



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 6, 2026 – 03:00 AM UTC

PDB ID : 4V4K / pdb_00004v4k
Title : Bacteriophage P22 Portal Protein bound to middle Tail Factor GP4. This file contain the second biological assembly
Authors : Olia, A.S.; Cingolani, G.
Deposited on : 2010-04-19
Resolution : 3.25 Å(reported)

This is a wwPDB X-ray Structure Validation Summary Report for a publicly released PDB entry.

We welcome your comments at validation@mail.wwpdb.org

A user guide is available at

<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

The types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

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

MolProbity	:	4-5-2 with Phenix2.0
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	2.0
EDS	:	3.0
Percentile statistics	:	20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4	:	9.0.010 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.49

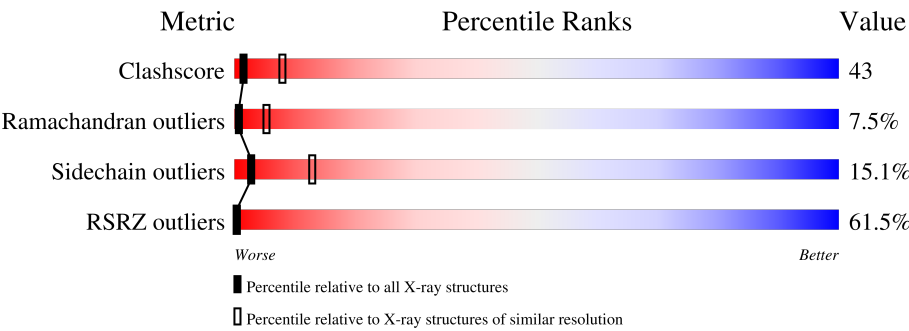
1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.25 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
Clashscore	190562	1660 (3.30-3.22)
Ramachandran outliers	187476	1630 (3.30-3.22)
Sidechain outliers	187428	1629 (3.30-3.22)
RSRZ outliers	180081	1605 (3.30-3.22)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments of the lower bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions $\leq 5\%$. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	602	62% 37% 44% 12% • 5%
1	B	602	61% 37% 43% 12% • 5%
1	C	602	62% 37% 44% 12% • 5%
1	D	602	58% 37% 43% 13% • 5%
1	E	602	59% 37% 44% 12% • 5%
1	F	602	59% 37% 43% 13% • 5%
1	G	602	60% 37% 44% 13% • 5%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	H	602	
1	I	602	
1	J	602	
1	K	602	
1	L	602	
1	M	602	
1	N	602	
1	O	602	
1	P	602	
1	Q	602	
1	R	602	
1	S	602	
1	T	602	
1	U	602	
1	V	602	
1	W	602	
1	X	602	
2	Y	166	
2	Z	166	
2	a	166	
2	b	166	
2	c	166	
2	d	166	
2	e	166	
2	f	166	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	g	166	
2	h	166	
2	i	166	
2	j	166	
2	k	166	
2	l	166	
2	m	166	
2	n	166	
2	o	166	
2	p	166	
2	q	166	
2	r	166	
2	s	166	
2	t	166	
2	u	166	
2	v	166	

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 135120 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called PORTAL PROTEIN.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	M	569	Total	C	N	O	S	Se	0	0	0
			4564	2871	786	887	4	16			
1	N	569	Total	C	N	O	S	Se	0	0	0
			4564	2871	786	887	4	16			
1	O	569	Total	C	N	O	S	Se	0	0	0
			4564	2871	786	887	4	16			
1	P	569	Total	C	N	O	S	Se	0	0	0
			4564	2871	786	887	4	16			
1	Q	569	Total	C	N	O	S	Se	0	0	0
			4564	2871	786	887	4	16			
1	R	569	Total	C	N	O	S	Se	0	0	0
			4564	2871	786	887	4	16			
1	S	569	Total	C	N	O	S	Se	0	0	0
			4564	2871	786	887	4	16			
1	T	569	Total	C	N	O	S	Se	0	0	0
			4564	2871	786	887	4	16			
1	U	569	Total	C	N	O	S	Se	0	0	0
			4564	2871	786	887	4	16			
1	V	569	Total	C	N	O	S	Se	0	0	0
			4564	2871	786	887	4	16			
1	W	569	Total	C	N	O	S	Se	0	0	0
			4564	2871	786	887	4	16			
1	X	569	Total	C	N	O	S	Se	0	0	0
			4564	2871	786	887	4	16			
1	A	569	Total	C	N	O	S	Se	0	0	0
			4553	2865	783	885	4	16			
1	B	569	Total	C	N	O	S	Se	0	0	0
			4553	2865	783	885	4	16			
1	C	569	Total	C	N	O	S	Se	0	0	0
			4553	2865	783	885	4	16			
1	D	569	Total	C	N	O	S	Se	0	0	0
			4553	2865	783	885	4	16			

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	E	569	Total	C	N	O	S	Se	0	0	0
			4553	2865	783	885	4	16			
1	F	569	Total	C	N	O	S	Se	0	0	0
			4553	2865	783	885	4	16			
1	G	569	Total	C	N	O	S	Se	0	0	0
			4553	2865	783	885	4	16			
1	H	569	Total	C	N	O	S	Se	0	0	0
			4553	2865	783	885	4	16			
1	I	569	Total	C	N	O	S	Se	0	0	0
			4553	2865	783	885	4	16			
1	J	569	Total	C	N	O	S	Se	0	0	0
			4553	2865	783	885	4	16			
1	K	569	Total	C	N	O	S	Se	0	0	0
			4553	2865	783	885	4	16			
1	L	569	Total	C	N	O	S	Se	0	0	0
			4553	2865	783	885	4	16			

- Molecule 2 is a protein called PACKAGED DNA STABILIZATION PROTEIN GP4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	k	145	Total	C	N	O	S	0	0	0
			1052	654	182	211	5			
2	l	145	Total	C	N	O	S	0	0	0
			1052	654	182	211	5			
2	m	145	Total	C	N	O	S	0	0	0
			1052	654	182	211	5			
2	n	145	Total	C	N	O	S	0	0	0
			1052	654	182	211	5			
2	o	145	Total	C	N	O	S	0	0	0
			1052	654	182	211	5			
2	p	145	Total	C	N	O	S	0	0	0
			1052	654	182	211	5			
2	q	145	Total	C	N	O	S	0	0	0
			1052	654	182	211	5			
2	r	145	Total	C	N	O	S	0	0	0
			1052	654	182	211	5			
2	s	145	Total	C	N	O	S	0	0	0
			1052	654	182	211	5			
2	t	145	Total	C	N	O	S	0	0	0
			1052	654	182	211	5			
2	u	145	Total	C	N	O	S	0	0	0
			1052	654	182	211	5			

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	v	145	Total	C	N	O	S	0	0	0
			1052	654	182	211	5			
2	Y	146	Total	C	N	O	S	0	0	0
			1048	652	179	212	5			
2	Z	146	Total	C	N	O	S	0	0	0
			1048	652	179	212	5			
2	a	146	Total	C	N	O	S	0	0	0
			1048	652	179	212	5			
2	b	146	Total	C	N	O	S	0	0	0
			1048	652	179	212	5			
2	c	146	Total	C	N	O	S	0	0	0
			1048	652	179	212	5			
2	d	146	Total	C	N	O	S	0	0	0
			1048	652	179	212	5			
2	e	146	Total	C	N	O	S	0	0	0
			1048	652	179	212	5			
2	f	146	Total	C	N	O	S	0	0	0
			1048	652	179	212	5			
2	g	146	Total	C	N	O	S	0	0	0
			1048	652	179	212	5			
2	h	146	Total	C	N	O	S	0	0	0
			1048	652	179	212	5			
2	i	146	Total	C	N	O	S	0	0	0
			1048	652	179	212	5			
2	j	146	Total	C	N	O	S	0	0	0
			1048	652	179	212	5			

There are 24 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
k	141	PRO	ALA	engineered mutation	UNP P26746
l	141	PRO	ALA	engineered mutation	UNP P26746
m	141	PRO	ALA	engineered mutation	UNP P26746
n	141	PRO	ALA	engineered mutation	UNP P26746
o	141	PRO	ALA	engineered mutation	UNP P26746
p	141	PRO	ALA	engineered mutation	UNP P26746
q	141	PRO	ALA	engineered mutation	UNP P26746
r	141	PRO	ALA	engineered mutation	UNP P26746
s	141	PRO	ALA	engineered mutation	UNP P26746
t	141	PRO	ALA	engineered mutation	UNP P26746
u	141	PRO	ALA	engineered mutation	UNP P26746
v	141	PRO	ALA	engineered mutation	UNP P26746
Y	150	PRO	ALA	engineered mutation	UNP P26746

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
Z	150	PRO	ALA	engineered mutation	UNP P26746
a	150	PRO	ALA	engineered mutation	UNP P26746
b	150	PRO	ALA	engineered mutation	UNP P26746
c	150	PRO	ALA	engineered mutation	UNP P26746
d	150	PRO	ALA	engineered mutation	UNP P26746
e	150	PRO	ALA	engineered mutation	UNP P26746
f	150	PRO	ALA	engineered mutation	UNP P26746
g	150	PRO	ALA	engineered mutation	UNP P26746
h	150	PRO	ALA	engineered mutation	UNP P26746
i	150	PRO	ALA	engineered mutation	UNP P26746
j	150	PRO	ALA	engineered mutation	UNP P26746

- Molecule 3 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	M	21	Total O 21 21	0	0
3	N	21	Total O 21 21	0	0
3	O	21	Total O 21 21	0	0
3	P	21	Total O 21 21	0	0
3	Q	21	Total O 21 21	0	0
3	R	21	Total O 21 21	0	0
3	S	21	Total O 21 21	0	0
3	T	21	Total O 21 21	0	0
3	U	21	Total O 21 21	0	0
3	V	21	Total O 21 21	0	0
3	W	21	Total O 21 21	0	0
3	X	21	Total O 21 21	0	0
3	A	22	Total O 22 22	0	0
3	B	22	Total O 22 22	0	0

Continued on next page...

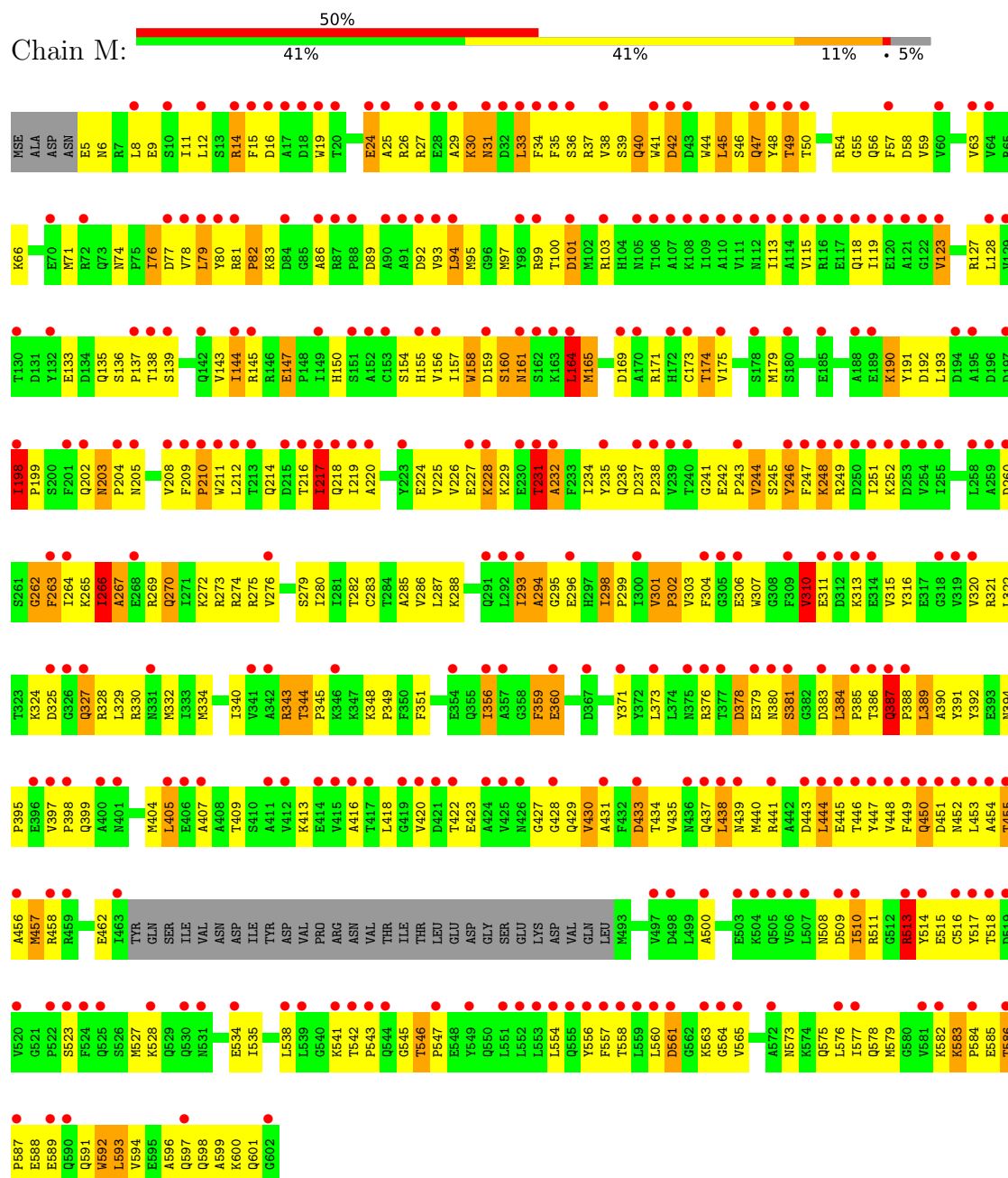
Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	C	22	Total 22	O 22	0	0
3	D	22	Total 22	O 22	0	0
3	E	22	Total 22	O 22	0	0
3	F	22	Total 22	O 22	0	0
3	G	22	Total 22	O 22	0	0
3	H	22	Total 22	O 22	0	0
3	I	22	Total 22	O 22	0	0
3	J	22	Total 22	O 22	0	0
3	K	22	Total 22	O 22	0	0
3	L	22	Total 22	O 22	0	0

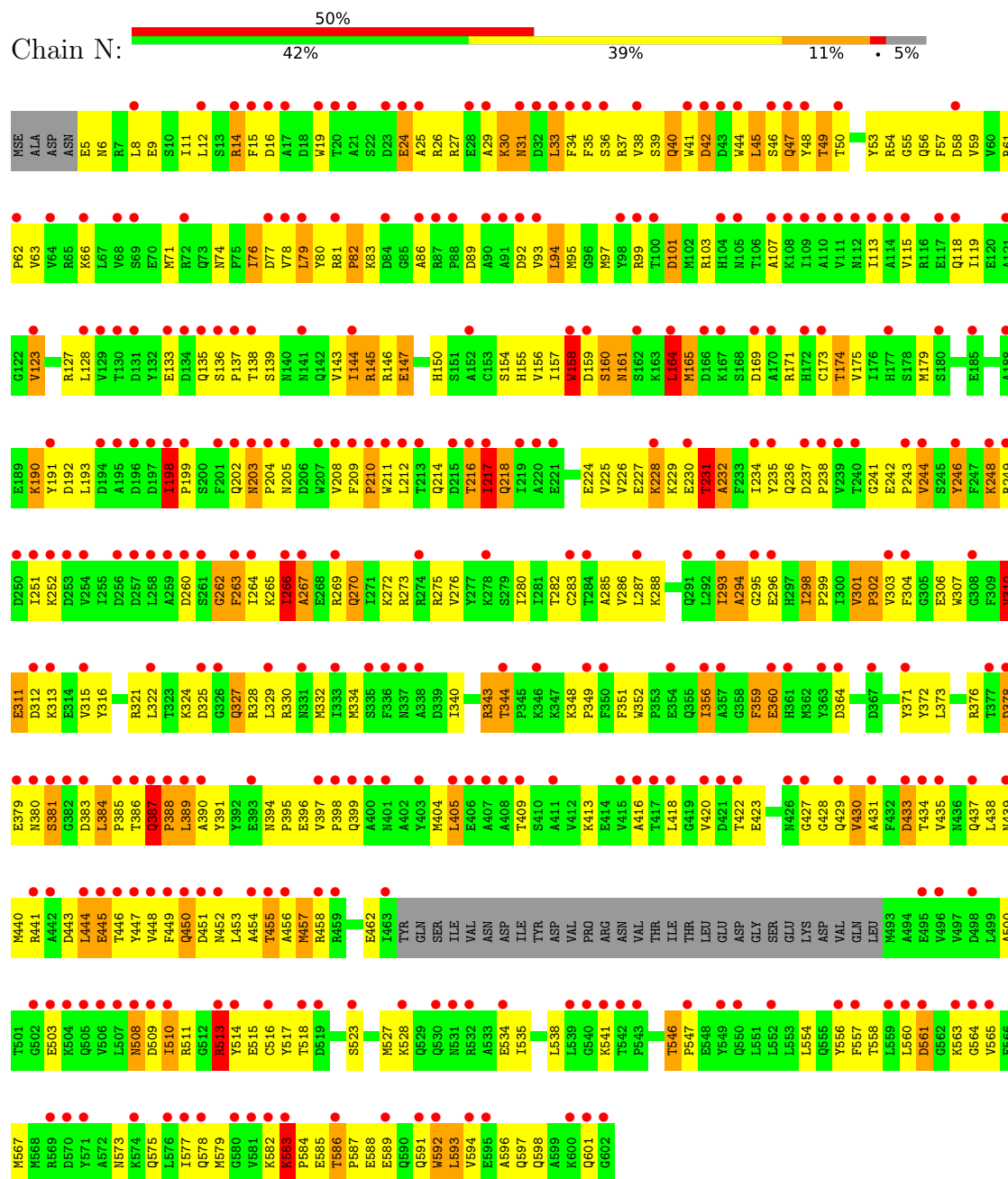
3 Residue-property plots

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density ($RSRZ > 2$). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

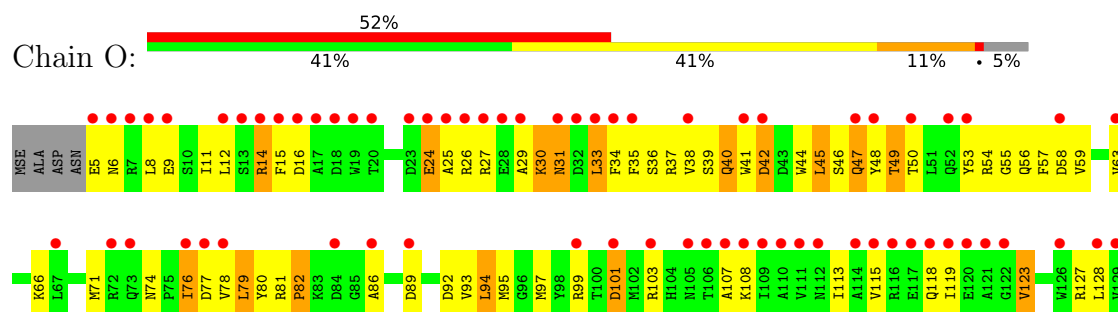
• Molecule 1: PORTAL PROTEIN

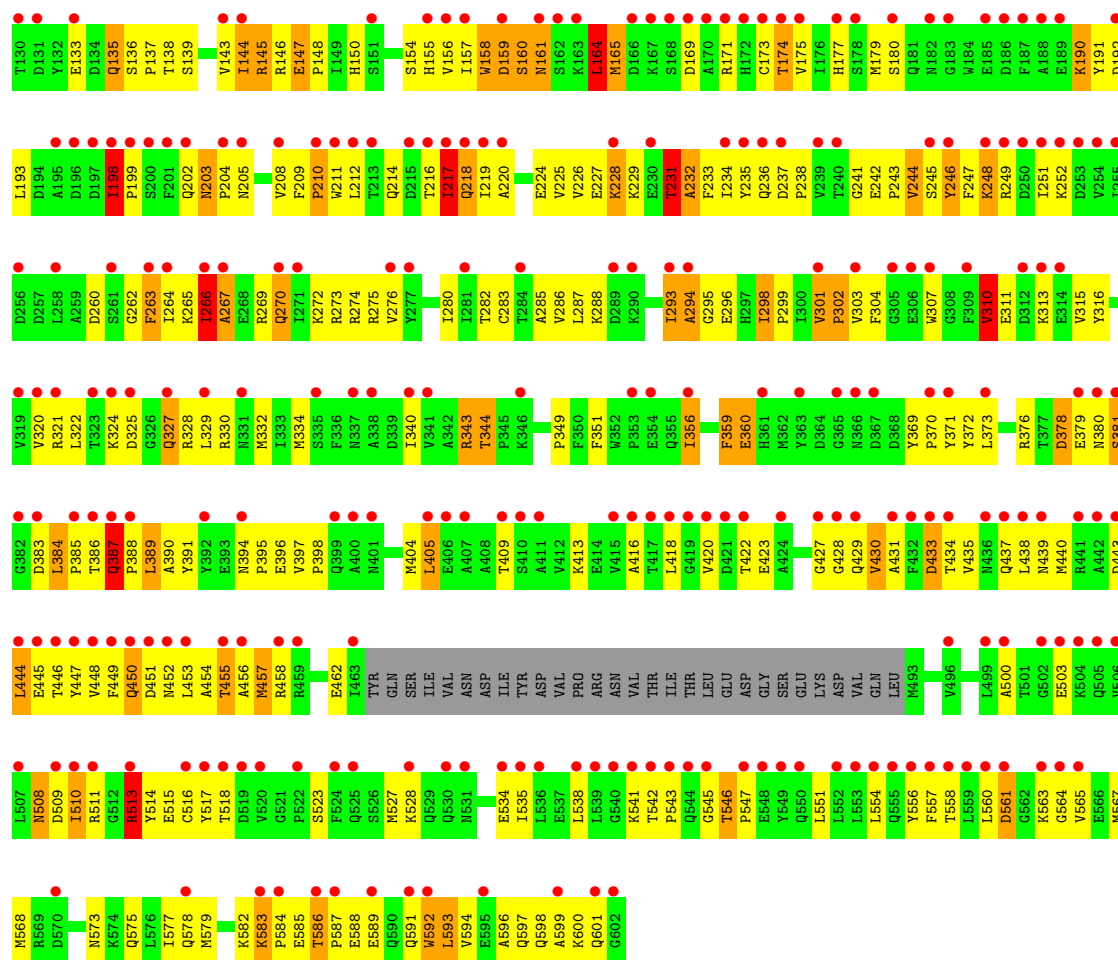


• Molecule 1: PORTAL PROTEIN

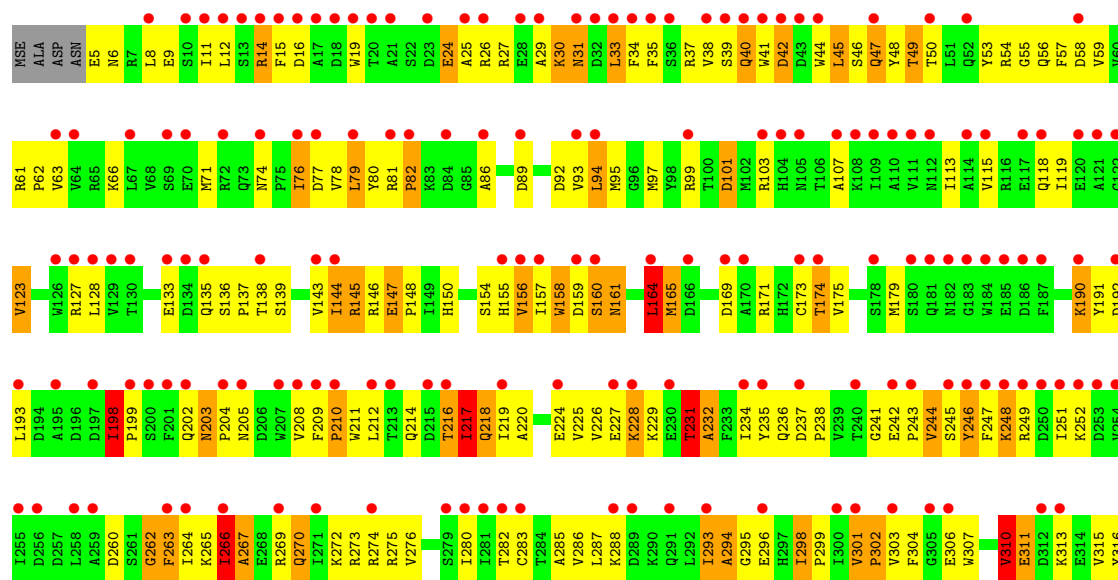


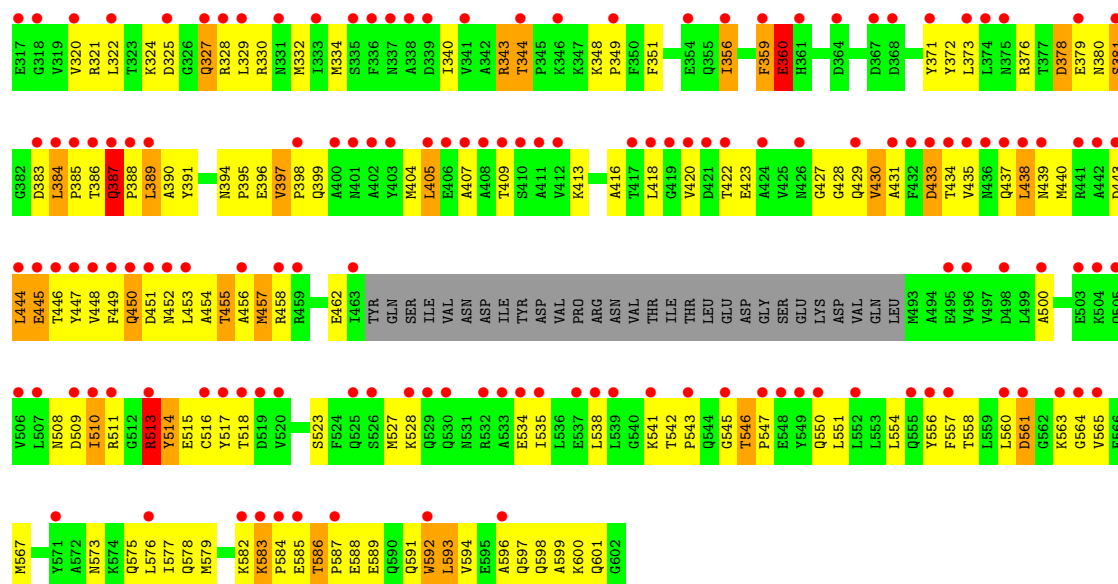
• Molecule 1: PORTAL PROTEIN



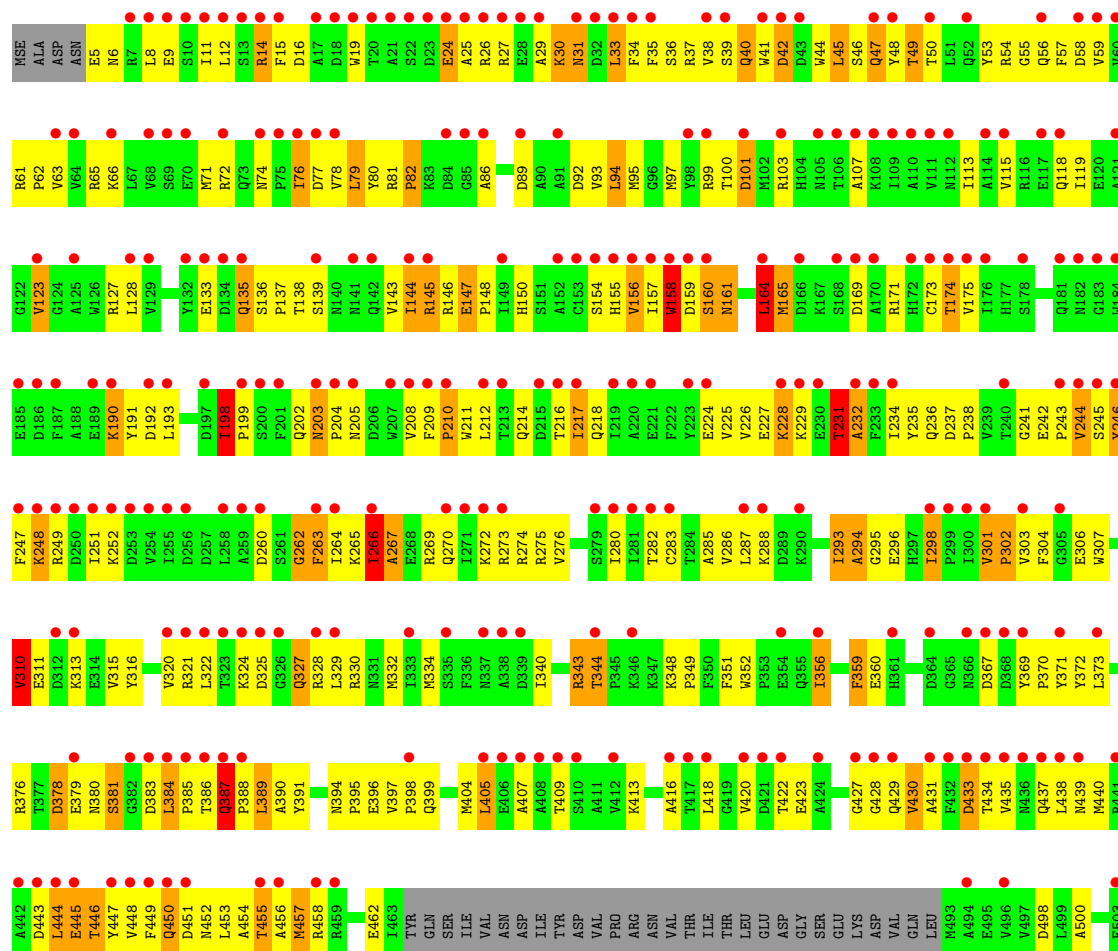
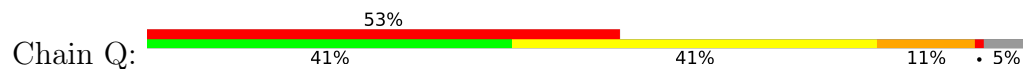


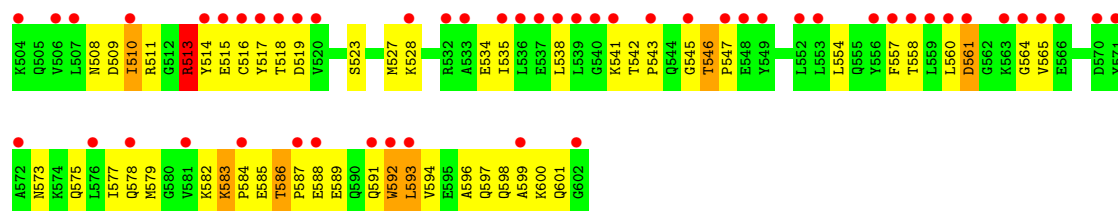
• Molecule 1: PORTAL PROTEIN



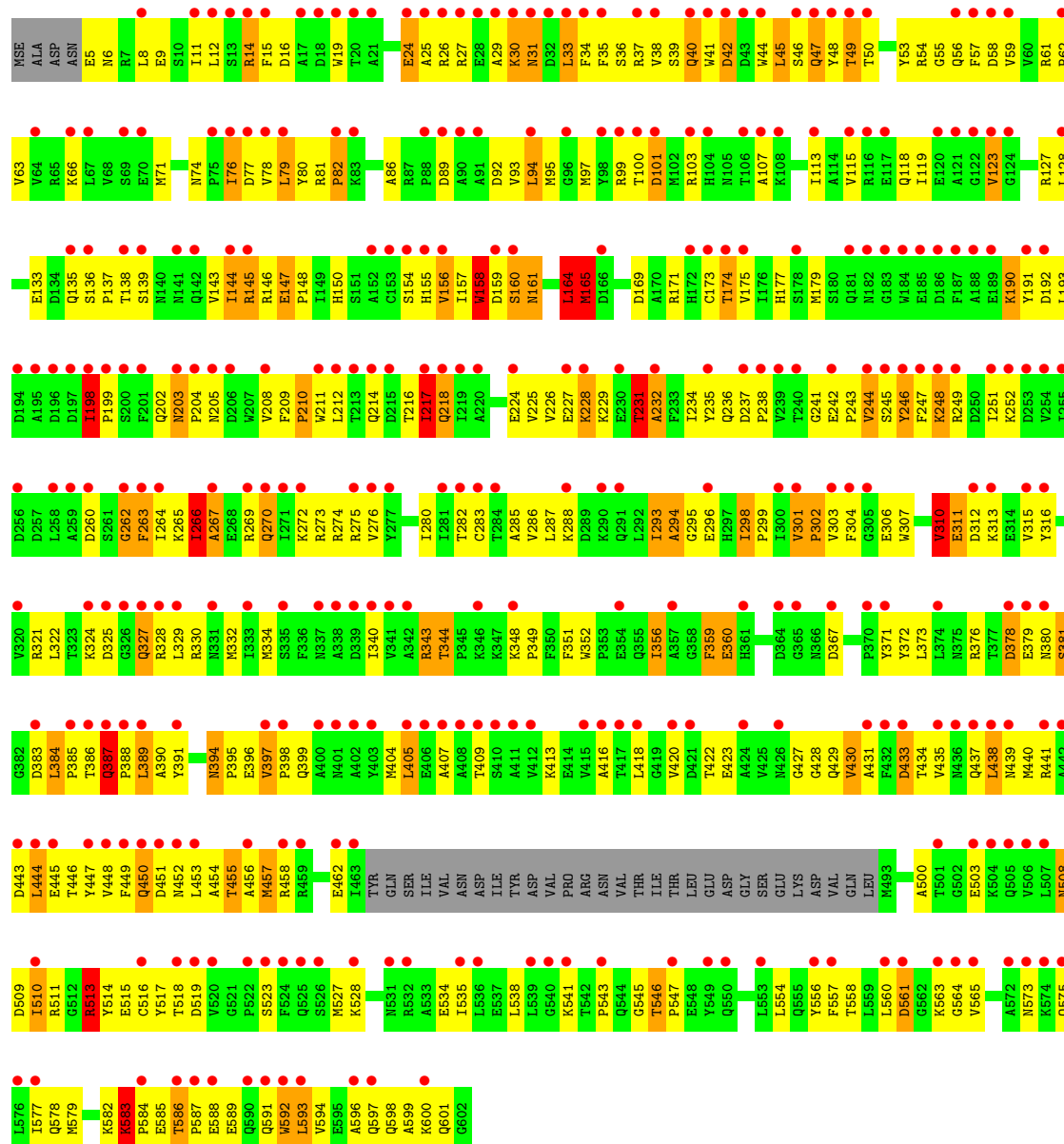


• Molecule 1: PORTAL PROTEIN



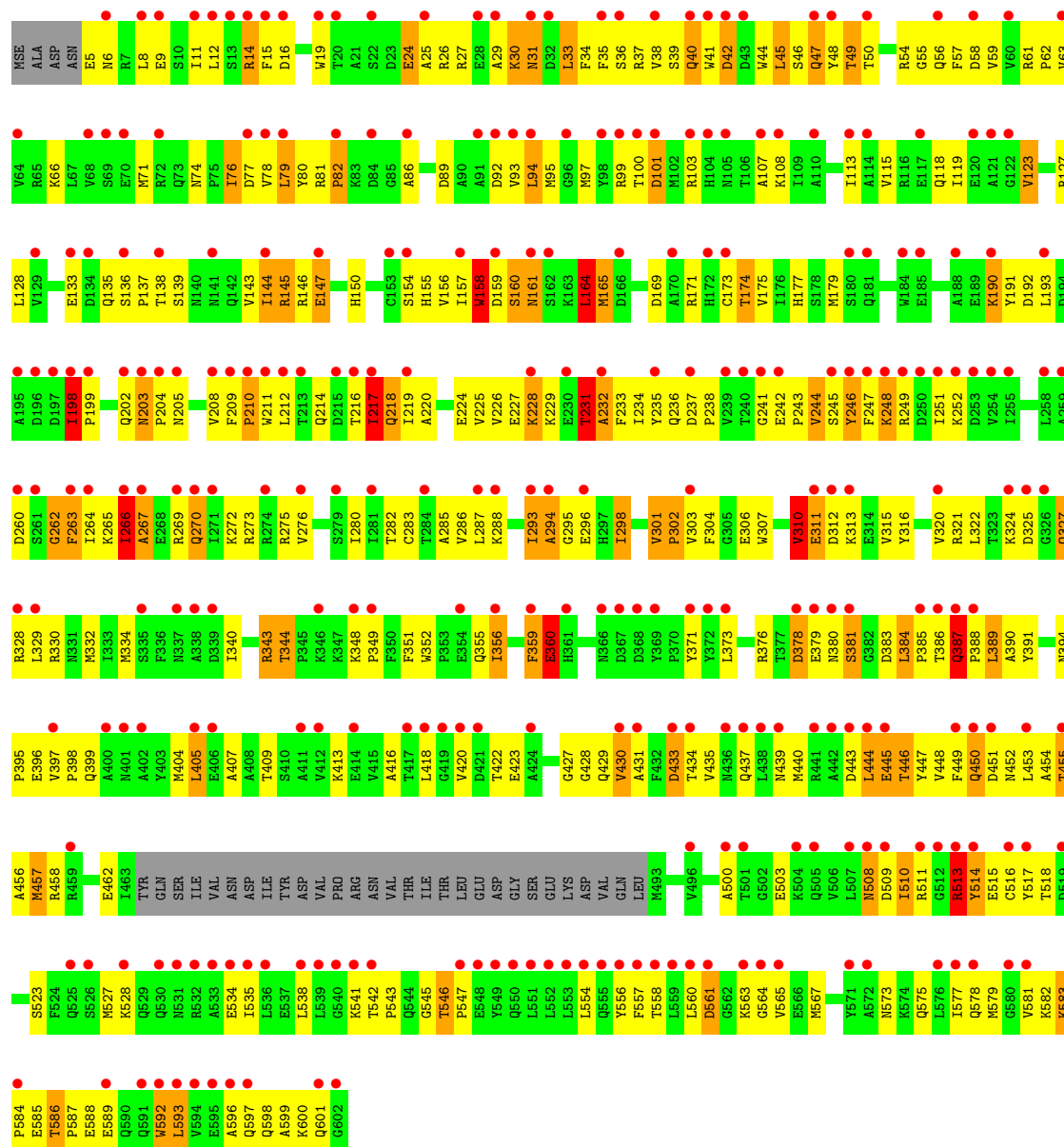


● Molecule 1: PORTAL PROTEIN



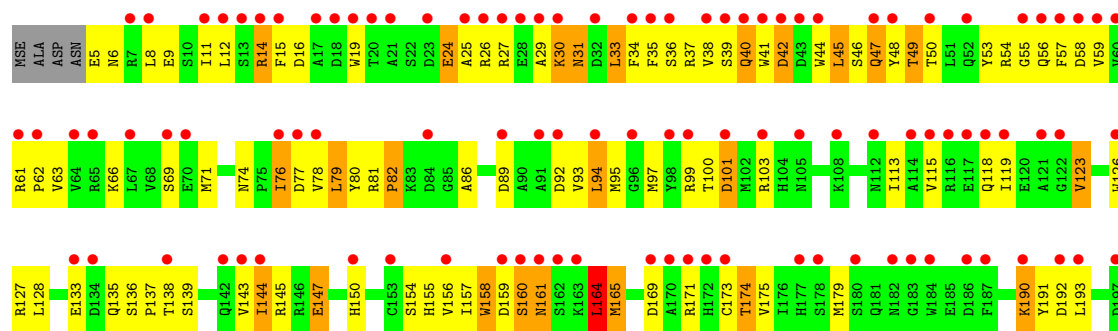
● Molecule 1: PORTAL PROTEIN

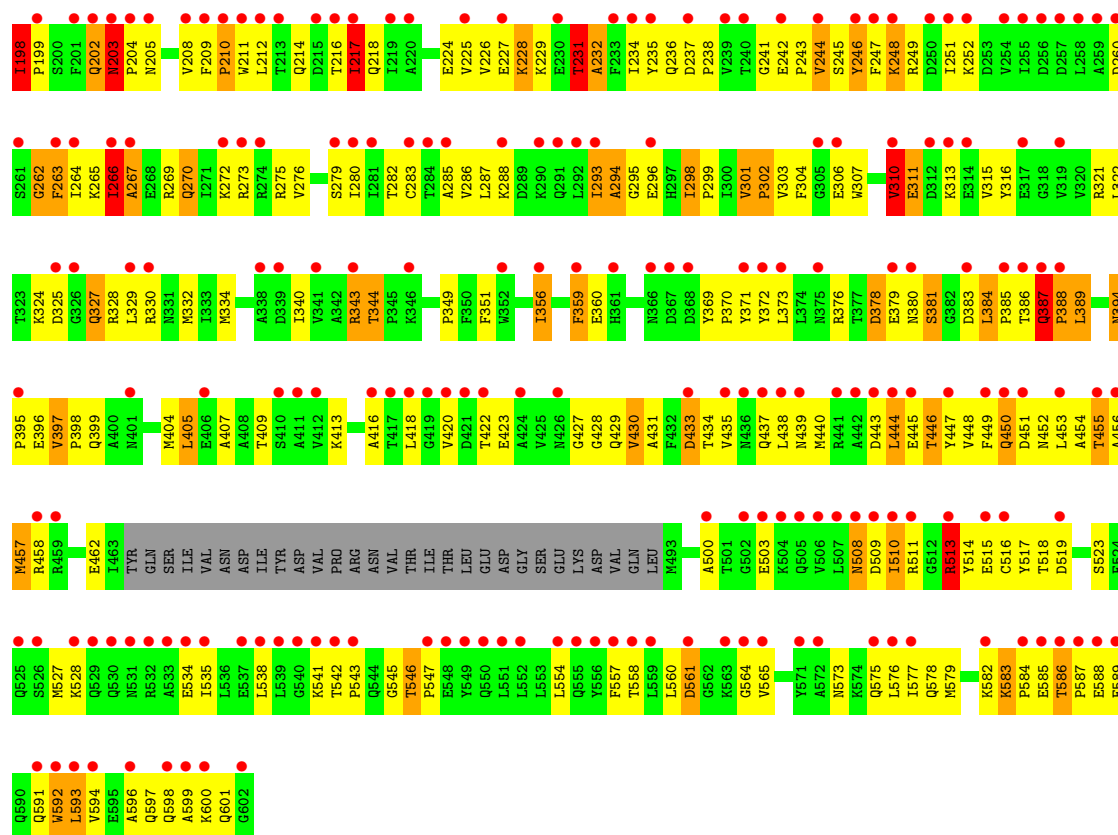




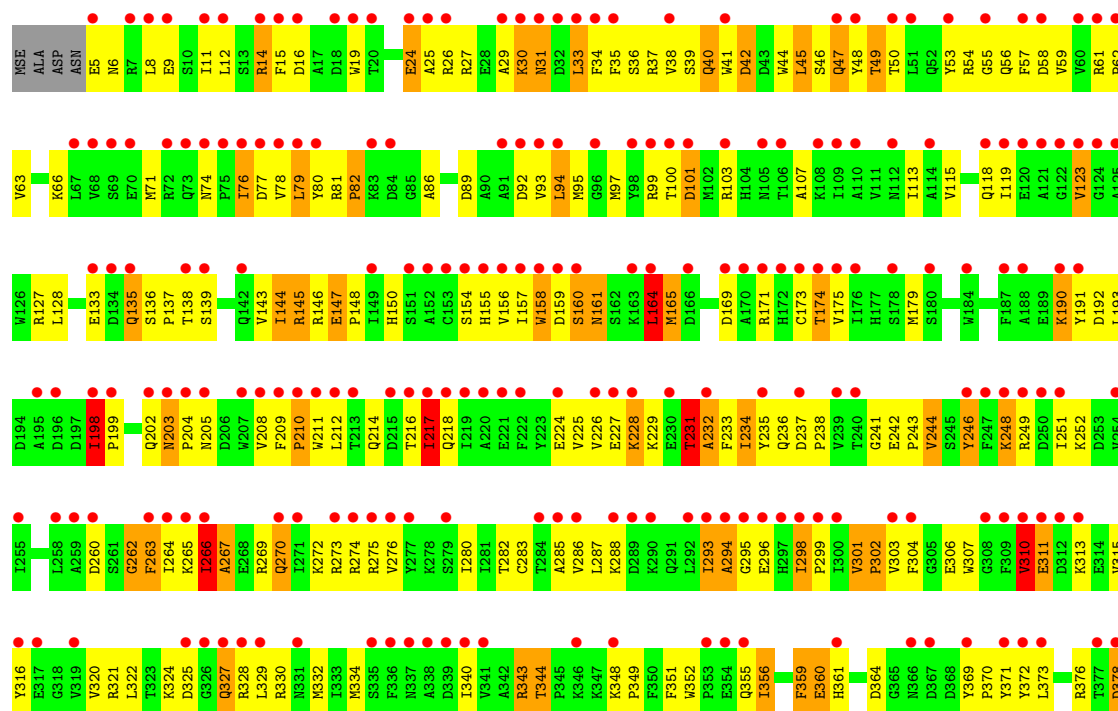
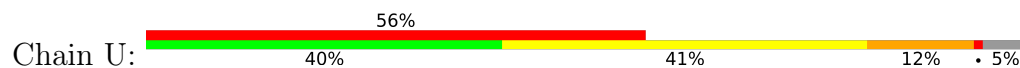
• Molecule 1: PORTAL PROTEIN

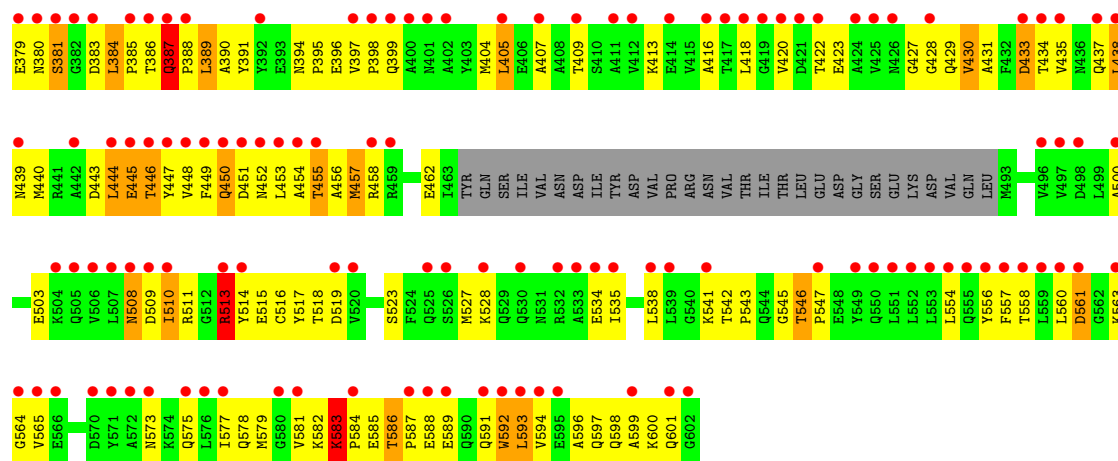
Chain T: 50% 42% 40% 11% 5%



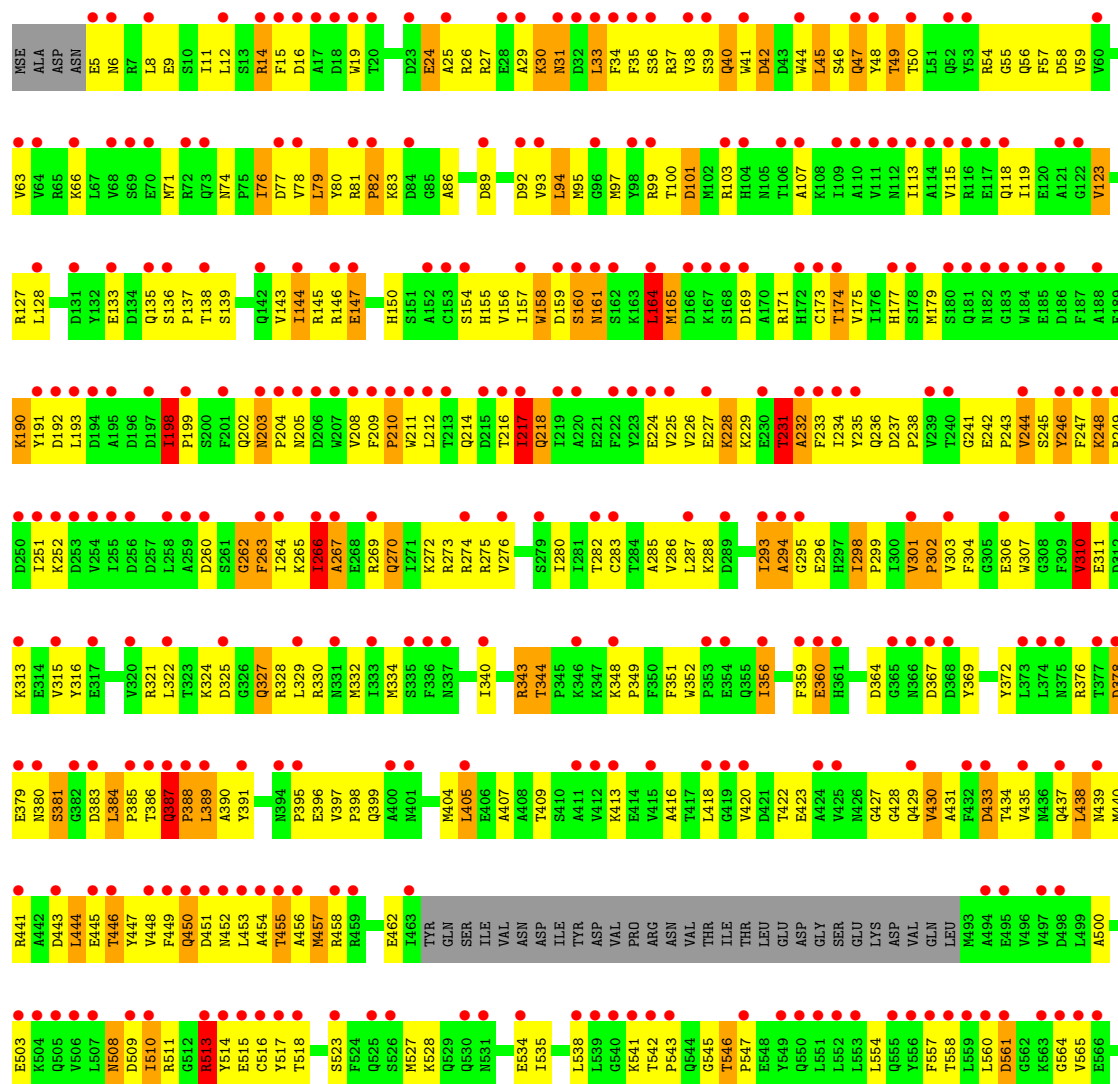


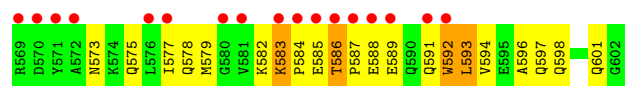
● Molecule 1: PORTAL PROTEIN



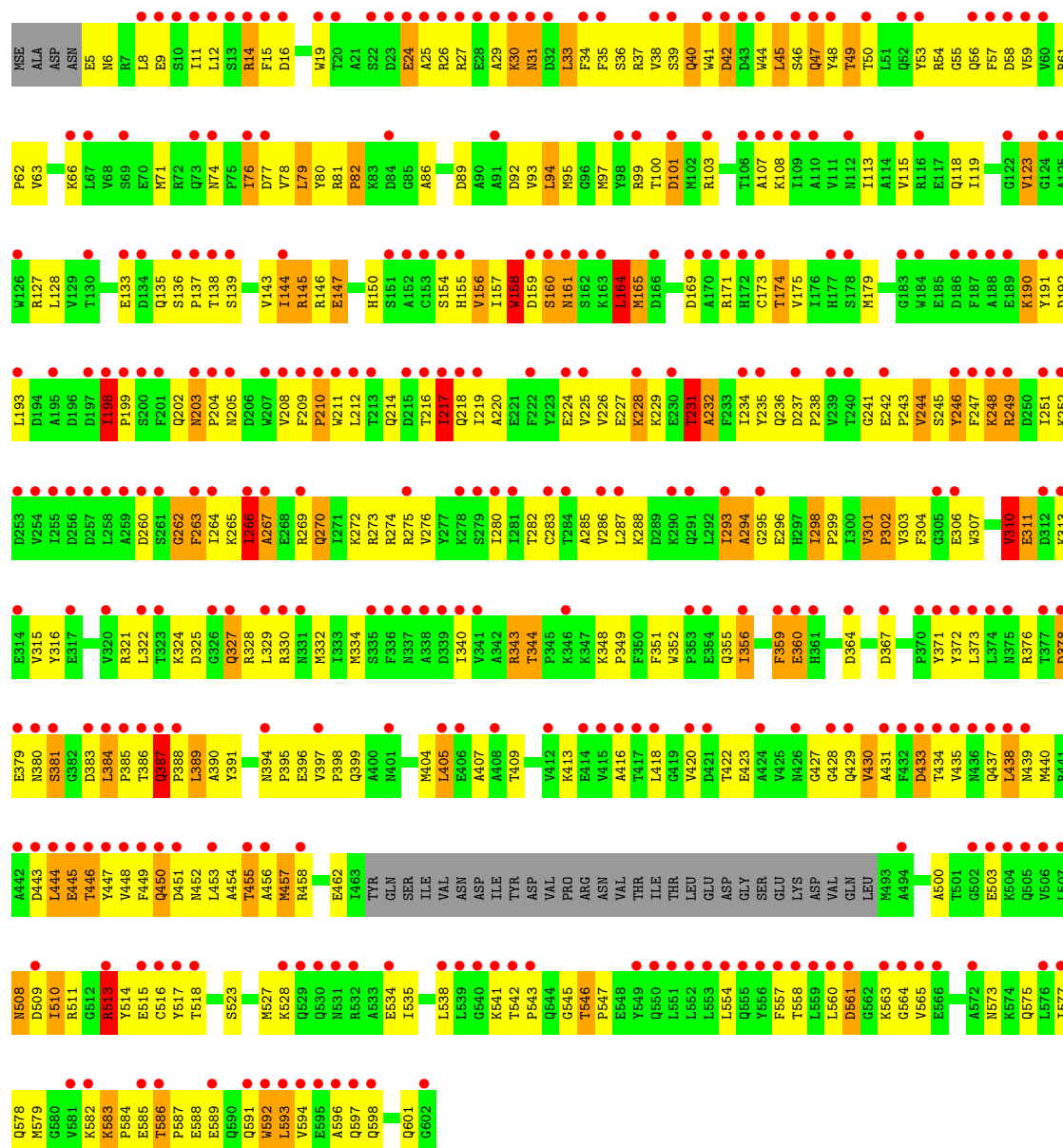


• Molecule 1: PORTAL PROTEIN

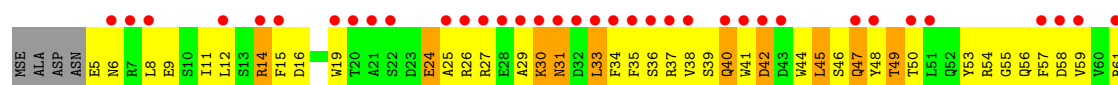


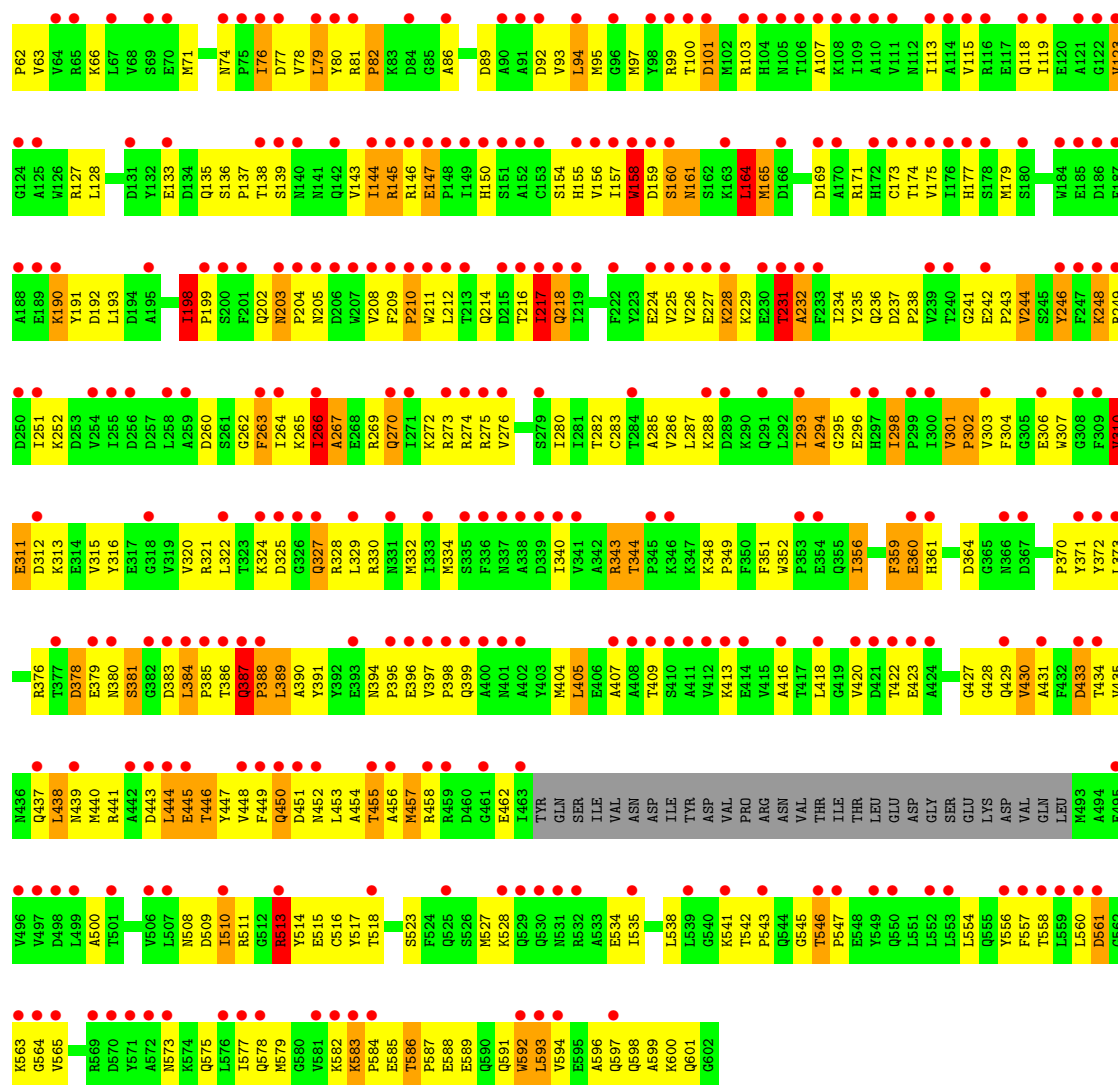


● Molecule 1: PORTAL PROTEIN



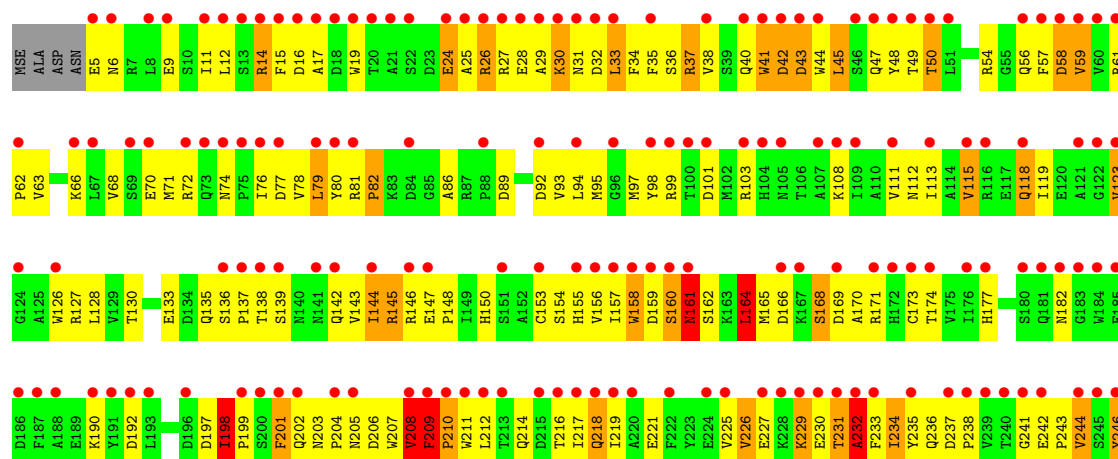
● Molecule 1: PORTAL PROTEIN

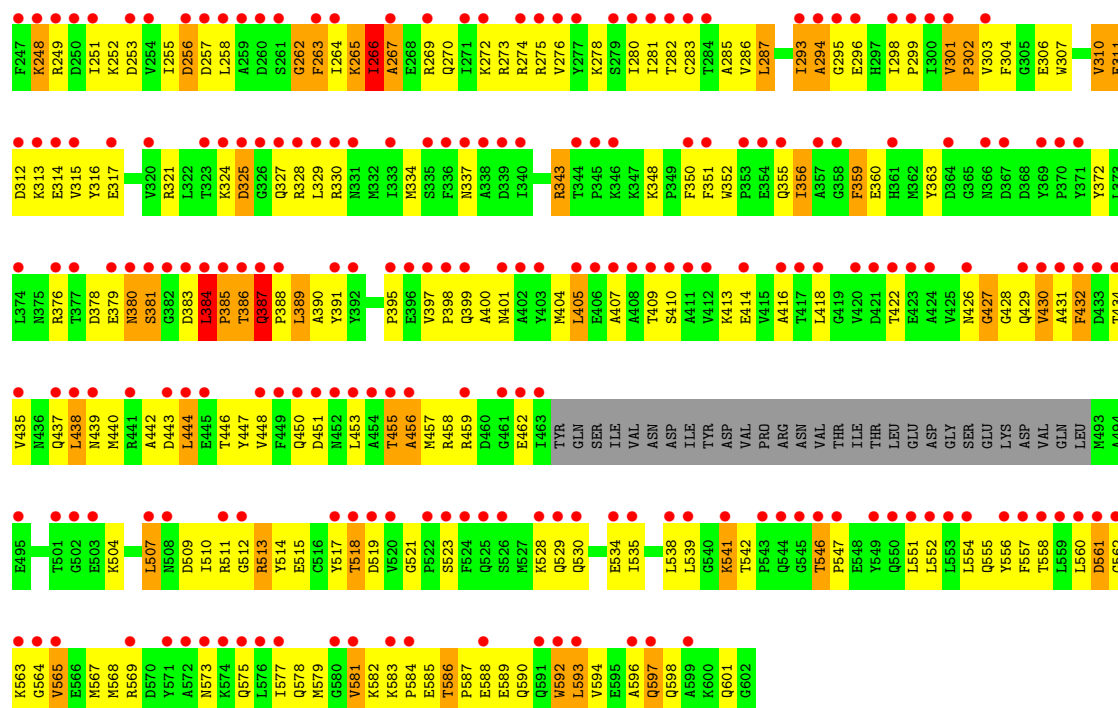




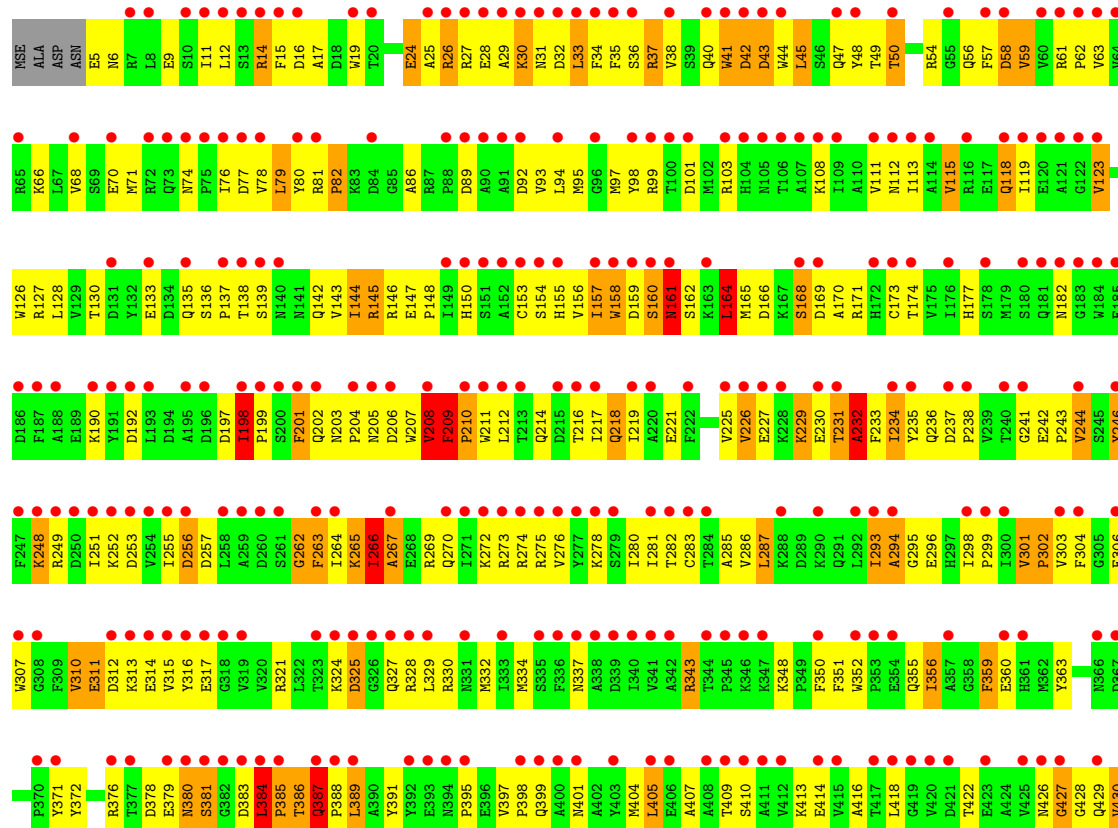
• Molecule 1: PORTAL PROTEIN

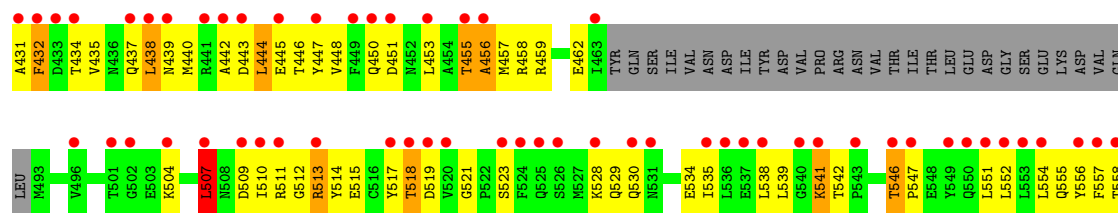
Chain A:



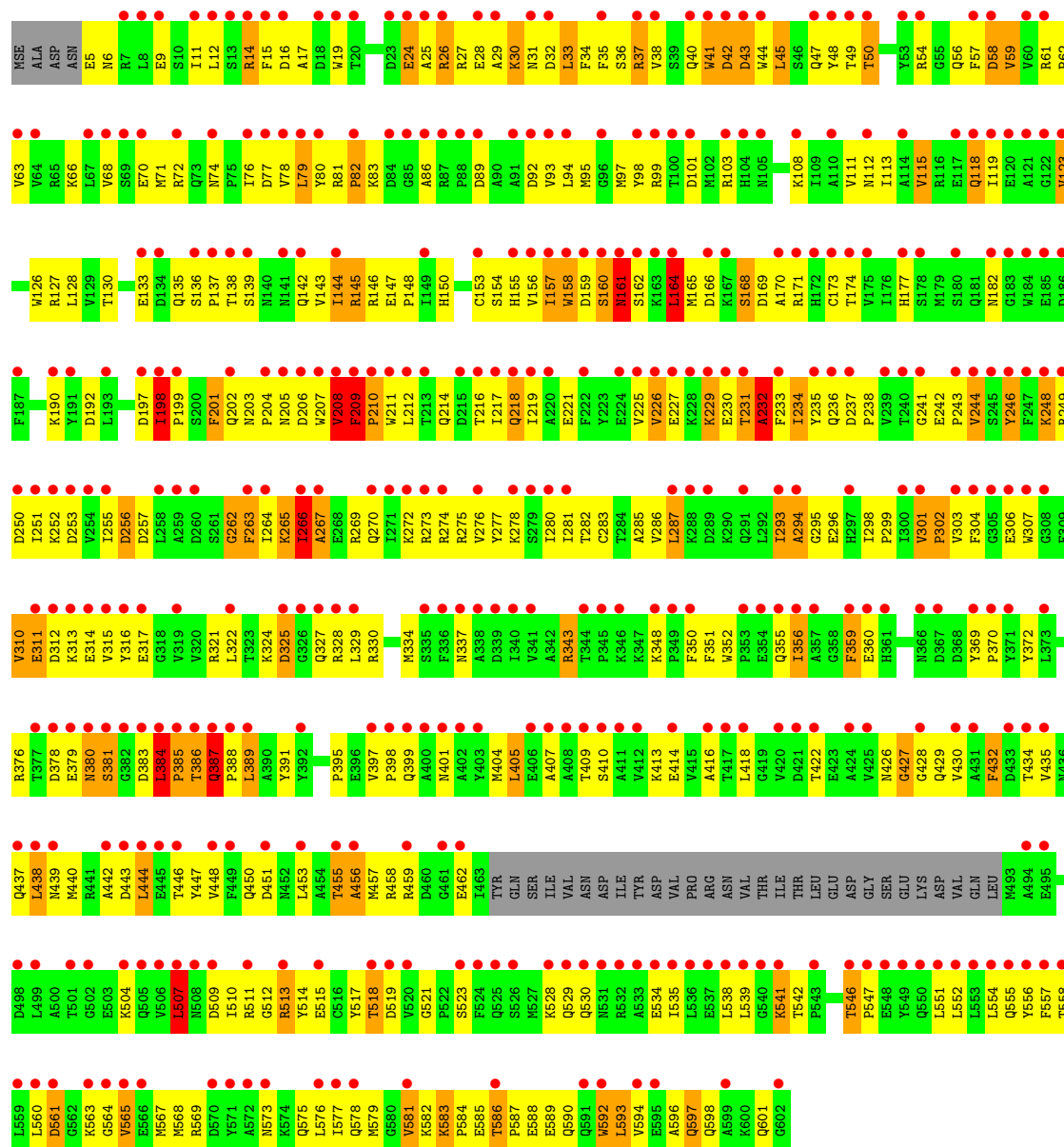


● Molecule 1: PORTAL PROTEIN

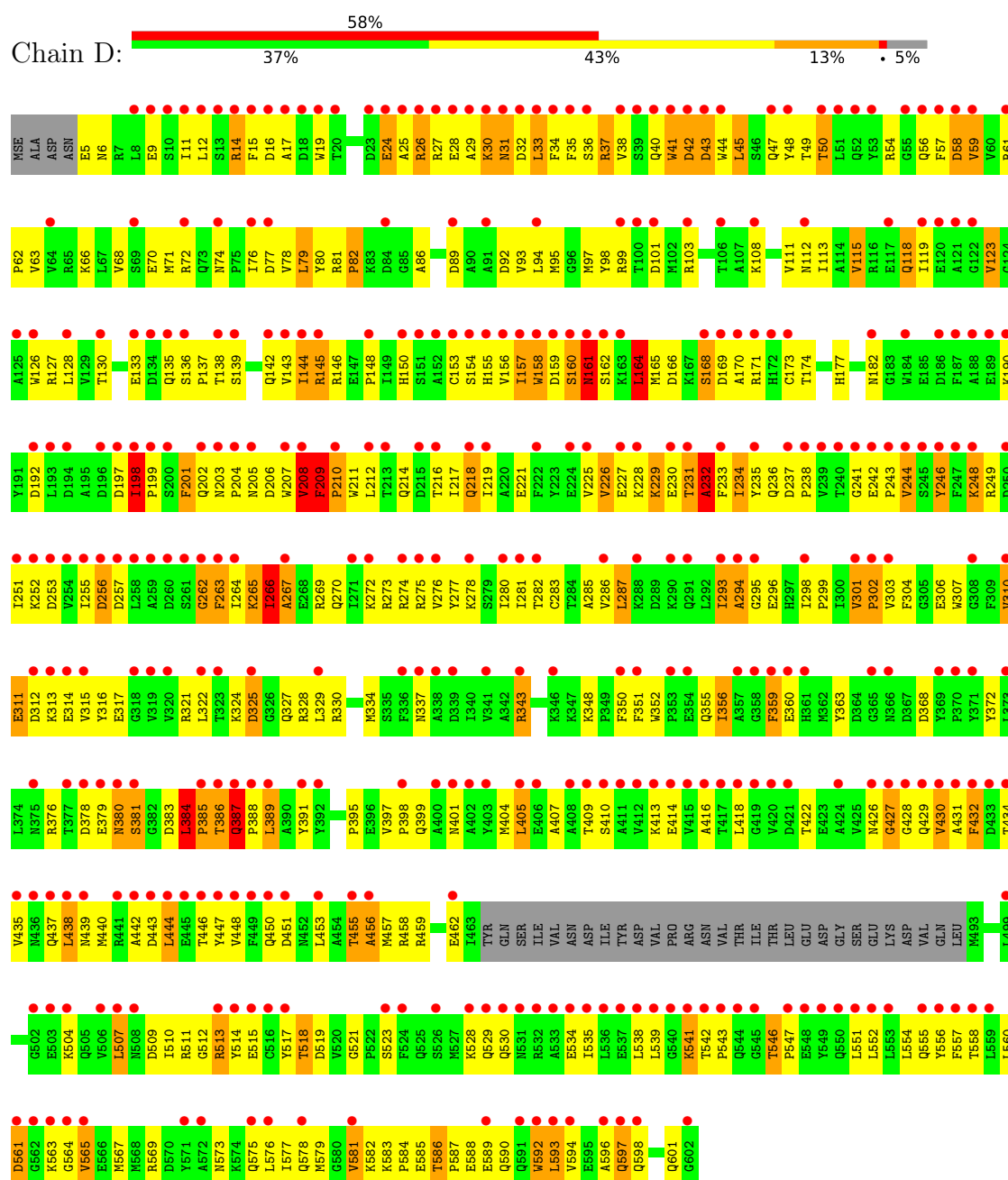




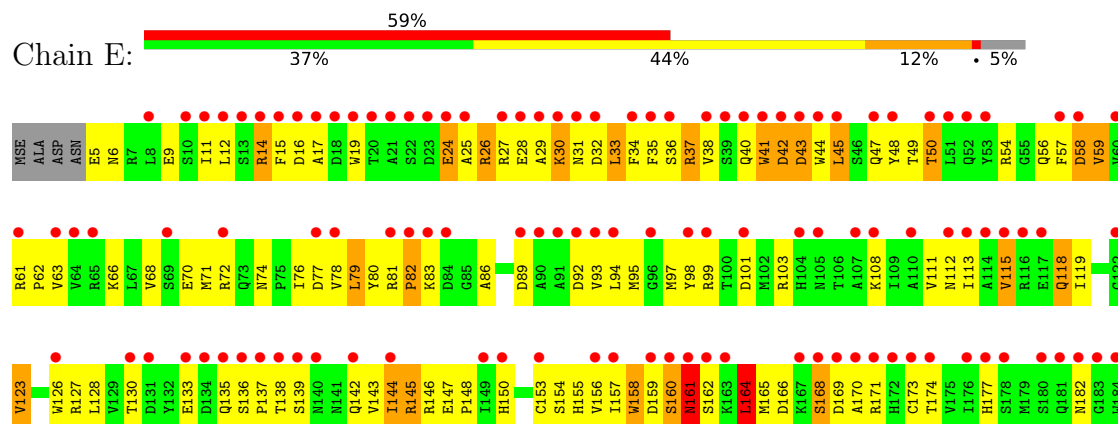
• Molecule 1: PORTAL PROTEIN

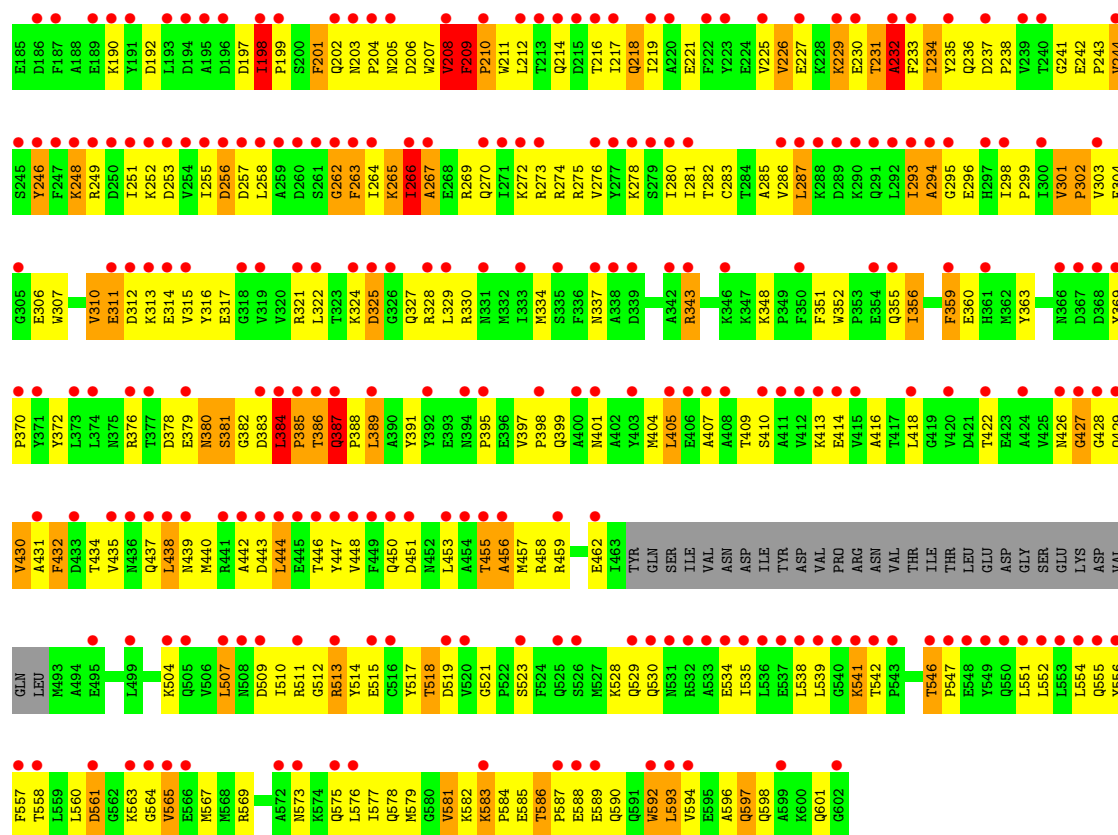


• Molecule 1: PORTAL PROTEIN

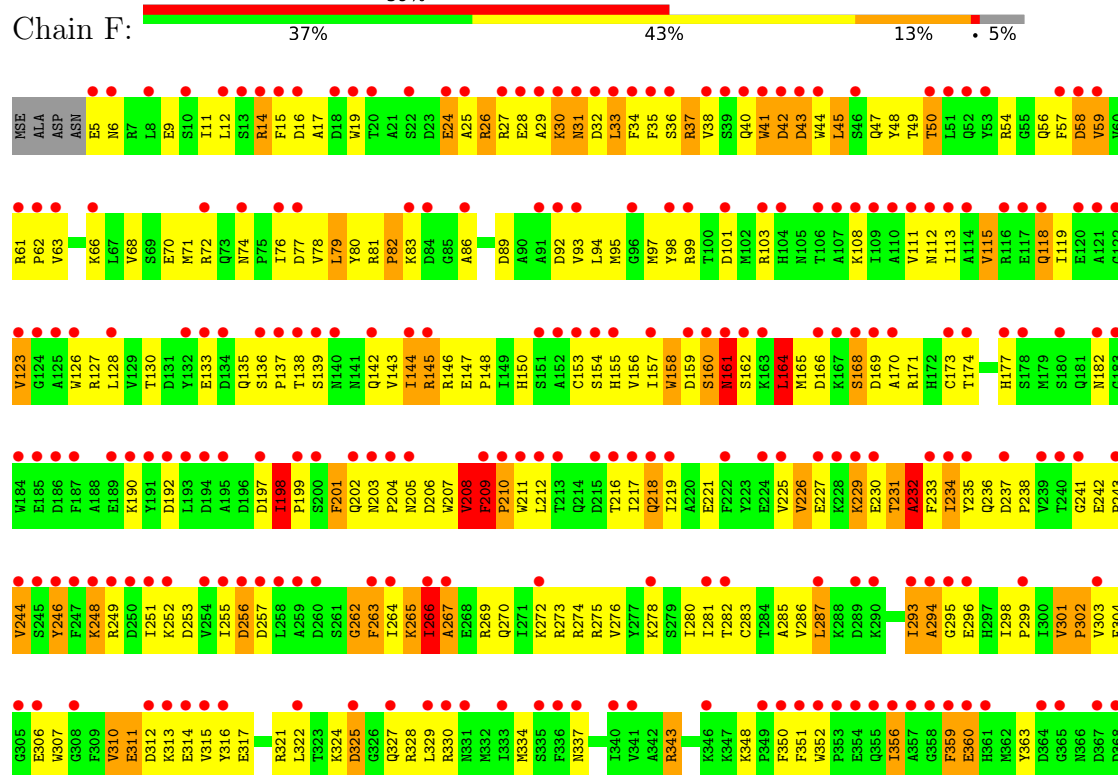


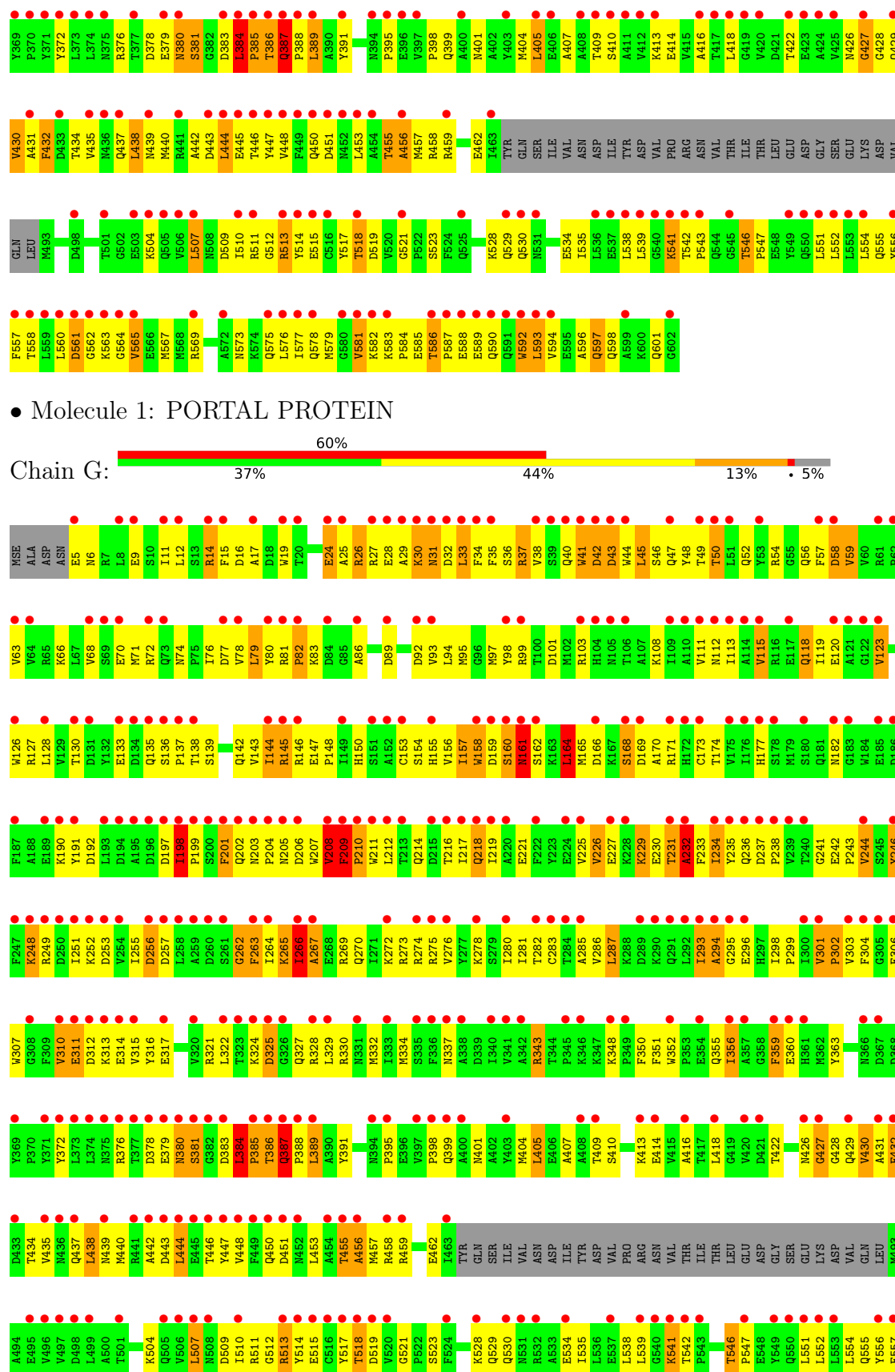
• Molecule 1: PORTAL PROTEIN

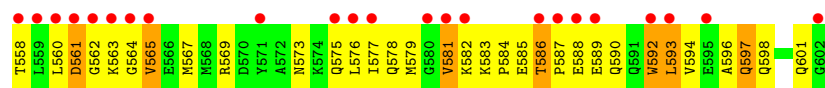




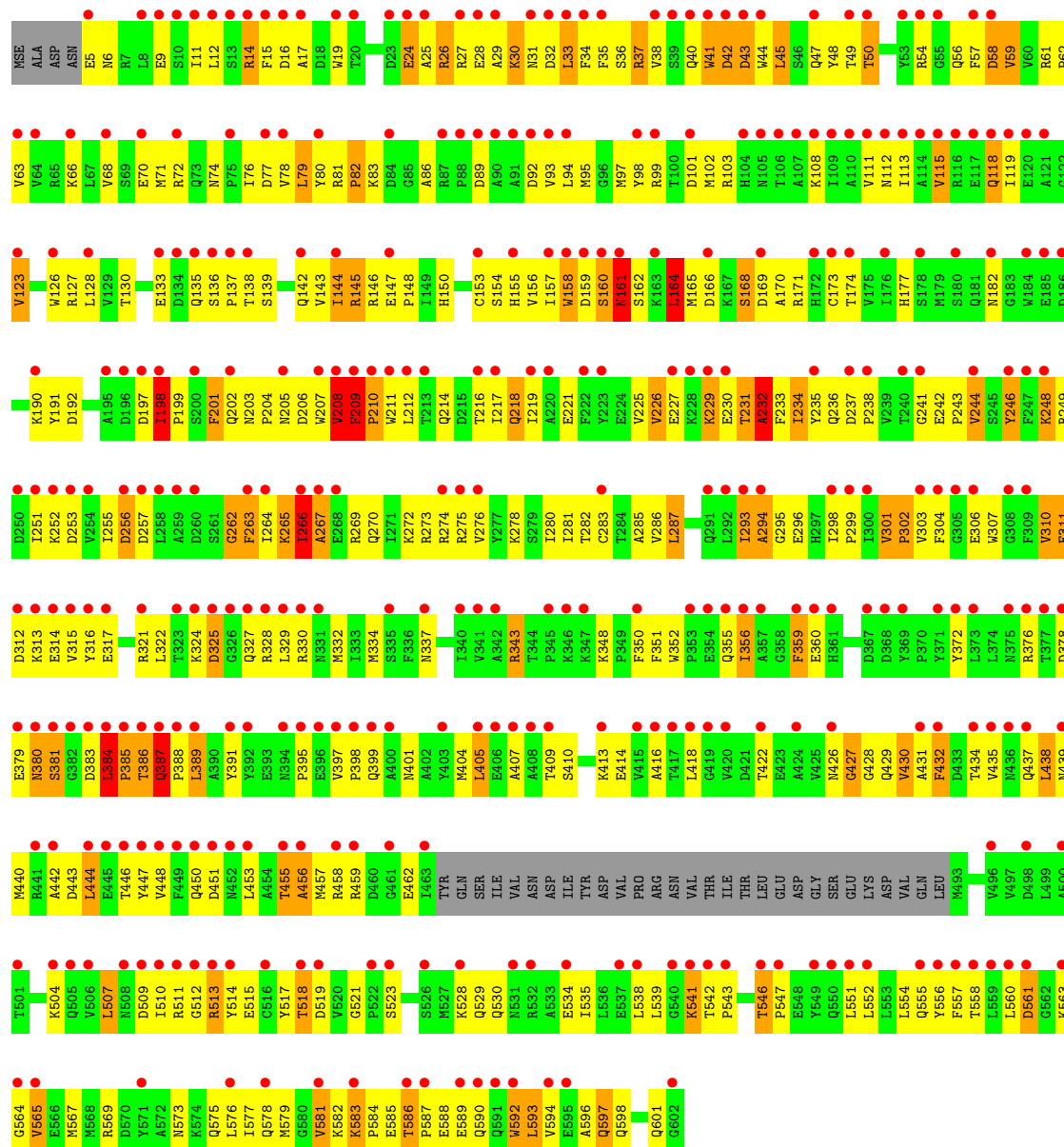
• Molecule 1: PORTAL PROTEIN



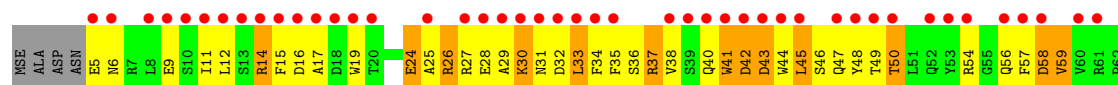
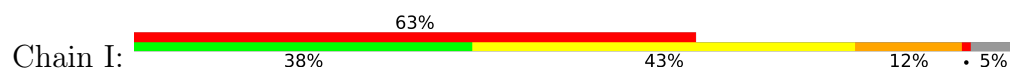


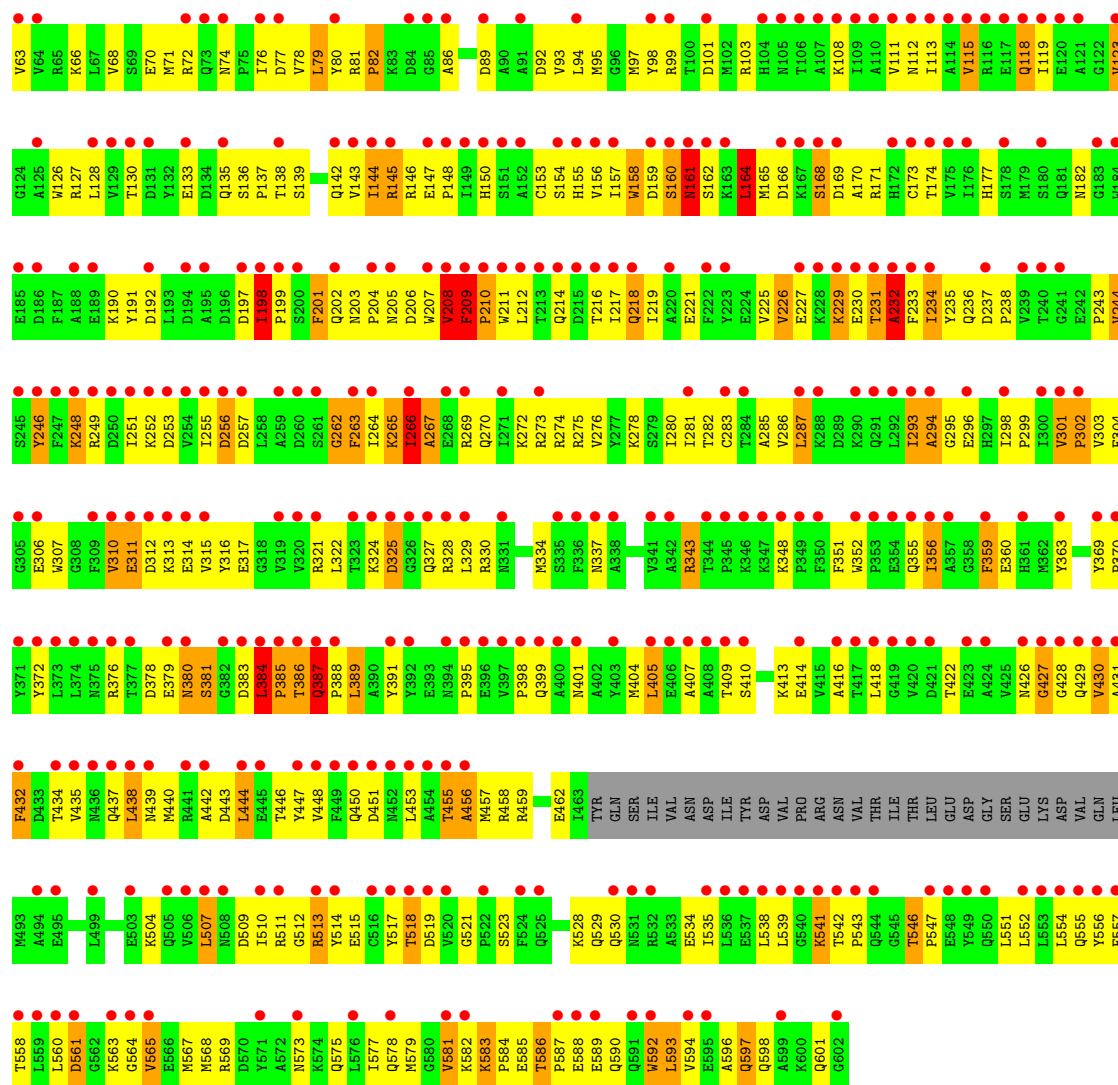


● Molecule 1: PORTAL PROTEIN



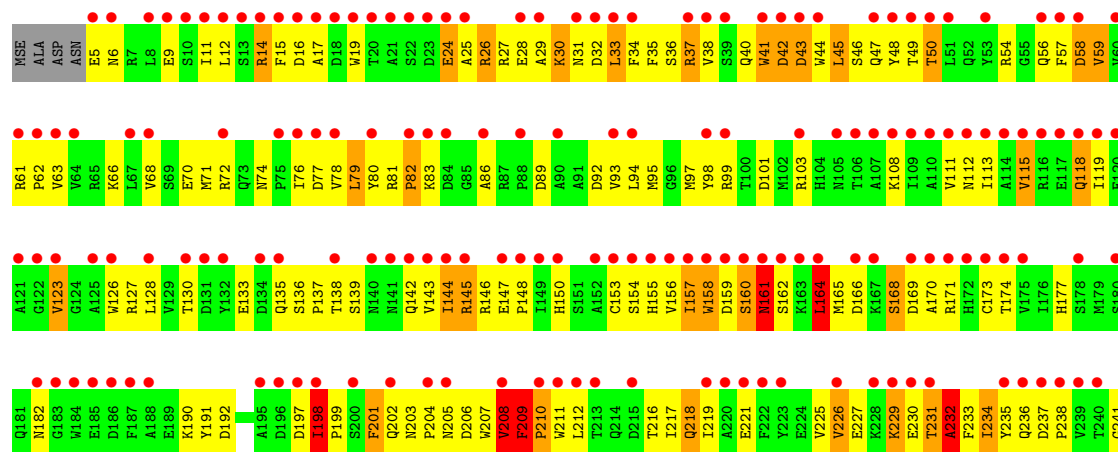
● Molecule 1: PORTAL PROTEIN

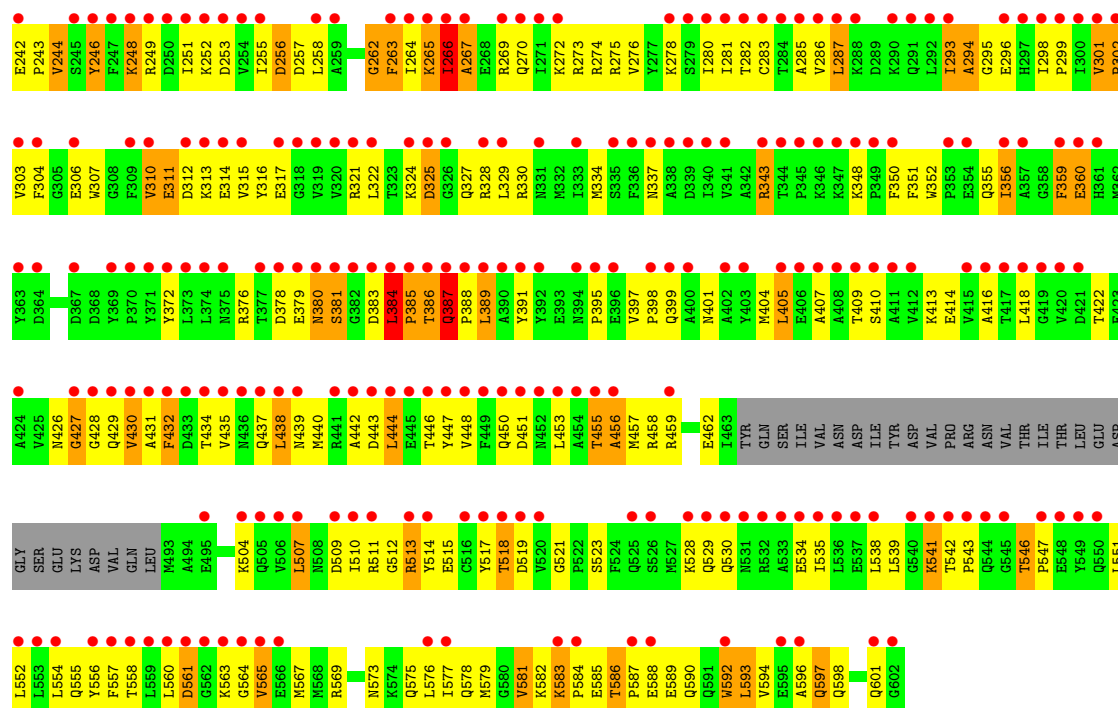




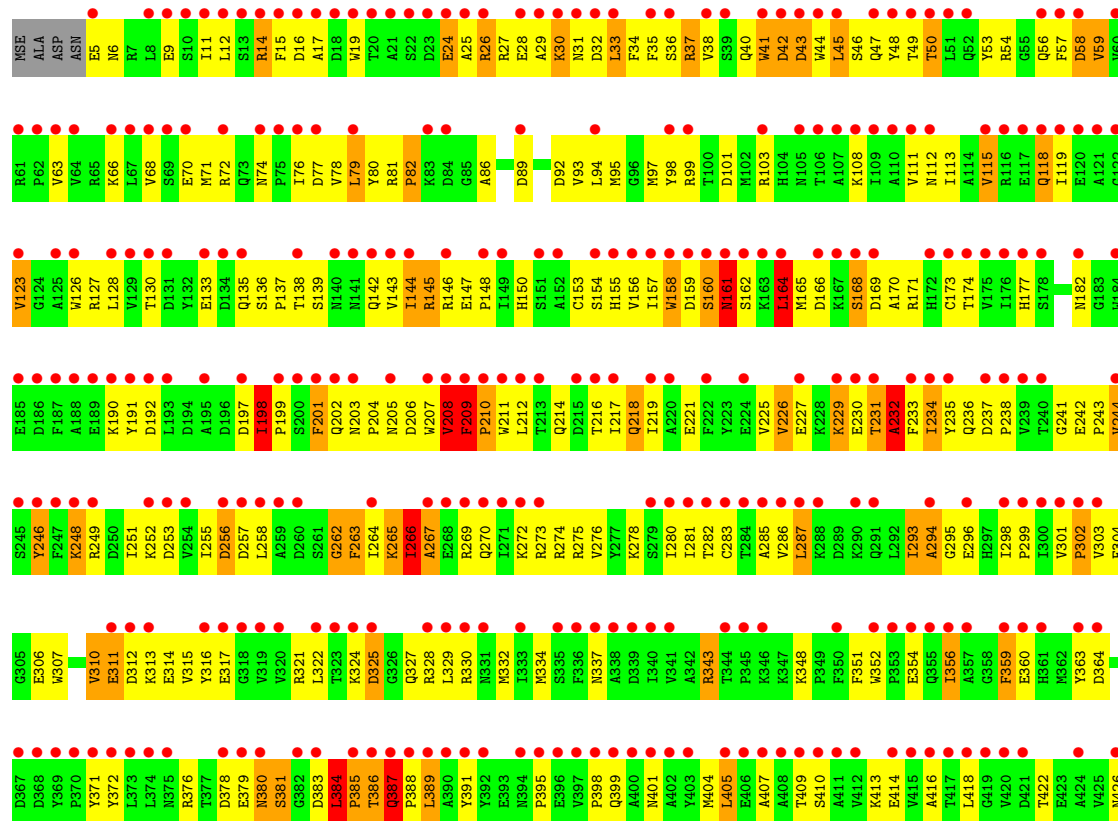
• Molecule 1: PORTAL PROTEIN

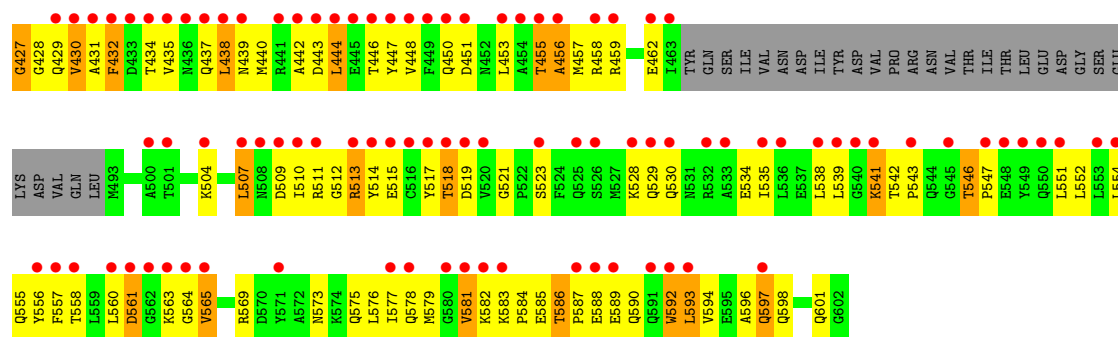
Chain J:



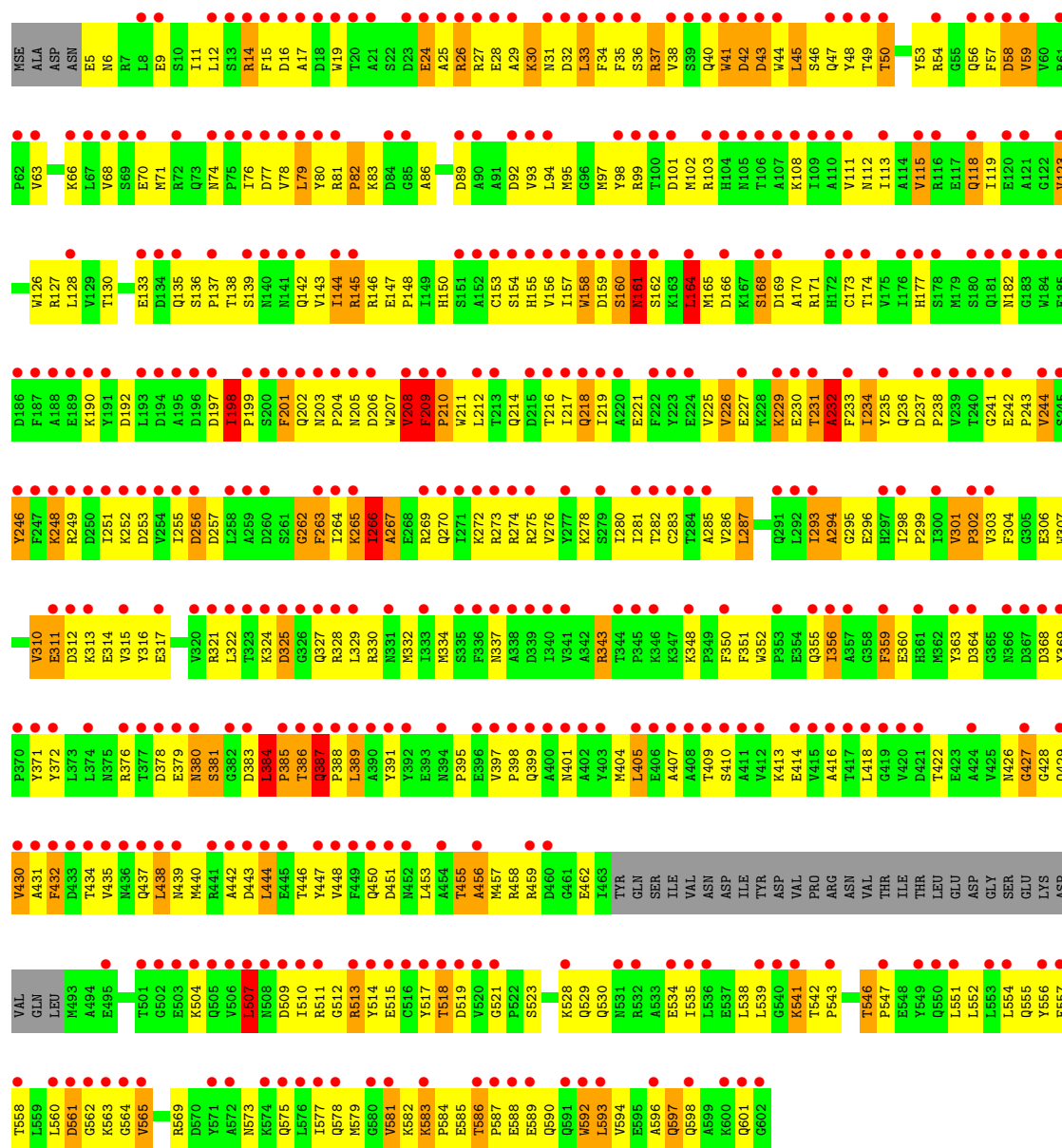


• Molecule 1: PORTAL PROTEIN

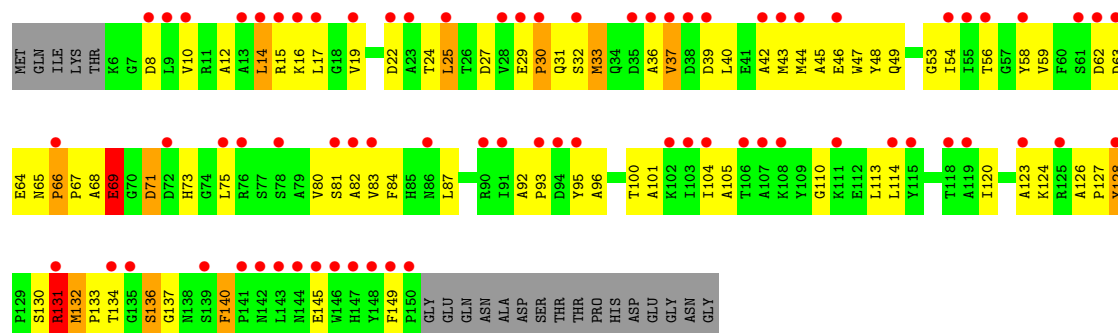
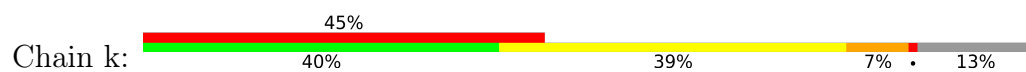




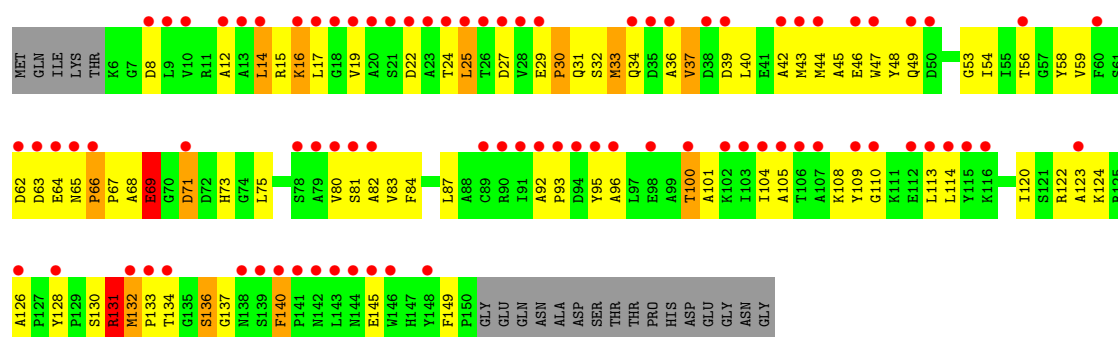
• Molecule 1: PORTAL PROTEIN



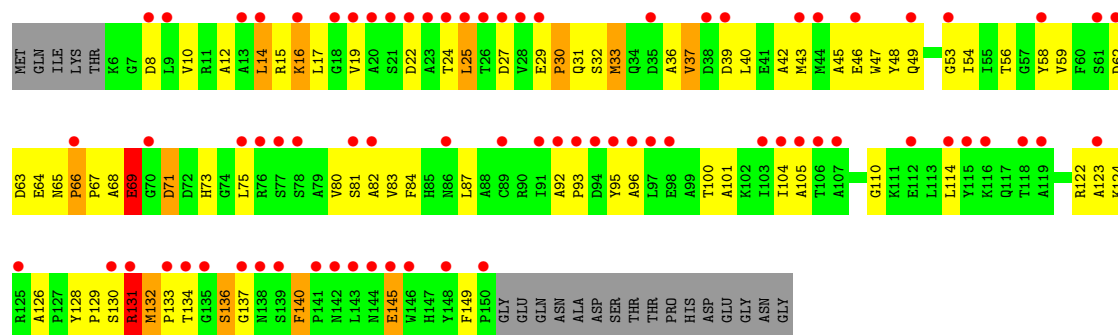
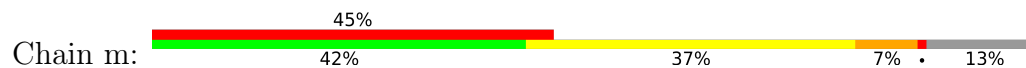
• Molecule 2: PACKAGED DNA STABILIZATION PROTEIN GP4



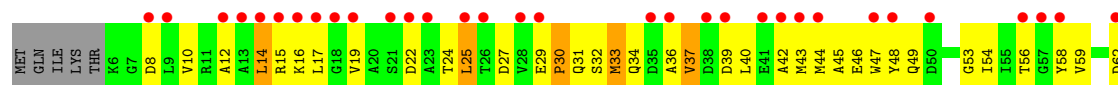
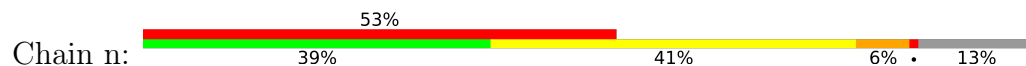
• Molecule 2: PACKAGED DNA STABILIZATION PROTEIN GP4

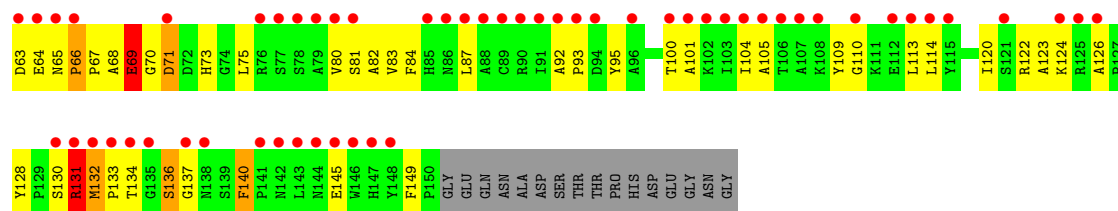


• Molecule 2: PACKAGED DNA STABILIZATION PROTEIN GP4

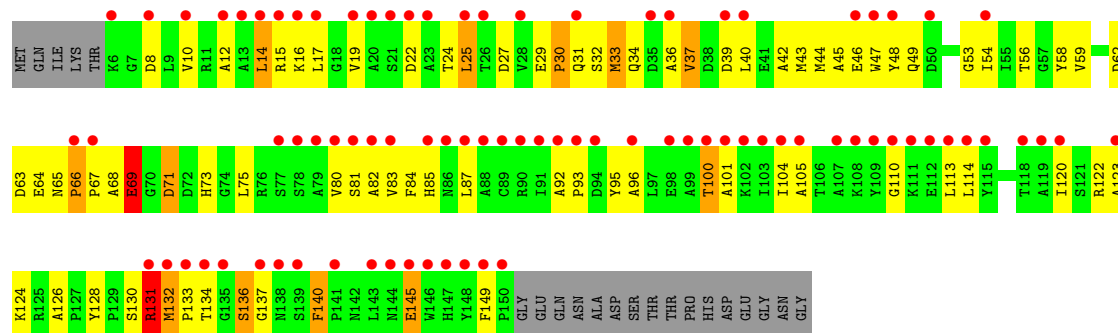


• Molecule 2: PACKAGED DNA STABILIZATION PROTEIN GP4

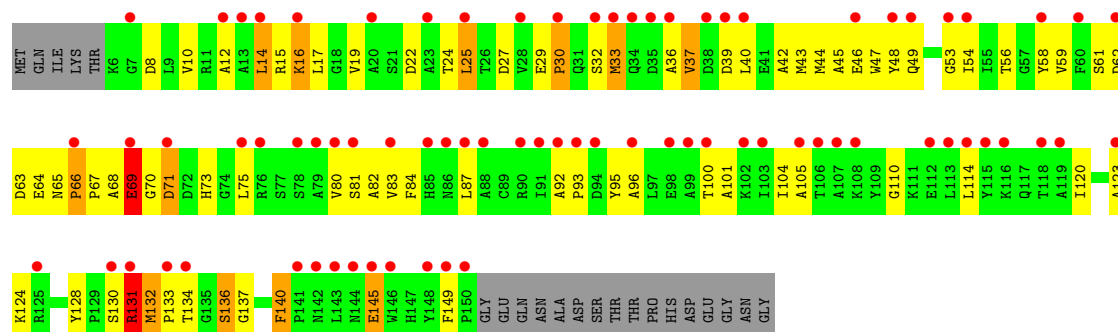




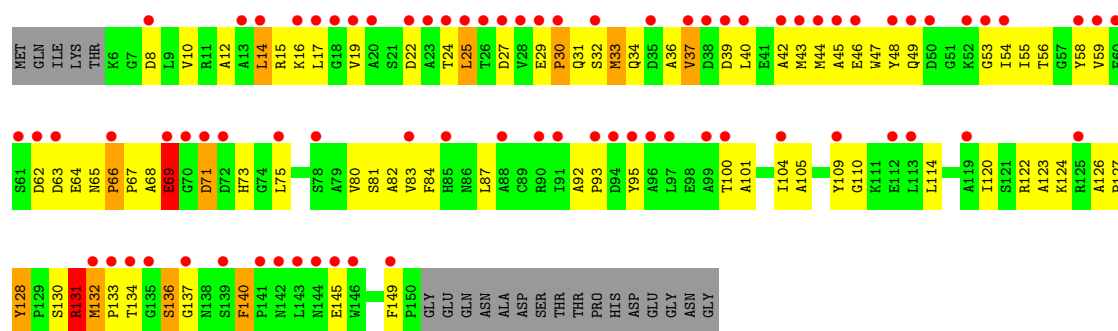
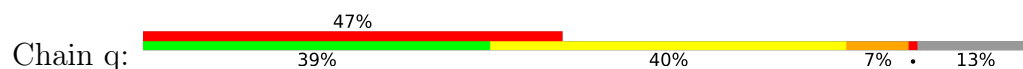
● Molecule 2: PACKAGED DNA STABILIZATION PROTEIN GP4



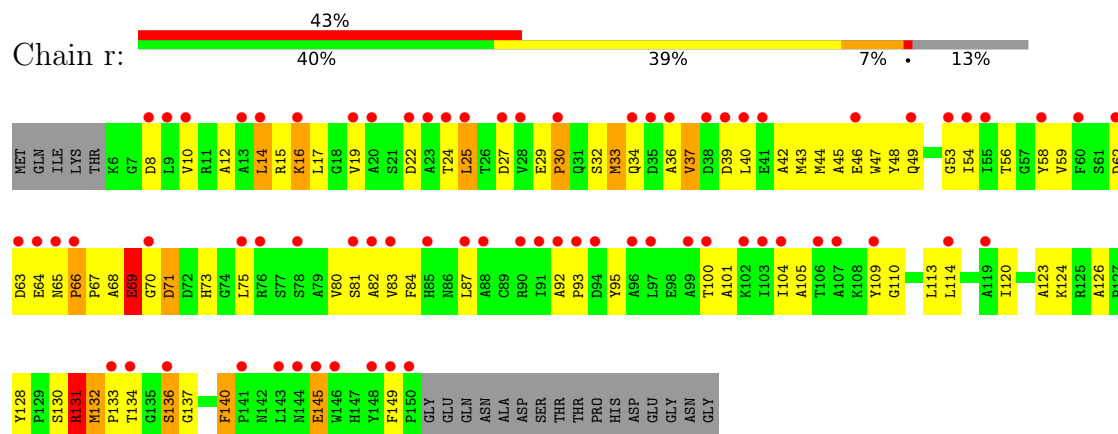
● Molecule 2: PACKAGED DNA STABILIZATION PROTEIN GP4



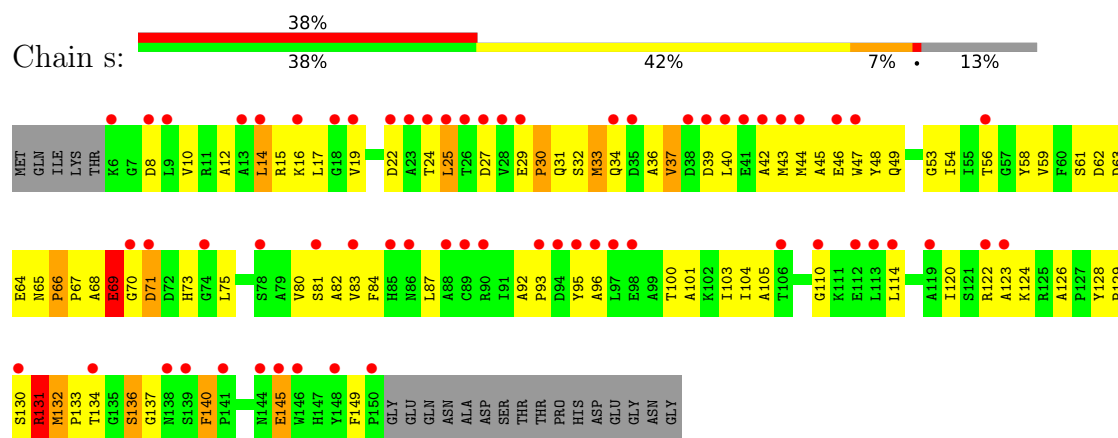
● Molecule 2: PACKAGED DNA STABILIZATION PROTEIN GP4



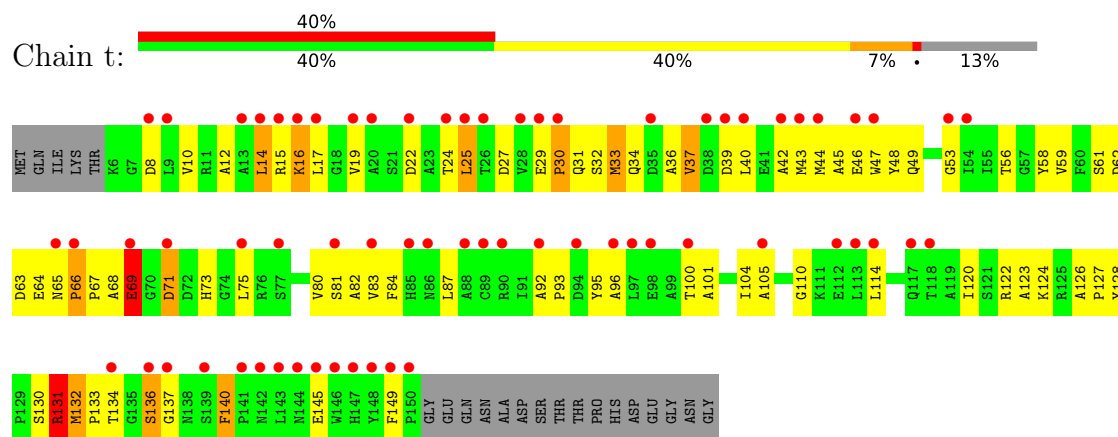
• Molecule 2: PACKAGED DNA STABILIZATION PROTEIN GP4



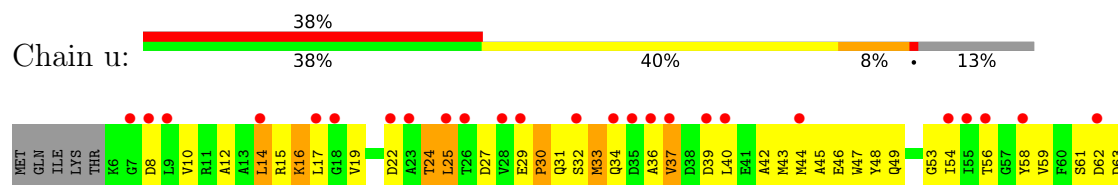
• Molecule 2: PACKAGED DNA STABILIZATION PROTEIN GP4

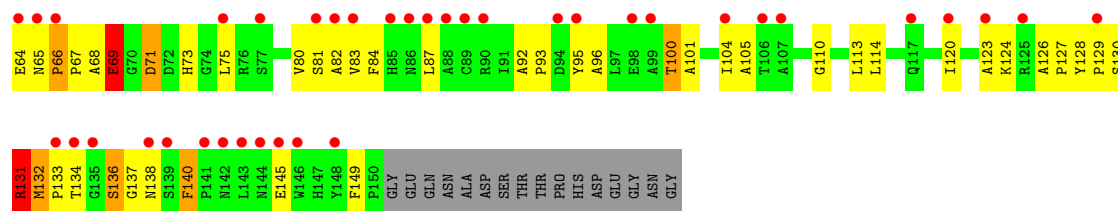


• Molecule 2: PACKAGED DNA STABILIZATION PROTEIN GP4



• Molecule 2: PACKAGED DNA STABILIZATION PROTEIN GP4





• Molecule 2: PACKAGED DNA STABILIZATION PROTEIN GP4



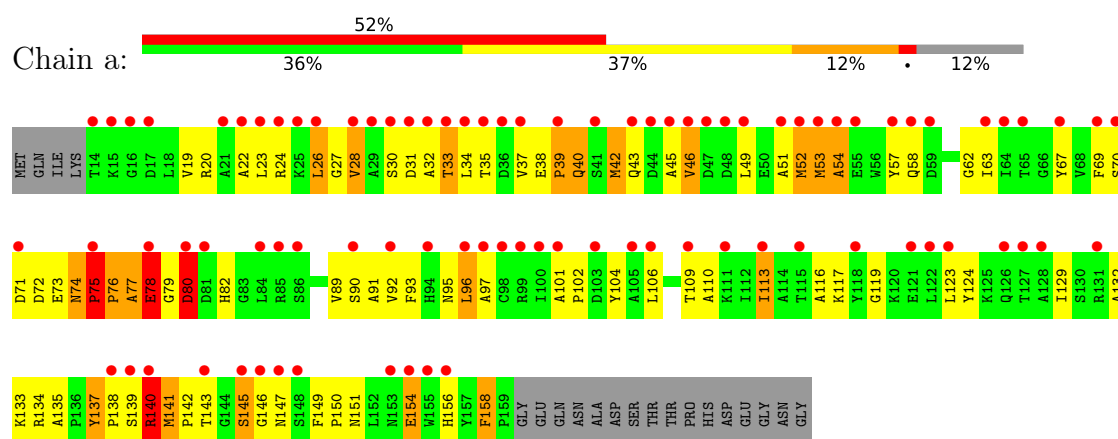
• Molecule 2: PACKAGED DNA STABILIZATION PROTEIN GP4



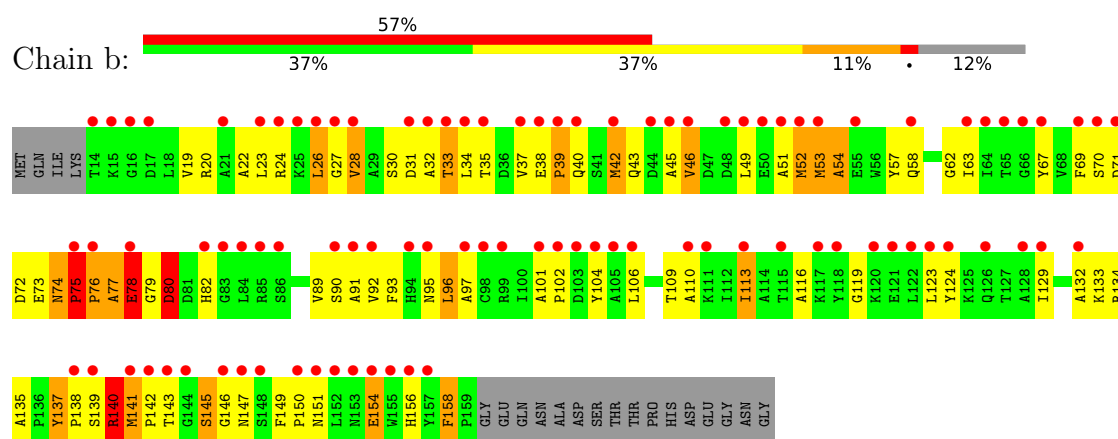
• Molecule 2: PACKAGED DNA STABILIZATION PROTEIN GP4



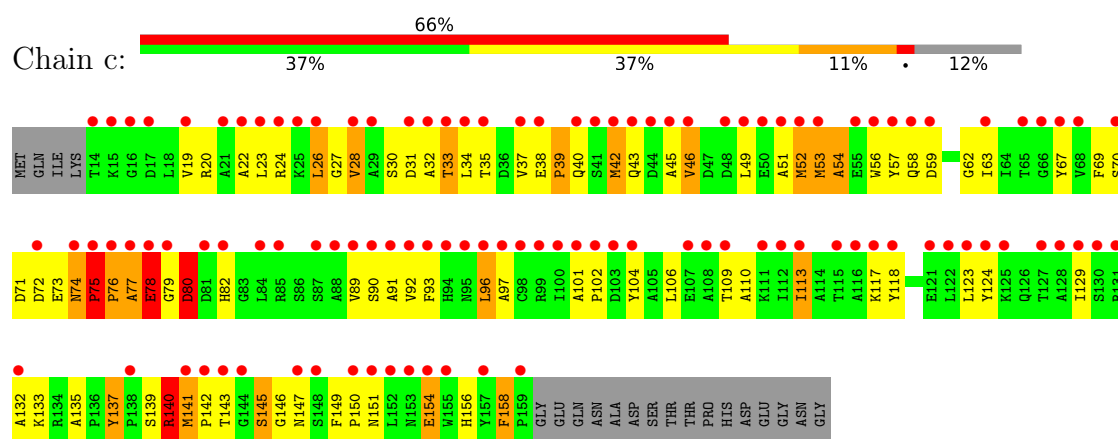
• Molecule 2: PACKAGED DNA STABILIZATION PROTEIN GP4



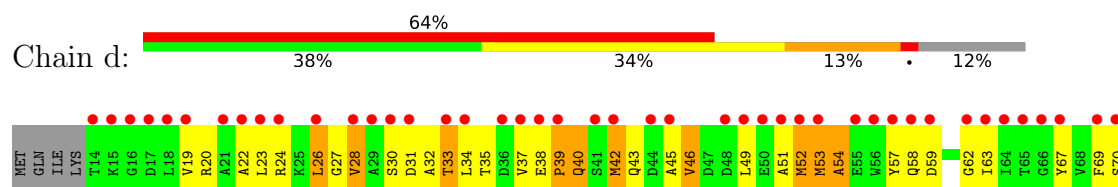
• Molecule 2: PACKAGED DNA STABILIZATION PROTEIN GP4

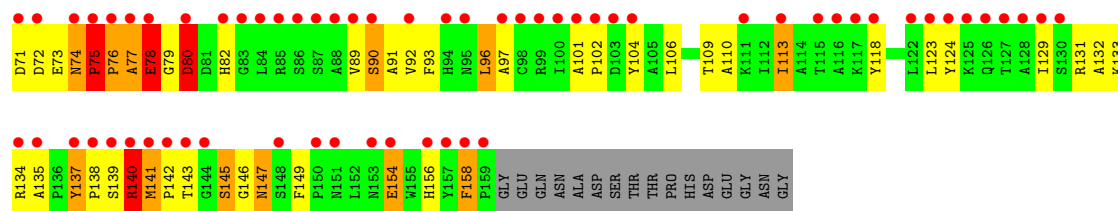


• Molecule 2: PACKAGED DNA STABILIZATION PROTEIN GP4

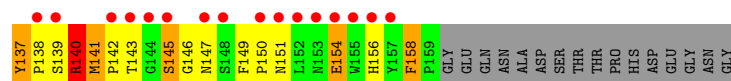
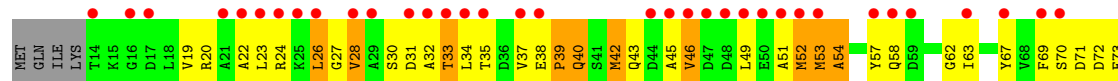
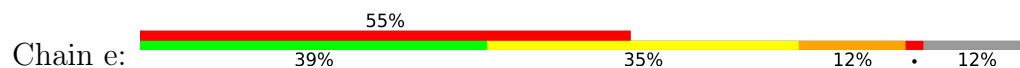


• Molecule 2: PACKAGED DNA STABILIZATION PROTEIN GP4

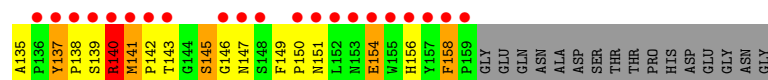
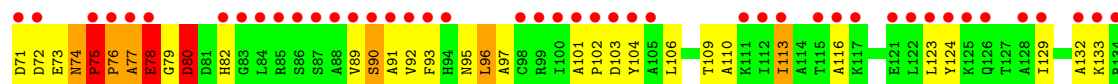
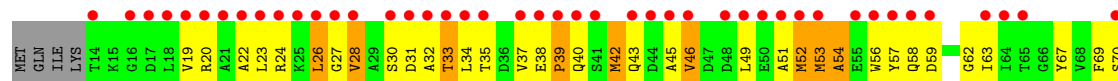




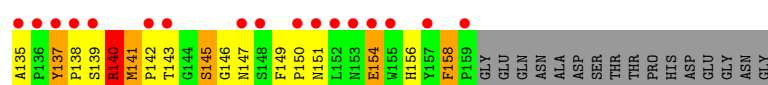
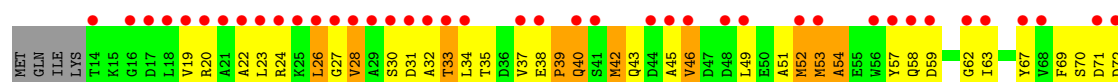
● Molecule 2: PACKAGED DNA STABILIZATION PROTEIN GP4



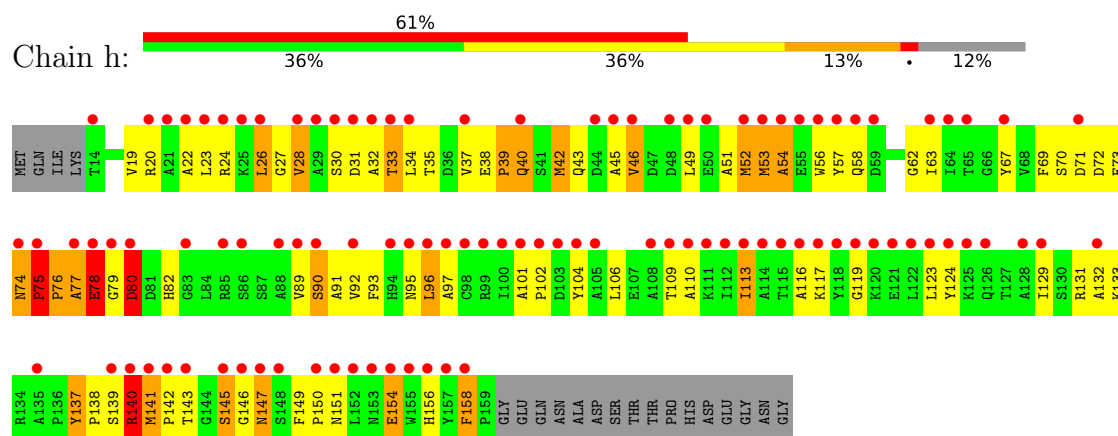
● Molecule 2: PACKAGED DNA STABILIZATION PROTEIN GP4



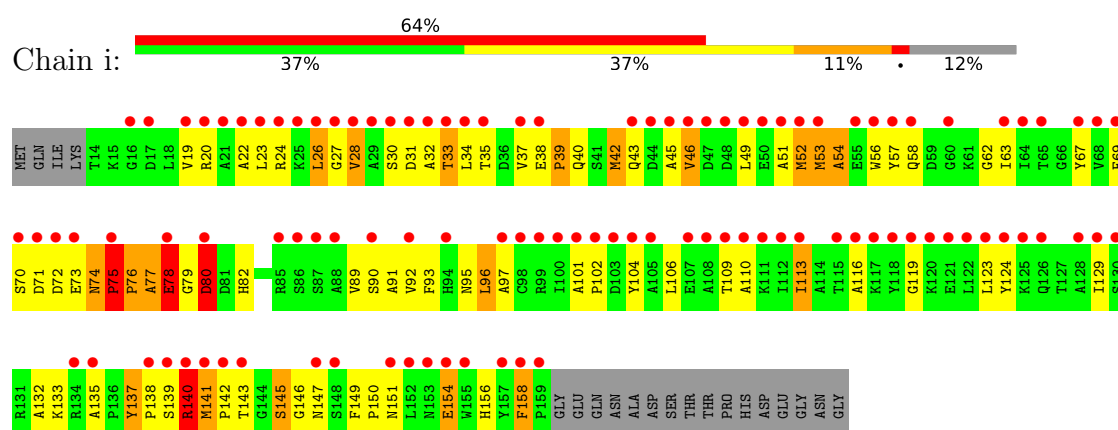
● Molecule 2: PACKAGED DNA STABILIZATION PROTEIN GP4



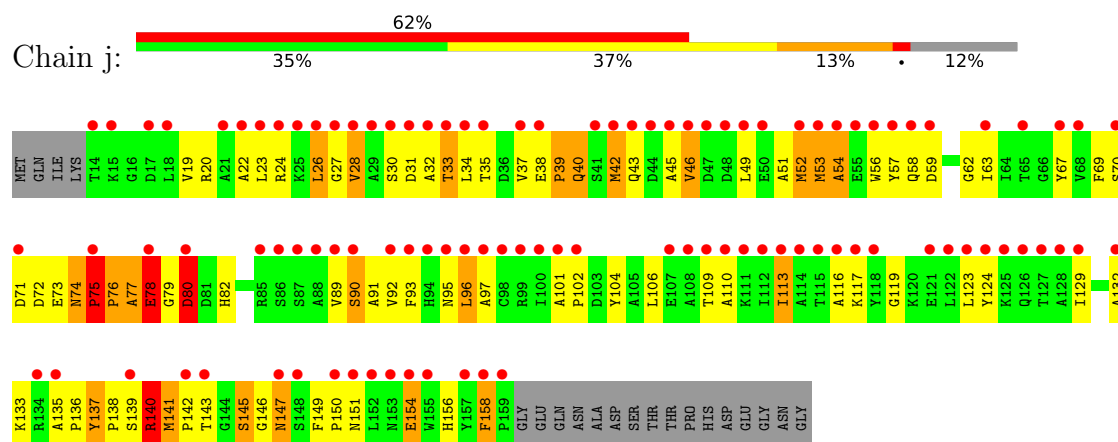
● Molecule 2: PACKAGED DNA STABILIZATION PROTEIN GP4



● Molecule 2: PACKAGED DNA STABILIZATION PROTEIN GP4



● Molecule 2: PACKAGED DNA STABILIZATION PROTEIN GP4



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	170.17Å 253.28Å 282.73Å 90.00° 90.68° 90.00°	Depositor
Resolution (Å)	19.99 – 3.25 19.99 – 3.25	Depositor EDS
% Data completeness (in resolution range)	59.7 (19.99-3.25) 91.9 (19.99-3.25)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.15	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.88 (at 3.26Å)	Xtriage
Refinement program	PHENIX (PHENIX.REFINE: 1.5_2)	Depositor
R, R_{free}	0.222 , 0.236 0.341 , (Not available)	Depositor DCC
R_{free} test set	No test flags present.	wwPDB-VP
Wilson B-factor (Å ²)	65.0	Xtriage
Anisotropy	1.241	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.32 , 115.4	EDS
L-test for twinning ²	$\langle L \rangle = 0.45$, $\langle L^2 \rangle = 0.27$	Xtriage
Estimated twinning fraction	0.025 for h,-k,-l	Xtriage
F_o, F_c correlation	0.80	EDS
Total number of atoms	135120	wwPDB-VP
Average B, all atoms (Å ²)	92.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.69% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality ⓘ

5.1 Standard geometry ⓘ

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 5$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	# Z >5	RMSZ	# Z >5
1	A	0.44	0/4635	0.91	14/6265 (0.2%)
1	B	0.44	0/4635	0.92	14/6265 (0.2%)
1	C	0.44	0/4635	0.91	14/6265 (0.2%)
1	D	0.44	0/4635	0.92	14/6265 (0.2%)
1	E	0.44	0/4635	0.92	14/6265 (0.2%)
1	F	0.44	0/4635	0.92	14/6265 (0.2%)
1	G	0.44	0/4635	0.92	14/6265 (0.2%)
1	H	0.44	0/4635	0.92	14/6265 (0.2%)
1	I	0.44	0/4635	0.91	14/6265 (0.2%)
1	J	0.44	0/4635	0.91	14/6265 (0.2%)
1	K	0.44	0/4635	0.92	14/6265 (0.2%)
1	L	0.44	0/4635	0.92	14/6265 (0.2%)
1	M	0.40	0/4646	0.87	20/6278 (0.3%)
1	N	0.41	0/4646	0.87	20/6278 (0.3%)
1	O	0.41	0/4646	0.87	20/6278 (0.3%)
1	P	0.41	0/4646	0.87	23/6278 (0.4%)
1	Q	0.40	0/4646	0.88	21/6278 (0.3%)
1	R	0.41	0/4646	0.87	24/6278 (0.4%)
1	S	0.40	0/4646	0.87	22/6278 (0.4%)
1	T	0.40	0/4646	0.87	21/6278 (0.3%)
1	U	0.40	0/4646	0.87	18/6278 (0.3%)
1	V	0.40	0/4646	0.88	19/6278 (0.3%)
1	W	0.40	0/4646	0.88	22/6278 (0.4%)
1	X	0.40	0/4646	0.87	21/6278 (0.3%)
2	Y	0.53	0/1067	1.17	10/1452 (0.7%)
2	Z	0.53	0/1067	1.18	10/1452 (0.7%)
2	a	0.53	0/1067	1.18	10/1452 (0.7%)
2	b	0.53	0/1067	1.17	10/1452 (0.7%)
2	c	0.53	0/1067	1.17	10/1452 (0.7%)
2	d	0.53	0/1067	1.18	10/1452 (0.7%)
2	e	0.53	0/1067	1.17	10/1452 (0.7%)
2	f	0.53	0/1067	1.17	10/1452 (0.7%)
2	g	0.53	0/1067	1.17	10/1452 (0.7%)
2	h	0.53	0/1067	1.17	10/1452 (0.7%)

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	# Z >5	RMSZ	# Z >5
2	i	0.53	0/1067	1.18	10/1452 (0.7%)
2	j	0.53	0/1067	1.18	10/1452 (0.7%)
2	k	0.45	0/1071	0.92	5/1455 (0.3%)
2	l	0.45	0/1071	0.92	3/1455 (0.2%)
2	m	0.45	0/1071	0.91	4/1455 (0.3%)
2	n	0.47	0/1071	0.91	3/1455 (0.2%)
2	o	0.47	0/1071	0.91	4/1455 (0.3%)
2	p	0.45	0/1071	0.91	4/1455 (0.3%)
2	q	0.43	0/1071	0.92	5/1455 (0.3%)
2	r	0.46	0/1071	0.91	4/1455 (0.3%)
2	s	0.43	0/1071	0.91	4/1455 (0.3%)
2	t	0.43	0/1071	0.91	3/1455 (0.2%)
2	u	0.44	0/1071	0.93	3/1455 (0.2%)
2	v	0.45	0/1071	0.91	4/1455 (0.3%)
All	All	0.44	0/137028	0.93	585/185400 (0.3%)

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	A	0	1
1	B	0	1
1	C	0	1
1	D	0	1
1	E	0	1
1	F	0	1
1	G	0	1
1	H	0	1
1	I	0	1
1	J	0	1
1	K	0	1
1	L	0	1
1	M	0	1
1	N	0	1
1	O	0	1
1	P	0	1
1	Q	0	1
1	R	0	1
1	S	0	1
1	T	0	1

Continued on next page...

Continued from previous page...

Mol	Chain	#Chirality outliers	#Planarity outliers
1	U	0	1
1	V	0	1
1	W	0	1
1	X	0	1
All	All	0	24

There are no bond length outliers.

The worst 5 of 585 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	i	80	ASP	N-CA-CB	18.38	141.54	110.49
2	d	80	ASP	N-CA-CB	18.36	141.52	110.49
2	Z	80	ASP	N-CA-CB	18.36	141.51	110.49
2	Y	80	ASP	N-CA-CB	18.35	141.50	110.49
2	e	80	ASP	N-CA-CB	18.35	141.50	110.49

There are no chirality outliers.

5 of 24 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	M	231	THR	Peptide
1	N	231	THR	Peptide
1	O	231	THR	Peptide
1	P	231	THR	Peptide
1	Q	231	THR	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	A	4553	0	4351	447	0
1	B	4553	0	4351	450	2
1	C	4553	0	4351	449	4
1	D	4553	0	4351	450	0
1	E	4553	0	4351	452	0
1	F	4553	0	4351	449	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	G	4553	0	4351	453	0
1	H	4553	0	4351	449	0
1	I	4553	0	4351	452	0
1	J	4553	0	4351	451	0
1	K	4553	0	4351	450	0
1	L	4553	0	4351	457	0
1	M	4564	0	4368	372	0
1	N	4564	0	4368	387	5
1	O	4564	0	4368	376	0
1	P	4564	0	4368	381	0
1	Q	4564	0	4368	384	0
1	R	4564	0	4368	392	0
1	S	4564	0	4368	376	3
1	T	4564	0	4368	373	0
1	U	4564	0	4368	386	0
1	V	4564	0	4368	377	0
1	W	4564	0	4368	381	1
1	X	4564	0	4368	378	0
2	Y	1048	0	957	117	0
2	Z	1048	0	957	115	0
2	a	1048	0	957	118	0
2	b	1048	0	957	116	0
2	c	1048	0	957	117	4
2	d	1048	0	957	120	0
2	e	1048	0	957	114	0
2	f	1048	0	957	117	0
2	g	1048	0	957	120	1
2	h	1048	0	957	122	0
2	i	1048	0	957	116	0
2	j	1048	0	957	126	0
2	k	1052	0	975	65	0
2	l	1052	0	975	67	0
2	m	1052	0	975	69	0
2	n	1052	0	975	74	0
2	o	1052	0	975	74	0
2	p	1052	0	975	65	0
2	q	1052	0	975	69	0
2	r	1052	0	975	70	3
2	s	1052	0	975	69	0
2	t	1052	0	975	71	0
2	u	1052	0	975	72	0
2	v	1052	0	975	73	3

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	A	22	0	0	1	0
3	B	22	0	0	1	0
3	C	22	0	0	1	0
3	D	22	0	0	1	0
3	E	22	0	0	2	0
3	F	22	0	0	2	0
3	G	22	0	0	1	0
3	H	22	0	0	2	0
3	I	22	0	0	1	0
3	J	22	0	0	2	0
3	K	22	0	0	2	0
3	L	22	0	0	2	0
3	M	21	0	0	4	0
3	N	21	0	0	4	0
3	O	21	0	0	3	0
3	P	21	0	0	5	0
3	Q	21	0	0	5	0
3	R	21	0	0	4	0
3	S	21	0	0	4	0
3	T	21	0	0	5	0
3	U	21	0	0	3	0
3	V	21	0	0	4	0
3	W	21	0	0	4	0
3	X	21	0	0	4	0
All	All	135120	0	127812	11405	13

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 43.

The worst 5 of 11405 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:Z:53:MET:HE1	2:Z:92:VAL:CG1	1.60	1.32
2:e:53:MET:HE1	2:e:92:VAL:CG1	1.60	1.32
2:c:53:MET:HE1	2:c:92:VAL:CG1	1.60	1.32
2:Y:53:MET:HE1	2:Y:92:VAL:CG1	1.60	1.31
2:d:53:MET:HE1	2:d:92:VAL:CG1	1.60	1.31

The worst 5 of 13 symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:230:GLU:OE2	2:c:74:ASN:ND2[2_546]	1.06	1.14
1:N:230:GLU:N	2:r:64:GLU:OE2[2_455]	1.79	0.41
1:N:260:ASP:OD2	1:B:601:GLN:NE2[2_555]	1.79	0.41
1:C:250:ASP:OD2	2:c:74:ASN:OD1[2_546]	1.83	0.37
1:N:228:LYS:CG	2:r:64:GLU:OE1[2_455]	1.89	0.31

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 X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	565/602 (94%)	434 (77%)	88 (16%)	43 (8%)	1	5
1	B	565/602 (94%)	434 (77%)	87 (15%)	44 (8%)	1	4
1	C	565/602 (94%)	434 (77%)	87 (15%)	44 (8%)	1	4
1	D	565/602 (94%)	434 (77%)	88 (16%)	43 (8%)	1	5
1	E	565/602 (94%)	434 (77%)	88 (16%)	43 (8%)	1	5
1	F	565/602 (94%)	434 (77%)	88 (16%)	43 (8%)	1	5
1	G	565/602 (94%)	434 (77%)	88 (16%)	43 (8%)	1	5
1	H	565/602 (94%)	434 (77%)	88 (16%)	43 (8%)	1	5
1	I	565/602 (94%)	434 (77%)	88 (16%)	43 (8%)	1	5
1	J	565/602 (94%)	434 (77%)	88 (16%)	43 (8%)	1	5
1	K	565/602 (94%)	434 (77%)	88 (16%)	43 (8%)	1	5
1	L	565/602 (94%)	434 (77%)	87 (15%)	44 (8%)	1	4
1	M	565/602 (94%)	446 (79%)	81 (14%)	38 (7%)	1	6
1	N	565/602 (94%)	447 (79%)	79 (14%)	39 (7%)	1	6
1	O	565/602 (94%)	447 (79%)	80 (14%)	38 (7%)	1	6
1	P	565/602 (94%)	448 (79%)	79 (14%)	38 (7%)	1	6
1	Q	565/602 (94%)	447 (79%)	79 (14%)	39 (7%)	1	6
1	R	565/602 (94%)	448 (79%)	79 (14%)	38 (7%)	1	6

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	S	565/602 (94%)	443 (78%)	83 (15%)	39 (7%)	1	6
1	T	565/602 (94%)	446 (79%)	81 (14%)	38 (7%)	1	6
1	U	565/602 (94%)	447 (79%)	79 (14%)	39 (7%)	1	6
1	V	565/602 (94%)	445 (79%)	82 (14%)	38 (7%)	1	6
1	W	565/602 (94%)	448 (79%)	78 (14%)	39 (7%)	1	6
1	X	565/602 (94%)	448 (79%)	78 (14%)	39 (7%)	1	6
2	Y	144/166 (87%)	115 (80%)	16 (11%)	13 (9%)	0	3
2	Z	144/166 (87%)	115 (80%)	16 (11%)	13 (9%)	0	3
2	a	144/166 (87%)	115 (80%)	16 (11%)	13 (9%)	0	3
2	b	144/166 (87%)	115 (80%)	16 (11%)	13 (9%)	0	3
2	c	144/166 (87%)	115 (80%)	16 (11%)	13 (9%)	0	3
2	d	144/166 (87%)	115 (80%)	16 (11%)	13 (9%)	0	3
2	e	144/166 (87%)	115 (80%)	16 (11%)	13 (9%)	0	3
2	f	144/166 (87%)	115 (80%)	16 (11%)	13 (9%)	0	3
2	g	144/166 (87%)	115 (80%)	16 (11%)	13 (9%)	0	3
2	h	144/166 (87%)	115 (80%)	16 (11%)	13 (9%)	0	3
2	i	144/166 (87%)	115 (80%)	16 (11%)	13 (9%)	0	3
2	j	144/166 (87%)	115 (80%)	16 (11%)	13 (9%)	0	3
2	k	143/166 (86%)	115 (80%)	17 (12%)	11 (8%)	1	5
2	l	143/166 (86%)	114 (80%)	18 (13%)	11 (8%)	1	5
2	m	143/166 (86%)	115 (80%)	18 (13%)	10 (7%)	1	6
2	n	143/166 (86%)	114 (80%)	18 (13%)	11 (8%)	1	5
2	o	143/166 (86%)	115 (80%)	17 (12%)	11 (8%)	1	5
2	p	143/166 (86%)	115 (80%)	17 (12%)	11 (8%)	1	5
2	q	143/166 (86%)	113 (79%)	19 (13%)	11 (8%)	1	5
2	r	143/166 (86%)	114 (80%)	18 (13%)	11 (8%)	1	5
2	s	143/166 (86%)	114 (80%)	18 (13%)	11 (8%)	1	5
2	t	143/166 (86%)	113 (79%)	19 (13%)	11 (8%)	1	5
2	u	143/166 (86%)	114 (80%)	18 (13%)	11 (8%)	1	5
2	v	143/166 (86%)	114 (80%)	19 (13%)	10 (7%)	1	6
All	All	17004/18432 (92%)	13318 (78%)	2419 (14%)	1267 (8%)	1	5

5 of 1267 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	M	82	PRO
1	M	263	PHE
1	M	294	ALA
1	M	462	GLU
1	M	514	TYR

5.3.2 Protein sidechains ⓘ

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	483/510 (95%)	409 (85%)	74 (15%)	3	12
1	B	483/510 (95%)	409 (85%)	74 (15%)	3	12
1	C	483/510 (95%)	409 (85%)	74 (15%)	3	12
1	D	483/510 (95%)	409 (85%)	74 (15%)	3	12
1	E	483/510 (95%)	409 (85%)	74 (15%)	3	12
1	F	483/510 (95%)	409 (85%)	74 (15%)	3	12
1	G	483/510 (95%)	409 (85%)	74 (15%)	3	12
1	H	483/510 (95%)	409 (85%)	74 (15%)	3	12
1	I	483/510 (95%)	409 (85%)	74 (15%)	3	12
1	J	483/510 (95%)	409 (85%)	74 (15%)	3	12
1	K	483/510 (95%)	409 (85%)	74 (15%)	3	12
1	L	483/510 (95%)	409 (85%)	74 (15%)	3	12
1	M	485/510 (95%)	413 (85%)	72 (15%)	3	13
1	N	485/510 (95%)	414 (85%)	71 (15%)	3	14
1	O	485/510 (95%)	414 (85%)	71 (15%)	3	14
1	P	485/510 (95%)	413 (85%)	72 (15%)	3	13
1	Q	485/510 (95%)	413 (85%)	72 (15%)	3	13
1	R	485/510 (95%)	413 (85%)	72 (15%)	3	13
1	S	485/510 (95%)	413 (85%)	72 (15%)	3	13

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	T	485/510 (95%)	412 (85%)	73 (15%)	3	13
1	U	485/510 (95%)	412 (85%)	73 (15%)	3	13
1	V	485/510 (95%)	413 (85%)	72 (15%)	3	13
1	W	485/510 (95%)	412 (85%)	73 (15%)	3	13
1	X	485/510 (95%)	412 (85%)	73 (15%)	3	13
2	Y	98/132 (74%)	83 (85%)	15 (15%)	3	12
2	Z	98/132 (74%)	83 (85%)	15 (15%)	3	12
2	a	98/132 (74%)	83 (85%)	15 (15%)	3	12
2	b	98/132 (74%)	83 (85%)	15 (15%)	3	12
2	c	98/132 (74%)	83 (85%)	15 (15%)	3	12
2	d	98/132 (74%)	83 (85%)	15 (15%)	3	12
2	e	98/132 (74%)	83 (85%)	15 (15%)	3	12
2	f	98/132 (74%)	83 (85%)	15 (15%)	3	12
2	g	98/132 (74%)	83 (85%)	15 (15%)	3	12
2	h	98/132 (74%)	83 (85%)	15 (15%)	3	12
2	i	98/132 (74%)	83 (85%)	15 (15%)	3	12
2	j	98/132 (74%)	83 (85%)	15 (15%)	3	12
2	k	100/132 (76%)	85 (85%)	15 (15%)	3	13
2	l	100/132 (76%)	84 (84%)	16 (16%)	2	11
2	m	100/132 (76%)	85 (85%)	15 (15%)	3	13
2	n	100/132 (76%)	85 (85%)	15 (15%)	3	13
2	o	100/132 (76%)	84 (84%)	16 (16%)	2	11
2	p	100/132 (76%)	85 (85%)	15 (15%)	3	13
2	q	100/132 (76%)	85 (85%)	15 (15%)	3	13
2	r	100/132 (76%)	85 (85%)	15 (15%)	3	13
2	s	100/132 (76%)	85 (85%)	15 (15%)	3	13
2	t	100/132 (76%)	85 (85%)	15 (15%)	3	13
2	u	100/132 (76%)	83 (83%)	17 (17%)	2	9
2	v	100/132 (76%)	84 (84%)	16 (16%)	2	11
All	All	13992/15408 (91%)	11873 (85%)	2119 (15%)	3	13

5 of 2119 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	K	507	LEU
1	L	301	VAL
1	K	438	LEU
2	h	78	GLU
1	W	389	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 447 such sidechains are listed below:

Mol	Chain	Res	Type
1	D	236	GLN
2	i	94	HIS
1	F	575	GLN
2	f	94	HIS
1	K	530	GLN

5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates [i](#)

There are no oligosaccharides in this entry.

5.6 Ligand geometry [i](#)

There are no ligands in this entry.

5.7 Other polymers [i](#)

There are no such residues in this entry.

5.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

6 Fit of model and data ⓘ

6.1 Protein, DNA and RNA chains ⓘ

In the following table, the column labelled ‘#RSRZ > 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95th percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled ‘Q < 0.9’ lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å ²)	Q<0.9
1	A	552/602 (91%)	2.87	373 (67%) 0 0	22, 84, 246, 379	0
1	B	552/602 (91%)	2.80	368 (66%) 0 0	23, 85, 246, 379	0
1	C	552/602 (91%)	2.82	376 (68%) 0 0	26, 84, 246, 379	0
1	D	552/602 (91%)	2.66	352 (63%) 0 0	25, 84, 246, 379	0
1	E	552/602 (91%)	2.74	357 (64%) 0 0	25, 84, 246, 379	0
1	F	552/602 (91%)	2.73	358 (64%) 0 0	25, 84, 246, 379	0
1	G	552/602 (91%)	2.79	360 (65%) 0 0	30, 84, 246, 379	0
1	H	552/602 (91%)	2.66	341 (61%) 0 0	26, 86, 246, 379	0
1	I	552/602 (91%)	2.80	378 (68%) 0 0	31, 88, 246, 379	0
1	J	552/602 (91%)	3.06	397 (71%) 0 0	33, 96, 246, 379	0
1	K	552/602 (91%)	2.99	393 (71%) 0 0	26, 88, 245, 379	0
1	L	552/602 (91%)	3.01	391 (70%) 0 0	24, 87, 246, 379	0
1	M	552/602 (91%)	2.34	298 (53%) 0 0	15, 70, 209, 327	0
1	N	551/602 (91%)	2.41	304 (55%) 0 0	18, 69, 209, 327	0
1	O	552/602 (91%)	2.49	316 (57%) 0 0	15, 69, 209, 326	0
1	P	552/602 (91%)	2.50	309 (55%) 0 0	15, 70, 209, 327	0
1	Q	552/602 (91%)	2.47	317 (57%) 0 0	15, 69, 209, 326	0
1	R	550/602 (91%)	2.55	322 (58%) 0 0	17, 69, 209, 327	0
1	S	552/602 (91%)	2.37	284 (51%) 0 0	19, 70, 209, 326	0
1	T	552/602 (91%)	2.35	303 (54%) 0 0	15, 70, 208, 326	0
1	U	552/602 (91%)	2.59	338 (61%) 0 0	17, 69, 209, 327	0
1	V	552/602 (91%)	2.42	312 (56%) 0 0	18, 70, 209, 327	0
1	W	552/602 (91%)	2.51	312 (56%) 0 0	17, 70, 209, 326	0
1	X	552/602 (91%)	2.45	322 (58%) 0 0	14, 69, 209, 326	0

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å ²)	Q<0.9
2	Y	146/166 (87%)	2.72	99 (67%) 0 0	38, 75, 221, 304	0
2	Z	146/166 (87%)	2.76	102 (69%) 0 0	37, 73, 222, 304	0
2	a	146/166 (87%)	2.61	87 (59%) 0 0	37, 77, 221, 304	0
2	b	146/166 (87%)	2.71	94 (64%) 0 0	39, 72, 221, 304	0
2	c	146/166 (87%)	2.93	109 (74%) 0 0	39, 72, 222, 305	0
2	d	146/166 (87%)	2.98	106 (72%) 0 0	35, 72, 222, 304	0
2	e	146/166 (87%)	2.69	91 (62%) 0 0	37, 73, 220, 304	0
2	f	146/166 (87%)	2.94	109 (74%) 0 0	39, 74, 222, 304	0
2	g	146/166 (87%)	3.02	101 (69%) 0 0	44, 75, 221, 304	0
2	h	146/166 (87%)	2.82	102 (69%) 0 0	38, 75, 221, 305	0
2	i	146/166 (87%)	3.03	106 (72%) 0 0	40, 75, 221, 304	0
2	j	146/166 (87%)	2.83	103 (70%) 0 0	36, 75, 222, 305	0
2	k	145/166 (87%)	2.29	74 (51%) 0 0	25, 57, 204, 314	0
2	l	145/166 (87%)	2.40	84 (57%) 0 0	28, 56, 204, 314	0
2	m	145/166 (87%)	2.35	75 (51%) 0 0	30, 59, 203, 314	0
2	n	145/166 (87%)	2.69	88 (60%) 0 0	23, 55, 204, 314	0
2	o	145/166 (87%)	2.41	85 (58%) 0 0	26, 57, 204, 314	0
2	p	145/166 (87%)	2.30	77 (53%) 0 0	27, 57, 204, 314	0
2	q	145/166 (87%)	2.41	78 (53%) 0 0	29, 58, 204, 314	0
2	r	145/166 (87%)	2.27	72 (49%) 0 0	26, 57, 205, 314	0
2	s	145/166 (87%)	2.10	63 (43%) 0 1	30, 58, 204, 314	0
2	t	145/166 (87%)	2.17	66 (45%) 0 0	31, 58, 204, 314	0
2	u	145/166 (87%)	2.12	63 (43%) 0 1	30, 58, 204, 314	0
2	v	145/166 (87%)	2.23	75 (51%) 0 0	29, 57, 204, 314	0
All	All	16737/18432 (90%)	2.63	10290 (61%) 0 0	14, 76, 228, 379	0

The worst 5 of 10290 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	c	28	VAL	16.7
1	E	32	ASP	15.0
1	F	32	ASP	14.7
1	L	183	GLY	13.9
1	E	13	SER	12.8

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

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

6.3 Carbohydrates [i](#)

There are no oligosaccharides in this entry.

6.4 Ligands [i](#)

There are no ligands in this entry.

6.5 Other polymers [i](#)

There are no such residues in this entry.