



wwPDB/EMDataBank EM Map/Model Validation Summary Report ⓘ

Jul 12, 2018 – 06:08 PM EDT

PDB ID : 3J2N
EMDB ID: : EMD-1086
Title : The X-ray structure of the gp15 hexamer and the model of the gp18 protein
fitted into the cryo-EM reconstruction of the contracted T4 tail
Authors : Fokine, A.; Zhang, Z.; Kanamaru, S.; Bowman, V.D.; Aksyuk, A.; Arisaka, F.;
Rao, V.B.; Rossmann, M.G.
Deposited on : 2012-11-10
Resolution : 16.00 Å(reported)

This is a wwPDB/EMDataBank EM Map/Model Validation Summary Report
for a publicly released PDB/EMDB entry.

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

A user guide is available at

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

with specific help available everywhere you see the ⓘ symbol.

MolProbity : 4.02b-467
Percentile statistics : 20171227.v01 (using entries in the PDB archive December 27th 2017)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP) : rb-20031172

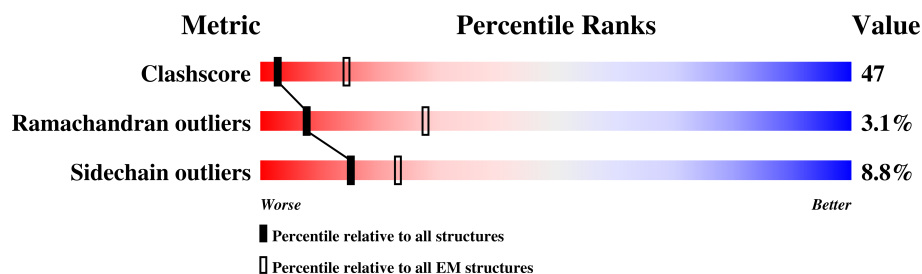
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

ELECTRON MICROSCOPY

The reported resolution of this entry is 16.00 Å.

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



Metric	Whole archive (#Entries)	EM structures (#Entries)
Clashscore	136327	1886
Ramachandran outliers	132723	1663
Sidechain outliers	132532	1531

The table below summarises the geometric issues observed across the polymeric chains. The red, orange, yellow and green segments on the bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria. 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 $\leq 5\%$

Mol	Chain	Length	Quality of chain
1	A	272	58% 16% • 22%
1	B	272	60% 15% • 22%
1	C	272	58% 16% • 22%
1	D	272	61% 13% • 22%
1	E	272	60% 14% • 22%
1	F	272	60% 14% • 22%
2	U	659	34% 49% 9% 8%
2	V	659	33% 50% 9% 8%
2	W	659	33% 50% 9% 8%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain			
2	X	659	<div><div></div></div> 34%	<div><div></div></div> 49%	<div><div></div></div> 9%	<div><div></div></div> 8%
2	Y	659	<div><div></div></div> 34%	<div><div></div></div> 50%	<div><div></div></div> 9%	<div><div></div></div> 8%
2	Z	659	<div><div></div></div> 33%	<div><div></div></div> 50%	<div><div></div></div> 9%	<div><div></div></div> 8%

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 38334 atoms, of which 0 are hydrogens and 0 are deuteriums.

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

- Molecule 1 is a protein called Tail connector protein Gp15.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	211	Total	C	N	O	S	0	0
			1742	1123	283	328	8		
1	B	211	Total	C	N	O	S	0	0
			1742	1123	283	328	8		
1	C	211	Total	C	N	O	S	0	0
			1742	1123	283	328	8		
1	D	211	Total	C	N	O	S	0	0
			1742	1123	283	328	8		
1	E	211	Total	C	N	O	S	0	0
			1742	1123	283	328	8		
1	F	211	Total	C	N	O	S	0	0
			1742	1123	283	328	8		

- Molecule 2 is a protein called Tail sheath protein Gp18.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	U	609	Total	C	N	O	S	0	0
			4647	2929	785	923	10		
2	V	609	Total	C	N	O	S	0	0
			4647	2929	785	923	10		
2	W	609	Total	C	N	O	S	0	0
			4647	2929	785	923	10		
2	X	609	Total	C	N	O	S	0	0
			4647	2929	785	923	10		
2	Y	609	Total	C	N	O	S	0	0
			4647	2929	785	923	10		
2	Z	609	Total	C	N	O	S	0	0
			4647	2929	785	923	10		

There are 48 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
U	100	GLU	ASP	SEE REMARK 999	UNP P13332
U	148	ALA	GLY	SEE REMARK 999	UNP P13332

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
U	150	ILE	ASN	SEE REMARK 999	UNP P13332
U	151	ILE	TYR	SEE REMARK 999	UNP P13332
U	301	GLY	GLU	SEE REMARK 999	UNP P13332
U	399	VAL	ALA	SEE REMARK 999	UNP P13332
U	454	TYR	HIS	SEE REMARK 999	UNP P13332
U	510	PRO	ARG	ENGINEERED MUTATION	UNP P13332
V	100	GLU	ASP	SEE REMARK 999	UNP P13332
V	148	ALA	GLY	SEE REMARK 999	UNP P13332
V	150	ILE	ASN	SEE REMARK 999	UNP P13332
V	151	ILE	TYR	SEE REMARK 999	UNP P13332
V	301	GLY	GLU	SEE REMARK 999	UNP P13332
V	399	VAL	ALA	SEE REMARK 999	UNP P13332
V	454	TYR	HIS	SEE REMARK 999	UNP P13332
V	510	PRO	ARG	ENGINEERED MUTATION	UNP P13332
W	100	GLU	ASP	SEE REMARK 999	UNP P13332
W	148	ALA	GLY	SEE REMARK 999	UNP P13332
W	150	ILE	ASN	SEE REMARK 999	UNP P13332
W	151	ILE	TYR	SEE REMARK 999	UNP P13332
W	301	GLY	GLU	SEE REMARK 999	UNP P13332
W	399	VAL	ALA	SEE REMARK 999	UNP P13332
W	454	TYR	HIS	SEE REMARK 999	UNP P13332
W	510	PRO	ARG	ENGINEERED MUTATION	UNP P13332
X	100	GLU	ASP	SEE REMARK 999	UNP P13332
X	148	ALA	GLY	SEE REMARK 999	UNP P13332
X	150	ILE	ASN	SEE REMARK 999	UNP P13332
X	151	ILE	TYR	SEE REMARK 999	UNP P13332
X	301	GLY	GLU	SEE REMARK 999	UNP P13332
X	399	VAL	ALA	SEE REMARK 999	UNP P13332
X	454	TYR	HIS	SEE REMARK 999	UNP P13332
X	510	PRO	ARG	ENGINEERED MUTATION	UNP P13332
Y	100	GLU	ASP	SEE REMARK 999	UNP P13332
Y	148	ALA	GLY	SEE REMARK 999	UNP P13332
Y	150	ILE	ASN	SEE REMARK 999	UNP P13332
Y	151	ILE	TYR	SEE REMARK 999	UNP P13332
Y	301	GLY	GLU	SEE REMARK 999	UNP P13332
Y	399	VAL	ALA	SEE REMARK 999	UNP P13332
Y	454	TYR	HIS	SEE REMARK 999	UNP P13332
Y	510	PRO	ARG	ENGINEERED MUTATION	UNP P13332
Z	100	GLU	ASP	SEE REMARK 999	UNP P13332
Z	148	ALA	GLY	SEE REMARK 999	UNP P13332
Z	150	ILE	ASN	SEE REMARK 999	UNP P13332
Z	151	ILE	TYR	SEE REMARK 999	UNP P13332

Continued on next page...

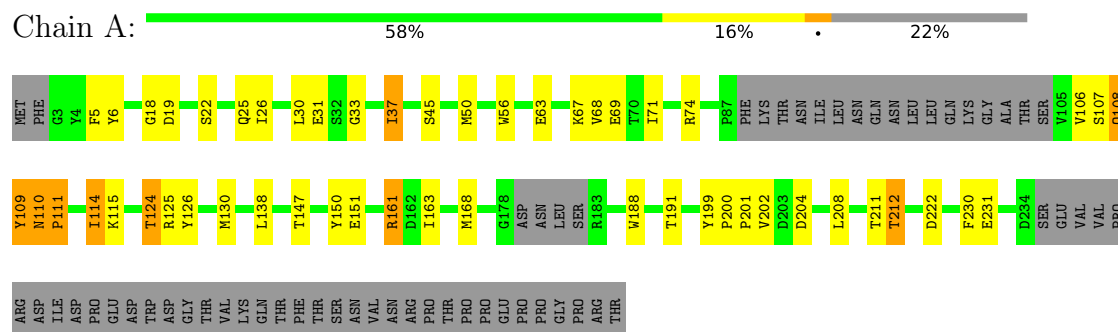
Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
Z	301	GLY	GLU	SEE REMARK 999	UNP P13332
Z	399	VAL	ALA	SEE REMARK 999	UNP P13332
Z	454	TYR	HIS	SEE REMARK 999	UNP P13332
Z	510	PRO	ARG	ENGINEERED MUTATION	UNP P13332

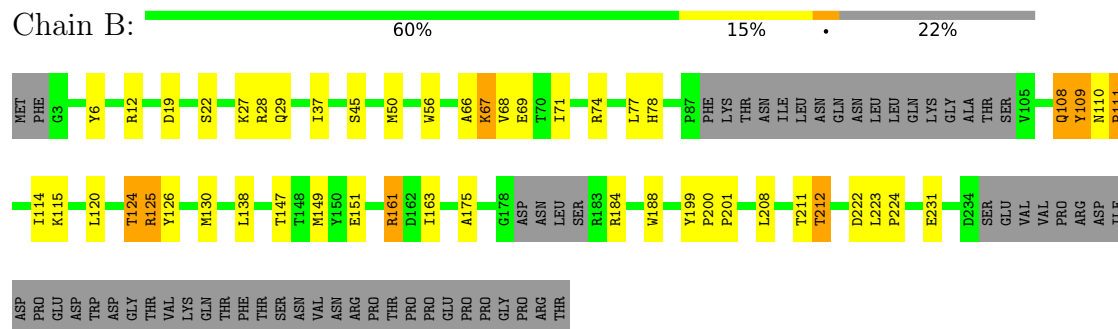
3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA 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.

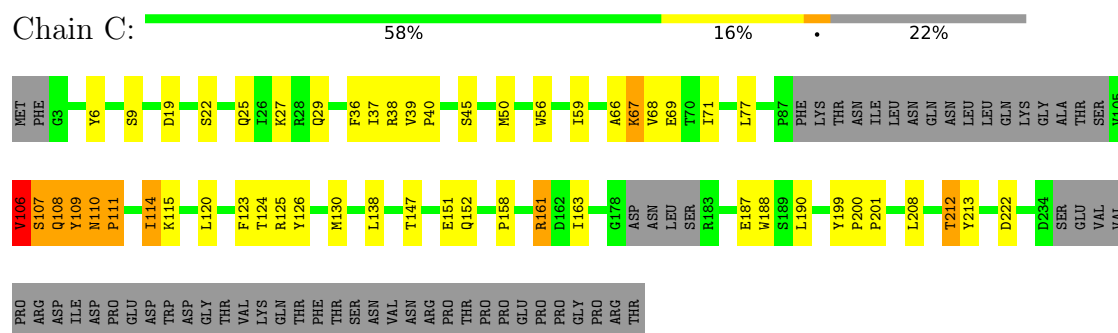
- Molecule 1: Tail connector protein Gp15



- Molecule 1: Tail connector protein Gp15

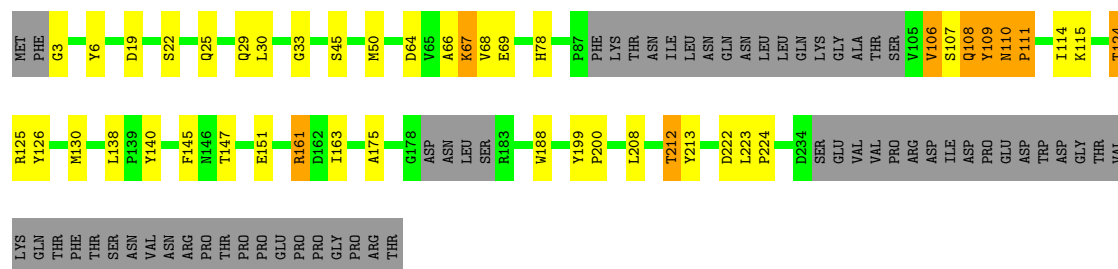


- Molecule 1: Tail connector protein Gp15



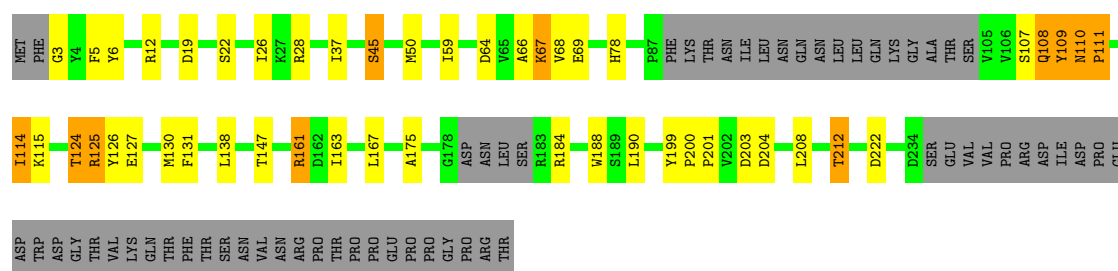
- Molecule 1: Tail connector protein Gp15

Chain D: 



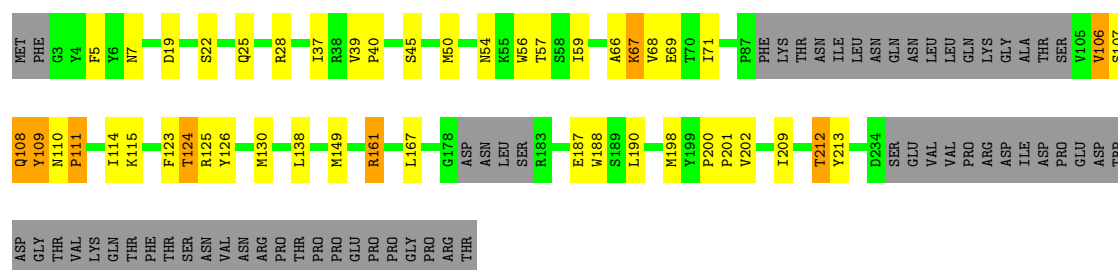
• Molecule 1: Tail connector protein Gp15

Chain E: 



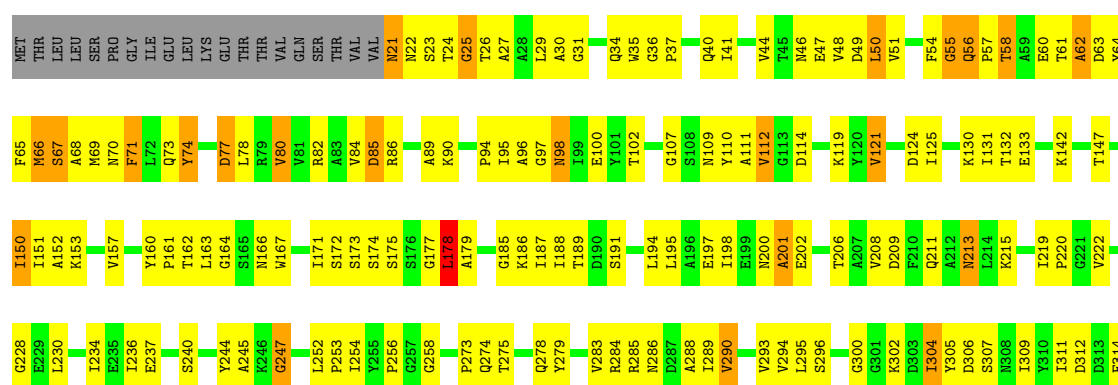
• Molecule 1: Tail connector protein Gp15

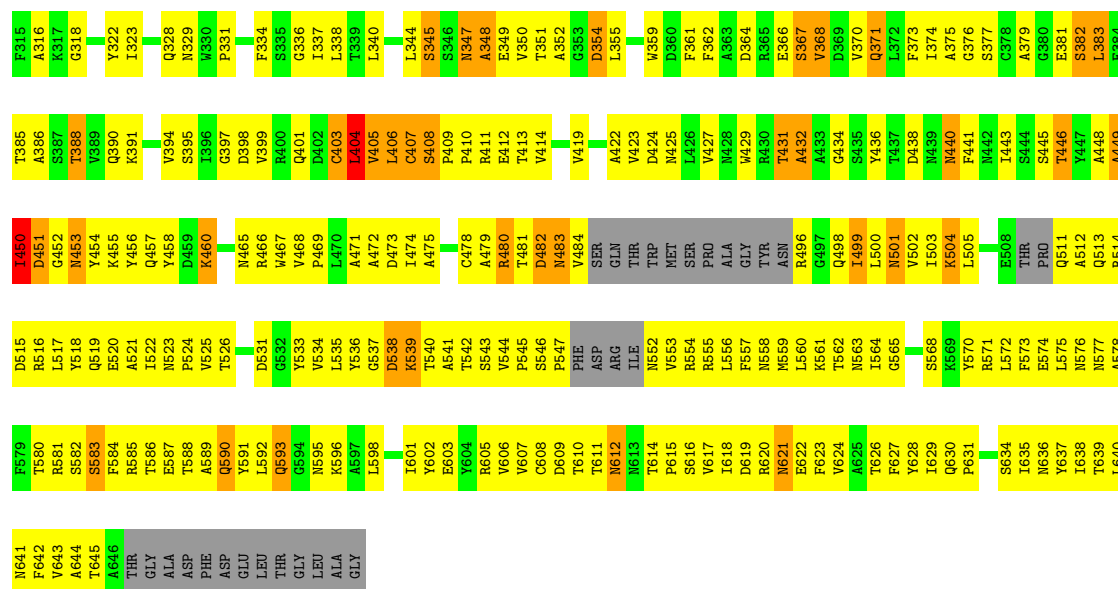
Chain F: 



• Molecule 2: Tail sheath protein Gp18

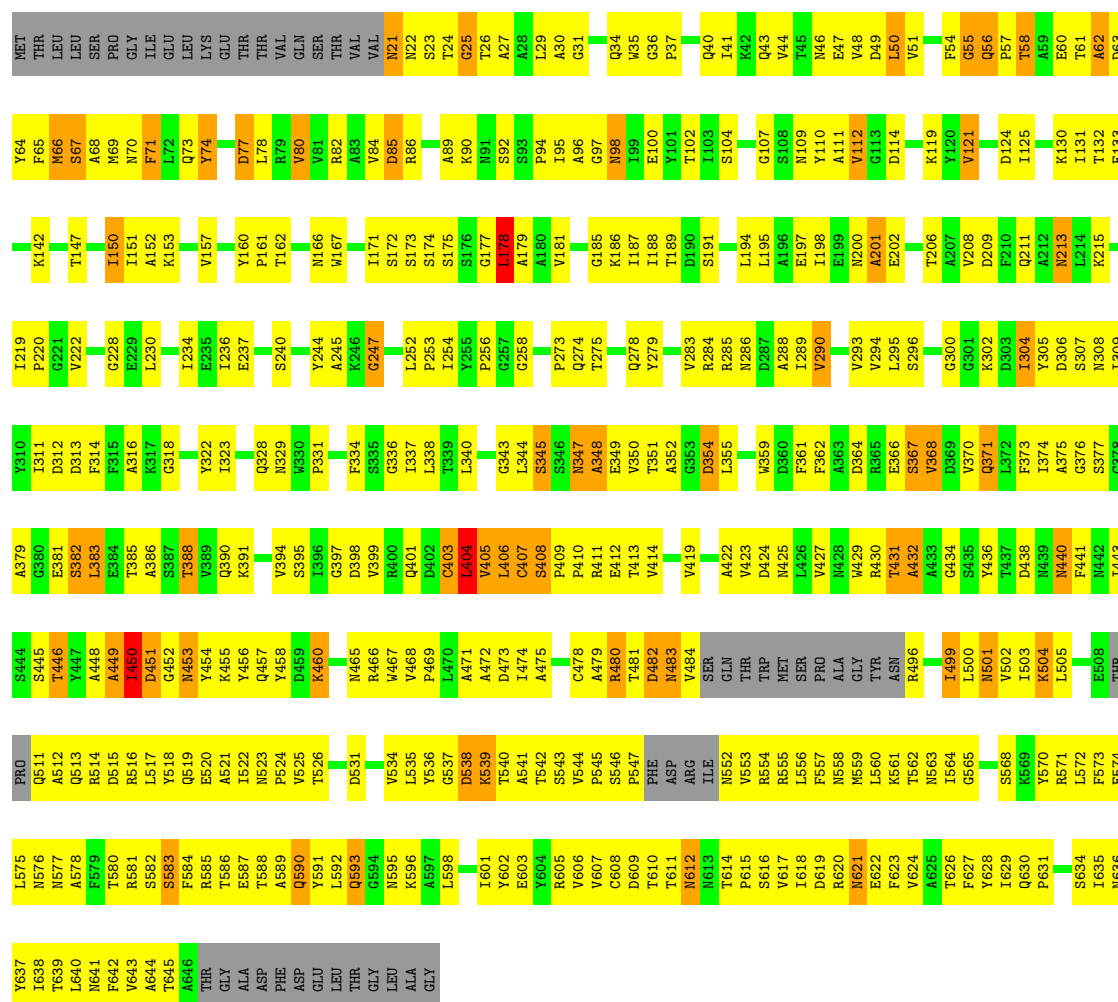
Chain U: 



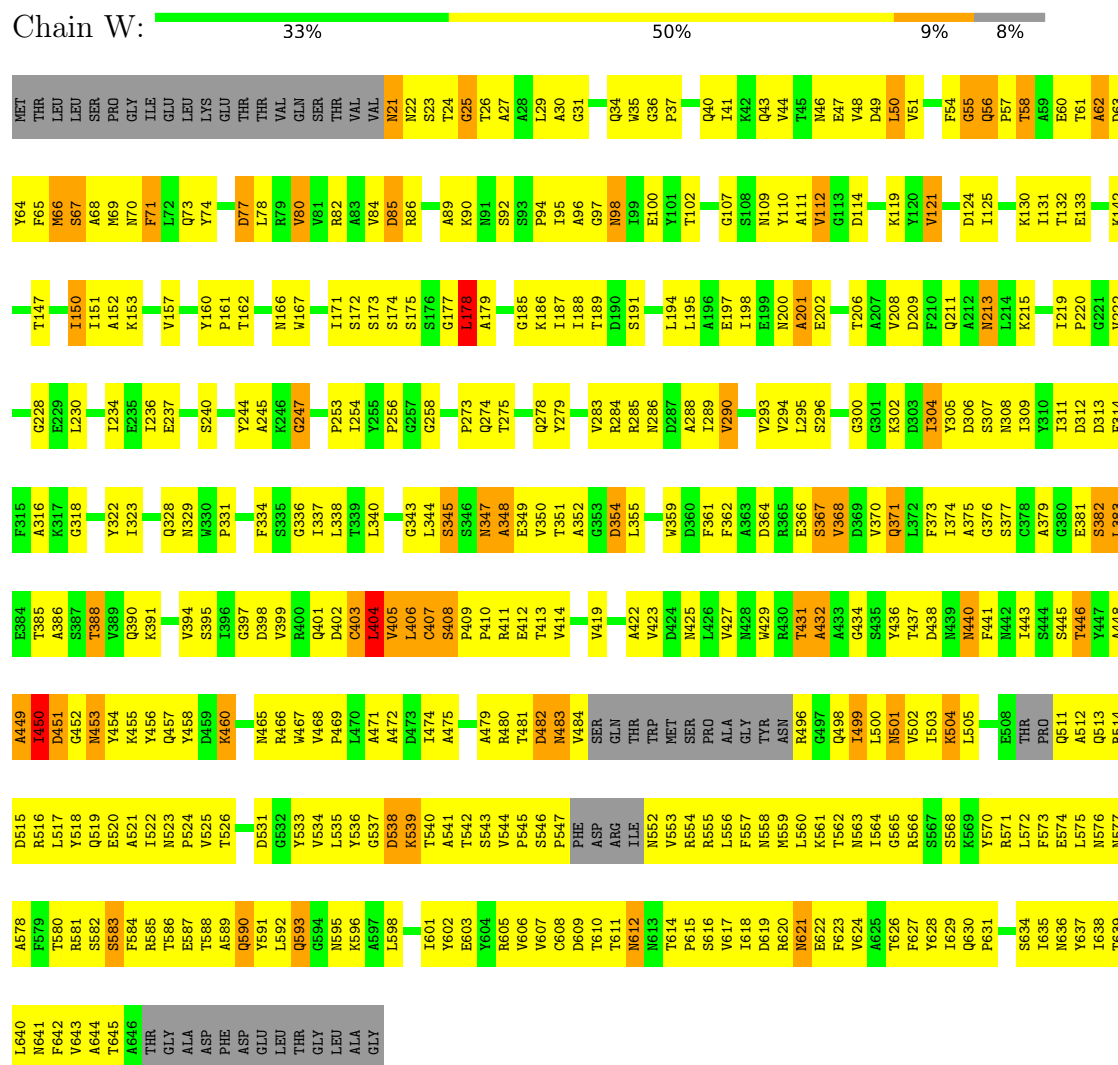


● Molecule 2: Tail sheath protein Gp18

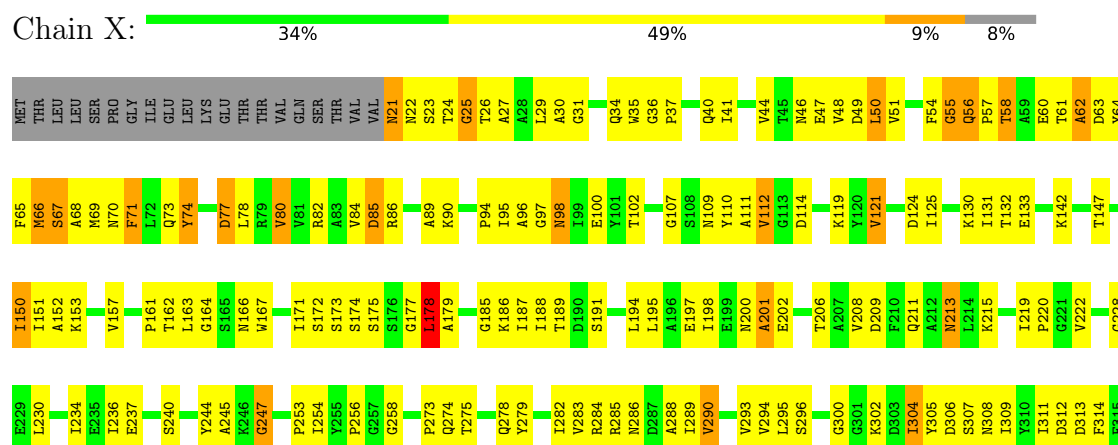
Chain V:



• Molecule 2: Tail sheath protein Gp18



• Molecule 2: Tail sheath protein Gp18





● Molecule 2: Tail sheath protein Gp18

Chain Z:  33% 50% 9% 8%

F642	T580	R516	D451	A386	A316	G228	Y64	MET
V643	R581	L517	C452	S387	K317	E229	F65	THR
A644	S582	Y518	N453	T388	G318	L230	M66	LEU
T645	S583	Q519	Y454	V389	Y322	I234	S67	LEU
A646	F584	E520	K455	Q390	I323	E235	A68	SER
THR	R585	A521	Y456	K391	Y322	E235	M69	PRO
GLY	T586	I522	Q457	V394	Q328	I236	N70	GLY
ALA	E587	N523	Y458	S395	N329	E237	F71	ILE
ASP	T588	P524	A459	S395	V330	S240	L72	GLU
PHE	A589	V525	K460	I396	V330	S240	Q73	LEU
ASP	O590	T526		G397	P331		Y74	LYS
GLU	Y591		M465	D398	F334	Y244		GLU
LEU	L592	D531	R466	V399	S335	A245	D77	THR
THR	O593	G532	W467	R400	S335	K246	L78	THR
GLY	Y594	Y533	V468	Q401	G336	G247	R79	VAL
LEU	V594	V534	P469	D402	I337		V80	GLN
ALA	K596	L535	L470	C403	L338		V81	SER
A597		Y536	A471	L403	T339	P253	R82	THR
	L598	G537	D472	V405	L340	I254	A83	VAL
		D538	A473	L406	L344	T255	V84	VAL
I601	Y602	T540	I474	C407	S345	G257	D85	N21
E603	E603	A541	A475	S408	S346	G258	R86	N22
Y604		T542		P409	S346		G87	S23
R605		S543	C478	P410	N347	P273	A89	T24
V606	R605	V544	A479	R411	A348	Q274	K90	G25
V607	V607	P545	R480	E412	E349	T275		T26
C608	V607	P545	T481	T413	V350		P94	A27
D609	P608	P547	D482	V414	T351	Q278	I95	A28
T610	P609	PHE	N483		A352	Y279	A96	L29
		ASP	V484	V419	G353		I98	A30
N612	T611	ARG	SER	A422	D354	V283	N98	G31
R613		ILE	GLN	V423	L355	R284	T99	
T614		N552	THR	D424	W359	R285	E100	Q34
P615		V553	TRP	N425	D360	N286	Y101	W35
S616		R554	MET	L426	F361	A288	T102	G36
V617		R555	SER	V427	F362	I289	G107	P37
T618		L556	PRO	N428	A363	V290	S108	Q40
D619		F557	ALA	W429	D364		N109	I41
R620		N558	GLY	R430	R365	V293	Y110	K42
N621		M559	TYR	T431	R365	V294	A111	Q43
		L560	ASN	A431	E366	V294	N200	Q43
E622		K561	R496	A432	S367	L295	A201	V44
F623		T562	G497	A433	V368	S296	E202	V44
V624		S661	Q498	G434	D369		G113	T45
A625		I564	I499	S435	D369	G300	D114	M46
F627		N501	L500	Y436	Q371	G301	K119	E47
Y628		V502	N501	T437	L372	K302	V120	V48
		I503	D438	D438	F373	I304	W121	D49
T629		K504	N439	M439	I374		D209	L50
P631		L505	F441	M440	A375	D306	Q211	V51
		R571	I442	N442	G376	S307	K215	F54
S634		L572	S508	I443	S377	N308	I214	G55
T635		F573	THR	S444	A379	I309	I131	Q56
N636		E574	PRO	T446	G380	Y310	T132	P57
Y637		L575	O511	S445	E381	I311	E133	T58
T638		N576	A512	T447	S382	D312	K142	E59
T639		N577	Q513	A448	L383	D313	G221	A60
		A578	R514	A449	F384	F314	V222	T61
L640		F579	D515	T449	T395	E215	T147	A62
N641								D63

4 Experimental information

Property	Value	Source
Reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, C6	Depositor
Number of particles used	1965	Depositor
Resolution determination method	FSC 0.5 CUT-OFF	Depositor
CTF correction method	Each particle	Depositor
Microscope	FEI/PHILIPS CM300FEG/T	Depositor
Voltage (kV)	300	Depositor
Electron dose ($e^-/\text{\AA}^2$)	20	Depositor
Minimum defocus (nm)	500	Depositor
Maximum defocus (nm)	3400	Depositor
Magnification	47000	Depositor
Image detector	KODAK SO-163 FILM	Depositor

5 Model quality

5.1 Standard geometry

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

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	# Z >2	RMSZ	# Z >2
1	A	0.49	0/1787	0.64	0/2421
1	B	0.47	0/1787	0.63	0/2421
1	C	0.48	0/1787	0.64	1/2421 (0.0%)
1	D	0.49	0/1787	0.63	0/2421
1	E	0.49	0/1787	0.63	0/2421
1	F	0.50	0/1787	0.65	0/2421
2	U	0.60	0/4729	0.90	21/6427 (0.3%)
2	V	0.60	0/4729	0.89	21/6427 (0.3%)
2	W	0.60	0/4729	0.90	21/6427 (0.3%)
2	X	0.59	0/4729	0.90	21/6427 (0.3%)
2	Y	0.60	0/4729	0.90	21/6427 (0.3%)
2	Z	0.60	0/4729	0.90	21/6427 (0.3%)
All	All	0.57	0/39096	0.83	127/53088 (0.2%)

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

Mol	Chain	#Chirality outliers	#Planarity outliers
2	U	0	4
2	V	0	4
2	W	0	4
2	X	0	4
2	Y	0	4
2	Z	0	4
All	All	0	24

There are no bond length outliers.

The worst 5 of 127 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	Y	55	GLY	N-CA-C	18.05	158.21	113.10

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	V	55	GLY	N-CA-C	18.03	158.18	113.10
2	X	55	GLY	N-CA-C	18.03	158.16	113.10
2	W	55	GLY	N-CA-C	18.01	158.13	113.10
2	Z	55	GLY	N-CA-C	18.01	158.12	113.10

There are no chirality outliers.

5 of 24 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	U	21	ASN	Peptide
2	U	450	ILE	Peptide
2	U	451	ASP	Peptide
2	U	452	GLY	Peptide
2	V	21	ASN	Peptide

5.2 Too-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 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	1742	0	1673	29	0
1	B	1742	0	1673	33	0
1	C	1742	0	1673	30	0
1	D	1742	0	1673	29	0
1	E	1742	0	1673	32	0
1	F	1742	0	1673	28	0
2	U	4647	0	4563	565	0
2	V	4647	0	4564	571	0
2	W	4647	0	4564	568	0
2	X	4647	0	4564	565	0
2	Y	4647	0	4564	557	0
2	Z	4647	0	4564	561	0
All	All	38334	0	37421	3566	0

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 47.

The worst 5 of 3566 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:U:404:LEU:CG	2:U:554:ARG:HH12	1.22	1.52
2:V:404:LEU:CG	2:V:554:ARG:HH12	1.22	1.51
2:Z:404:LEU:CG	2:Z:554:ARG:HH12	1.23	1.51
2:Y:404:LEU:CG	2:Y:554:ARG:HH12	1.22	1.50
2:X:496:ARG:N	2:X:534:VAL:CB	1.73	1.50

There are no symmetry-related clashes.

5.3 Torsion angles [i](#)

5.3.1 Protein backbone [i](#)

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all PDB entries followed by that with respect to all EM entries.

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

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	205/272 (75%)	184 (90%)	18 (9%)	3 (2%)	11	51
1	B	205/272 (75%)	184 (90%)	17 (8%)	4 (2%)	8	45
1	C	205/272 (75%)	184 (90%)	17 (8%)	4 (2%)	8	45
1	D	205/272 (75%)	184 (90%)	17 (8%)	4 (2%)	8	45
1	E	205/272 (75%)	183 (89%)	18 (9%)	4 (2%)	8	45
1	F	205/272 (75%)	184 (90%)	18 (9%)	3 (2%)	11	51
2	U	601/659 (91%)	479 (80%)	101 (17%)	21 (4%)	4	32
2	V	601/659 (91%)	479 (80%)	101 (17%)	21 (4%)	4	32
2	W	601/659 (91%)	479 (80%)	101 (17%)	21 (4%)	4	32
2	X	601/659 (91%)	479 (80%)	101 (17%)	21 (4%)	4	32
2	Y	601/659 (91%)	481 (80%)	99 (16%)	21 (4%)	4	32
2	Z	601/659 (91%)	479 (80%)	101 (17%)	21 (4%)	4	32
All	All	4836/5586 (87%)	3979 (82%)	709 (15%)	148 (3%)	8	35

5 of 148 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	108	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	67	LYS
1	B	108	GLN
1	C	67	LYS
1	C	108	GLN

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 PDB entries followed by that with respect to all EM entries.

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

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	192/250 (77%)	176 (92%)	16 (8%)	12	41
1	B	192/250 (77%)	181 (94%)	11 (6%)	23	53
1	C	192/250 (77%)	176 (92%)	16 (8%)	12	41
1	D	192/250 (77%)	180 (94%)	12 (6%)	20	50
1	E	192/250 (77%)	180 (94%)	12 (6%)	20	50
1	F	192/250 (77%)	179 (93%)	13 (7%)	17	48
2	U	494/536 (92%)	448 (91%)	46 (9%)	10	34
2	V	494/536 (92%)	447 (90%)	47 (10%)	9	33
2	W	494/536 (92%)	447 (90%)	47 (10%)	9	33
2	X	494/536 (92%)	448 (91%)	46 (9%)	10	34
2	Y	494/536 (92%)	446 (90%)	48 (10%)	9	32
2	Z	494/536 (92%)	446 (90%)	48 (10%)	9	32
All	All	4116/4716 (87%)	3754 (91%)	362 (9%)	15	37

5 of 362 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	V	450	ILE
2	W	395	SER
2	Z	347	ASN
2	V	501	ASN
2	W	102	THR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 99 such sidechains are listed below:

Mol	Chain	Res	Type
2	W	513	GLN
2	X	440	ASN
2	Z	511	GLN
2	W	612	ASN
2	X	43	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 carbohydrates 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.