



wwPDB X-ray Structure Validation Summary Report

Jun 16, 2014 – 08:44 PM BST

PDB ID : 4V4K
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 validation summary report for a publicly released PDB entry.
We welcome your comments at validation@mail.wwpdb.org
A user guide is available at <http://wwpdb.org/ValidationPDFNotes.html>

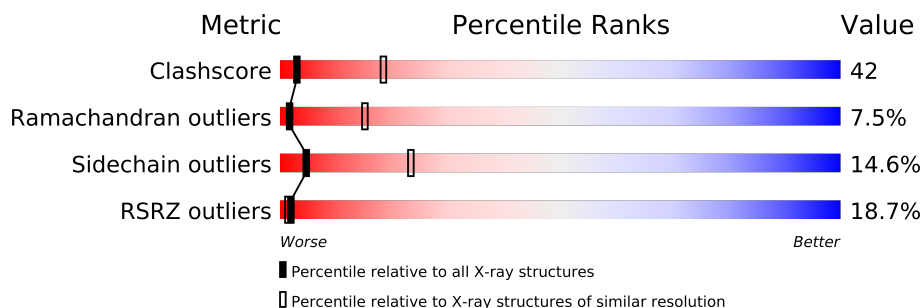
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity : 4.02b-467
Mogul : 1.16 November 2013
Xtriage (Phenix) : dev-1323
EDS : stable23397
Percentile statistics : 21963
Refmac : 5.8.0049
CCP4 : 6.3.0 (Settle)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP) : stable23397

1 Overall quality at a glance

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	79885	1374 (3.32-3.20)
Ramachandran outliers	78287	1348 (3.32-3.20)
Sidechain outliers	78261	1346 (3.32-3.20)
RSRZ outliers	66119	1086 (3.32-3.20)

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 on the lower bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	602	
1	B	602	
1	C	602	
1	D	602	
1	E	602	
1	F	602	
1	G	602	
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	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
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	
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 hydrogen and 0 are deuterium.

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	UNP P26746

Continued on next page...

Continued from previous page...

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

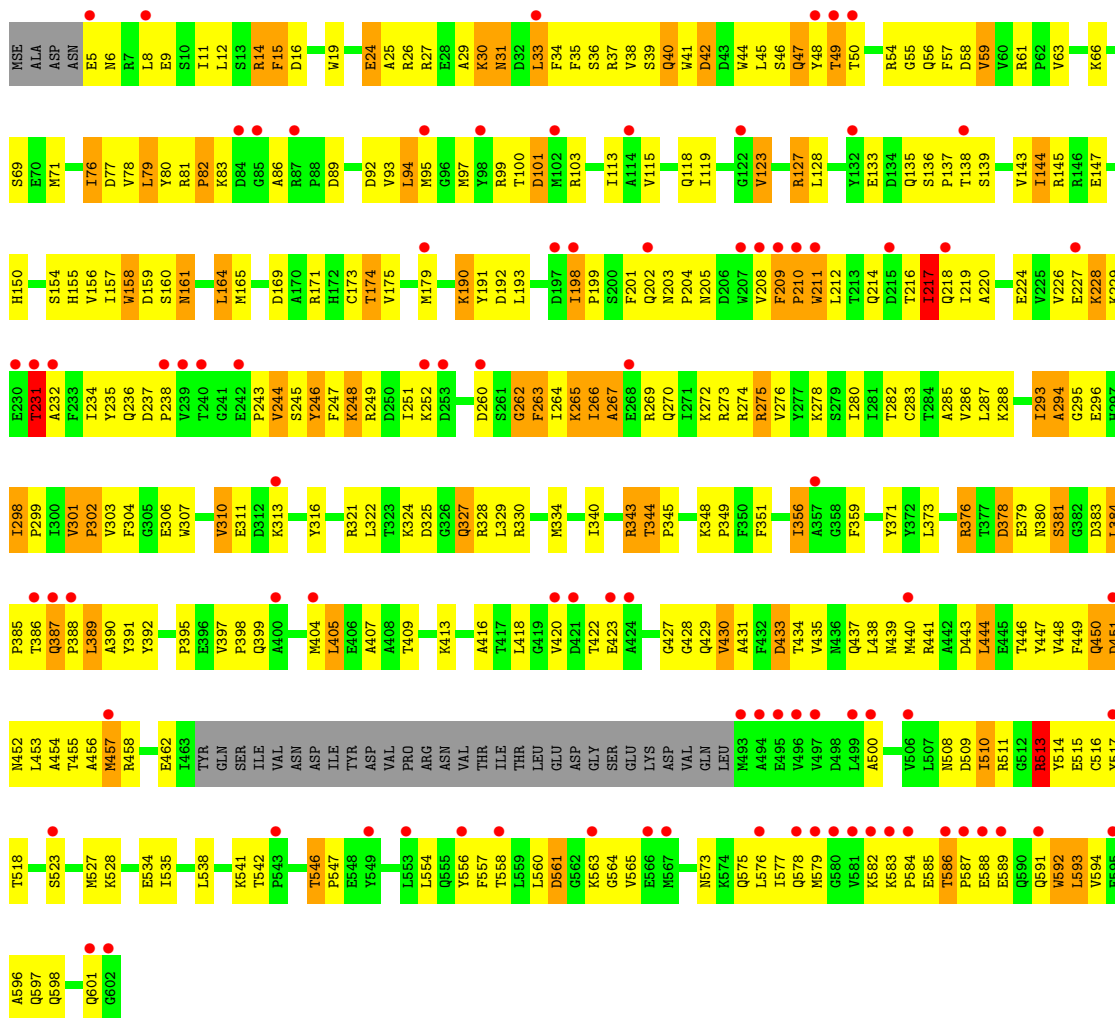
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	C	22	Total 22	O 22	0	0
3	D	22	Total 22	O 22	0	0
3	E	22	Total 22	O 22	0	0
3	F	22	Total 22	O 22	0	0
3	G	22	Total 22	O 22	0	0
3	H	22	Total 22	O 22	0	0
3	I	22	Total 22	O 22	0	0
3	J	22	Total 22	O 22	0	0
3	K	22	Total 22	O 22	0	0
3	L	22	Total 22	O 22	0	0

3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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.

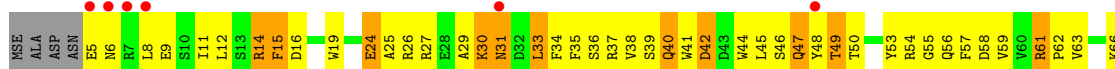
• Molecule 1: PORTAL PROTEIN

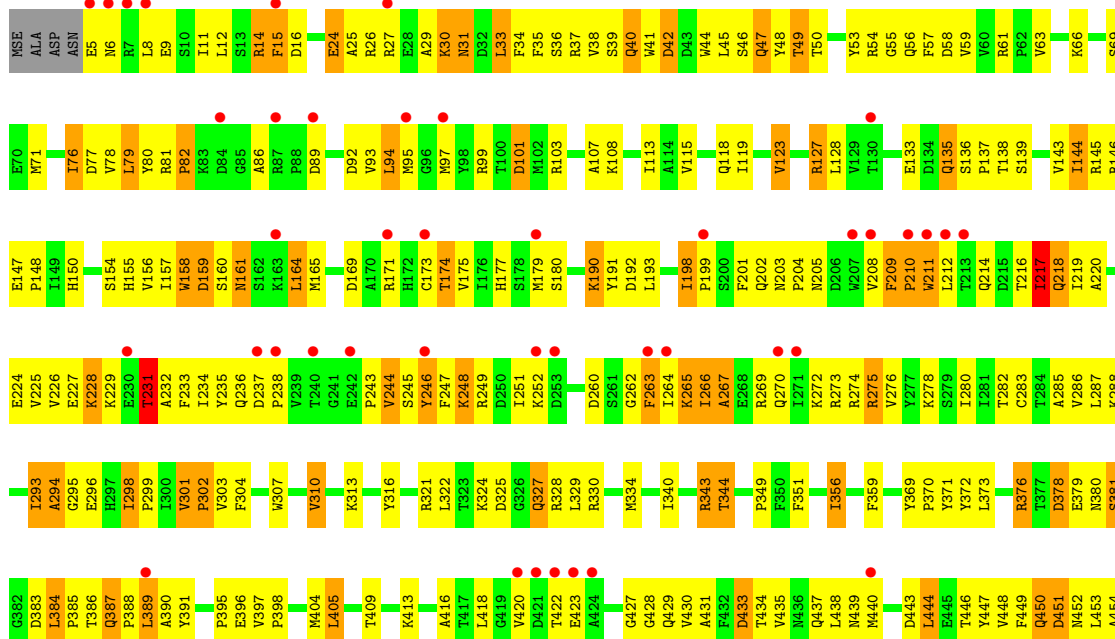
Chain M: 

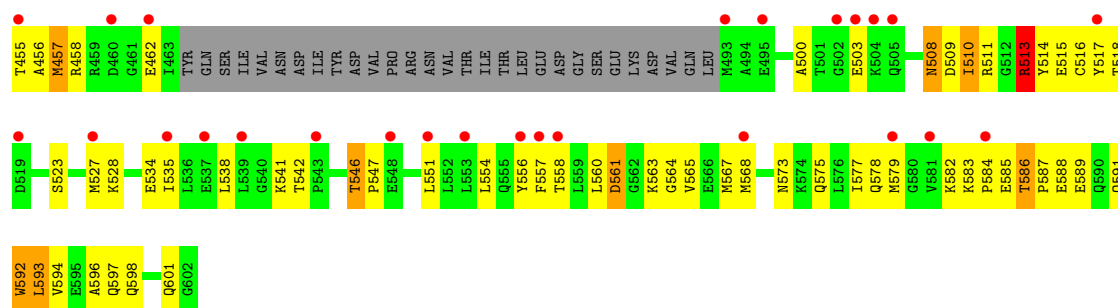


• Molecule 1: PORTAL PROTEIN

Chain N: 

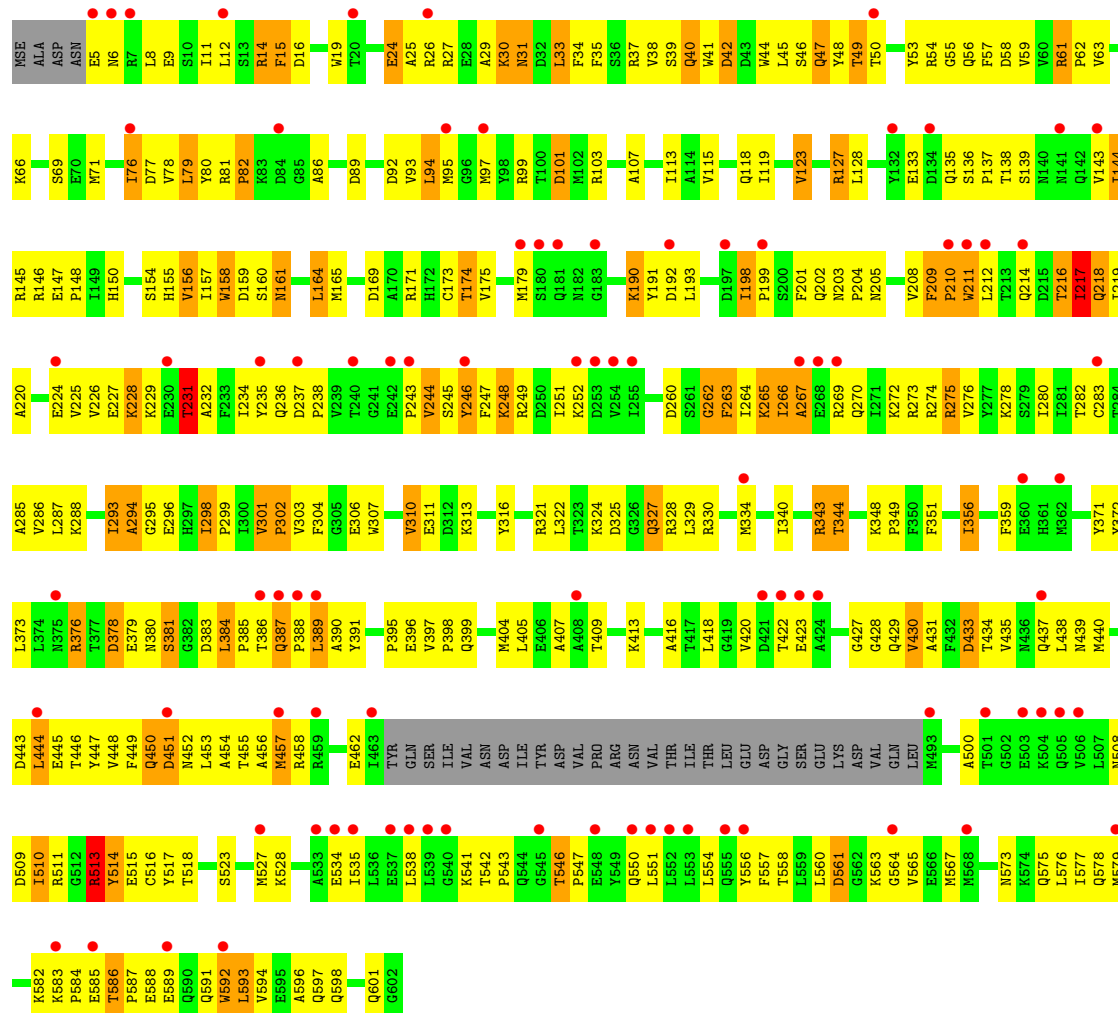






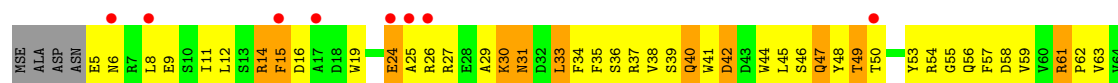
• Molecule 1: PORTAL PROTEIN

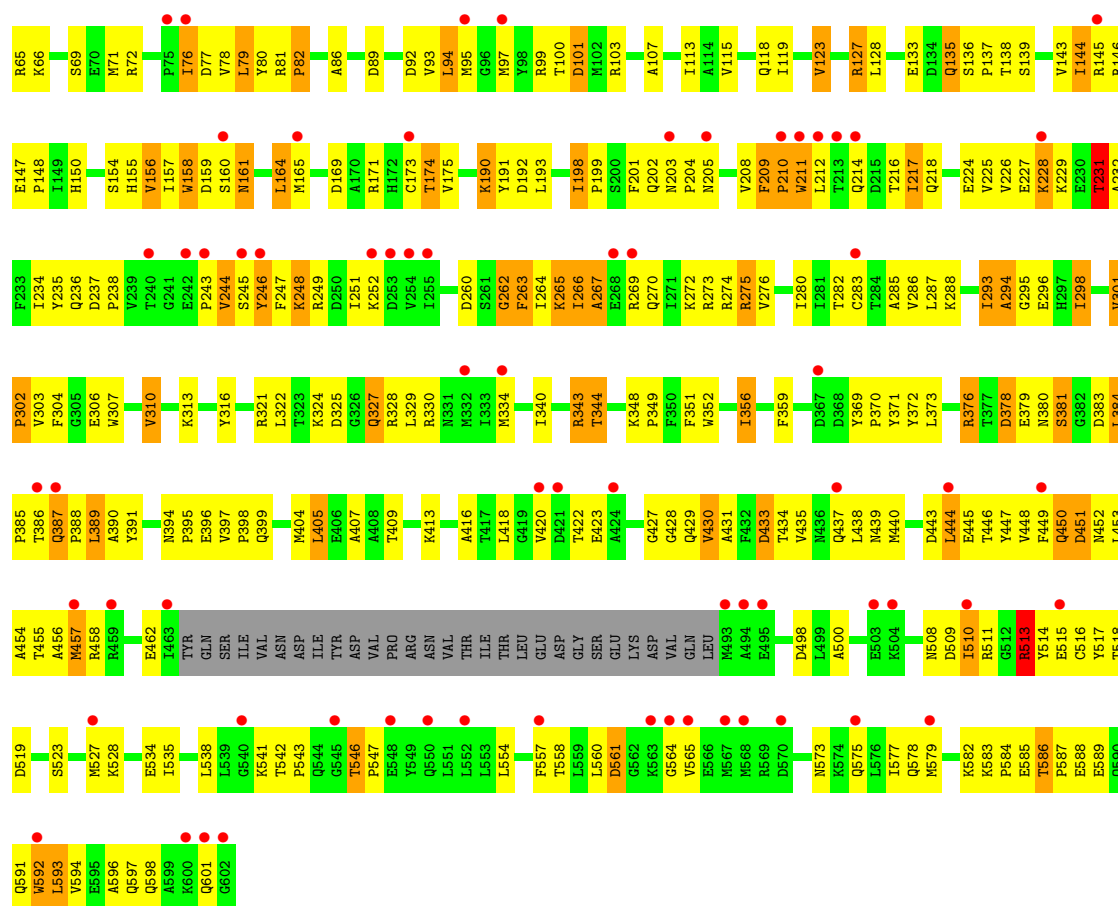
Chain P:



• Molecule 1: PORTAL PROTEIN

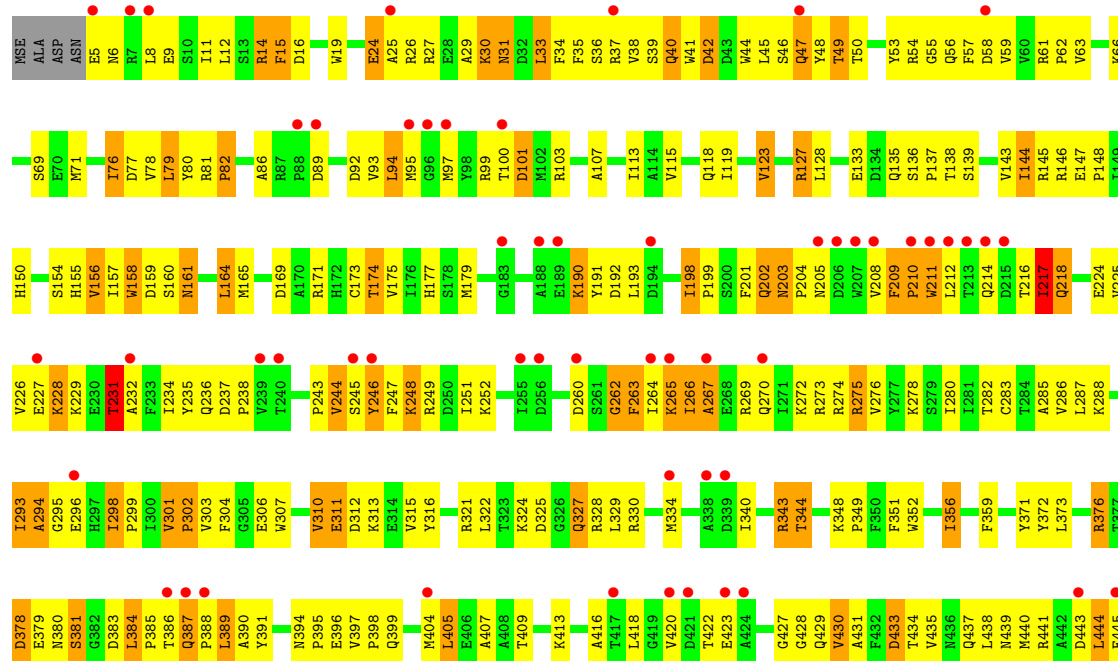
Chain Q:

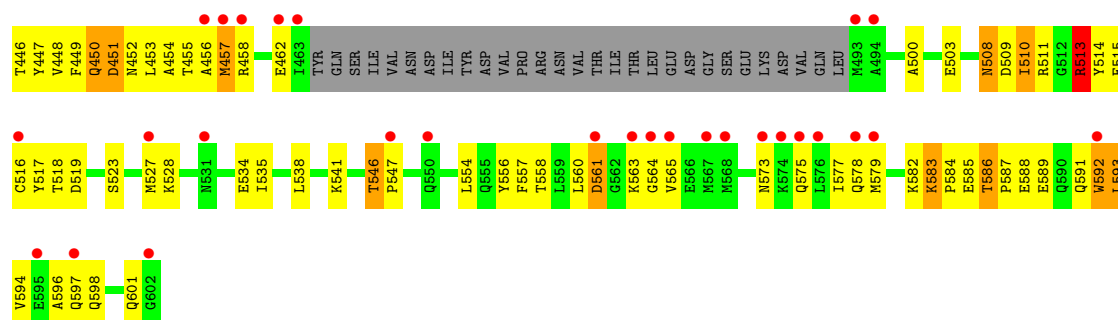




• Molecule 1: PORTAL PROTEIN

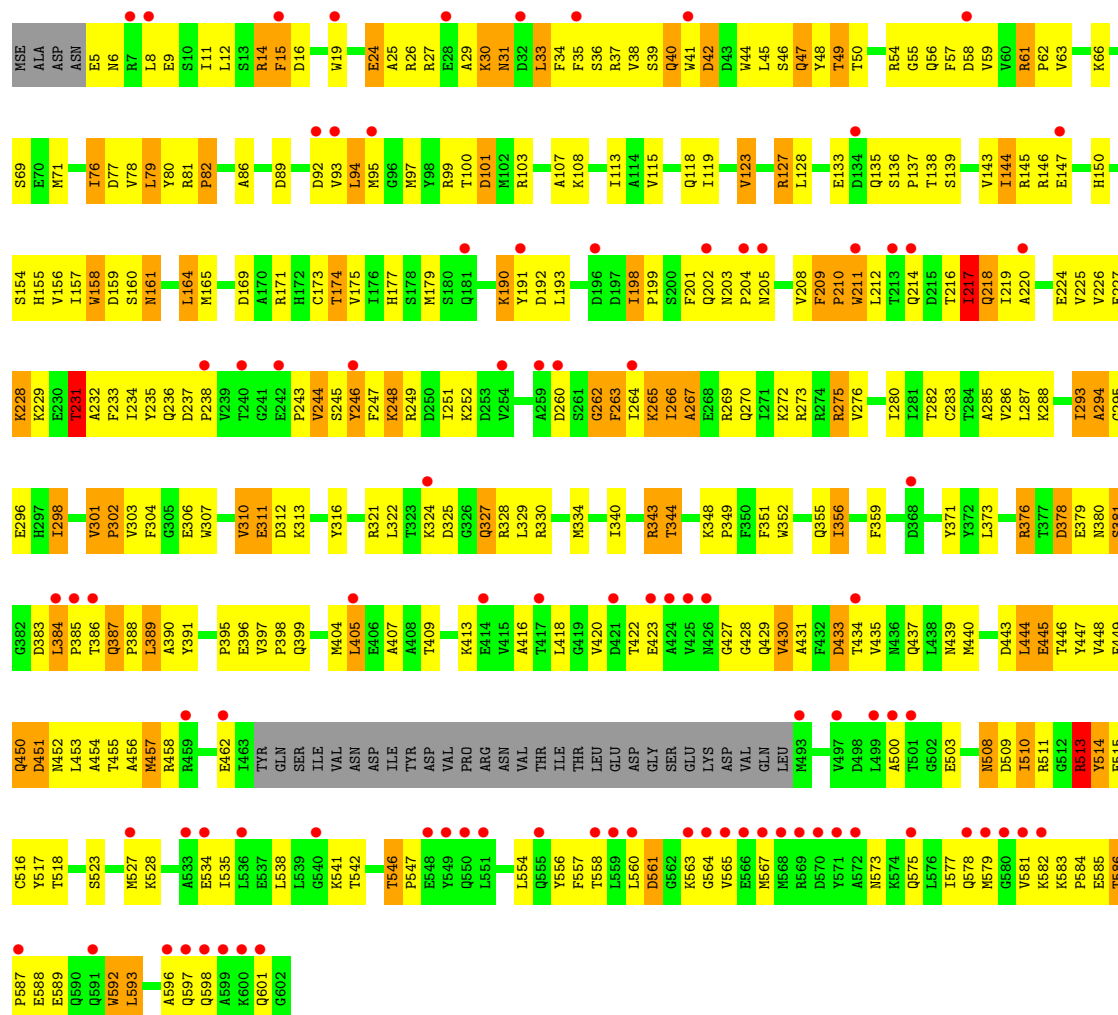
Chain R:





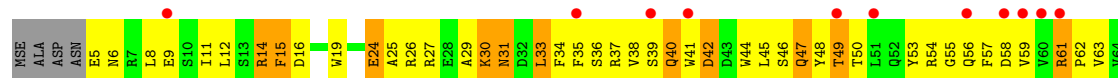
• Molecule 1: PORTAL PROTEIN

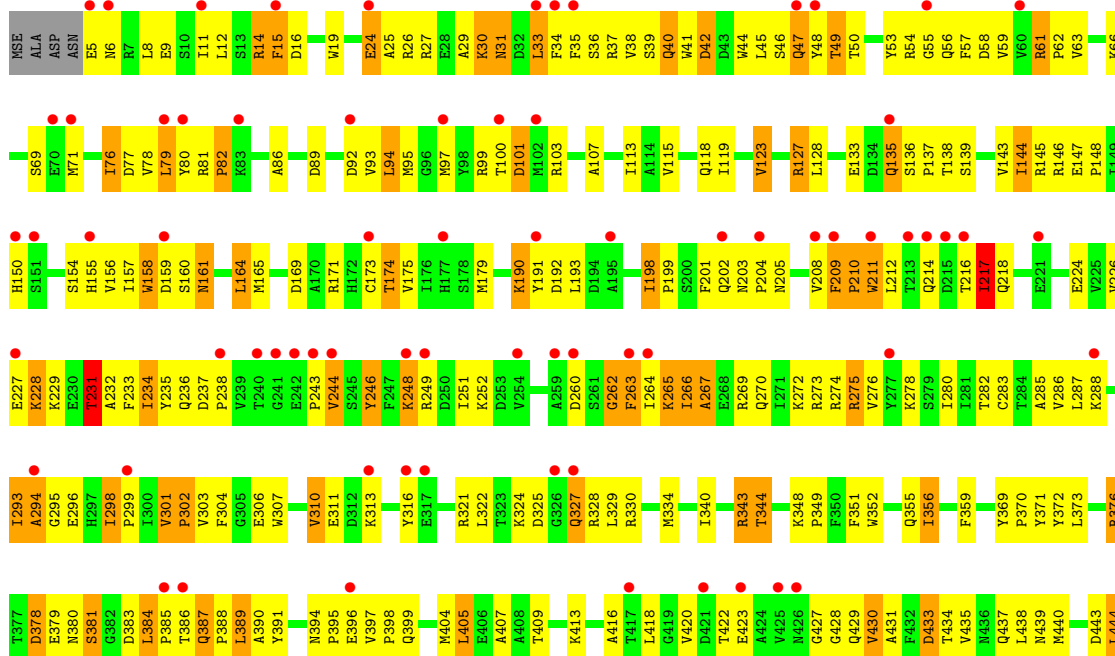
Chain S:

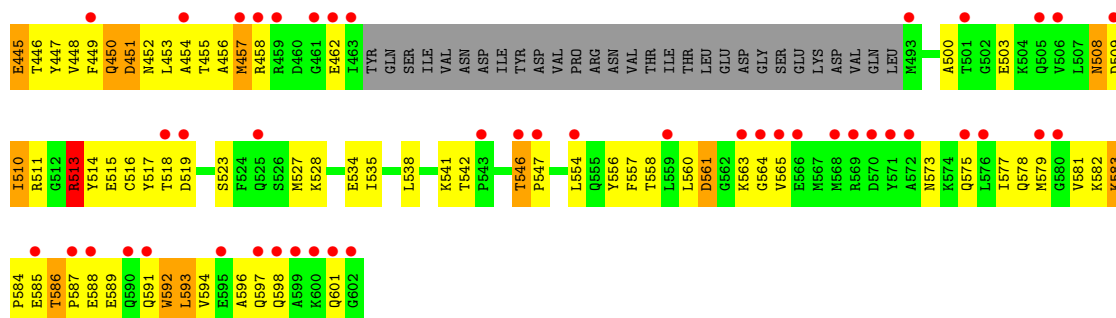


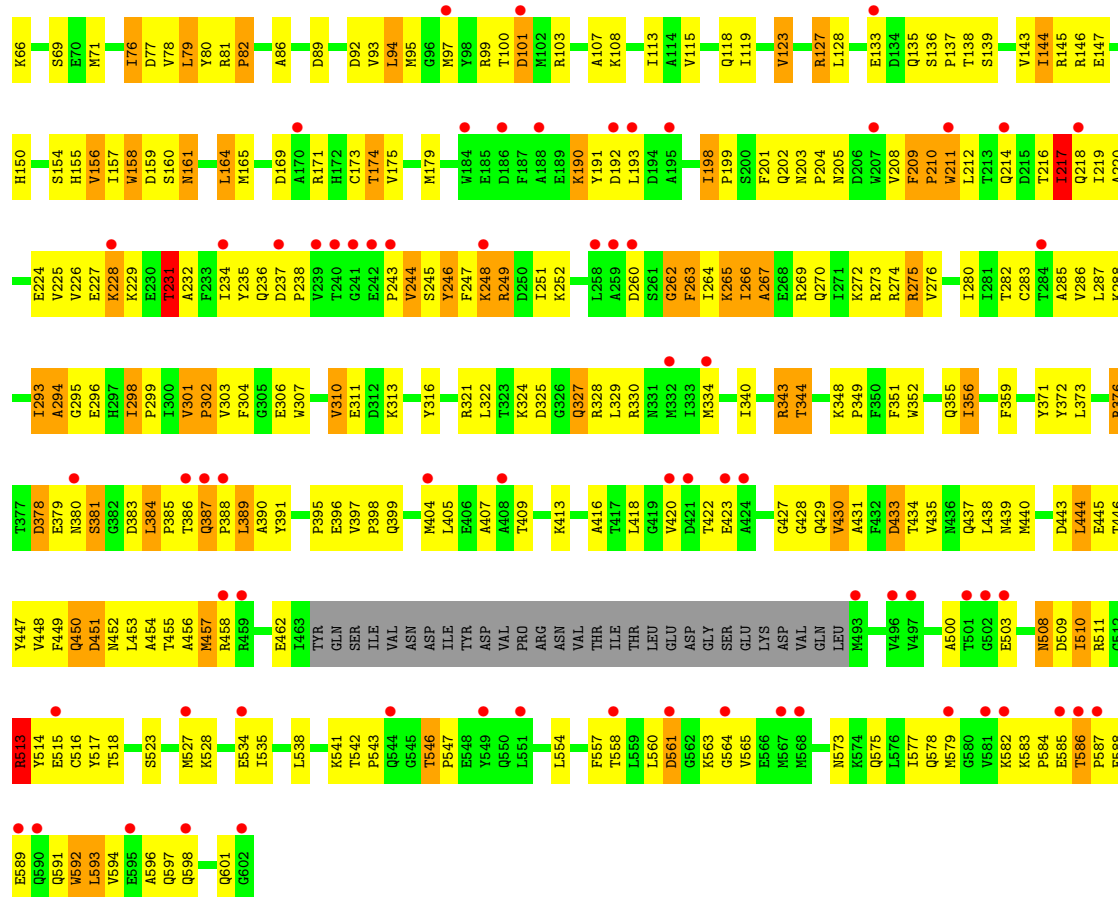
• Molecule 1: PORTAL PROTEIN

Chain T:



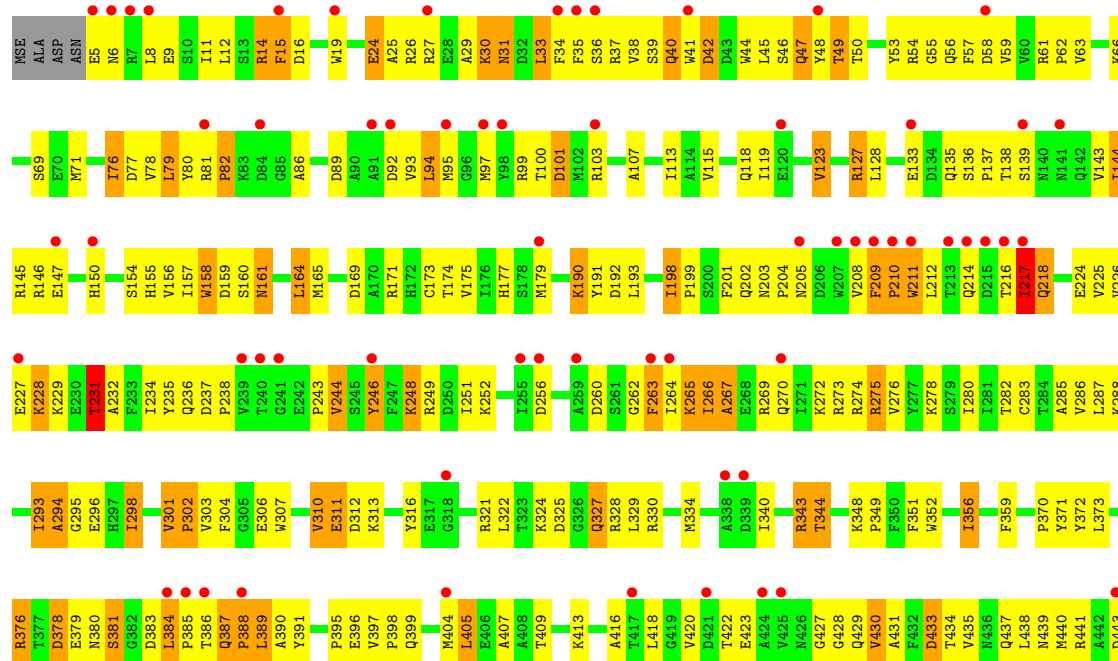


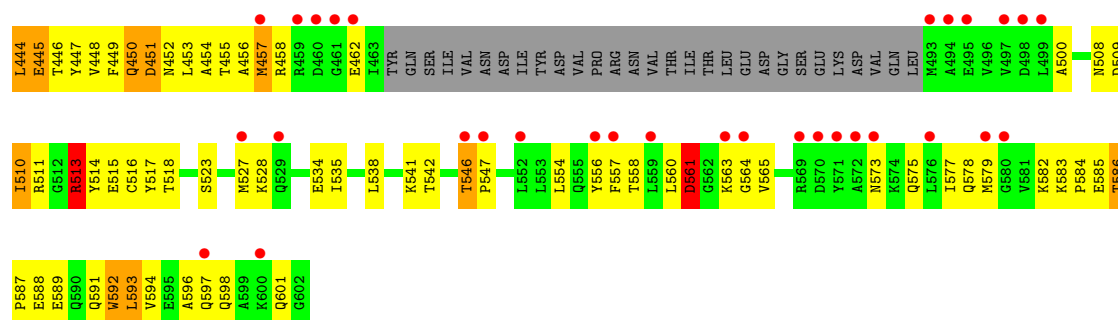




Molecule 1: PORTAL PROTEIN

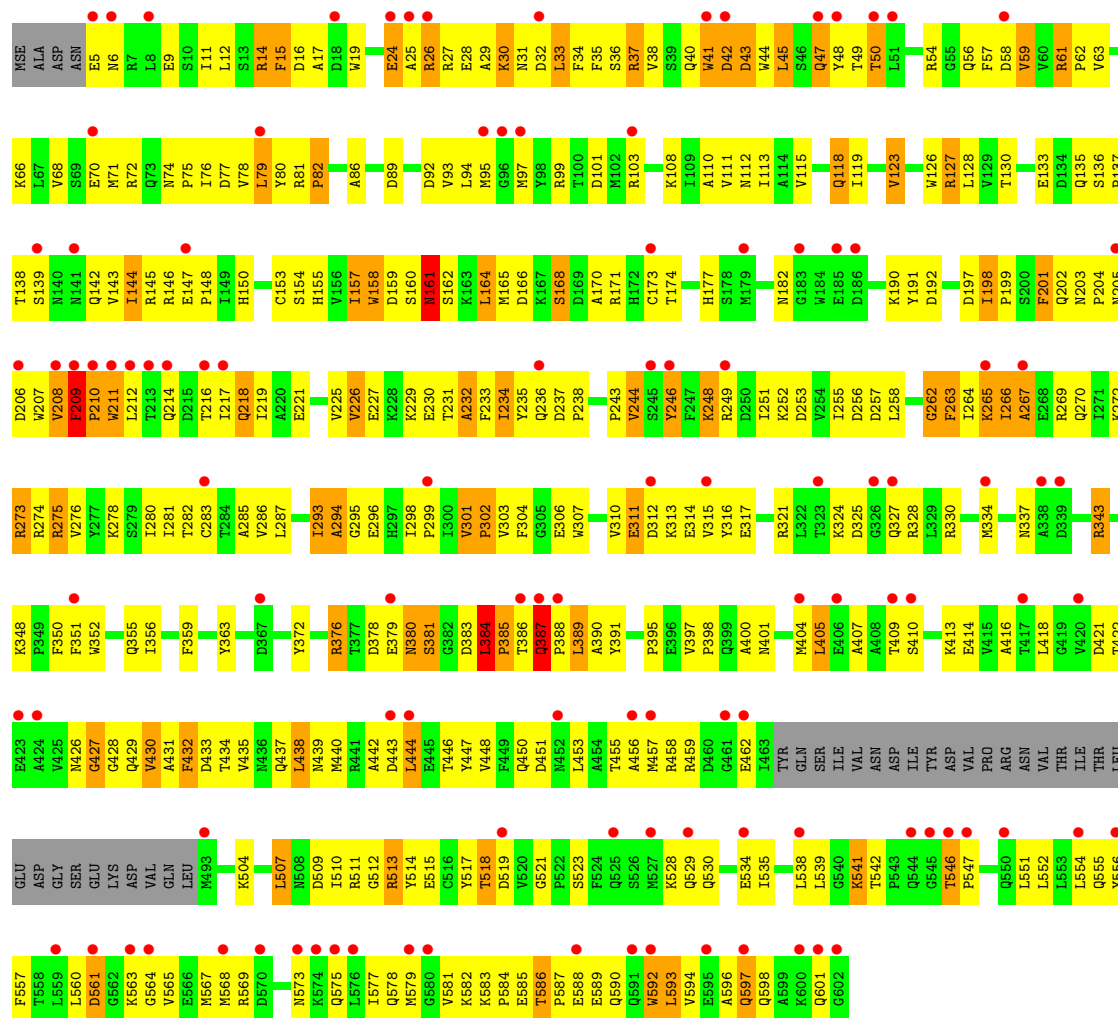
Chain X:





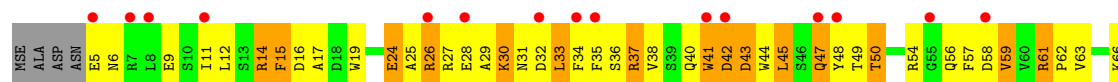
• Molecule 1: PORTAL PROTEIN

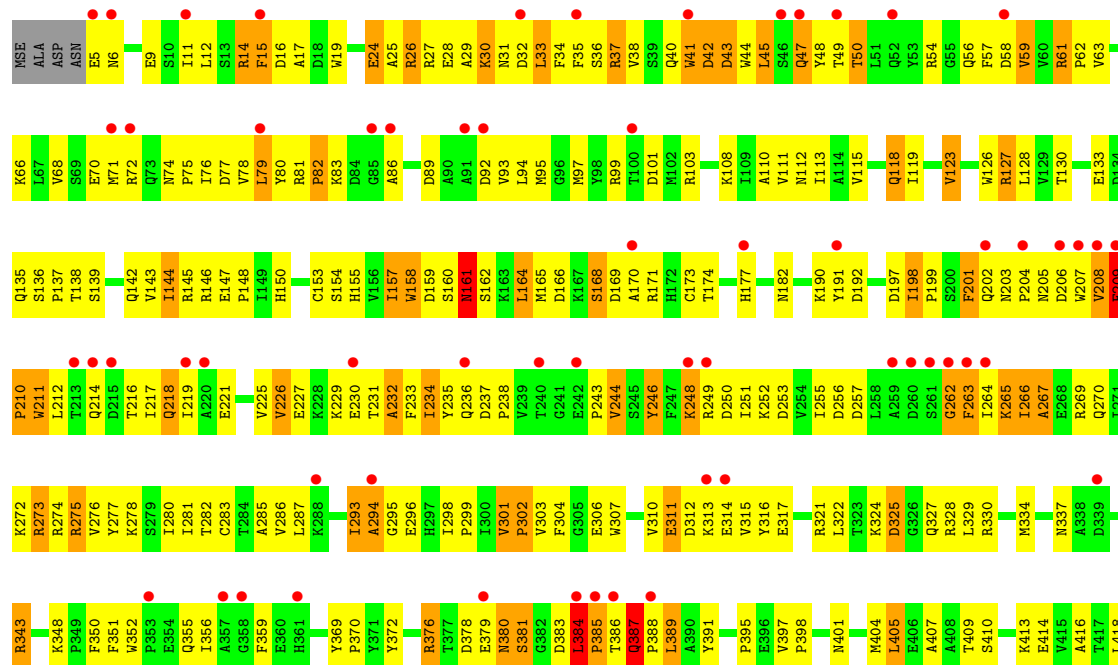
Chain A:

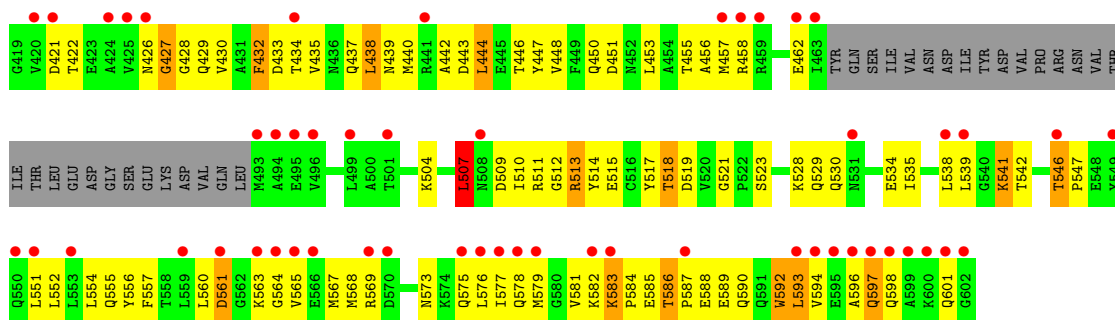


• Molecule 1: PORTAL PROTEIN

Chain B:

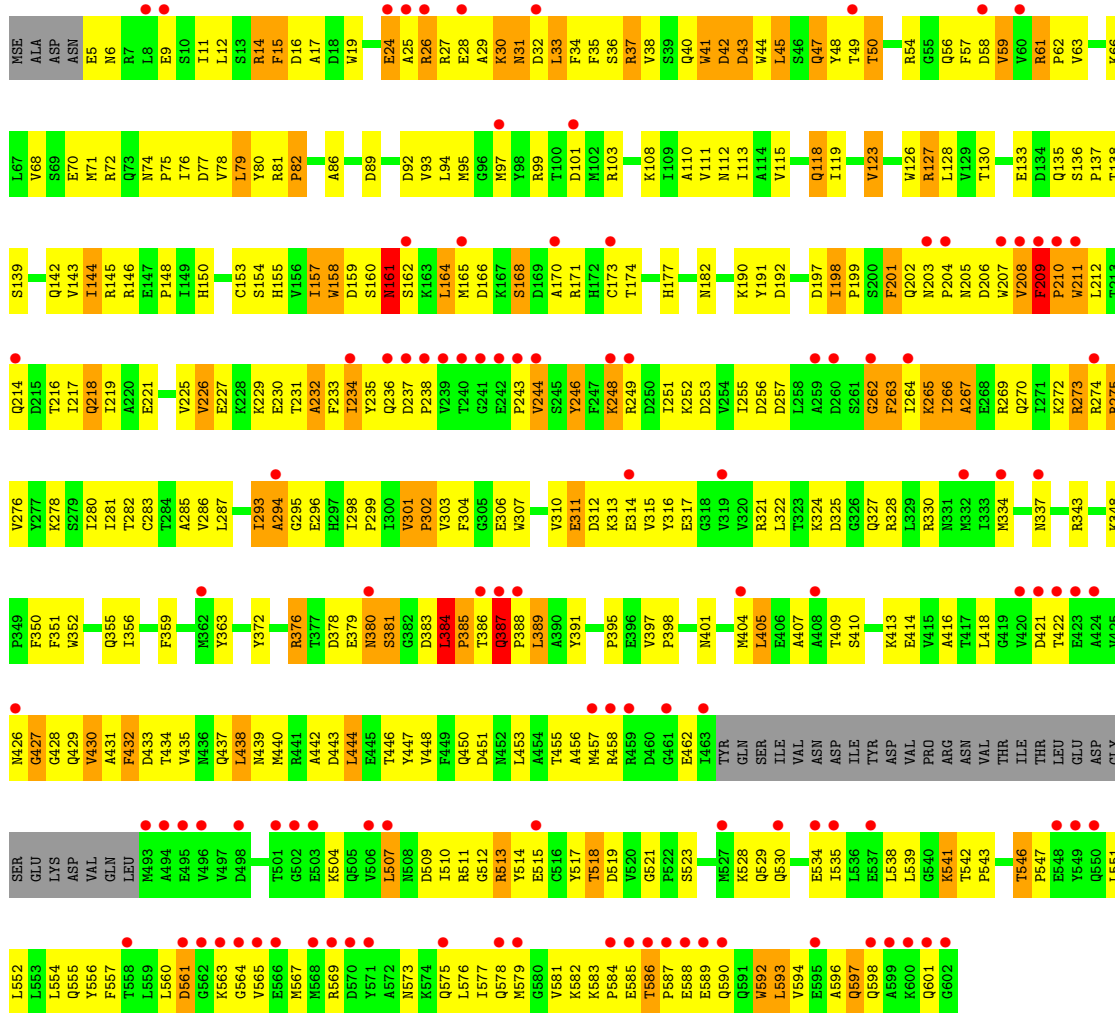






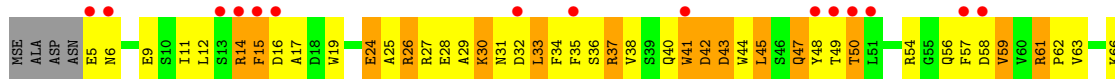
• Molecule 1: PORTAL PROTEIN

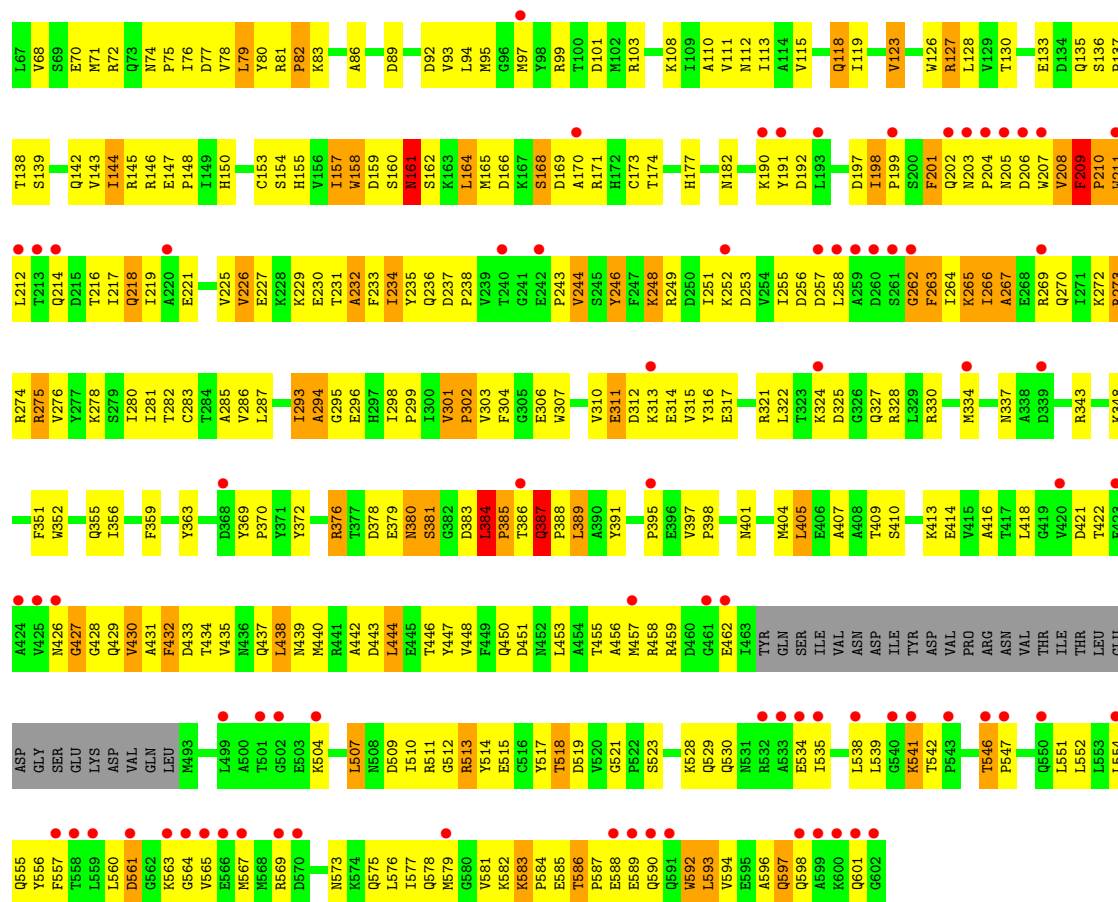
Chain D:



• Molecule 1: PORTAL PROTEIN

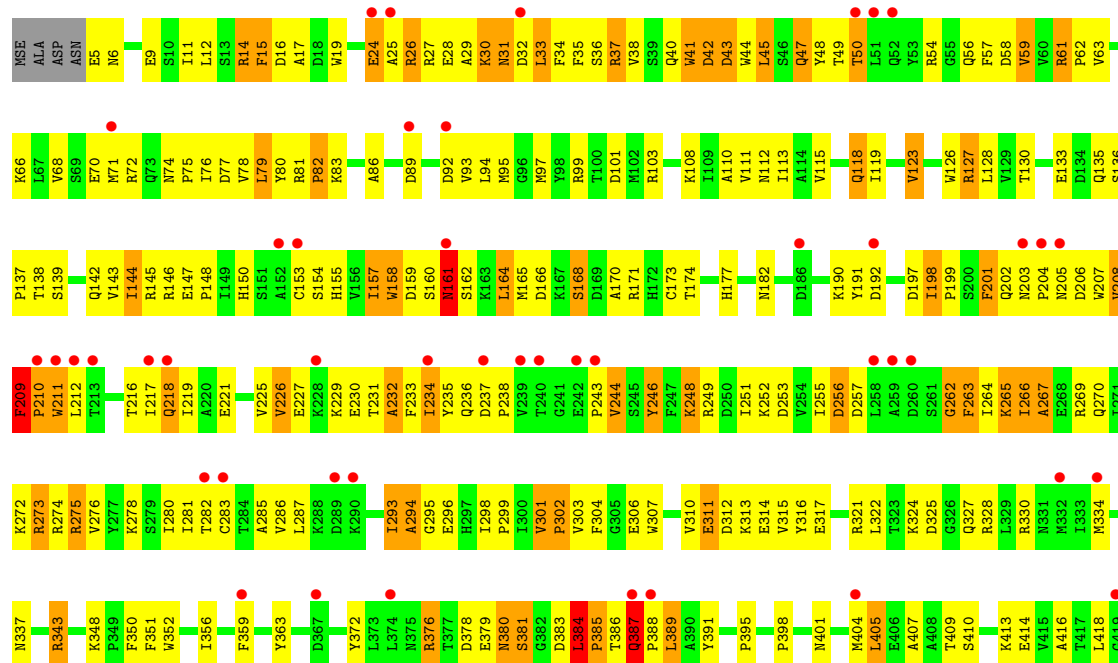
Chain E:

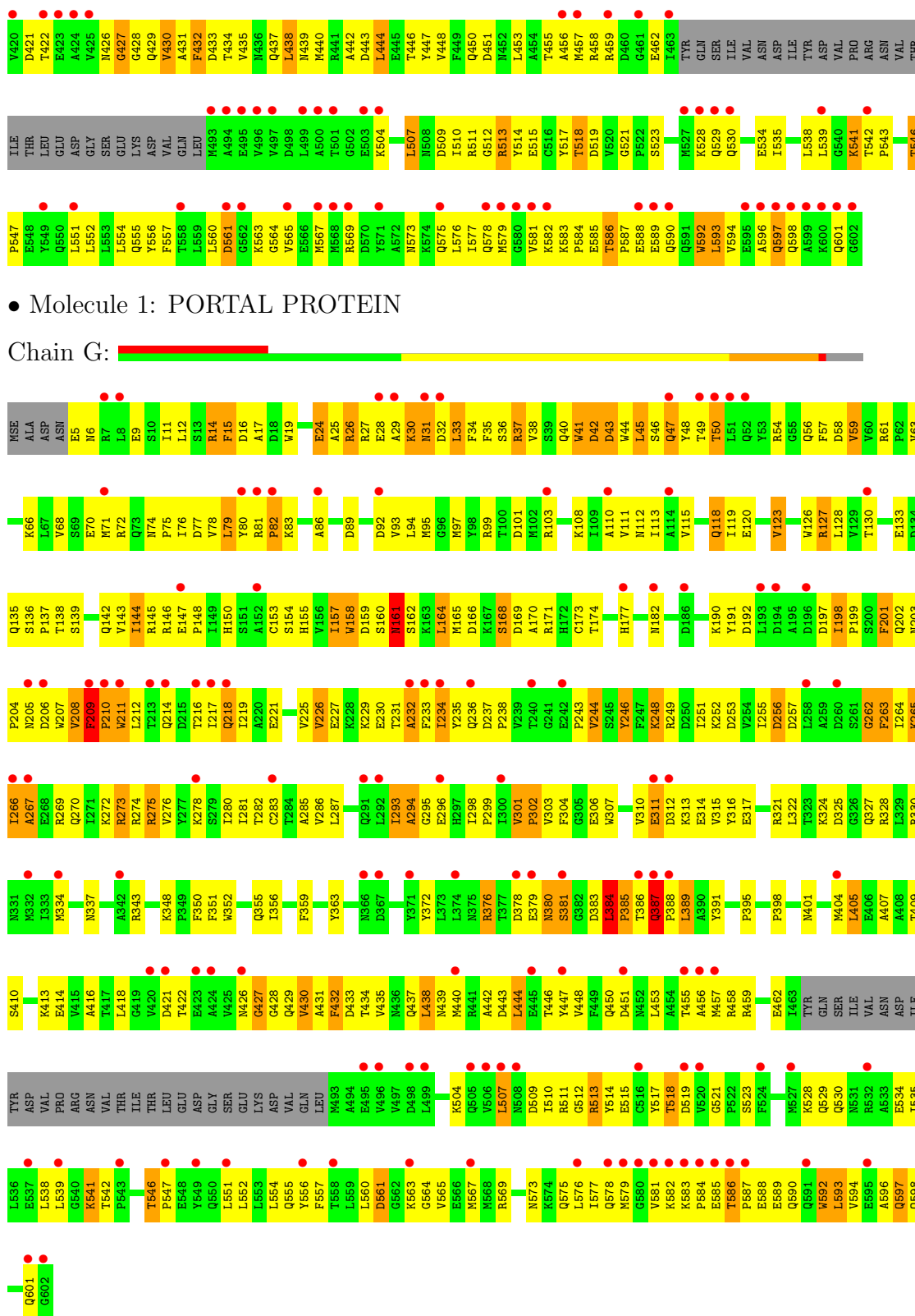




• Molecule 1: PORTAL PROTEIN

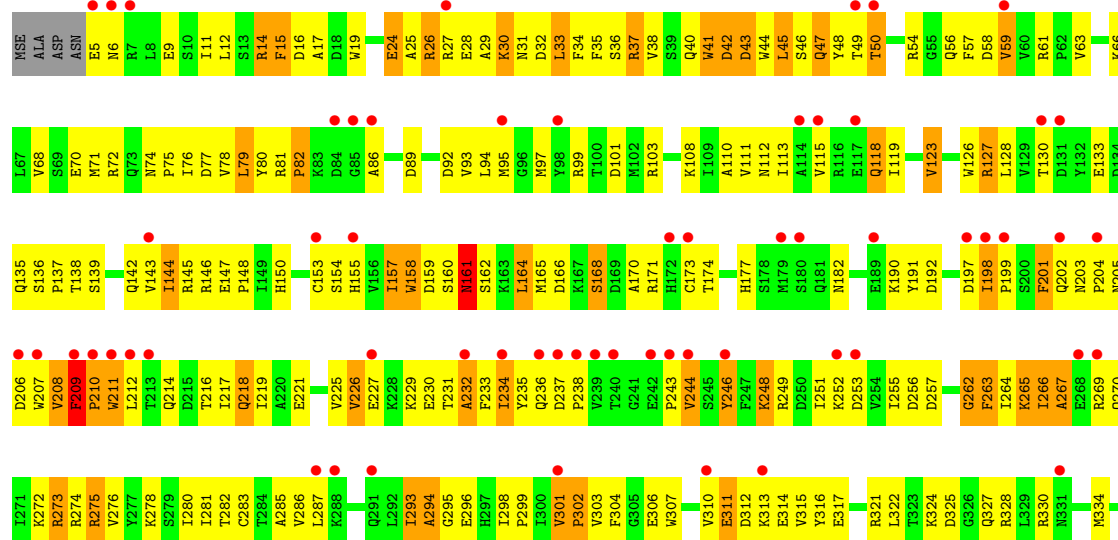
Chain F:



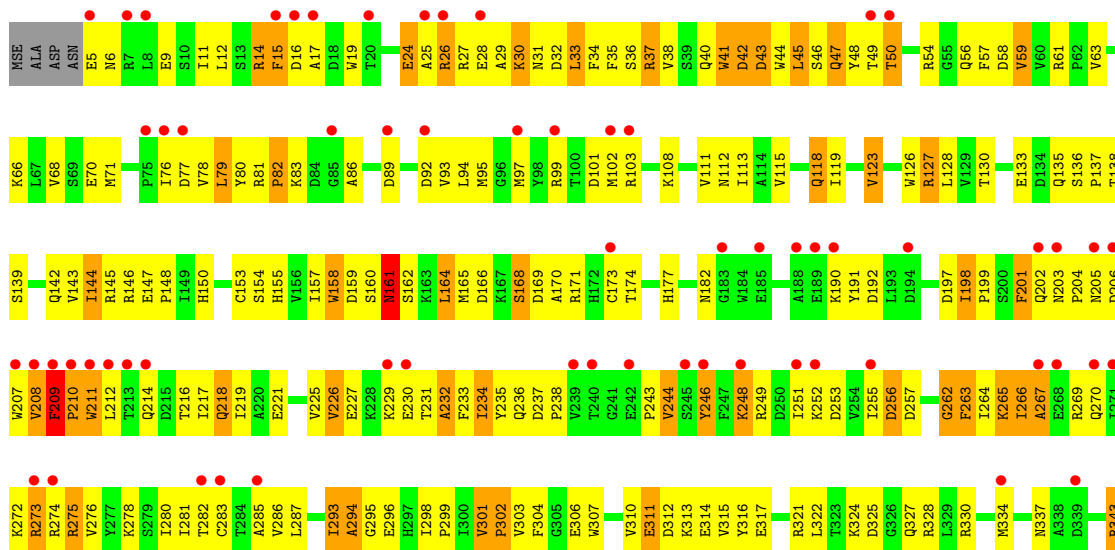


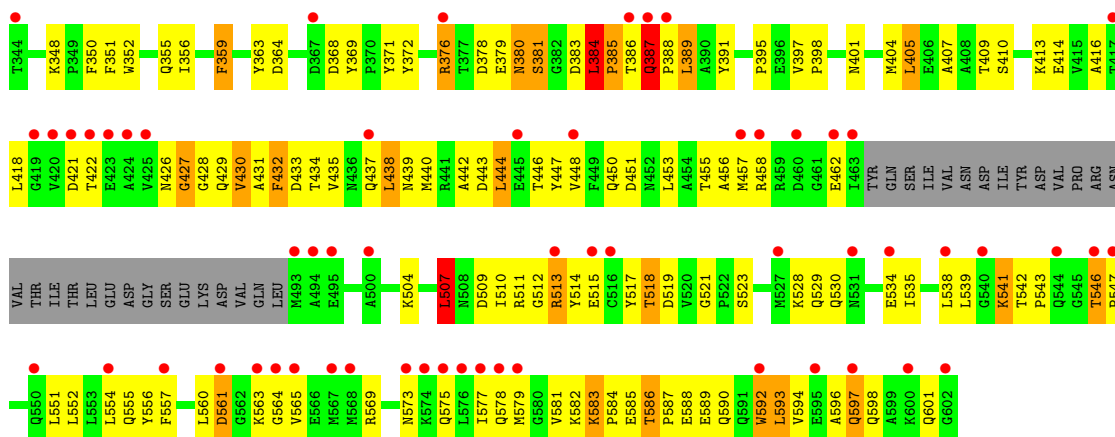
• Molecule 1: PORTAL PROTEIN

Chain H:



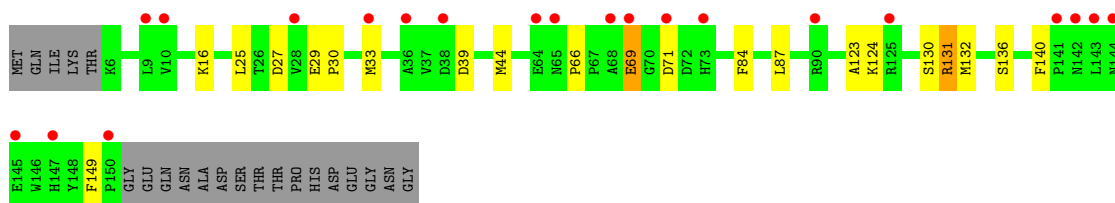






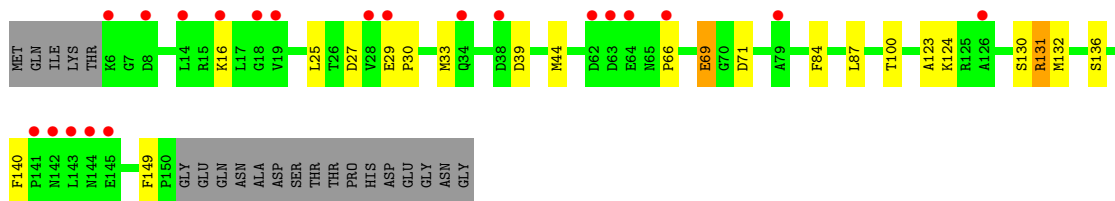
• Molecule 2: PACKAGED DNA STABILIZATION PROTEIN GP4

Chain k:



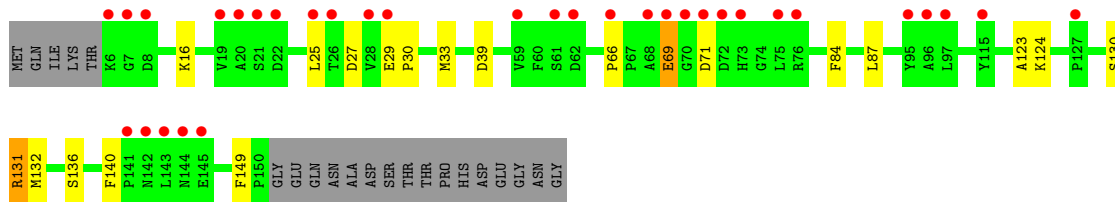
• Molecule 2: PACKAGED DNA STABILIZATION PROTEIN GP4

Chain l:



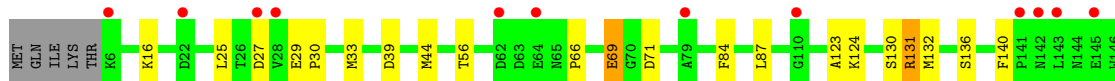
• Molecule 2: PACKAGED DNA STABILIZATION PROTEIN GP4

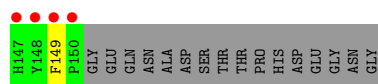
Chain m:



• Molecule 2: PACKAGED DNA STABILIZATION PROTEIN GP4

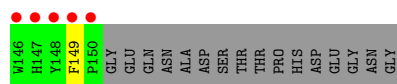
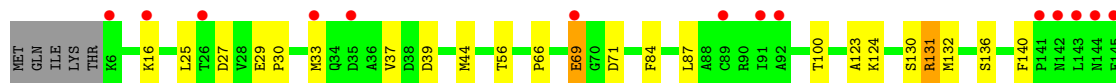
Chain n:





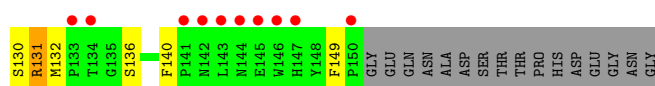
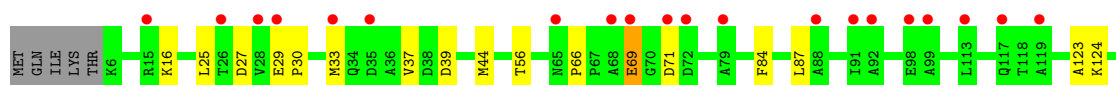
• Molecule 2: PACKAGED DNA STABILIZATION PROTEIN GP4

Chain o:



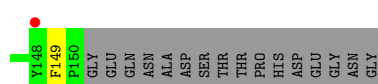
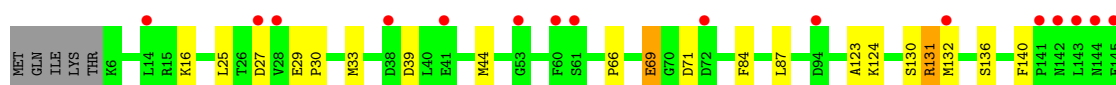
• Molecule 2: PACKAGED DNA STABILIZATION PROTEIN GP4

Chain p:



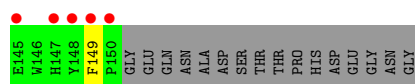
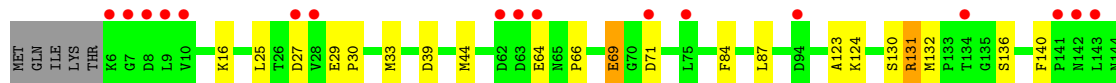
• Molecule 2: PACKAGED DNA STABILIZATION PROTEIN GP4

Chain q:



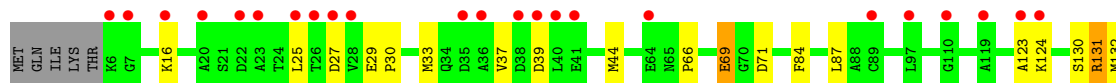
• Molecule 2: PACKAGED DNA STABILIZATION PROTEIN GP4

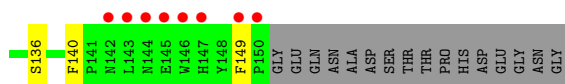
Chain r:



• Molecule 2: PACKAGED DNA STABILIZATION PROTEIN GP4

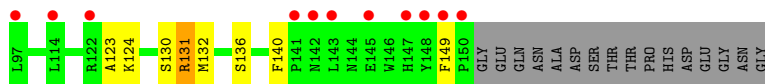
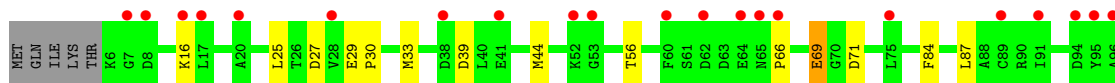
Chain s:





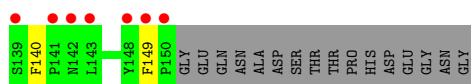
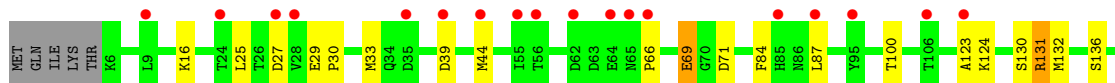
• Molecule 2: PACKAGED DNA STABILIZATION PROTEIN GP4

Chain t:



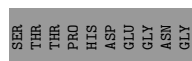
• Molecule 2: PACKAGED DNA STABILIZATION PROTEIN GP4

Chain u:



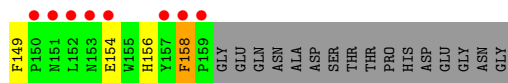
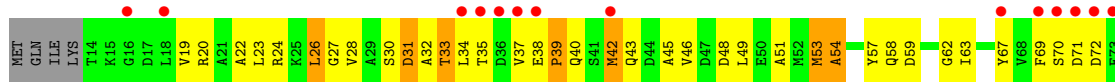
• Molecule 2: PACKAGED DNA STABILIZATION PROTEIN GP4

Chain v:



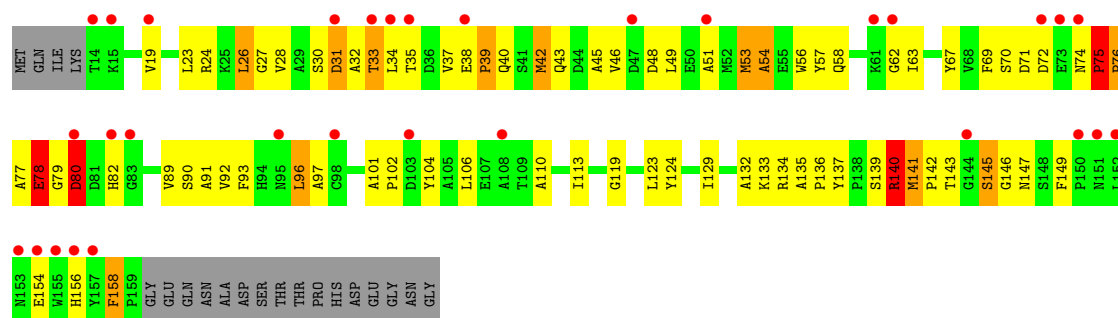
• Molecule 2: PACKAGED DNA STABILIZATION PROTEIN GP4

Chain Y:



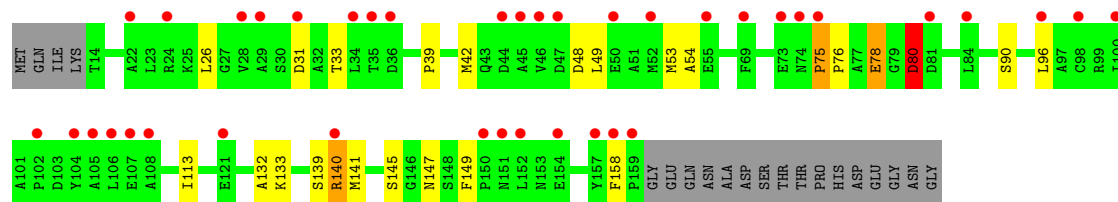
• Molecule 2: PACKAGED DNA STABILIZATION PROTEIN GP4

Chain Z:



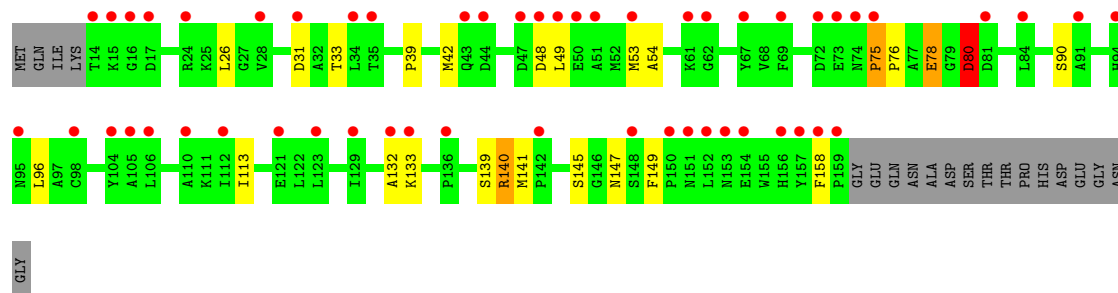
• Molecule 2: PACKAGED DNA STABILIZATION PROTEIN GP4

Chain a:



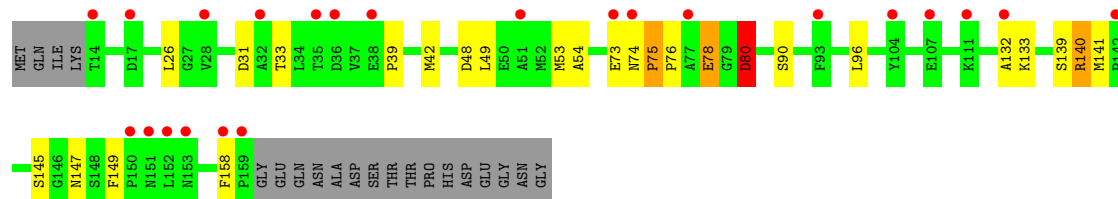
• Molecule 2: PACKAGED DNA STABILIZATION PROTEIN GP4

Chain b:



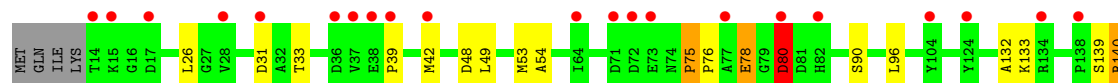
• Molecule 2: PACKAGED DNA STABILIZATION PROTEIN GP4

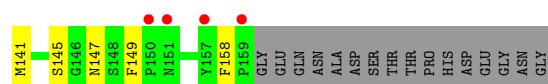
Chain c:



• Molecule 2: PACKAGED DNA STABILIZATION PROTEIN GP4

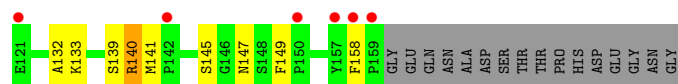
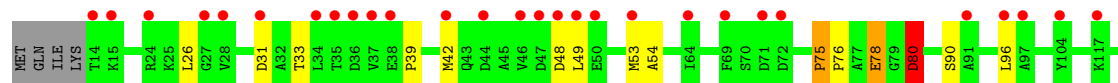
Chain d:





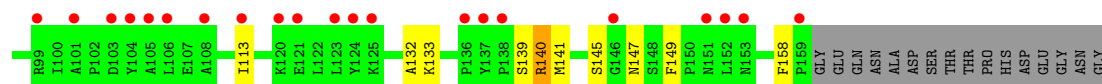
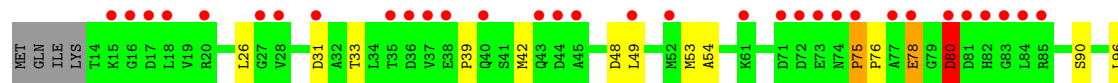
• Molecule 2: PACKAGED DNA STABILIZATION PROTEIN GP4

Chain e:



• Molecule 2: PACKAGED DNA STABILIZATION PROTEIN GP4

Chain f:



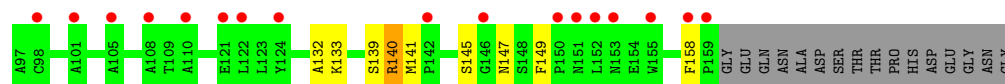
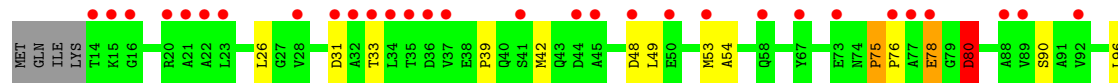
• Molecule 2: PACKAGED DNA STABILIZATION PROTEIN GP4

Chain g:



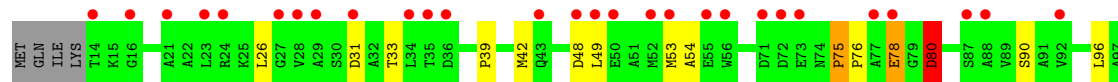
• Molecule 2: PACKAGED DNA STABILIZATION PROTEIN GP4

Chain h:



• Molecule 2: PACKAGED DNA STABILIZATION PROTEIN GP4

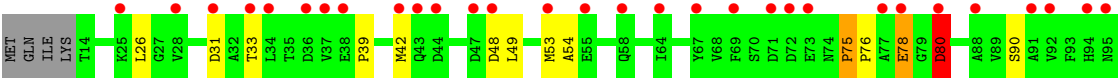
Chain i:





● Molecule 2: PACKAGED DNA STABILIZATION PROTEIN GP4

Chain j: 



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 78.75 – 3.23	Depositor EDS
% Data completeness (in resolution range)	59.7 (19.99-3.25) 91.6 (78.75-3.23)	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: 1.6.4_486)	Depositor
R, R_{free}	0.222 , 0.236 0.293 , (Not available)	Depositor DCC
R_{free} test set	No test flags present.	DCC
Wilson B-factor (Å ²)	65.0	Xtriage
Anisotropy	1.241	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.31 , 70.8	EDS
Estimated twinning fraction	0.025 for h,-k,-l	Xtriage
L-test for twinning	$\langle L \rangle = 0.45$, $\langle L^2 \rangle = 0.27$	Xtriage
Outliers	36 of 349411 reflections (0.010%)	Xtriage
F_o, F_c correlation	0.77	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.

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	0.35	0/4635	0.54	0/6265
1	B	0.35	0/4635	0.54	0/6265
1	C	0.35	0/4635	0.54	0/6265
1	D	0.35	0/4635	0.54	0/6265
1	E	0.35	0/4635	0.54	0/6265
1	F	0.35	0/4635	0.54	0/6265
1	G	0.35	0/4635	0.54	0/6265
1	H	0.35	0/4635	0.54	0/6265
1	I	0.35	0/4635	0.54	0/6265
1	J	0.35	0/4635	0.54	0/6265
1	K	0.35	0/4635	0.54	0/6265
1	L	0.35	0/4635	0.54	0/6265
1	M	0.31	0/4646	0.51	0/6278
1	N	0.32	0/4646	0.51	0/6278
1	O	0.32	0/4646	0.50	0/6278
1	P	0.32	0/4646	0.51	0/6278
1	Q	0.32	0/4646	0.51	0/6278
1	R	0.32	0/4646	0.51	0/6278
1	S	0.31	0/4646	0.51	0/6278
1	T	0.31	0/4646	0.50	0/6278
1	U	0.31	0/4646	0.51	0/6278
1	V	0.32	0/4646	0.51	0/6278
1	W	0.31	0/4646	0.55	2/6278 (0.0%)
1	X	0.31	0/4646	0.50	0/6278
2	Y	0.37	0/1067	0.84	4/1452 (0.3%)
2	Z	0.37	0/1067	0.85	4/1452 (0.3%)
2	a	0.37	0/1067	0.84	4/1452 (0.3%)
2	b	0.37	0/1067	0.84	4/1452 (0.3%)
2	c	0.37	0/1067	0.84	4/1452 (0.3%)
2	d	0.37	0/1067	0.84	4/1452 (0.3%)
2	e	0.37	0/1067	0.84	4/1452 (0.3%)
2	f	0.37	0/1067	0.84	4/1452 (0.3%)
2	g	0.37	0/1067	0.84	4/1452 (0.3%)
2	h	0.37	0/1067	0.84	4/1452 (0.3%)

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	# Z >5	RMSZ	# Z >5
2	i	0.37	0/1067	0.85	4/1452 (0.3%)
2	j	0.37	0/1067	0.84	4/1452 (0.3%)
2	k	0.32	0/1071	0.52	0/1455
2	l	0.33	0/1071	0.52	0/1455
2	m	0.33	0/1071	0.51	0/1455
2	n	0.33	0/1071	0.51	0/1455
2	o	0.33	0/1071	0.51	0/1455
2	p	0.33	0/1071	0.51	0/1455
2	q	0.32	0/1071	0.51	0/1455
2	r	0.36	0/1071	0.51	0/1455
2	s	0.32	0/1071	0.52	0/1455
2	t	0.32	0/1071	0.51	0/1455
2	u	0.31	0/1071	0.51	0/1455
2	v	0.32	0/1071	0.51	0/1455
All	All	0.34	0/137028	0.56	50/185400 (0.0%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	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 50 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	i	80	ASP	N-CA-CB	17.19	141.54	110.60
2	d	80	ASP	N-CA-CB	17.18	141.52	110.60
2	Z	80	ASP	N-CA-CB	17.17	141.51	110.60
2	Y	80	ASP	N-CA-CB	17.17	141.50	110.60
2	e	80	ASP	N-CA-CB	17.17	141.50	110.60

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 Close contacts ⓘ

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	4553	0	4351	430	0
1	B	4553	0	4351	437	2
1	C	4553	0	4351	432	4
1	D	4553	0	4351	423	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	E	4553	0	4351	430	0
1	F	4553	0	4351	428	0
1	G	4553	0	4351	430	0
1	H	4553	0	4351	424	0
1	I	4553	0	4351	433	0
1	J	4553	0	4351	430	0
1	K	4553	0	4351	431	0
1	L	4553	0	4351	441	0
1	M	4564	0	4368	365	0
1	N	4564	0	4368	375	5
1	O	4564	0	4368	363	0
1	P	4564	0	4368	371	0
1	Q	4564	0	4368	373	0
1	R	4564	0	4368	385	0
1	S	4564	0	4368	364	3
1	T	4564	0	4368	366	0
1	U	4564	0	4368	374	0
1	V	4564	0	4368	368	0
1	W	4564	0	4368	364	1
1	X	4564	0	4368	364	0
2	Y	1048	0	957	111	0
2	Z	1048	0	957	108	0
2	a	1048	0	957	0	0
2	b	1048	0	957	0	0
2	c	1048	0	957	0	4
2	d	1048	0	957	0	0
2	e	1048	0	957	0	0
2	f	1048	0	957	0	0
2	g	1048	0	957	0	1
2	h	1048	0	957	0	0
2	i	1048	0	957	0	0
2	j	1048	0	957	0	0
2	k	1052	0	975	0	0
2	l	1052	0	975	0	0
2	m	1052	0	975	0	0
2	n	1052	0	975	0	0
2	o	1052	0	975	0	0
2	p	1052	0	975	0	0
2	q	1052	0	975	0	0
2	r	1052	0	975	0	3
2	s	1052	0	975	0	0
2	t	1052	0	975	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	u	1052	0	975	0	0
2	v	1052	0	975	0	3
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	9241	13

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 42.

The worst 5 of 9241 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
2:Z:32:ALA:O	2:Z:33:THR:CG2	1.82	1.27
2:Z:32:ALA:O	2:Z:33:THR:HG23	1.12	1.26
2:Y:32:ALA:O	2:Y:33:THR:CG2	1.82	1.26
2:Y:32:ALA:O	2:Y:33:THR:HG23	1.12	1.24
2:Z:28:VAL:HG21	2:Z:96:LEU:CD1	1.68	1.23

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	Distance(Å)	Clash(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

5.3.1 Protein backbone ⓘ

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

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

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	565/602 (94%)	434 (77%)	88 (16%)	43 (8%)	2	15
1	B	565/602 (94%)	434 (77%)	87 (15%)	44 (8%)	1	13
1	C	565/602 (94%)	434 (77%)	87 (15%)	44 (8%)	1	13
1	D	565/602 (94%)	434 (77%)	88 (16%)	43 (8%)	2	15
1	E	565/602 (94%)	434 (77%)	88 (16%)	43 (8%)	2	15
1	F	565/602 (94%)	434 (77%)	88 (16%)	43 (8%)	2	15
1	G	565/602 (94%)	434 (77%)	88 (16%)	43 (8%)	2	15
1	H	565/602 (94%)	434 (77%)	88 (16%)	43 (8%)	2	15
1	I	565/602 (94%)	434 (77%)	88 (16%)	43 (8%)	2	15
1	J	565/602 (94%)	434 (77%)	88 (16%)	43 (8%)	2	15
1	K	565/602 (94%)	434 (77%)	88 (16%)	43 (8%)	2	15
1	L	565/602 (94%)	434 (77%)	87 (15%)	44 (8%)	1	13
1	M	565/602 (94%)	446 (79%)	81 (14%)	38 (7%)	2	18
1	N	565/602 (94%)	447 (79%)	79 (14%)	39 (7%)	2	17
1	O	565/602 (94%)	447 (79%)	80 (14%)	38 (7%)	2	18
1	P	565/602 (94%)	448 (79%)	79 (14%)	38 (7%)	2	18
1	Q	565/602 (94%)	447 (79%)	79 (14%)	39 (7%)	2	17
1	R	565/602 (94%)	448 (79%)	79 (14%)	38 (7%)	2	18
1	S	565/602 (94%)	443 (78%)	83 (15%)	39 (7%)	2	17

Continued on next page...

Continued from previous page...

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

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%)	412 (85%)	71 (15%)	4	23
1	B	483/510 (95%)	412 (85%)	71 (15%)	4	23
1	C	483/510 (95%)	412 (85%)	71 (15%)	4	23
1	D	483/510 (95%)	412 (85%)	71 (15%)	4	23
1	E	483/510 (95%)	412 (85%)	71 (15%)	4	23
1	F	483/510 (95%)	412 (85%)	71 (15%)	4	23
1	G	483/510 (95%)	412 (85%)	71 (15%)	4	23
1	H	483/510 (95%)	412 (85%)	71 (15%)	4	23
1	I	483/510 (95%)	412 (85%)	71 (15%)	4	23
1	J	483/510 (95%)	412 (85%)	71 (15%)	4	23
1	K	483/510 (95%)	412 (85%)	71 (15%)	4	23
1	L	483/510 (95%)	412 (85%)	71 (15%)	4	23
1	M	485/510 (95%)	412 (85%)	73 (15%)	4	21
1	N	485/510 (95%)	412 (85%)	73 (15%)	4	21
1	O	485/510 (95%)	413 (85%)	72 (15%)	4	22
1	P	485/510 (95%)	413 (85%)	72 (15%)	4	22
1	Q	485/510 (95%)	412 (85%)	73 (15%)	4	21
1	R	485/510 (95%)	411 (85%)	74 (15%)	4	21
1	S	485/510 (95%)	411 (85%)	74 (15%)	4	21
1	T	485/510 (95%)	412 (85%)	73 (15%)	4	21
1	U	485/510 (95%)	413 (85%)	72 (15%)	4	22

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	V	485/510 (95%)	413 (85%)	72 (15%)	4	22
1	W	485/510 (95%)	413 (85%)	72 (15%)	4	22
1	X	485/510 (95%)	411 (85%)	74 (15%)	4	21
2	Y	98/132 (74%)	85 (87%)	13 (13%)	6	28
2	Z	98/132 (74%)	85 (87%)	13 (13%)	6	28
2	a	98/132 (74%)	84 (86%)	14 (14%)	5	24
2	b	98/132 (74%)	84 (86%)	14 (14%)	5	24
2	c	98/132 (74%)	85 (87%)	13 (13%)	6	28
2	d	98/132 (74%)	85 (87%)	13 (13%)	6	28
2	e	98/132 (74%)	85 (87%)	13 (13%)	6	28
2	f	98/132 (74%)	84 (86%)	14 (14%)	5	24
2	g	98/132 (74%)	84 (86%)	14 (14%)	5	24
2	h	98/132 (74%)	85 (87%)	13 (13%)	6	28
2	i	98/132 (74%)	85 (87%)	13 (13%)	6	28
2	j	98/132 (74%)	84 (86%)	14 (14%)	5	24
2	k	100/132 (76%)	88 (88%)	12 (12%)	7	34
2	l	100/132 (76%)	87 (87%)	13 (13%)	6	30
2	m	100/132 (76%)	88 (88%)	12 (12%)	7	34
2	n	100/132 (76%)	87 (87%)	13 (13%)	6	30
2	o	100/132 (76%)	85 (85%)	15 (15%)	4	21
2	p	100/132 (76%)	86 (86%)	14 (14%)	5	25
2	q	100/132 (76%)	88 (88%)	12 (12%)	7	34
2	r	100/132 (76%)	88 (88%)	12 (12%)	7	34
2	s	100/132 (76%)	87 (87%)	13 (13%)	6	30
2	t	100/132 (76%)	87 (87%)	13 (13%)	6	30
2	u	100/132 (76%)	87 (87%)	13 (13%)	6	30
2	v	100/132 (76%)	87 (87%)	13 (13%)	6	30
All	All	13992/15408 (91%)	11950 (85%)	2042 (15%)	5	23

5 of 2042 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	n	136	SER

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	301	VAL
2	Y	80	ASP
2	p	130	SER
1	A	43	ASP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 319 such sidechains are listed below:

Mol	Chain	Res	Type
1	E	161	ASN
1	F	530	GLN
1	L	214	GLN
1	E	214	GLN
1	F	112	ASN

5.3.3 RNA ⓘ

There are no RNA chains in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

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

5.5 Carbohydrates ⓘ

There are no carbohydrates in this entry.

5.6 Ligand geometry ⓘ

There are no ligands in this entry.

5.7 Other polymers ⓘ

There are no such residues in this entry.

5.8 Polymer linkage issues

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	569/602 (94%)	1.32	111 (19%) 2 1	22, 84, 245, 379	0
1	B	569/602 (94%)	1.32	116 (20%) 1 1	23, 84, 245, 379	0
1	C	569/602 (94%)	1.36	113 (19%) 2 1	26, 83, 245, 379	0
1	D	569/602 (94%)	1.33	111 (19%) 2 1	25, 84, 245, 379	0
1	E	569/602 (94%)	1.22	94 (16%) 2 1	25, 84, 245, 379	0
1	F	569/602 (94%)	1.22	99 (17%) 2 1	25, 84, 245, 379	0
1	G	569/602 (94%)	1.31	122 (21%) 1 1	30, 84, 245, 379	0
1	H	569/602 (94%)	1.39	115 (20%) 2 1	26, 86, 245, 379	0
1	I	569/602 (94%)	1.35	116 (20%) 1 1	31, 88, 245, 379	0
1	J	569/602 (94%)	1.44	119 (20%) 1 1	33, 96, 245, 379	0
1	K	569/602 (94%)	1.46	147 (25%) 1 1	26, 88, 245, 379	0
1	L	569/602 (94%)	1.42	121 (21%) 1 1	24, 87, 245, 379	0
1	M	569/602 (94%)	1.13	87 (15%) 3 2	15, 69, 205, 327	0
1	N	569/602 (94%)	1.05	73 (12%) 4 2	18, 69, 204, 327	0
1	O	569/602 (94%)	0.98	68 (11%) 5 2	15, 68, 205, 326	0
1	P	569/602 (94%)	1.03	90 (15%) 3 1	15, 69, 205, 327	0
1	Q	569/602 (94%)	1.02	76 (13%) 4 2	15, 69, 206, 326	0
1	R	569/602 (94%)	1.11	83 (14%) 3 2	17, 69, 205, 327	0
1	S	569/602 (94%)	1.07	90 (15%) 3 1	19, 68, 206, 326	0
1	T	569/602 (94%)	1.15	104 (18%) 2 1	15, 69, 205, 326	0
1	U	569/602 (94%)	1.24	117 (20%) 1 1	17, 69, 205, 327	0
1	V	569/602 (94%)	1.09	96 (16%) 2 1	18, 70, 204, 327	0
1	W	569/602 (94%)	1.06	79 (13%) 4 2	17, 69, 205, 326	0
1	X	569/602 (94%)	1.12	94 (16%) 2 1	14, 69, 205, 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%)	1.55	39 (26%)	1	1	38, 75, 221, 304	0
2	Z	146/166 (87%)	1.55	31 (21%)	1	1	37, 73, 222, 304	0
2	a	146/166 (87%)	1.52	39 (26%)	1	1	37, 77, 221, 304	0
2	b	146/166 (87%)	1.89	53 (36%)	1	0	39, 72, 221, 304	0
2	c	146/166 (87%)	1.17	23 (15%)	3	1	39, 72, 222, 305	0
2	d	146/166 (87%)	1.29	25 (17%)	2	1	35, 72, 222, 304	0
2	e	146/166 (87%)	1.45	34 (23%)	1	1	37, 73, 220, 304	0
2	f	146/166 (87%)	1.81	53 (36%)	1	0	39, 74, 222, 304	0
2	g	146/166 (87%)	1.84	45 (30%)	1	1	44, 75, 221, 304	0
2	h	146/166 (87%)	1.69	47 (32%)	1	1	38, 75, 221, 305	0
2	i	146/166 (87%)	1.58	44 (30%)	1	1	40, 75, 221, 304	0
2	j	146/166 (87%)	1.94	50 (34%)	1	0	36, 75, 222, 305	0
2	k	145/166 (87%)	1.07	21 (14%)	3	2	25, 57, 204, 314	0
2	l	145/166 (87%)	1.12	21 (14%)	3	2	28, 56, 204, 314	0
2	m	145/166 (87%)	1.56	33 (22%)	1	1	30, 59, 203, 314	0
2	n	145/166 (87%)	1.11	16 (11%)	6	2	23, 55, 204, 314	0
2	o	145/166 (87%)	1.12	19 (13%)	4	2	26, 57, 204, 314	0
2	p	145/166 (87%)	1.28	30 (20%)	1	1	27, 57, 204, 314	0
2	q	145/166 (87%)	1.06	17 (11%)	5	2	29, 58, 204, 314	0
2	r	145/166 (87%)	1.06	22 (15%)	3	2	26, 57, 205, 314	0
2	s	145/166 (87%)	1.34	31 (21%)	1	1	30, 58, 204, 314	0
2	t	145/166 (87%)	1.40	32 (22%)	1	1	31, 58, 204, 314	0
2	u	145/166 (87%)	1.10	25 (17%)	2	1	30, 58, 204, 314	0
2	v	145/166 (87%)	0.77	10 (6%)	17	4	29, 57, 204, 314	0
All	All	17148/18432 (93%)	1.25	3201 (18%)	2	1	14, 76, 227, 379	0

The worst 5 of 3201 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	N	493	MSE	23.4
1	D	240	THR	18.0
2	s	150	PRO	17.5
1	H	493	MSE	16.3
1	K	493	MSE	16.3

6.2 Non-standard residues in protein, DNA, RNA chains ⓘ

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

6.3 Carbohydrates ⓘ

There are no carbohydrates in this entry.

6.4 Ligands ⓘ

There are no ligands in this entry.

6.5 Other polymers ⓘ

There are no such residues in this entry.