYS	PRO	GLN
LYS	ASN	ASP	ASP	ARG	LYS	ALA	ASP	LEU	LEU	GLN	PRO	ALA	GLU	GLN	VAL	LYS	GLU	ALA
LYS	LYS	PHE	ASP	GLU	GLU	LYS	CYS	LEU	GLU	ALA	LYS	LYS	ALA	LYS	LYS	GLU	ALA	LEU
LYS	LYS	LYS	PRO	LYS	CYS	GLU	VAL	THR	GLN	LYS	ASP	LYS	GLU	LYS	ALA	CYS	ARG	ASP
ALA	ASP	ASN	THR	ALA	GLU	SER	SER	PRO	GLN	THR	LEU	SER	LYS	VAL	TYR	GLU	LYS	CYS
GLU	ASP	GLY	ASP	VAL	LYS	LEU	ARG	GLU	VAL	GLU	GLN	VAL	LYS	LEU	LEU	LYS	LYS	LEU
LYS	ASP	GLY	GLN	ALA	LEU	LYS	ALA	ALA	LEU	ALA	LYS	LYS	GLU	ALA	ASP	LEU	LEU	LYS
GLN	LYS	ASP	GLU	ASP	LEU	ALA	ARG	LYS	ASP	GLU	LYS	ALA	CYS	LYS	CYS	LEU	GLU	ASN
ASP	GLU	ILE	ALA	CYS	THR	TYR	ASN	LYS	CYS	LYS	VAL	TYR	GLU	GLU	VAL	THR	GLN	ALA
GLU	LYS	ALA	ILE	LEU	PRO	LYS	GLU	LEU	LEU	LYS	LEU	LEU	LYS	SER	SER	PRO	GLN	LYS
THR	SER	THR	GLU	ALA	GLU	ASP	LYS	LEU	LYS	GLU	ALA	ASP	LEU	VAL	ARG	GLU	VAL	THR
SER	LYS	ILE	GLN	MET	ALA	CYS	GLU	GLU	ASN	CYS	LYS	CYS	LEU	LYS	ALA	ALA	LEU	ASP
P1677	LYS	ASN	CYS	ALA	ARG	LEU	LYS	GLU	ALA	GLU	GLU	VAL	THR	ALA	ARG	LYS	ASP	GLU
	SER	ALA	LEU	LYS	LYS	SER	LYS	ALA	LYS	LYS	SER	SER	PRO	TYR	ASN	LYS	CYS	GLU
I1683	THR	THR	GLU	THR	PHE	GLN	GLU	LYS	THR	LEU	LEU	GLN	GLU	LEU	GLU	LEU	LEU	ARG
G1684	ALA	ASN	GLY	ASP	LEU	ALA	CYS	GLU	GLU	LEU	LYS	ALA	ALA	ASP	LYS	LEU	LYS	ASN



• Molecule 3: Cag pathogenicity island protein (Cag7)




CYS VAL SER GLN ALA LYS THR GLU ALA GLU LYS LYS GLU CYS GLU LYS LEU LEU THR PRO GLU ALA ARG LYS LEU LEU GLU GLU ALA LYS LYS SER VAL LYS ALA TYR LEU ASP CYS VAL SER GLN ALA LYS THR GLU ALA GLU LYS LYS GLU CYS GLU LYS LEU LEU THR PRO GLU ALA
ARG LYS LEU LEU GLU GLU UNK ALA LYS GLU SER VAL LYS ALA TYR LEU ASP CYS VAL SER GLN ALA LYS ASN GLU ALA GLU LYS LYS GLU CYS GLU LYS LEU LEU THR LEU GLU SER LYS LYS LEU GLU GLU ALA LYS LYS SER VAL LYS ALA TYR LEU ASP CYS VAL SER GLN ALA
LYS THR GLU ALA GLU LYS LYS GLU CYS GLU LYS LEU LEU THR PRO GLU ALA LYS LYS LEU LEU GLU GLN GLN ALA LEU ASP CYS LEU LYS ASN ALA LYS THR GLU ALA ASP LYS LYS ARG CYS VAL LYS ASP LEU PRO LYS ASP LEU GLN LYS LYS VAL LEU ALA LYS GLU SER LEU LYS
ALA TYR LYS ASP CYS VAL SER LYS ALA ARG ASN GLU LYS GLU LYS LYS GLU CYS GLU LYS LEU LEU THR PRO GLU ALA LYS LYS LEU LEU GLU GLU ALA LYS LYS SER VAL LYS ALA TYR LEU ASP CYS VAL SER GLN ALA LYS THR GLU ALA GLU LYS LYS GLU CYS GLU LYS LEU LEU
THR PRO GLU ALA ARG LYS LEU LEU GLU GLU ALA LYS GLU SER VAL LYS ALA TYR LYS ASP CYS VAL SER LYS ALA ARG ASN GLU LYS GLU LYS LYS GLU CYS GLU LYS LEU LEU THR PRO GLU ALA LYS LYS LEU LEU GLU GLN GLN VAL LEU ASP CYS LEU LYS ASN ALA LYS THR GLU
ALA ASP LYS LYS ARG CYS VAL LYS ASP LEU PRO LYS ASP LEU GLN LYS LYS VAL LEU ALA LYS GLU SER VAL LYS ALA TYR LEU ASP CYS VAL SER ARG ALA ARG ASN GLU LYS GLU LYS LYS GLU CYS GLU LYS LEU LEU THR PRO GLU ALA LYS LYS LEU LEU GLU GLU ALA LYS GLU
SER LEU LYS ALA TYR LYS ASP CYS LEU SER GLN ALA ARG ASN GLU GLU GLU ARG ARG ALA CYS GLU LYS LEU LEU THR PRO GLU ALA ARG LYS LEU LEU GLU GLN GLU VAL LYS LYS SER ILE LYS ALA TYR LEU ASP CYS VAL SER ARG ALA ARG ASN GLU LYS GLU LYS LYS GLU CYS
GLU LYS LEU LEU THR PRO GLU ALA ARG LYS PHE LEU ALA LYS GLN VAL LEU ASN CYS LEU GLU LYS ALA GLY ASN GLU GLU ARG LYS ALA CYS LEU LYS ASN LEU PRO LYS ASP LEU GLN GLU ASN ILE LEU ALA LYS GLU SER LEU LYS ALA TYR LYS ASP CYS LEU SER GLN ALA
ARG ASN GLU GLU GLU ARG ARG ALA CYS GLU LYS LEU LEU THR PRO GLU ALA ARG LYS LEU LEU GLU GLN GLU VAL LYS LYS SER VAL LYS ALA TYR LEU ASP CYS VAL SER ARG ALA ARG ASN GLU LYS GLU LYS GLU CYS GLU LYS LEU LEU THR PRO GLU ALA ARG LYS PHE LEU
ALA LYS GLU LEU GLN GLN LYS ASP LYS ALA ILE LYS ASP CYS LEU LYS ASN ALA ASP PRO ASN ASP ARG ALA ALA ILE MET LYS CYS LEU ASP GLY LEU SER ASP GLU GLU LYS LEU LYS TYR LEU GLN GLU ALA ARG GLU LYS ALA VAL ALA ASP CYS LEU ALA MET ALA LYS THR ASP
GLU GLU LYS ARG LYS CYS GLN ASN LEU TYR SER ASP LEU ILE GLN GLU ILE GLN ASN LYS ARG THR GLN ASN LYS GLN ASN GLN LEU SER LYS THR GLU ARG LEU HIS GLN ALA SER GLU CYS LEU ASP ASN LEU ASP ASP PRO THR ASP GLN GLU ALA ILE GLU GLN CYS LEU GLU GLY
LEU SER ASP SER GLU ARG ALA LEU ILE LEU GLY ILE LYS ARG GLN ALA ASP GLU VAL ASP LEU ILE TYR SER ASP LEU ARG ASN ARG LYS THR PHE ASP ASN MET ALA ALA LYS GLY TYR PRO LEU LEU PRO MET ASP PHE LYS ASN GLY GLY ASP ILE ALA THR ILE ASN ALA THR ASN
VAL ASP ALA ASP LYS ILE ALA SER ASP ASN PRO ILE TYR ALA SER ILE GLU PRO ASP ILE ALA LYS GLN TYR GLU THR GLU LYS THR ILE LYS ASP LYS ASN LEU GLU ALA LYS LEU ALA LYS ALA LEU GLY GLY ASN LYS LYS ASP ASP ASP LYS GLU LYS SER LYS LYS SER THR ALA
GLU ALA LYS ALA GLU ASN ASN LYS ILE ASP LYS ASP VAL ALA GLU THR ALA LYS ASN ILE SER GLU ILE ALA LEU LYS ASN LYS LYS GLU LYS SER GLY GLU PHE VAL ASP GLU ASN GLY ASN PRO ILE ASP ASP LYS LYS LYS ALA GLU LYS GLN ASP GLU THR SER P1677 K1685
F1690 I1698 E1699 I1700 V1706 D1707 A1708 T1709 L1710 I1713 V1714 A1719 K1720 D1721 V1722 T1728 M1729 L1732 D1733 K1734 G1735 T1736 G1740 G1747 G1748 I1757 V1758 F1759 A1762 I1763 D1766 I1769 I1770 P1771 L1772 L1781 G1782 E1783 A1784 G1785 V1786 D1787 G1788 Y1789 N1792
R1797 I1816 I1817 A1818 L1819 D1820 K1821 L1822 I1823 GLY LEU GLY LYS GLY ARG SER GLU ARG THR PRO GLU PHE ASN TYR ALA LEU GLY GLN ALA ILE ASN GLY SER MET Q1849 S1850 S1851 A1852 N1856 Q1857 I1858 L1859 M1863 N1864 I1865 F1869 Y1870 K1871 N1872 E1873 S1876 I1879 L1880
D1883 F1887 T1896 N1897 K1898 S1899 V1900 V1901 D1902 E1903 I1904 I1905 K1906 Q1907 S1908 T1909 K1910 THR LEU SER ARG GLU HIS GLU GLU ILE THR THR SER PRO LYS GLY GLY ASN

4 Experimental information (i)

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, C14	Depositor
Number of particles used	7337	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 ($e^-/Å^2$)	59.7	Depositor
Minimum defocus (nm)	Not provided	
Maximum defocus (nm)	Not provided	
Magnification	Not provided	
Image detector	GATAN K2 SUMMIT (4k x 4k)	Depositor
Maximum map value	0.049	Depositor
Minimum map value	-0.036	Depositor
Average map value	0.000	Depositor
Map value standard deviation	0.002	Depositor
Recommended contour level	0.015	Depositor
Map size (Å)	510.0, 510.0, 510.0	wwPDB
Map dimensions	510, 510, 510	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	1.0, 1.0, 1.0	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	AT	0.41	0/1525	0.71	4/2049 (0.2%)
1	BT	0.41	0/1525	0.71	4/2049 (0.2%)
1	CT	0.41	0/1525	0.71	4/2049 (0.2%)
1	DT	0.41	0/1525	0.71	4/2049 (0.2%)
1	ET	0.41	0/1525	0.71	4/2049 (0.2%)
1	FT	0.41	0/1525	0.71	4/2049 (0.2%)
1	GT	0.41	0/1525	0.71	4/2049 (0.2%)
1	HT	0.41	0/1525	0.71	4/2049 (0.2%)
1	IT	0.41	0/1525	0.71	4/2049 (0.2%)
1	JT	0.41	0/1525	0.71	4/2049 (0.2%)
1	KT	0.41	0/1525	0.71	4/2049 (0.2%)
1	LT	0.41	0/1525	0.71	4/2049 (0.2%)
1	MT	0.41	0/1525	0.71	4/2049 (0.2%)
1	NT	0.41	0/1525	0.71	4/2049 (0.2%)
2	AX	0.54	0/1301	0.69	0/1752
2	BX	0.54	0/1301	0.69	0/1752
2	CX	0.54	0/1301	0.69	0/1752
2	DX	0.54	0/1301	0.69	0/1752
2	EX	0.54	0/1301	0.69	0/1752
2	FX	0.54	0/1301	0.69	1/1752 (0.1%)
2	GX	0.54	0/1301	0.69	0/1752
2	HX	0.54	0/1301	0.69	0/1752
2	IX	0.54	0/1301	0.69	0/1752
2	JX	0.54	0/1301	0.69	0/1752
2	KX	0.54	0/1301	0.69	1/1752 (0.1%)
2	LX	0.54	0/1301	0.69	1/1752 (0.1%)
2	MX	0.54	0/1301	0.69	1/1752 (0.1%)
2	NX	0.54	0/1301	0.69	1/1752 (0.1%)
3	AY	0.55	0/1602	0.72	3/2170 (0.1%)
3	BY	0.55	0/1602	0.72	4/2170 (0.2%)
3	CY	0.55	0/1602	0.72	4/2170 (0.2%)
3	DY	0.55	0/1602	0.72	3/2170 (0.1%)
3	EY	0.55	0/1602	0.72	3/2170 (0.1%)
3	FY	0.55	0/1602	0.72	4/2170 (0.2%)

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	#|Z| >5	RMSZ	#|Z| >5
3	GY	0.55	0/1602	0.72	4/2170 (0.2%)
3	HY	0.55	0/1602	0.72	4/2170 (0.2%)
3	IY	0.55	0/1602	0.72	4/2170 (0.2%)
3	JY	0.55	0/1602	0.72	4/2170 (0.2%)
3	KY	0.55	0/1602	0.72	4/2170 (0.2%)
3	LY	0.55	0/1602	0.72	3/2170 (0.1%)
3	MY	0.55	0/1602	0.72	4/2170 (0.2%)
3	NY	0.55	0/1602	0.72	4/2170 (0.2%)
All	All	0.50	0/61992	0.71	113/83594 (0.1%)

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
2	AX	0	1
2	BX	0	1
2	CX	0	1
2	DX	0	1
2	EX	0	1
2	FX	0	1
2	GX	0	1
2	HX	0	1
2	IX	0	1
2	JX	0	1
2	KX	0	1
2	LX	0	1
2	MX	0	1
2	NX	0	1
All	All	0	14

There are no bond length outliers.

The worst 5 of 113 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	MY	1772	LEU	CA-CB-CG	7.79	133.22	115.30
3	CY	1772	LEU	CA-CB-CG	7.79	133.21	115.30
3	KY	1772	LEU	CA-CB-CG	7.78	133.20	115.30
3	AY	1772	LEU	CA-CB-CG	7.78	133.19	115.30
3	HY	1772	LEU	CA-CB-CG	7.78	133.20	115.30

There are no chirality outliers.

5 of 14 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	AX	479	ILE	Peptide
2	BX	479	ILE	Peptide
2	CX	479	ILE	Peptide
2	DX	479	ILE	Peptide
2	EX	479	ILE	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	AT	1502	0	1525	21	0
1	BT	1502	0	1525	23	0
1	CT	1502	0	1525	23	0
1	DT	1502	0	1525	23	0
1	ET	1502	0	1525	20	0
1	FT	1502	0	1525	27	0
1	GT	1502	0	1525	21	0
1	HT	1502	0	1525	23	0
1	IT	1502	0	1525	23	0
1	JT	1502	0	1525	23	0
1	KT	1502	0	1525	26	0
1	LT	1502	0	1525	26	0
1	MT	1502	0	1525	24	0
1	NT	1502	0	1525	20	0
2	AX	1275	0	1309	26	0
2	BX	1275	0	1309	28	0
2	CX	1275	0	1309	26	0
2	DX	1275	0	1309	24	0
2	EX	1275	0	1309	27	0
2	FX	1275	0	1309	29	0
2	GX	1275	0	1309	31	0
2	HX	1275	0	1309	29	0
2	IX	1275	0	1309	25	0
2	JX	1275	0	1309	27	0
2	KX	1275	0	1309	21	0
2	LX	1275	0	1309	26	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	MX	1275	0	1309	27	0
2	NX	1275	0	1309	28	0
3	AY	1578	0	1637	45	0
3	BY	1578	0	1637	45	0
3	CY	1578	0	1637	48	0
3	DY	1578	0	1637	43	0
3	EY	1578	0	1637	42	0
3	FY	1578	0	1637	45	0
3	GY	1578	0	1637	44	0
3	HY	1578	0	1637	44	0
3	IY	1578	0	1637	46	0
3	JY	1578	0	1637	42	0
3	KY	1578	0	1637	42	0
3	LY	1578	0	1637	41	0
3	MY	1578	0	1637	43	0
3	NY	1578	0	1637	42	0
All	All	60970	0	62594	984	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 8.

The worst 5 of 984 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:BX:444:PRO:O	2:CX:410:ARG:NH2	2.22	0.73
2:HX:444:PRO:O	2:IX:410:ARG:NH2	2.21	0.73
3:AY:1722:VAL:HB	3:AY:1732:LEU:HB2	1.72	0.72
3:DY:1722:VAL:HB	3:DY:1732:LEU:HB2	1.72	0.72
3:FY:1722:VAL:HB	3:FY:1732:LEU:HB2	1.72	0.72

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	AT	171/278 (62%)	160 (94%)	11 (6%)	0	100	100
1	BT	171/278 (62%)	160 (94%)	11 (6%)	0	100	100
1	CT	171/278 (62%)	160 (94%)	11 (6%)	0	100	100
1	DT	171/278 (62%)	160 (94%)	11 (6%)	0	100	100
1	ET	171/278 (62%)	160 (94%)	11 (6%)	0	100	100
1	FT	171/278 (62%)	160 (94%)	11 (6%)	0	100	100
1	GT	171/278 (62%)	160 (94%)	11 (6%)	0	100	100
1	HT	171/278 (62%)	160 (94%)	11 (6%)	0	100	100
1	IT	171/278 (62%)	160 (94%)	11 (6%)	0	100	100
1	JT	171/278 (62%)	160 (94%)	11 (6%)	0	100	100
1	KT	171/278 (62%)	160 (94%)	11 (6%)	0	100	100
1	LT	171/278 (62%)	160 (94%)	11 (6%)	0	100	100
1	MT	171/278 (62%)	160 (94%)	11 (6%)	0	100	100
1	NT	171/278 (62%)	160 (94%)	11 (6%)	0	100	100
2	AX	153/522 (29%)	145 (95%)	8 (5%)	0	100	100
2	BX	153/522 (29%)	145 (95%)	8 (5%)	0	100	100
2	CX	153/522 (29%)	145 (95%)	8 (5%)	0	100	100
2	DX	153/522 (29%)	145 (95%)	8 (5%)	0	100	100
2	EX	153/522 (29%)	145 (95%)	8 (5%)	0	100	100
2	FX	153/522 (29%)	145 (95%)	8 (5%)	0	100	100
2	GX	153/522 (29%)	145 (95%)	8 (5%)	0	100	100
2	HX	153/522 (29%)	145 (95%)	8 (5%)	0	100	100
2	IX	153/522 (29%)	145 (95%)	8 (5%)	0	100	100
2	JX	153/522 (29%)	145 (95%)	8 (5%)	0	100	100
2	KX	153/522 (29%)	145 (95%)	8 (5%)	0	100	100
2	LX	153/522 (29%)	145 (95%)	8 (5%)	0	100	100
2	MX	153/522 (29%)	145 (95%)	8 (5%)	0	100	100
2	NX	153/522 (29%)	145 (95%)	8 (5%)	0	100	100
3	AY	205/1927 (11%)	195 (95%)	10 (5%)	0	100	100
3	BY	205/1927 (11%)	195 (95%)	10 (5%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
3	CY	205/1927 (11%)	195 (95%)	10 (5%)	0	100	100
3	DY	205/1927 (11%)	195 (95%)	10 (5%)	0	100	100
3	EY	205/1927 (11%)	195 (95%)	10 (5%)	0	100	100
3	FY	205/1927 (11%)	195 (95%)	10 (5%)	0	100	100
3	GY	205/1927 (11%)	195 (95%)	10 (5%)	0	100	100
3	HY	205/1927 (11%)	195 (95%)	10 (5%)	0	100	100
3	IY	205/1927 (11%)	195 (95%)	10 (5%)	0	100	100
3	JY	205/1927 (11%)	195 (95%)	10 (5%)	0	100	100
3	KY	205/1927 (11%)	195 (95%)	10 (5%)	0	100	100
3	LY	205/1927 (11%)	195 (95%)	10 (5%)	0	100	100
3	MY	205/1927 (11%)	195 (95%)	10 (5%)	0	100	100
3	NY	205/1927 (11%)	195 (95%)	10 (5%)	0	100	100
All	All	7406/38178 (19%)	7000 (94%)	406 (6%)	0	100	100

There are no Ramachandran outliers to report.

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	AT	169/254 (66%)	167 (99%)	2 (1%)	71	88
1	BT	169/254 (66%)	167 (99%)	2 (1%)	71	88
1	CT	169/254 (66%)	167 (99%)	2 (1%)	71	88
1	DT	169/254 (66%)	167 (99%)	2 (1%)	71	88
1	ET	169/254 (66%)	167 (99%)	2 (1%)	71	88
1	FT	169/254 (66%)	167 (99%)	2 (1%)	71	88
1	GT	169/254 (66%)	167 (99%)	2 (1%)	71	88
1	HT	169/254 (66%)	167 (99%)	2 (1%)	71	88
1	IT	169/254 (66%)	167 (99%)	2 (1%)	71	88

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	JT	169/254 (66%)	167 (99%)	2 (1%)	71	88
1	KT	169/254 (66%)	167 (99%)	2 (1%)	71	88
1	LT	169/254 (66%)	167 (99%)	2 (1%)	71	88
1	MT	169/254 (66%)	167 (99%)	2 (1%)	71	88
1	NT	169/254 (66%)	167 (99%)	2 (1%)	71	88
2	AX	139/470 (30%)	139 (100%)	0	100	100
2	BX	139/470 (30%)	139 (100%)	0	100	100
2	CX	139/470 (30%)	139 (100%)	0	100	100
2	DX	139/470 (30%)	139 (100%)	0	100	100
2	EX	139/470 (30%)	139 (100%)	0	100	100
2	FX	139/470 (30%)	139 (100%)	0	100	100
2	GX	139/470 (30%)	139 (100%)	0	100	100
2	HX	139/470 (30%)	139 (100%)	0	100	100
2	IX	139/470 (30%)	139 (100%)	0	100	100
2	JX	139/470 (30%)	139 (100%)	0	100	100
2	KX	139/470 (30%)	139 (100%)	0	100	100
2	LX	139/470 (30%)	139 (100%)	0	100	100
2	MX	139/470 (30%)	139 (100%)	0	100	100
2	NX	139/470 (30%)	139 (100%)	0	100	100
3	AY	177/1724 (10%)	176 (99%)	1 (1%)	86	94
3	BY	177/1724 (10%)	176 (99%)	1 (1%)	86	94
3	CY	177/1724 (10%)	176 (99%)	1 (1%)	86	94
3	DY	177/1724 (10%)	176 (99%)	1 (1%)	86	94
3	EY	177/1724 (10%)	176 (99%)	1 (1%)	86	94
3	FY	177/1724 (10%)	176 (99%)	1 (1%)	86	94
3	GY	177/1724 (10%)	176 (99%)	1 (1%)	86	94
3	HY	177/1724 (10%)	176 (99%)	1 (1%)	86	94
3	IY	177/1724 (10%)	176 (99%)	1 (1%)	86	94
3	JY	177/1724 (10%)	176 (99%)	1 (1%)	86	94
3	KY	177/1724 (10%)	176 (99%)	1 (1%)	86	94
3	LY	177/1724 (10%)	176 (99%)	1 (1%)	86	94

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
3	MY	177/1724 (10%)	176 (99%)	1 (1%)	86	94
3	NY	177/1724 (10%)	176 (99%)	1 (1%)	86	94
All	All	6790/34272 (20%)	6748 (99%)	42 (1%)	86	94

5 of 42 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	JT	214	LYS
3	LY	1906	LYS
3	JY	1906	LYS
3	KY	1906	LYS
1	MT	214	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

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.

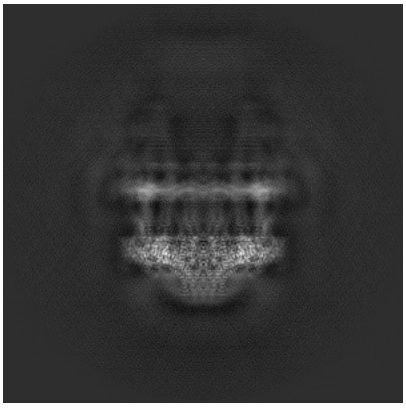
6 Map visualisation (i)

This section contains visualisations of the EMDB entry EMD-22076. These allow visual inspection of the internal detail of the map and identification of artifacts.

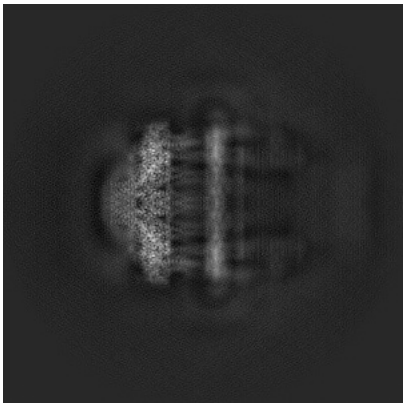
No raw map or half-maps were deposited for this entry and therefore no images, graphs, etc. pertaining to the raw map can be shown.

6.1 Orthogonal projections (i)

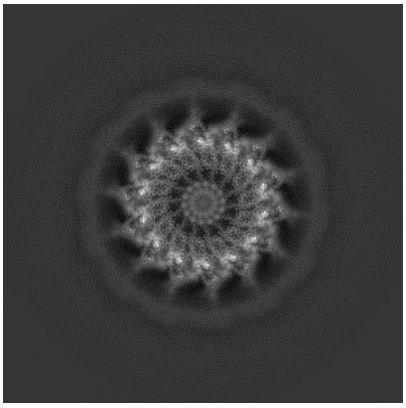
6.1.1 Primary map



X



Y

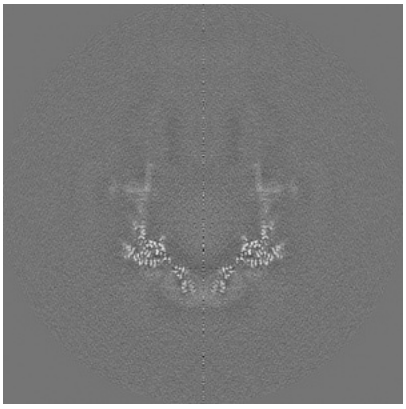


Z

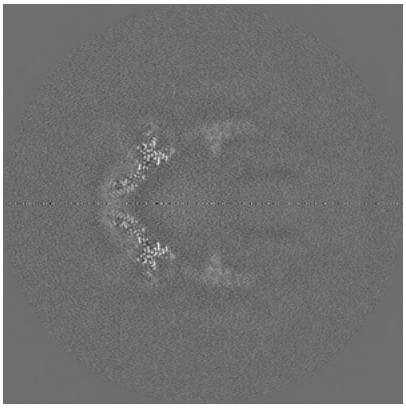
The images above show the map projected in three orthogonal directions.

6.2 Central slices (i)

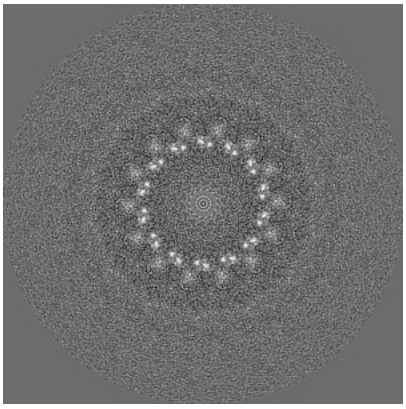
6.2.1 Primary map



X Index: 255



Y Index: 255

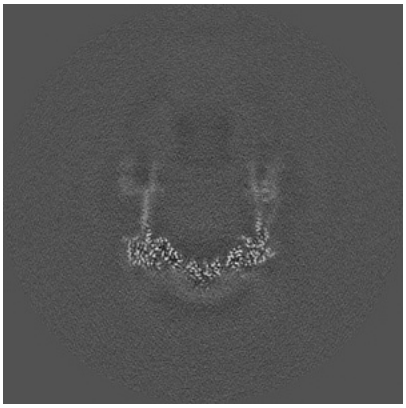


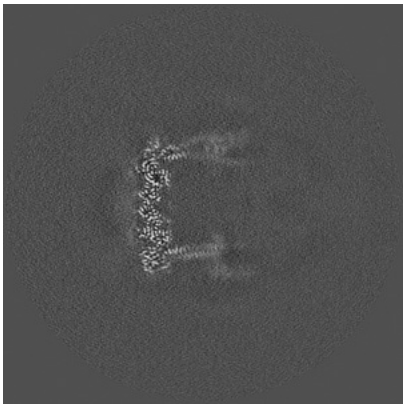
Z Index: 255

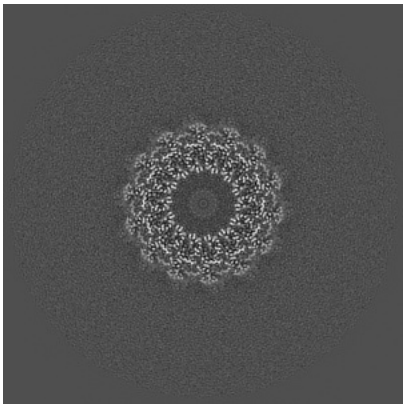
The images above show central slices of the map in three orthogonal directions.

6.3 Largest variance slices (i)

6.3.1 Primary map


X Index: 224

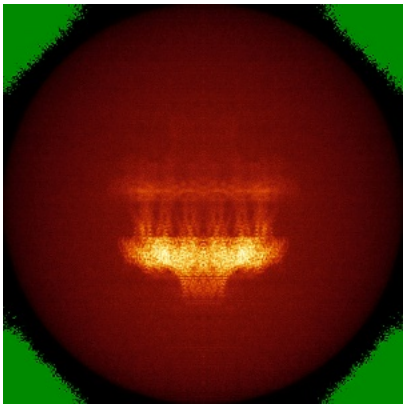

Y Index: 303

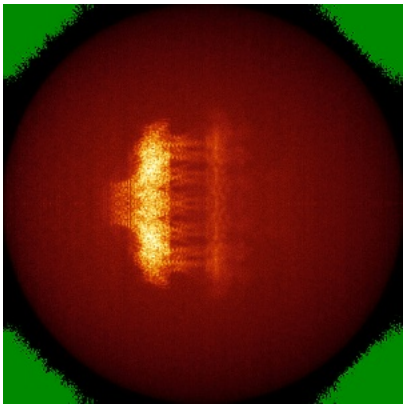

Z Index: 188

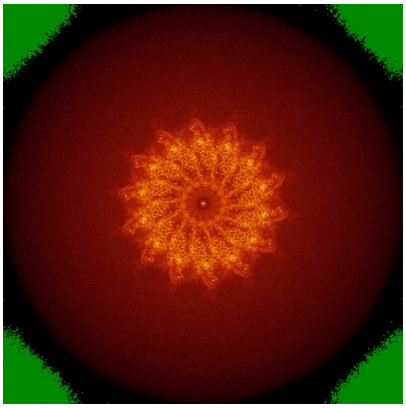
The images above show the largest variance slices of the map in three orthogonal directions.

6.4 Orthogonal standard-deviation projections (False-color) (i)

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 0.015. These images, in conjunction with the slice images, may facilitate assessment of whether an appropriate contour level has been provided.

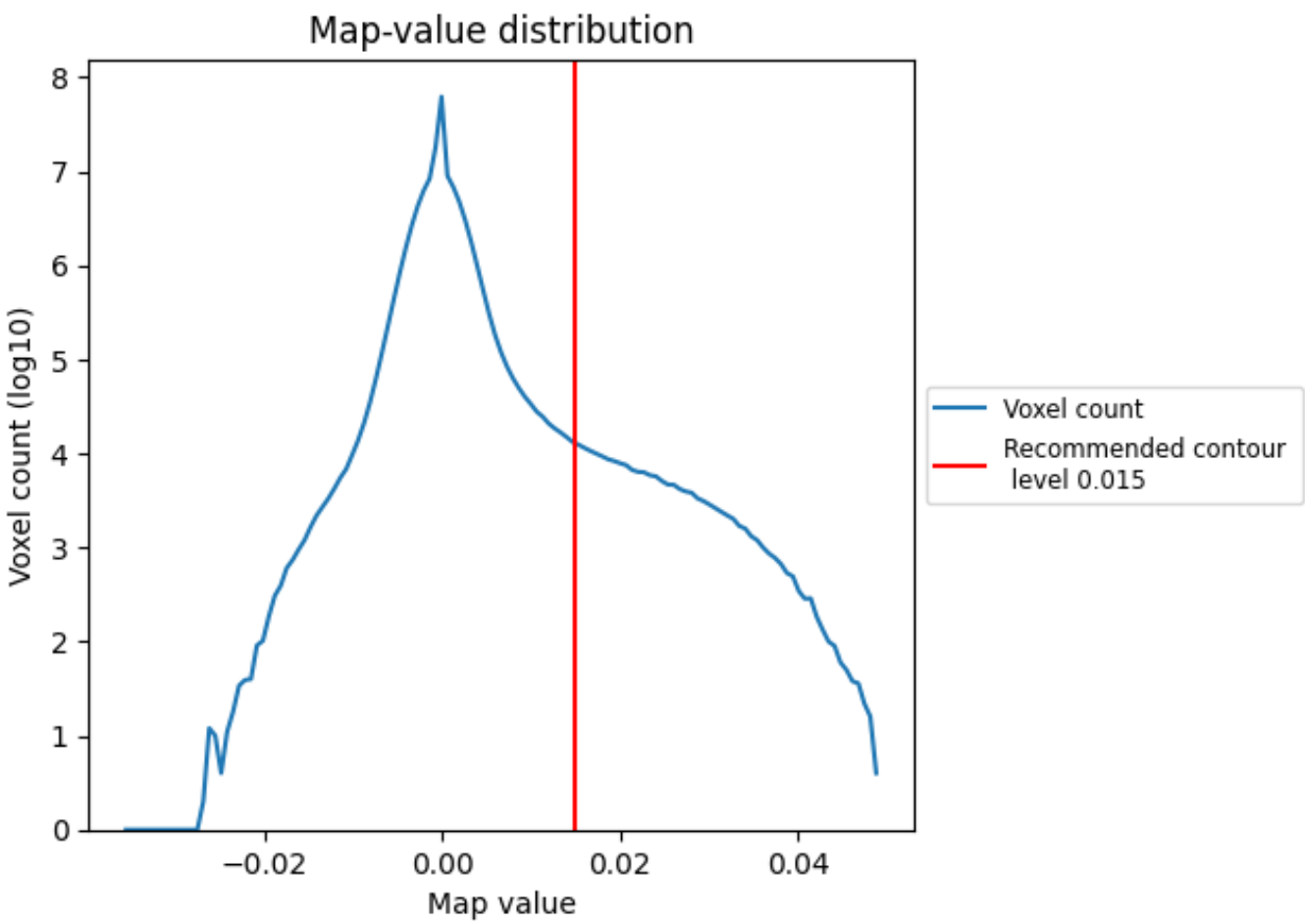
6.6 Mask visualisation (i)

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

7 Map analysis (i)

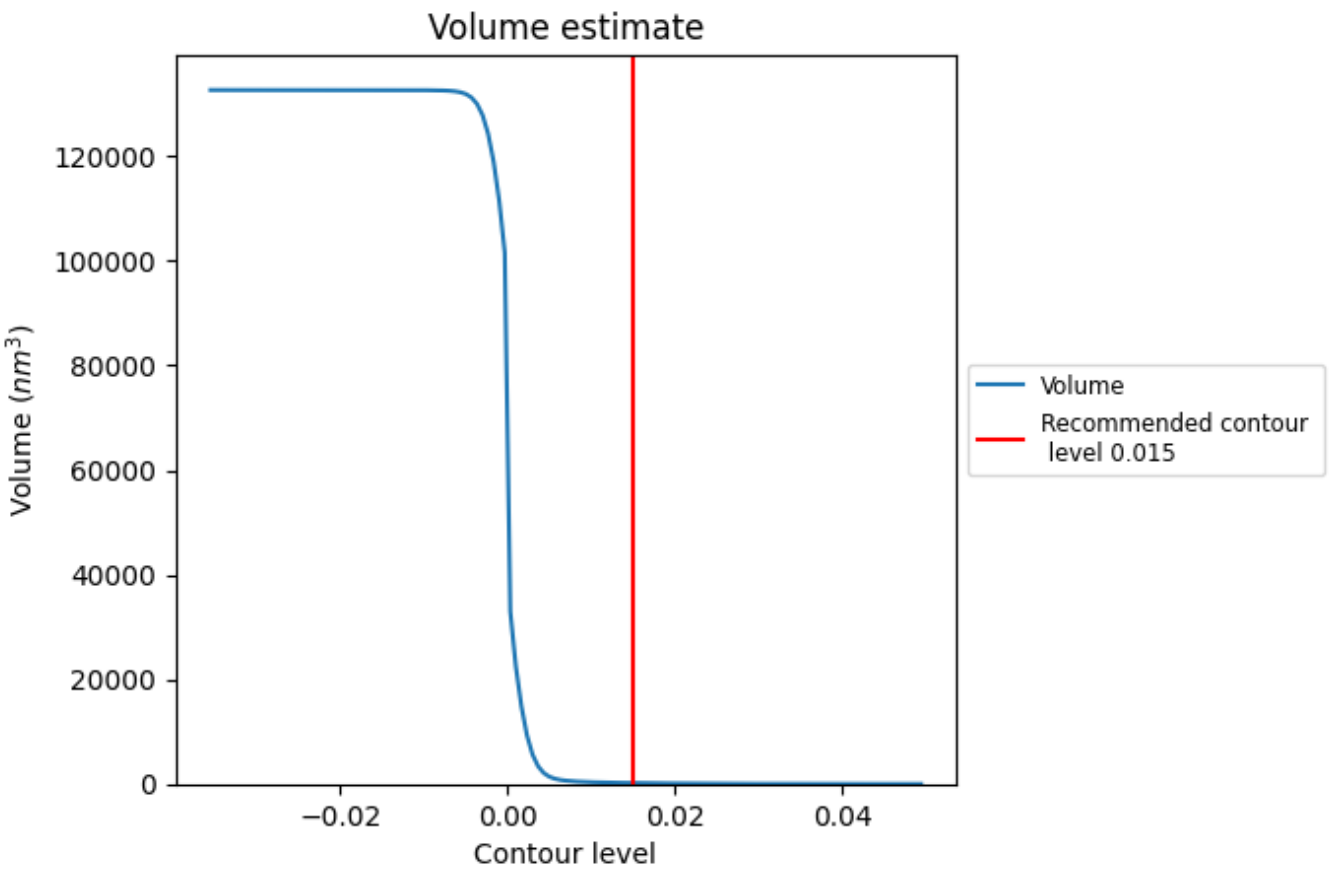
This section contains the results of statistical analysis of the map.

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.

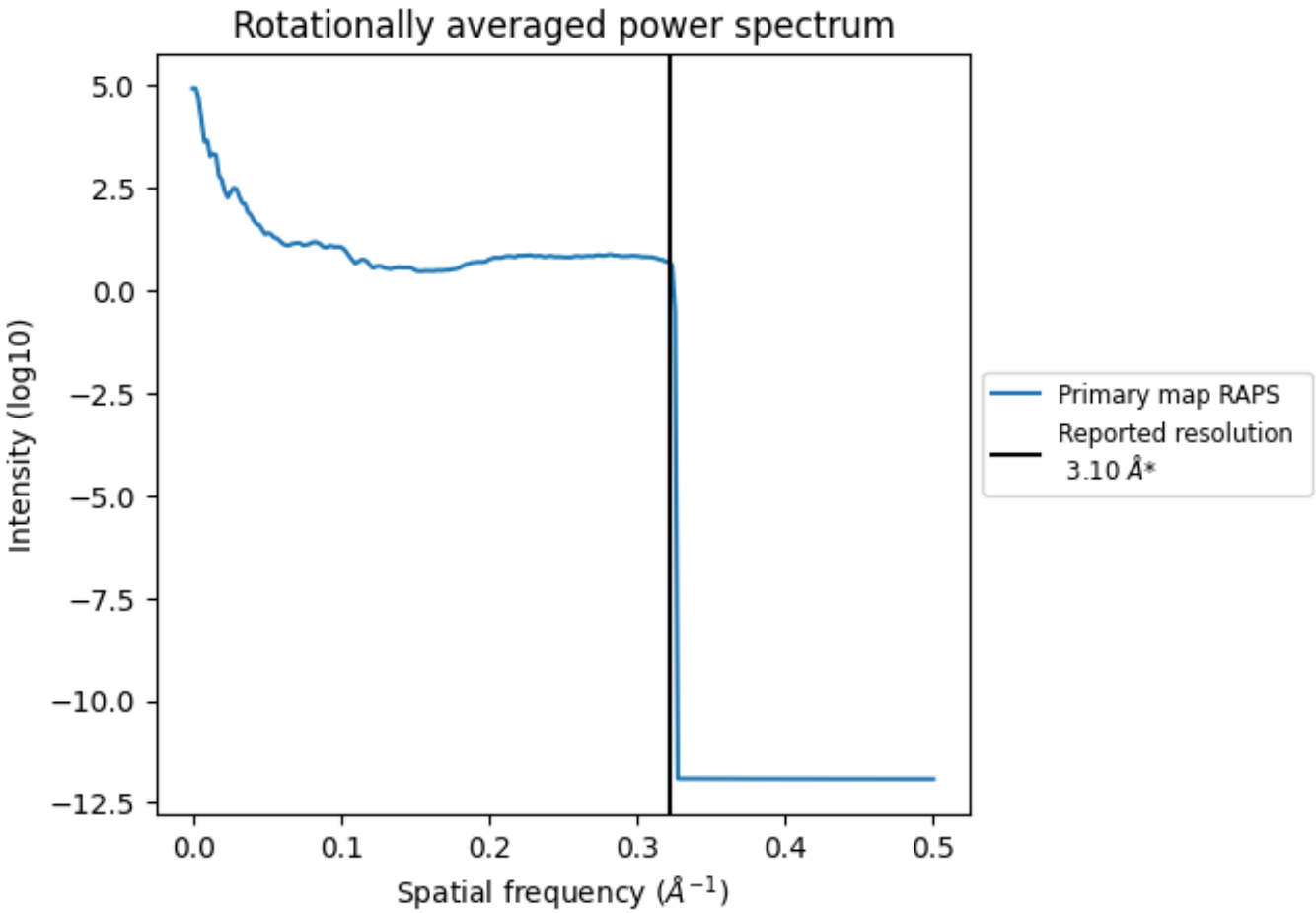
7.2 Volume estimate (i)



The volume at the recommended contour level is 180 nm^3; this corresponds to an approximate mass of 163 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)

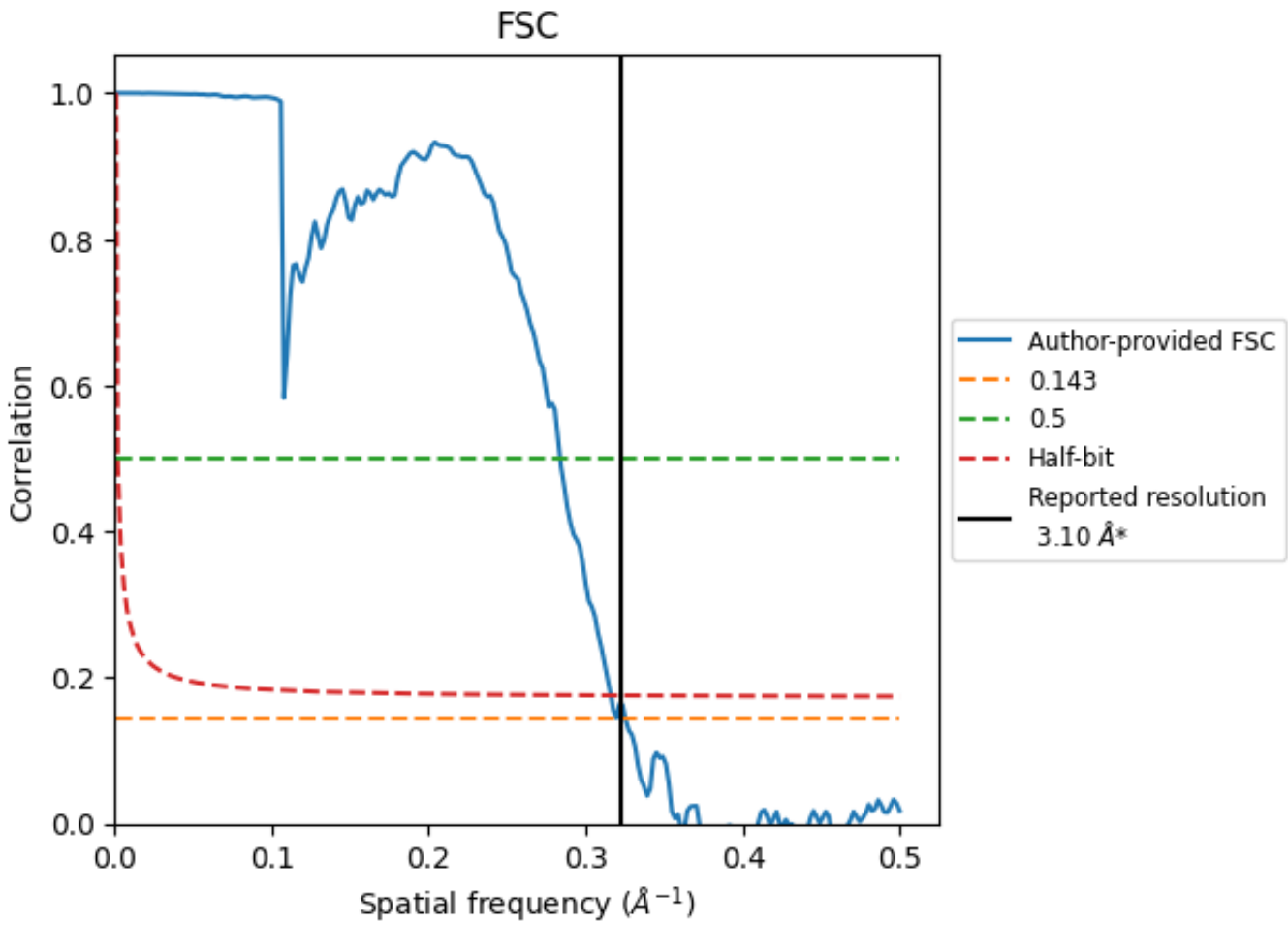


*Reported resolution corresponds to spatial frequency of 0.323 Å^{-1}

8 Fourier-Shell correlation (i)

Fourier-Shell Correlation (FSC) is the most commonly used method to estimate the resolution of single-particle and subtomogram-averaged maps. The shape of the curve depends on the imposed symmetry, mask and whether or not the two 3D reconstructions used were processed from a common reference. The reported resolution is shown as a black line. A curve is displayed for the half-bit criterion in addition to lines showing the 0.143 gold standard cut-off and 0.5 cut-off.

8.1 FSC (i)



*Reported resolution corresponds to spatial frequency of 0.323 Å^{-1}

8.2 Resolution estimates (i)

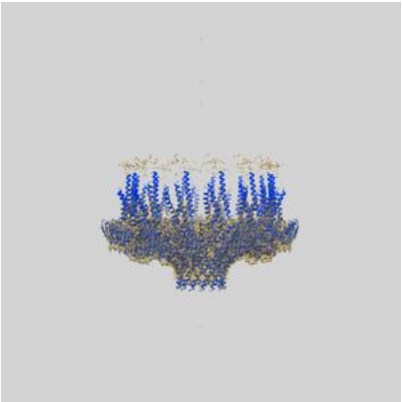
Resolution estimate (Å)	Estimation criterion (FSC cut-off)		
	0.143	0.5	Half-bit
Reported by author	3.10	-	-
Author-provided FSC curve	3.07	3.52	3.16
Unmasked-calculated*	-	-	-

*Resolution estimate based on FSC curve calculated by comparison of deposited half-maps.

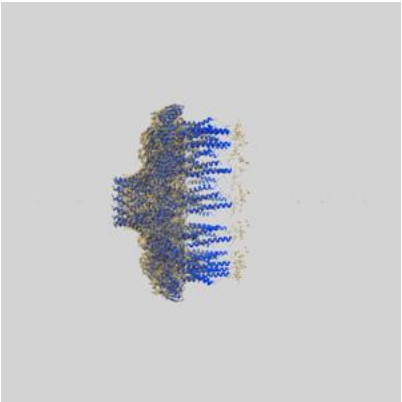
9 Map-model fit (i)

This section contains information regarding the fit between EMDB map EMD-22076 and PDB model 6X6K. Per-residue inclusion information can be found in section 3 on page 8.

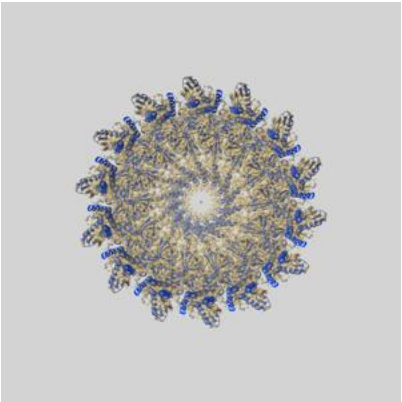
9.1 Map-model overlay (i)



X



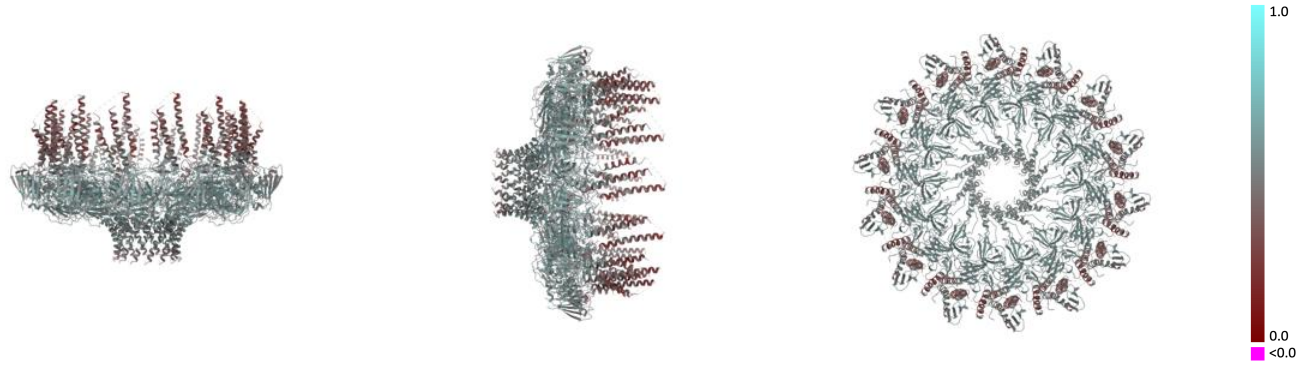
Y



Z

The images above show the 3D surface view of the map at the recommended contour level 0.015 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.

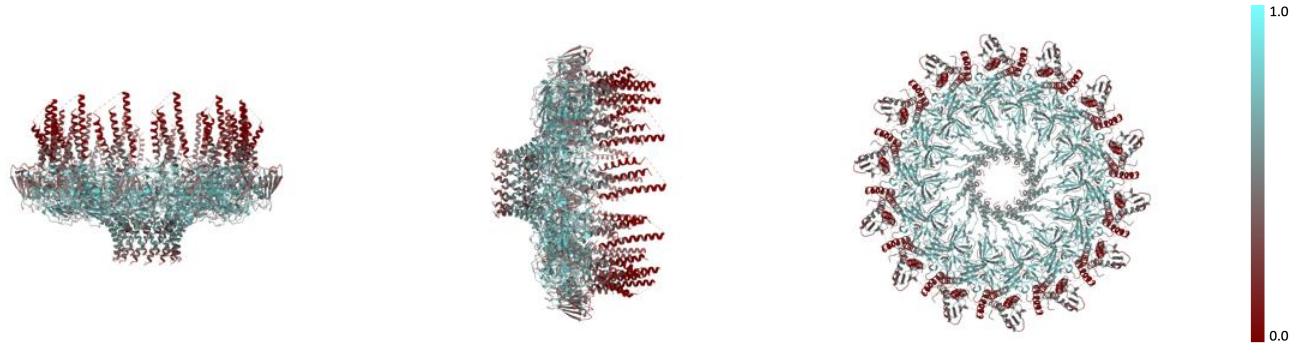
9.2 Q-score mapped to coordinate model (i)



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.

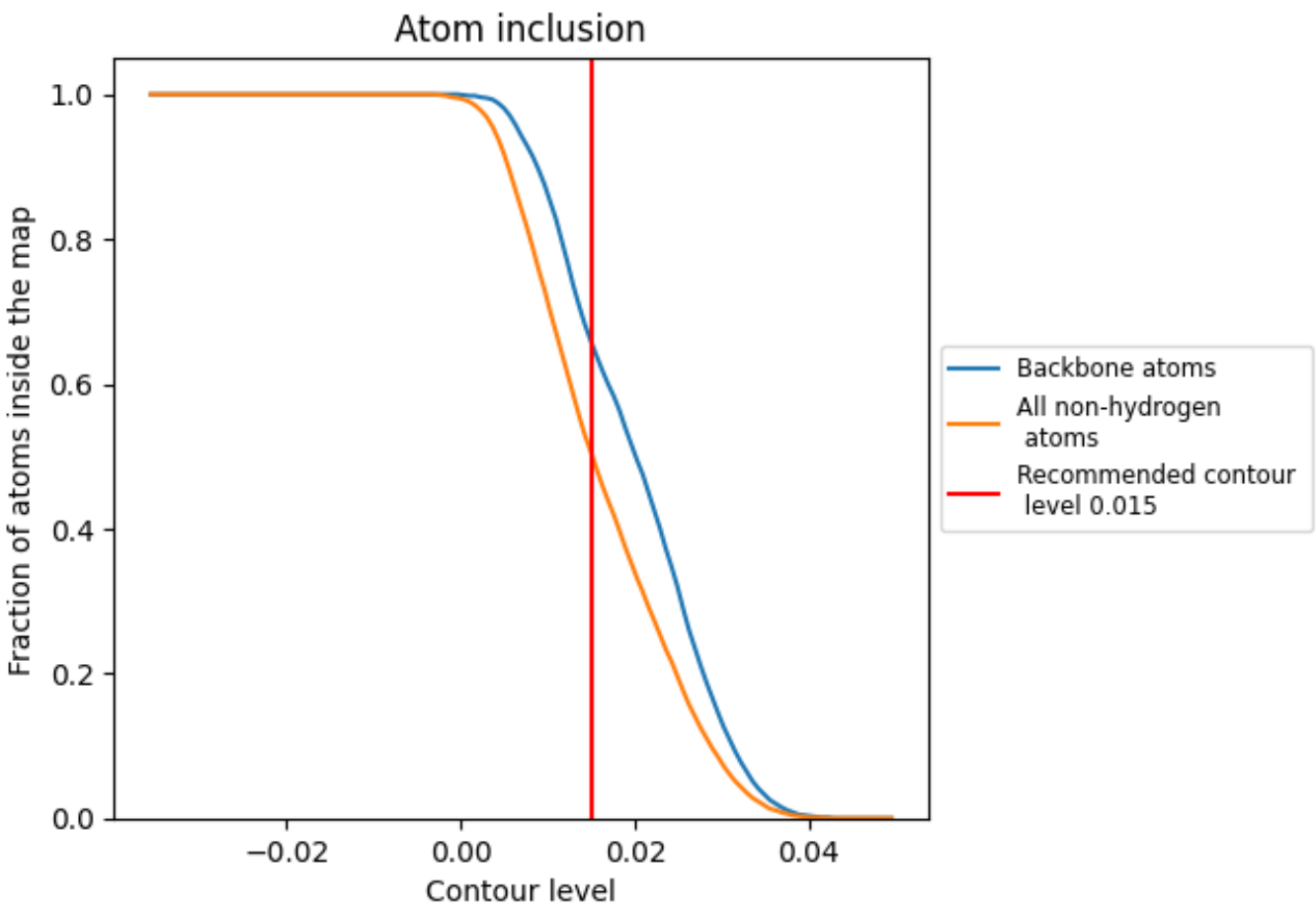
9.3 Atom inclusion mapped to coordinate model (i)



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 (0.015).

9.4 Atom inclusion (i)



At the recommended contour level, 66% of all backbone atoms, 50% of all non-hydrogen atoms, are inside the map.

9.5 Map-model fit summary (i)

The table lists the average atom inclusion at the recommended contour level (0.015) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	0.5050	0.5040
AT	0.3380	0.4550
AX	0.5770	0.5170
AY	0.6070	0.5400
BT	0.3390	0.4580
BX	0.5750	0.5160
BY	0.6070	0.5410
CT	0.3350	0.4550
CX	0.5740	0.5180
CY	0.6040	0.5410
DT	0.3390	0.4550
DX	0.5720	0.5190
DY	0.6020	0.5420
ET	0.3400	0.4550
EX	0.5790	0.5170
EY	0.6040	0.5430
FT	0.3360	0.4540
FX	0.5780	0.5190
FY	0.6050	0.5400
GT	0.3330	0.4530
GX	0.5800	0.5190
GY	0.6090	0.5410
HT	0.3370	0.4560
HX	0.5770	0.5170
HY	0.6060	0.5400
IT	0.3370	0.4560
IX	0.5710	0.5190
IY	0.6070	0.5400
JT	0.3350	0.4520
JX	0.5740	0.5170
JY	0.6050	0.5390
KT	0.3360	0.4550
KX	0.5730	0.5170
KY	0.6020	0.5400
LT	0.3370	0.4560

Continued on next page...

Continued from previous page...

Chain	Atom inclusion	Q-score
LX	0.5790	0.5170
LY	0.6050	0.5390
MT	0.3350	0.4570
MX	0.5770	0.5190
MY	0.6050	0.5390
NT	0.3390	0.4560
NX	0.5790	0.5170
NY	0.6090	0.5400