wwPDB X-ray Structure Validation Summary Report (i)

Sep 30, 2021 – 01:32 PM EDT

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 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 (i) symbol.

The following versions of software and data (see references (1)) were used in the production of this report:

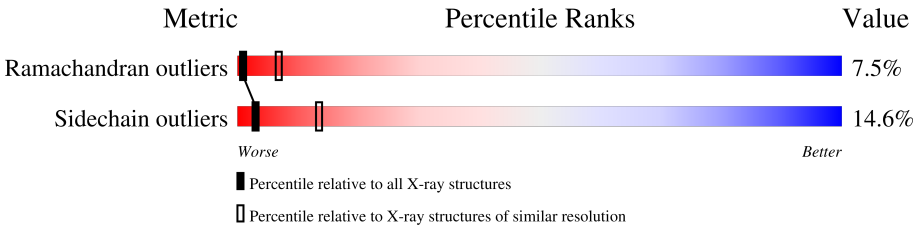
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.23.2
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.23.2

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))
Ramachandran outliers	138981	1229 (3.30-3.22)
Sidechain outliers	138945	1228 (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 <=5%

Mol	Chain	Length	Quality of chain
1	A	602	76% 18% • 5%
1	B	602	76% 18% • 5%
1	C	602	76% 18% • 5%
1	D	602	76% 18% • 5%
1	E	602	76% 18% • 5%
1	F	602	76% 18% • 5%
1	G	602	76% 18% • 5%
1	H	602	76% 18% • 5%
1	I	602	76% 18% • 5%
1	J	602	76% 18% • 5%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	K	602	76% 18% • 5%
1	L	602	76% 18% • 5%
1	M	602	77% 17% • 5%
1	N	602	76% 18% • 5%
1	O	602	77% 17% • 5%
1	P	602	77% 17% • 5%
1	Q	602	76% 17% • 5%
1	R	602	76% 17% • 5%
1	S	602	76% 18% • 5%
1	T	602	77% 17% • 5%
1	U	602	77% 17% • 5%
1	V	602	77% 17% • 5%
1	W	602	76% 17% • 5%
1	X	602	76% 17% • 5%
2	Y	166	73% 12% •• 12%
2	Z	166	73% 12% •• 12%
2	a	166	73% 13% •• 12%
2	b	166	73% 13% •• 12%
2	c	166	72% 13% •• 12%
2	d	166	73% 12% •• 12%
2	e	166	73% 12% •• 12%
2	f	166	73% 13% •• 12%
2	g	166	72% 13% •• 12%
2	h	166	73% 12% •• 12%
2	i	166	73% 12% •• 12%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	j	166	73% 13% 12%
2	k	166	75% 11% 13%
2	l	166	74% 12% 13%
2	m	166	75% 11% 13%
2	n	166	74% 12% 13%
2	o	166	73% 13% 13%
2	p	166	73% 13% 13%
2	q	166	75% 11% 13%
2	r	166	74% 12% 13%
2	s	166	74% 12% 13%
2	t	166	74% 12% 13%
2	u	166	74% 12% 13%
2	v	166	74% 12% 13%

2 Entry composition (i)

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...

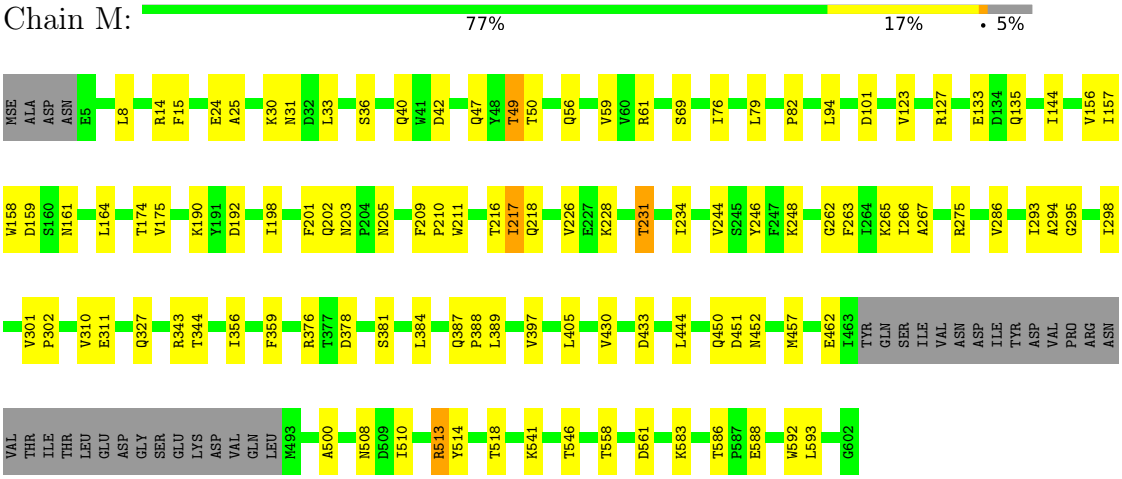
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	C	22	Total O 22 22	0	0
3	D	22	Total O 22 22	0	0
3	E	22	Total O 22 22	0	0
3	F	22	Total O 22 22	0	0
3	G	22	Total O 22 22	0	0
3	H	22	Total O 22 22	0	0
3	I	22	Total O 22 22	0	0
3	J	22	Total O 22 22	0	0
3	K	22	Total O 22 22	0	0
3	L	22	Total O 22 22	0	0

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. 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. 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.

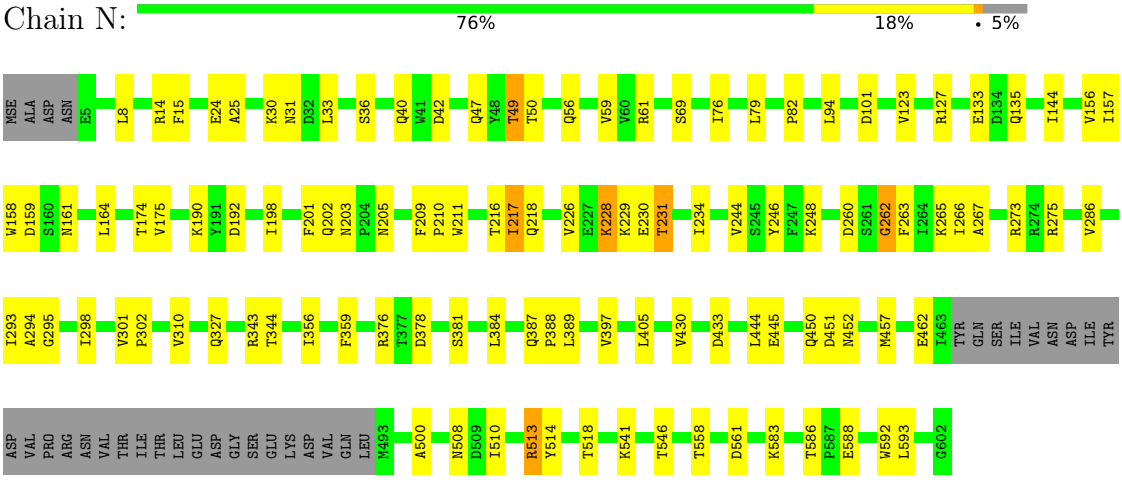
• Molecule 1: PORTAL PROTEIN

Chain M:



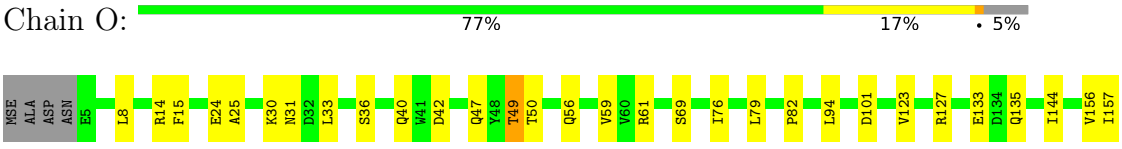
• Molecule 1: PORTAL PROTEIN

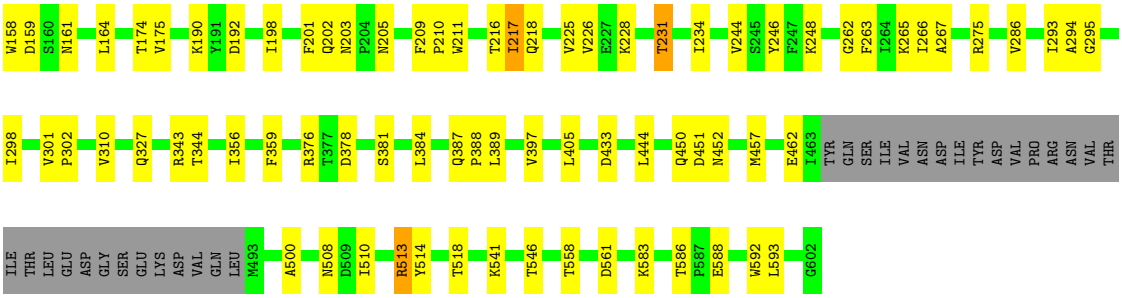
Chain N:



• Molecule 1: PORTAL PROTEIN

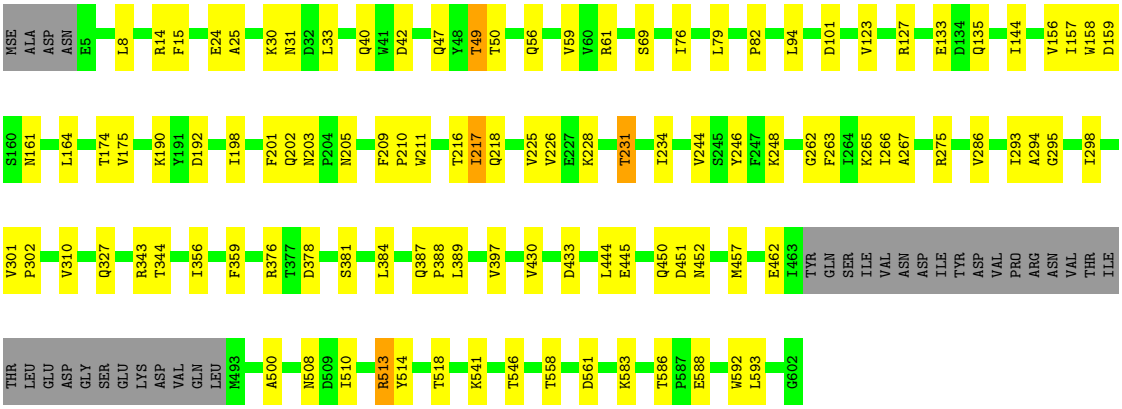
Chain O:





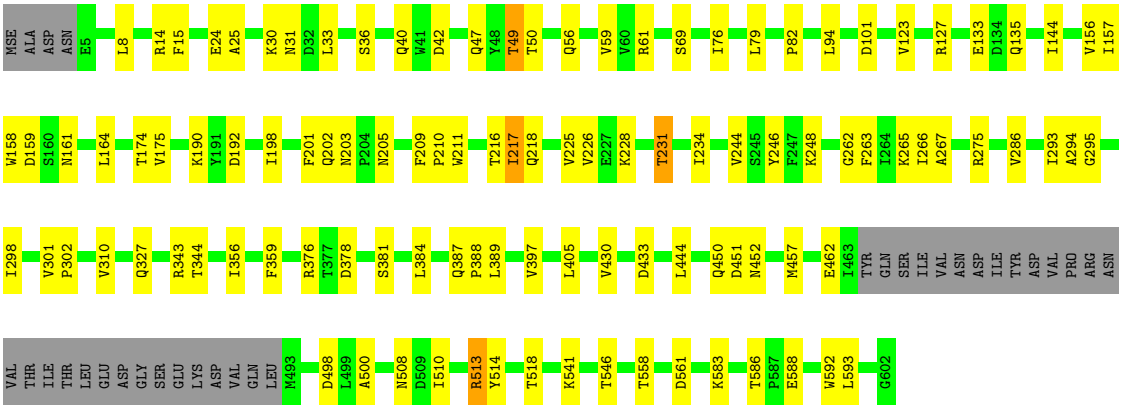
• Molecule 1: PORTAL PROTEIN

Chain P:



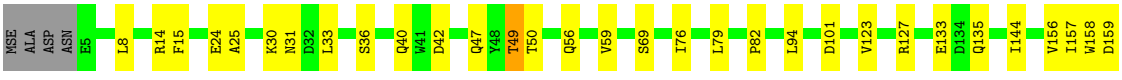
• Molecule 1: PORTAL PROTEIN

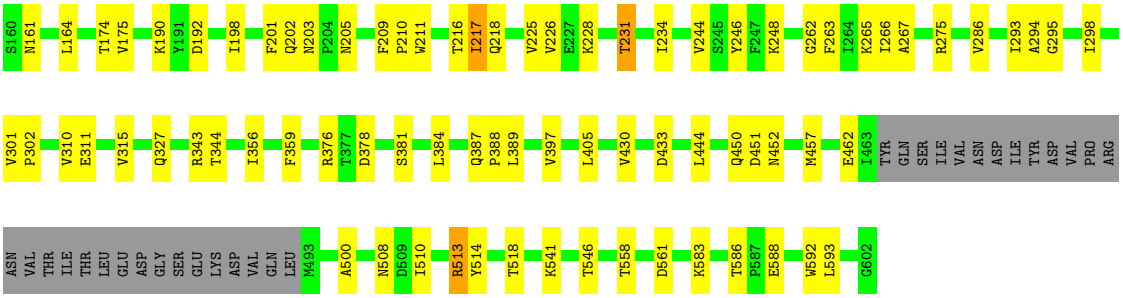
Chain Q:



• Molecule 1: PORTAL PROTEIN

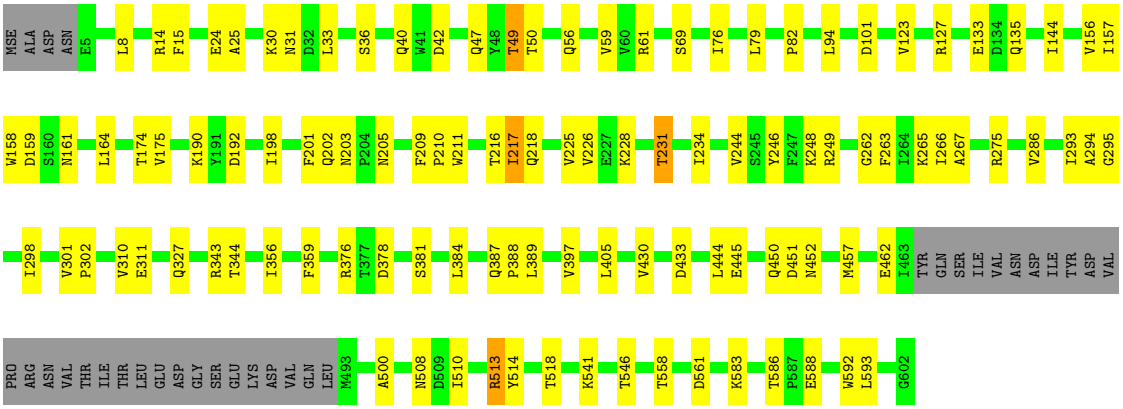
Chain R:





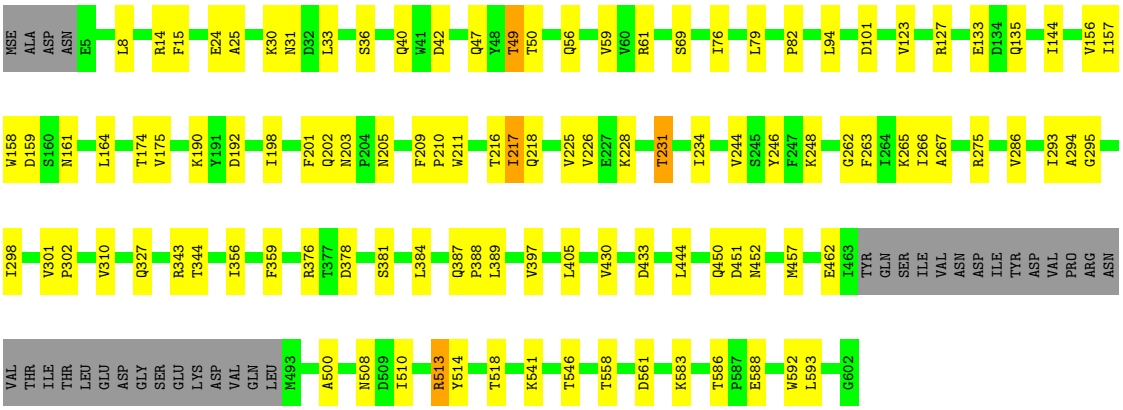
• Molecule 1: PORTAL PROTEIN

Chain S:



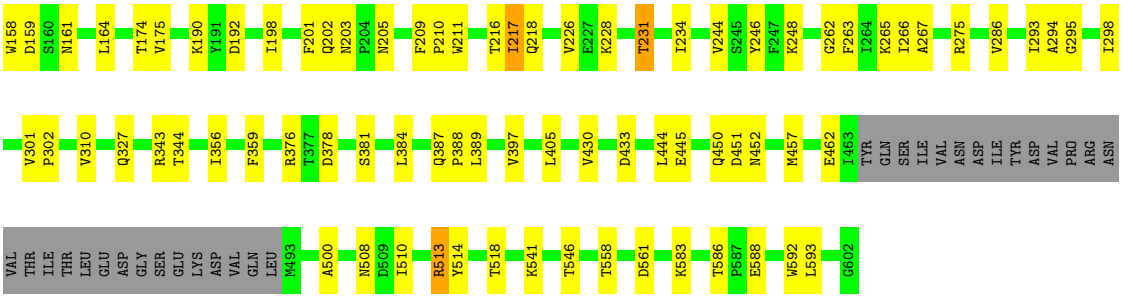
• Molecule 1: PORTAL PROTEIN

Chain T:



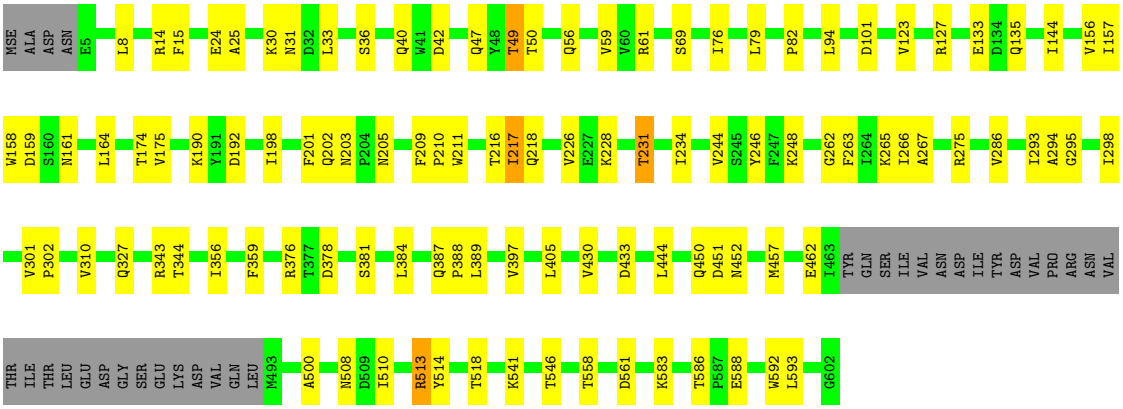
• Molecule 1: PORTAL PROTEIN

Chain U:



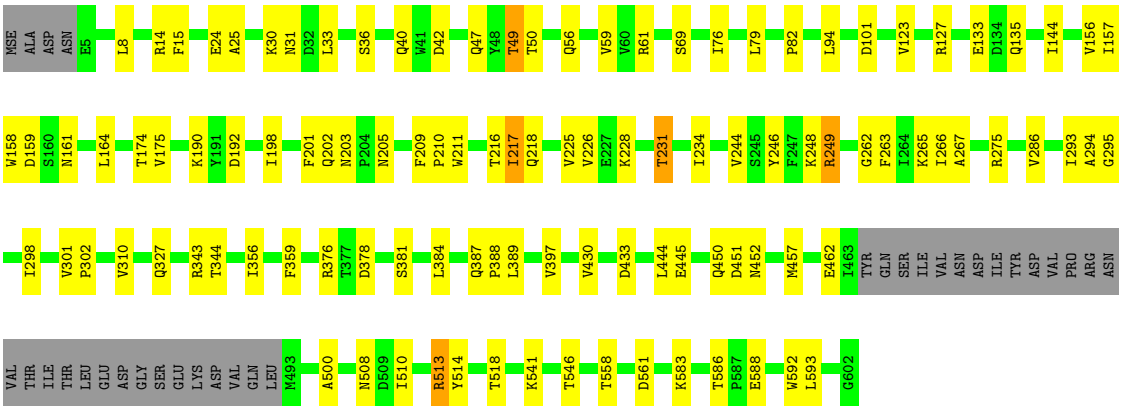
• Molecule 1: PORTAL PROTEIN

Chain V:



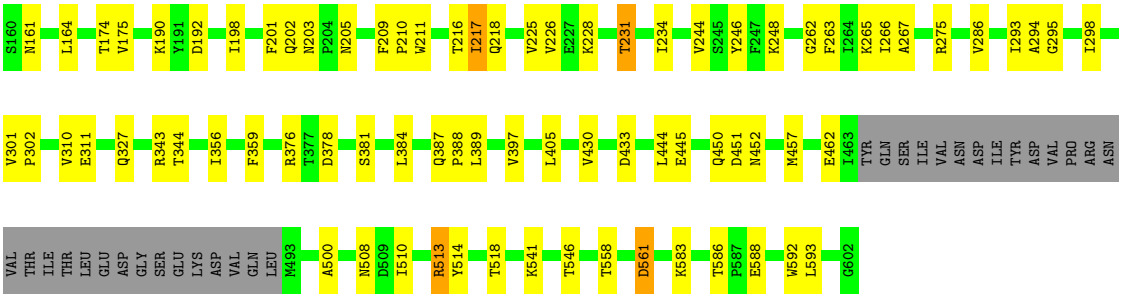
• Molecule 1: PORTAL PROTEIN

Chain W:



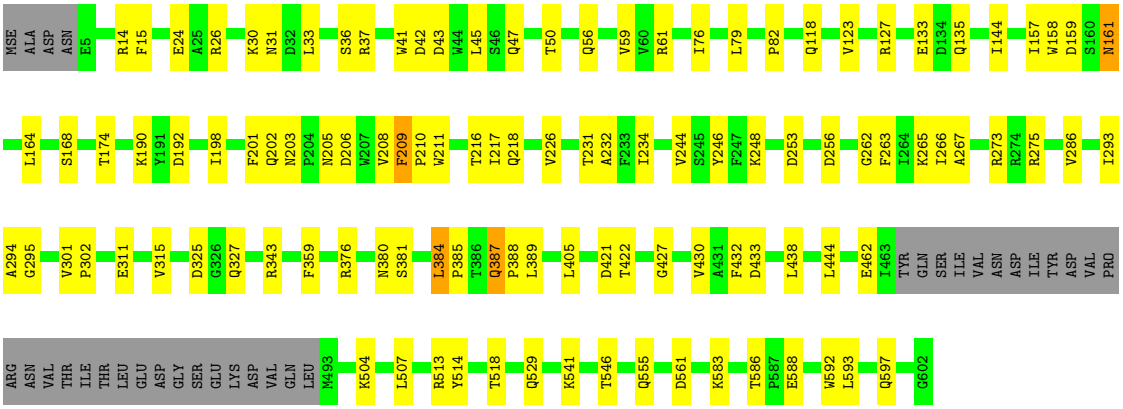
• Molecule 1: PORTAL PROTEIN

Chain X:



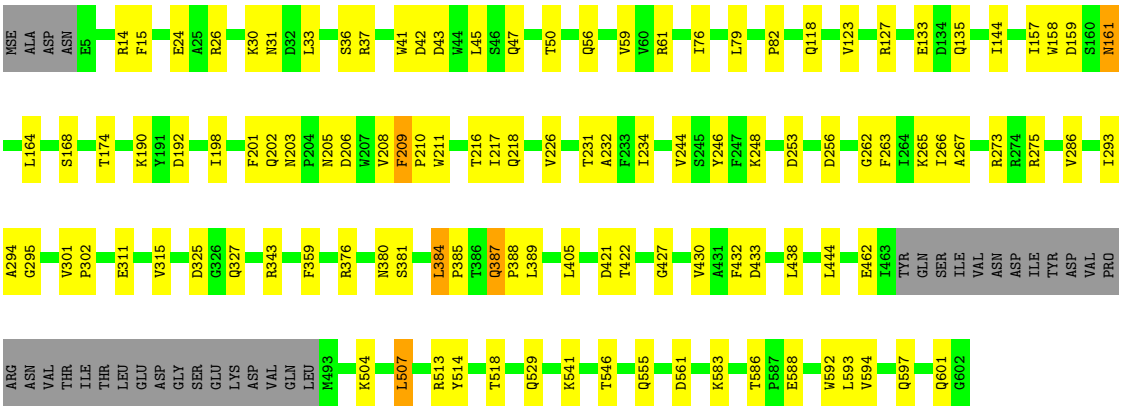
• Molecule 1: PORTAL PROTEIN

Chain A:



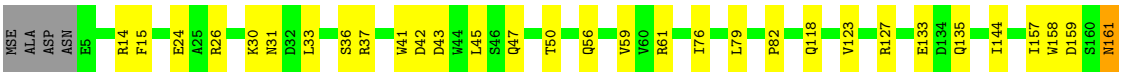
• Molecule 1: PORTAL PROTEIN

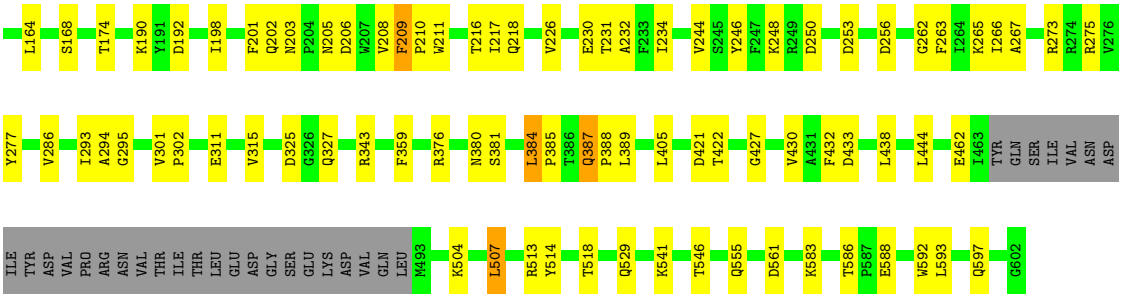
Chain B:



• Molecule 1: PORTAL PROTEIN

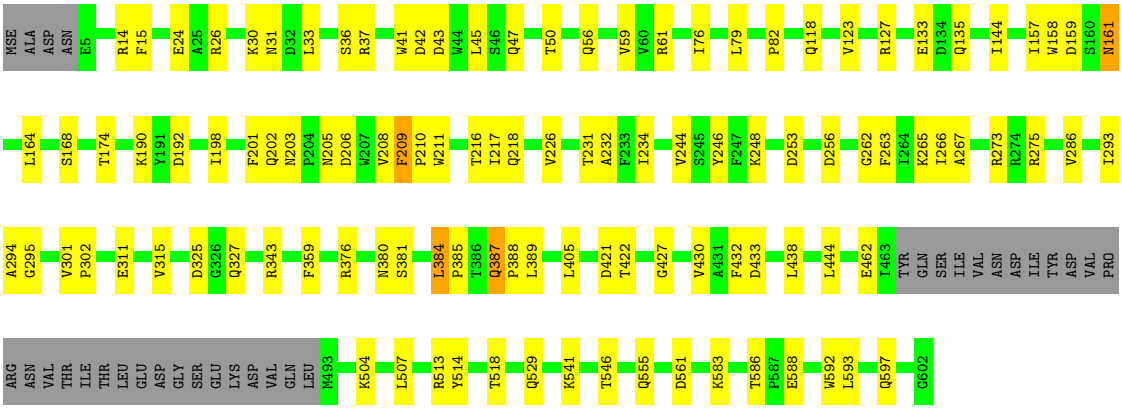
Chain C:





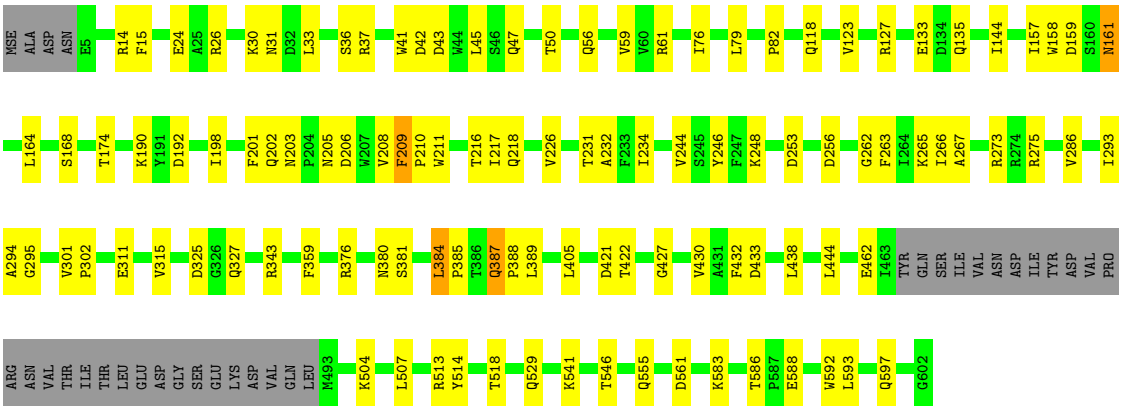
• Molecule 1: PORTAL PROTEIN

Chain D:



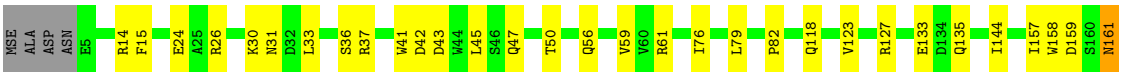
• Molecule 1: PORTAL PROTEIN

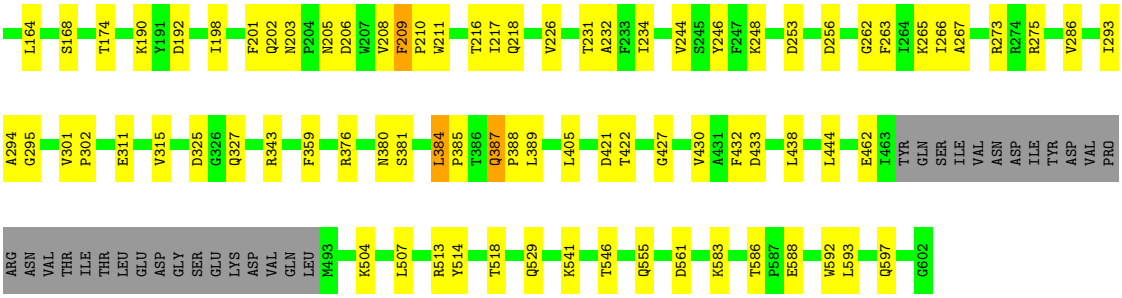
Chain E:



• Molecule 1: PORTAL PROTEIN

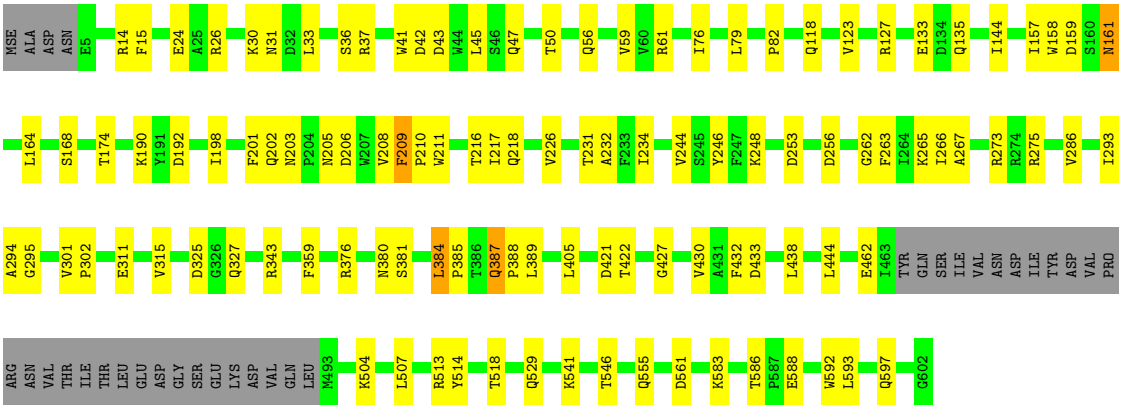
Chain F:





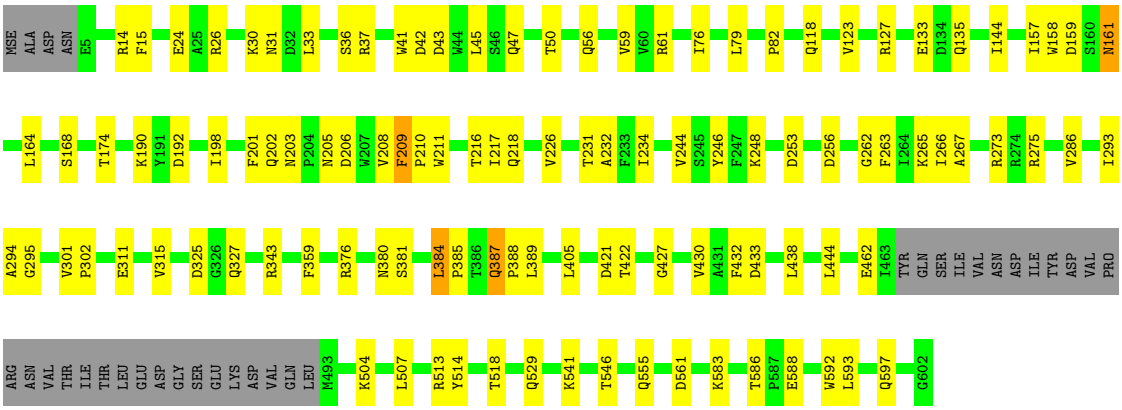
• Molecule 1: PORTAL PROTEIN

Chain G:



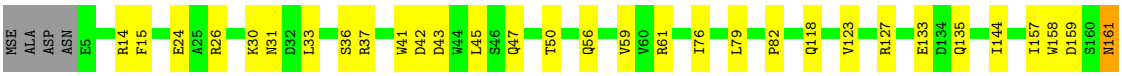
• Molecule 1: PORTAL PROTEIN

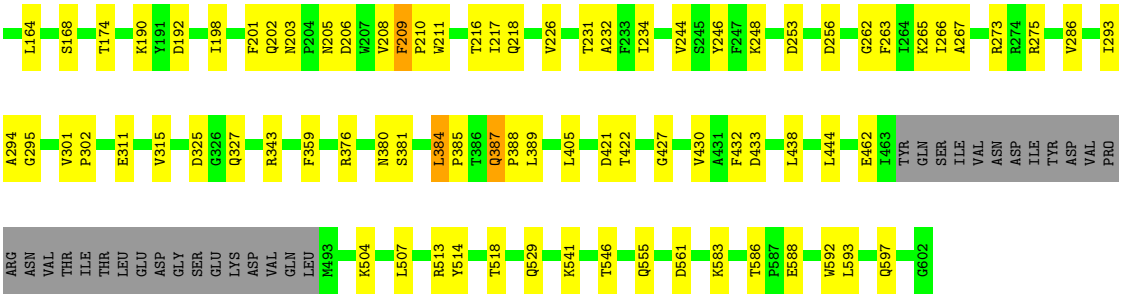
Chain H:



• Molecule 1: PORTAL PROTEIN

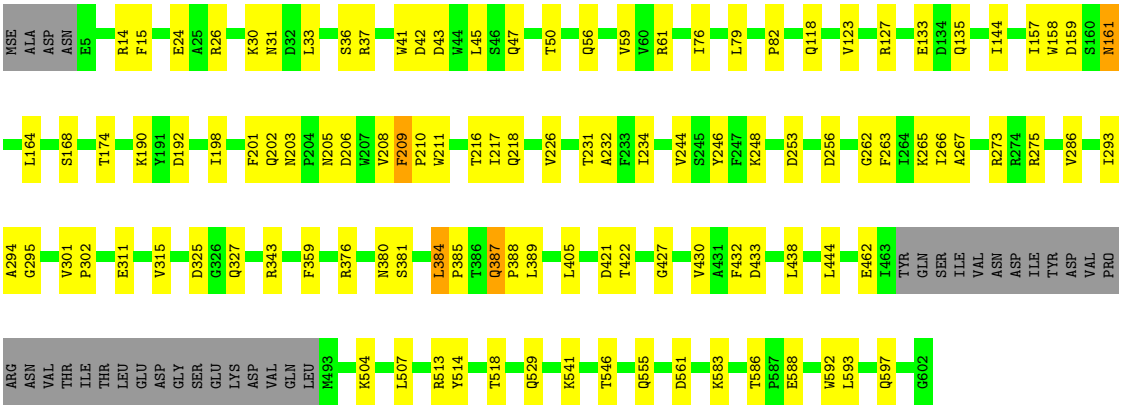
Chain I:





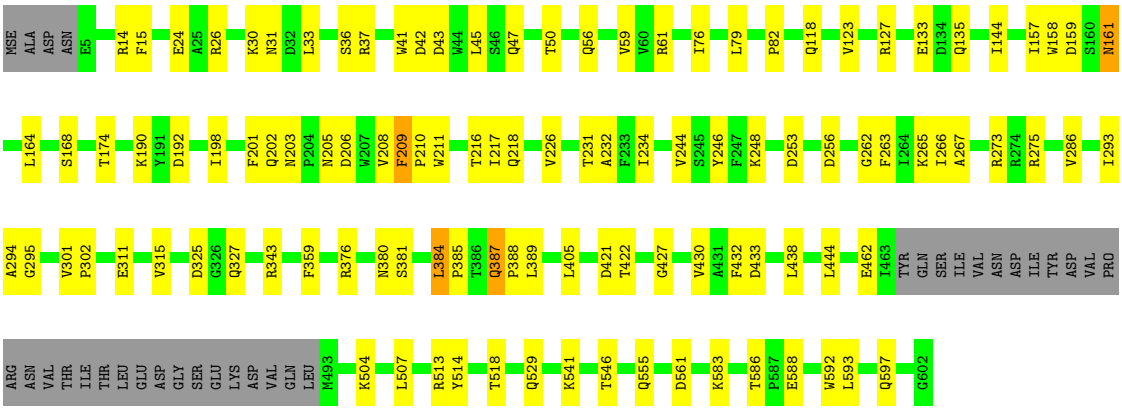
• Molecule 1: PORTAL PROTEIN

Chain J:



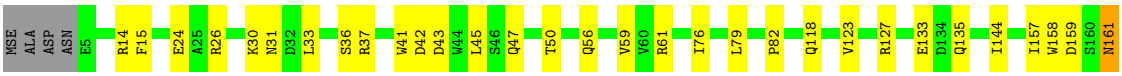
• Molecule 1: PORTAL PROTEIN

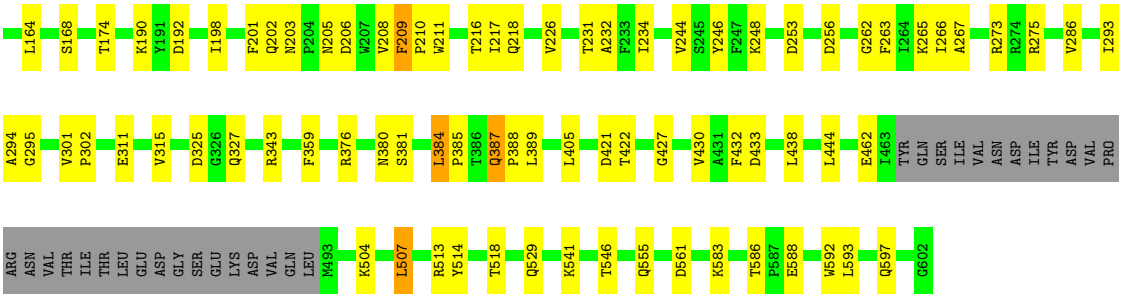
Chain K:



• Molecule 1: PORTAL PROTEIN

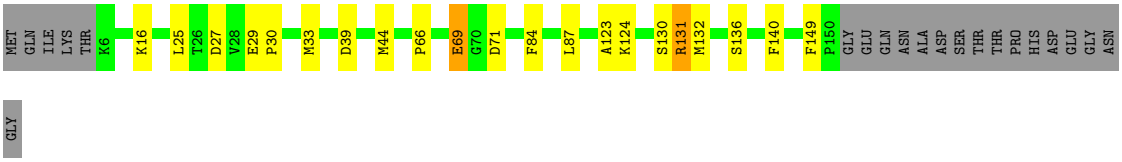
Chain L:





• Molecule 2: PACKAGED DNA STABILIZATION PROTEIN GP4

Chain k:



• Molecule 2: PACKAGED DNA STABILIZATION PROTEIN GP4

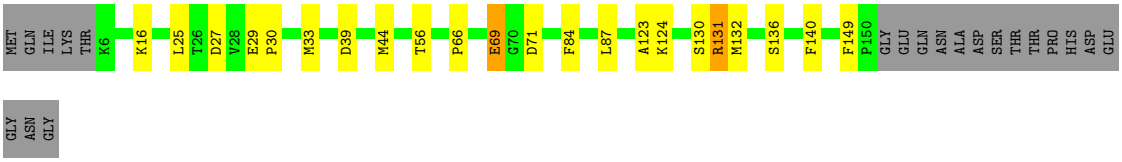
Chain l:

• Molecule 2: PACKAGED DNA STABILIZATION PROTEIN GP4

Chain m:

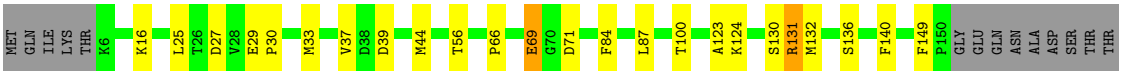
• Molecule 2: PACKAGED DNA STABILIZATION PROTEIN GP4

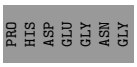
Chain n:



• Molecule 2: PACKAGED DNA STABILIZATION PROTEIN GP4

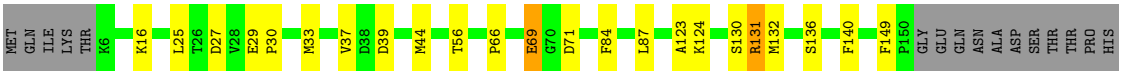
Chain o:





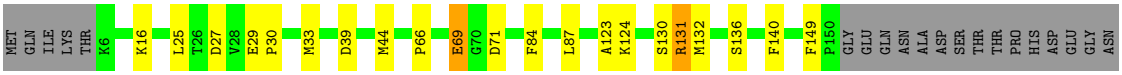
• Molecule 2: PACKAGED DNA STABILIZATION PROTEIN GP4

Chain p:



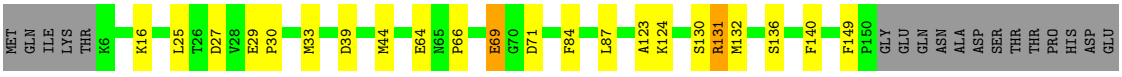
• Molecule 2: PACKAGED DNA STABILIZATION PROTEIN GP4

Chain q:



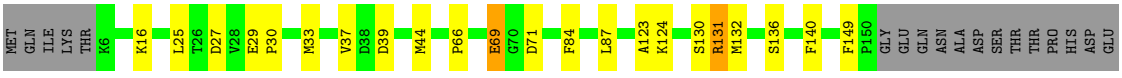
• Molecule 2: PACKAGED DNA STABILIZATION PROTEIN GP4

Chain r:



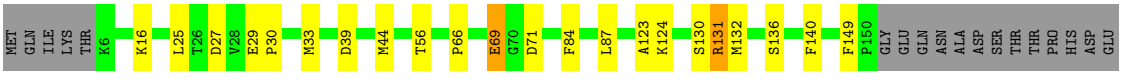
• Molecule 2: PACKAGED DNA STABILIZATION PROTEIN GP4

Chain s:



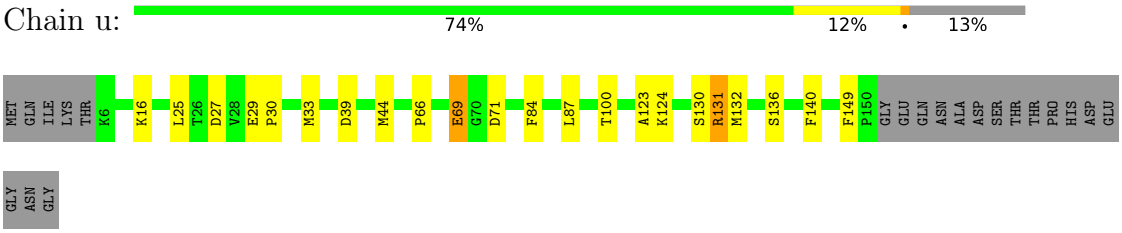
• Molecule 2: PACKAGED DNA STABILIZATION PROTEIN GP4

Chain t:



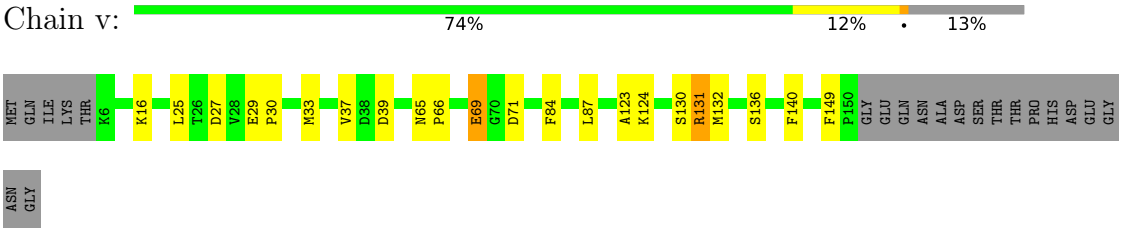
• Molecule 2: PACKAGED DNA STABILIZATION PROTEIN GP4

Chain u:



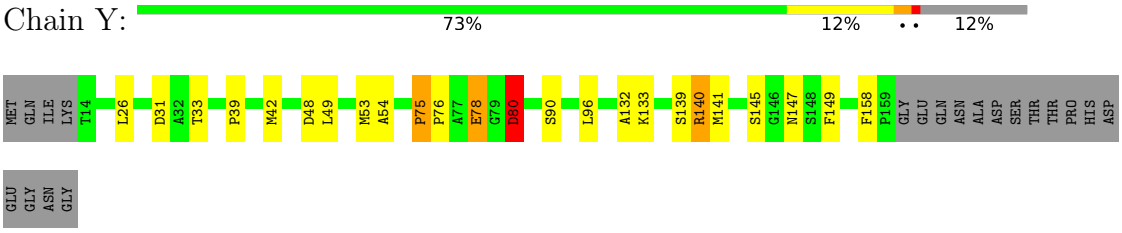
• Molecule 2: PACKAGED DNA STABILIZATION PROTEIN GP4

Chain v:



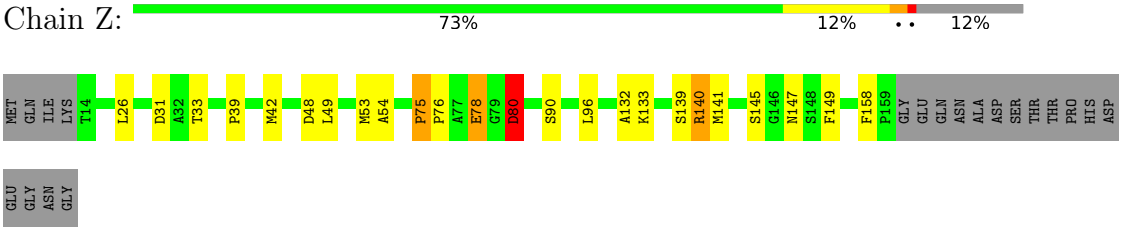
• Molecule 2: PACKAGED DNA STABILIZATION PROTEIN GP4

Chain Y:



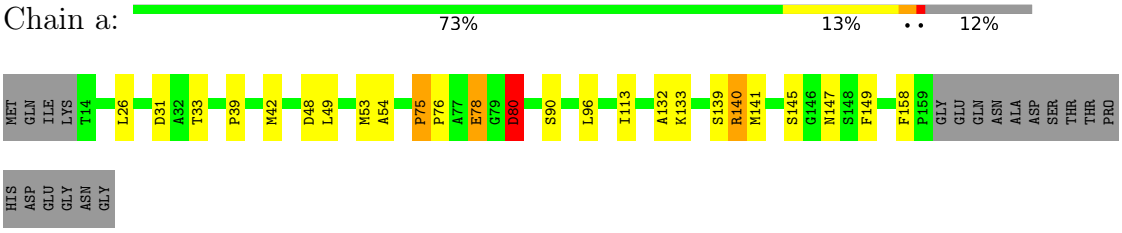
• Molecule 2: PACKAGED DNA STABILIZATION PROTEIN GP4

Chain Z:



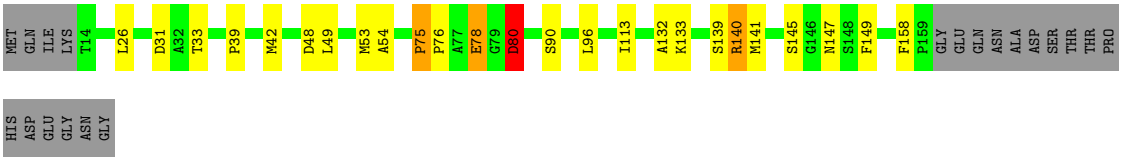
• Molecule 2: PACKAGED DNA STABILIZATION PROTEIN GP4

Chain a:



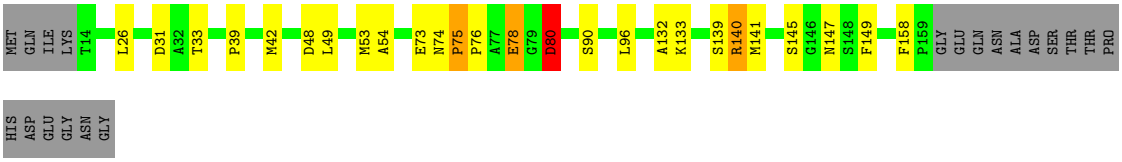
• Molecule 2: PACKAGED DNA STABILIZATION PROTEIN GP4

Chain b:



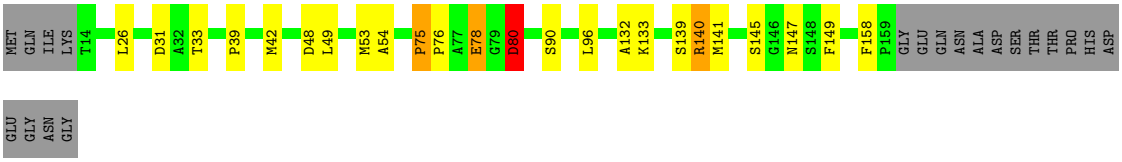
• Molecule 2: PACKAGED DNA STABILIZATION PROTEIN GP4

Chain c:



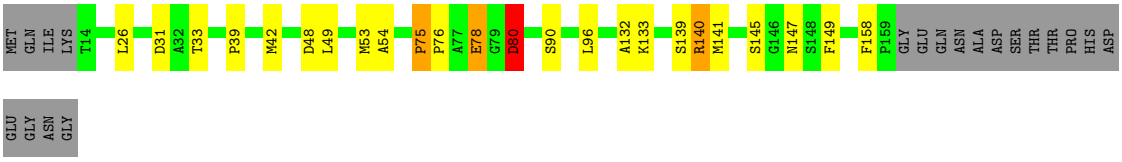
• Molecule 2: PACKAGED DNA STABILIZATION PROTEIN GP4

Chain d:



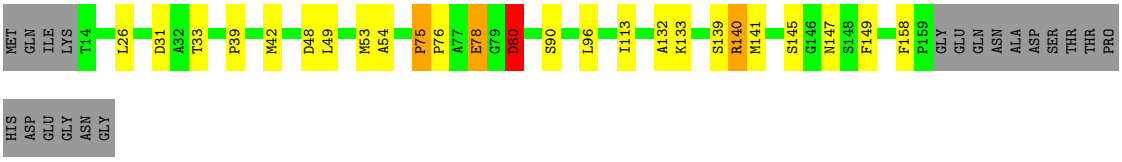
• Molecule 2: PACKAGED DNA STABILIZATION PROTEIN GP4

Chain e:



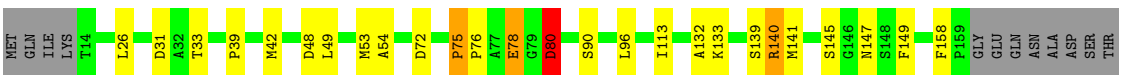
• Molecule 2: PACKAGED DNA STABILIZATION PROTEIN GP4

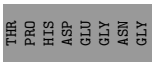
Chain f:



• Molecule 2: PACKAGED DNA STABILIZATION PROTEIN GP4

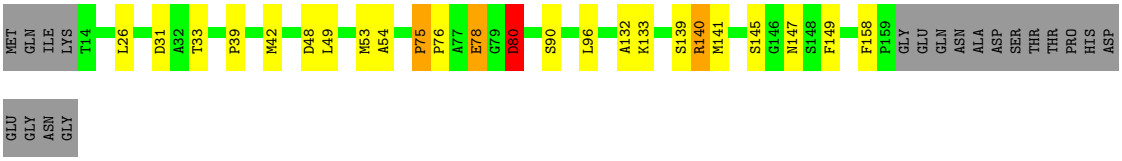
Chain g:





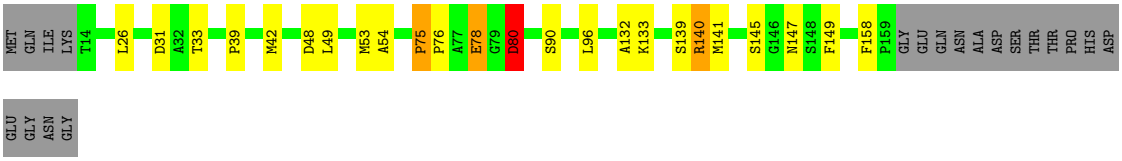
• Molecule 2: PACKAGED DNA STABILIZATION PROTEIN GP4

Chain h:



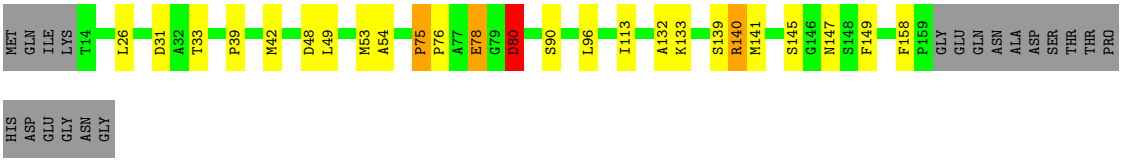
• Molecule 2: PACKAGED DNA STABILIZATION PROTEIN GP4

Chain i:



• Molecule 2: PACKAGED DNA STABILIZATION PROTEIN GP4

Chain j:



4 Data and refinement statistics (i)

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
$< I/\sigma(I) >$ [1]	1.88 (at 3.26Å)	Xtriage
Refinement program	PHENIX (PHENIX.REFINE: 1.5_2)	Depositor
R, R_{free}	0.222 , 0.236 0.318 , (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)	0.31 , 107.8	EDS
L-test for twinning[2]	$< \|L\| > = 0.45$, $< L^2 > = 0.27$	Xtriage
Estimated twinning fraction	0.025 for h,-k,-l	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.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

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	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(^{o})	Ideal(^{o})
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 Too-close contacts (i)

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

5.3 Torsion angles (i)

5.3.1 Protein backbone (i)

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all 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	6
1	B	565/602 (94%)	434 (77%)	87 (15%)	44 (8%)	1	6
1	C	565/602 (94%)	434 (77%)	87 (15%)	44 (8%)	1	6
1	D	565/602 (94%)	434 (77%)	88 (16%)	43 (8%)	1	6
1	E	565/602 (94%)	434 (77%)	88 (16%)	43 (8%)	1	6
1	F	565/602 (94%)	434 (77%)	88 (16%)	43 (8%)	1	6
1	G	565/602 (94%)	434 (77%)	88 (16%)	43 (8%)	1	6
1	H	565/602 (94%)	434 (77%)	88 (16%)	43 (8%)	1	6
1	I	565/602 (94%)	434 (77%)	88 (16%)	43 (8%)	1	6
1	J	565/602 (94%)	434 (77%)	88 (16%)	43 (8%)	1	6
1	K	565/602 (94%)	434 (77%)	88 (16%)	43 (8%)	1	6
1	L	565/602 (94%)	434 (77%)	87 (15%)	44 (8%)	1	6
1	M	565/602 (94%)	446 (79%)	81 (14%)	38 (7%)	1	8
1	N	565/602 (94%)	447 (79%)	79 (14%)	39 (7%)	1	8
1	O	565/602 (94%)	447 (79%)	80 (14%)	38 (7%)	1	8
1	P	565/602 (94%)	448 (79%)	79 (14%)	38 (7%)	1	8
1	Q	565/602 (94%)	447 (79%)	79 (14%)	39 (7%)	1	8
1	R	565/602 (94%)	448 (79%)	79 (14%)	38 (7%)	1	8
1	S	565/602 (94%)	443 (78%)	83 (15%)	39 (7%)	1	8
1	T	565/602 (94%)	446 (79%)	81 (14%)	38 (7%)	1	8
1	U	565/602 (94%)	447 (79%)	79 (14%)	39 (7%)	1	8
1	V	565/602 (94%)	445 (79%)	82 (14%)	38 (7%)	1	8
1	W	565/602 (94%)	448 (79%)	78 (14%)	39 (7%)	1	8
1	X	565/602 (94%)	448 (79%)	78 (14%)	39 (7%)	1	8
2	Y	144/166 (87%)	115 (80%)	16 (11%)	13 (9%)	1	4
2	Z	144/166 (87%)	115 (80%)	16 (11%)	13 (9%)	1	4
2	a	144/166 (87%)	115 (80%)	16 (11%)	13 (9%)	1	4
2	b	144/166 (87%)	115 (80%)	16 (11%)	13 (9%)	1	4
2	c	144/166 (87%)	115 (80%)	16 (11%)	13 (9%)	1	4
2	d	144/166 (87%)	115 (80%)	16 (11%)	13 (9%)	1	4
2	e	144/166 (87%)	115 (80%)	16 (11%)	13 (9%)	1	4
2	f	144/166 (87%)	115 (80%)	16 (11%)	13 (9%)	1	4

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	g	144/166 (87%)	115 (80%)	16 (11%)	13 (9%)	1	4
2	h	144/166 (87%)	115 (80%)	16 (11%)	13 (9%)	1	4
2	i	144/166 (87%)	115 (80%)	16 (11%)	13 (9%)	1	4
2	j	144/166 (87%)	115 (80%)	16 (11%)	13 (9%)	1	4
2	k	143/166 (86%)	115 (80%)	17 (12%)	11 (8%)	1	6
2	l	143/166 (86%)	114 (80%)	18 (13%)	11 (8%)	1	6
2	m	143/166 (86%)	115 (80%)	18 (13%)	10 (7%)	1	7
2	n	143/166 (86%)	114 (80%)	18 (13%)	11 (8%)	1	6
2	o	143/166 (86%)	115 (80%)	17 (12%)	11 (8%)	1	6
2	p	143/166 (86%)	115 (80%)	17 (12%)	11 (8%)	1	6
2	q	143/166 (86%)	113 (79%)	19 (13%)	11 (8%)	1	6
2	r	143/166 (86%)	114 (80%)	18 (13%)	11 (8%)	1	6
2	s	143/166 (86%)	114 (80%)	18 (13%)	11 (8%)	1	6
2	t	143/166 (86%)	113 (79%)	19 (13%)	11 (8%)	1	6
2	u	143/166 (86%)	114 (80%)	18 (13%)	11 (8%)	1	6
2	v	143/166 (86%)	114 (80%)	19 (13%)	10 (7%)	1	7
All	All	17004/18432 (92%)	13318 (78%)	2419 (14%)	1267 (8%)	1	7

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 (i)

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%)	3	13
1	B	483/510 (95%)	412 (85%)	71 (15%)	3	13
1	C	483/510 (95%)	412 (85%)	71 (15%)	3	13
1	D	483/510 (95%)	412 (85%)	71 (15%)	3	13
1	E	483/510 (95%)	412 (85%)	71 (15%)	3	13
1	F	483/510 (95%)	412 (85%)	71 (15%)	3	13
1	G	483/510 (95%)	412 (85%)	71 (15%)	3	13
1	H	483/510 (95%)	412 (85%)	71 (15%)	3	13
1	I	483/510 (95%)	412 (85%)	71 (15%)	3	13
1	J	483/510 (95%)	412 (85%)	71 (15%)	3	13
1	K	483/510 (95%)	412 (85%)	71 (15%)	3	13
1	L	483/510 (95%)	412 (85%)	71 (15%)	3	13
1	M	485/510 (95%)	412 (85%)	73 (15%)	3	12
1	N	485/510 (95%)	412 (85%)	73 (15%)	3	12
1	O	485/510 (95%)	413 (85%)	72 (15%)	3	13
1	P	485/510 (95%)	413 (85%)	72 (15%)	3	13
1	Q	485/510 (95%)	412 (85%)	73 (15%)	3	12
1	R	485/510 (95%)	411 (85%)	74 (15%)	2	12
1	S	485/510 (95%)	411 (85%)	74 (15%)	2	12
1	T	485/510 (95%)	412 (85%)	73 (15%)	3	12
1	U	485/510 (95%)	413 (85%)	72 (15%)	3	13
1	V	485/510 (95%)	413 (85%)	72 (15%)	3	13
1	W	485/510 (95%)	413 (85%)	72 (15%)	3	13
1	X	485/510 (95%)	411 (85%)	74 (15%)	2	12
2	Y	98/132 (74%)	85 (87%)	13 (13%)	4	16
2	Z	98/132 (74%)	85 (87%)	13 (13%)	4	16
2	a	98/132 (74%)	84 (86%)	14 (14%)	3	14
2	b	98/132 (74%)	84 (86%)	14 (14%)	3	14
2	c	98/132 (74%)	85 (87%)	13 (13%)	4	16
2	d	98/132 (74%)	85 (87%)	13 (13%)	4	16
2	e	98/132 (74%)	85 (87%)	13 (13%)	4	16
2	f	98/132 (74%)	84 (86%)	14 (14%)	3	14

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
2	g	98/132 (74%)	84 (86%)	14 (14%)	3	14
2	h	98/132 (74%)	85 (87%)	13 (13%)	4	16
2	i	98/132 (74%)	85 (87%)	13 (13%)	4	16
2	j	98/132 (74%)	84 (86%)	14 (14%)	3	14
2	k	100/132 (76%)	88 (88%)	12 (12%)	5	20
2	l	100/132 (76%)	87 (87%)	13 (13%)	4	17
2	m	100/132 (76%)	88 (88%)	12 (12%)	5	20
2	n	100/132 (76%)	87 (87%)	13 (13%)	4	17
2	o	100/132 (76%)	85 (85%)	15 (15%)	3	12
2	p	100/132 (76%)	86 (86%)	14 (14%)	3	15
2	q	100/132 (76%)	88 (88%)	12 (12%)	5	20
2	r	100/132 (76%)	88 (88%)	12 (12%)	5	20
2	s	100/132 (76%)	87 (87%)	13 (13%)	4	17
2	t	100/132 (76%)	87 (87%)	13 (13%)	4	17
2	u	100/132 (76%)	87 (87%)	13 (13%)	4	17
2	v	100/132 (76%)	87 (87%)	13 (13%)	4	17
All	All	13992/15408 (91%)	11950 (85%)	2042 (15%)	3	13

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

Mol	Chain	Res	Type
2	o	29	GLU
1	K	234	ILE
1	B	327	GLN
1	K	123	VAL
2	Y	145	SER

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

Mol	Chain	Res	Type
1	I	214	GLN
1	L	73	GLN
1	I	529	GLN
1	J	555	GLN
1	L	439	ASN

5.3.3 RNA

There are no RNA molecules 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 monosaccharides 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

Unable to reproduce the depositors R factor - this section is therefore empty.

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

Unable to reproduce the depositors R factor - this section is therefore empty.

6.3 Carbohydrates

Unable to reproduce the depositors R factor - this section is therefore empty.

6.4 Ligands

Unable to reproduce the depositors R factor - this section is therefore empty.

6.5 Other polymers

Unable to reproduce the depositors R factor - this section is therefore empty.