



wwPDB EM Validation Summary Report ⓘ

Nov 19, 2022 – 07:46 AM EST

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 EM Validation Summary Report for a publicly released PDB entry.

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

A user guide is available at

<https://www.wwpdb.org/validation/2017/EMValidationReportHelp>
with specific help available everywhere you see the ⓘ symbol.

The types of validation reports are described at

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

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

EMDB validation analysis : 0.0.1.dev43
MolProbity : 4.02b-467
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)
MapQ : 1.9.9
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.31.3

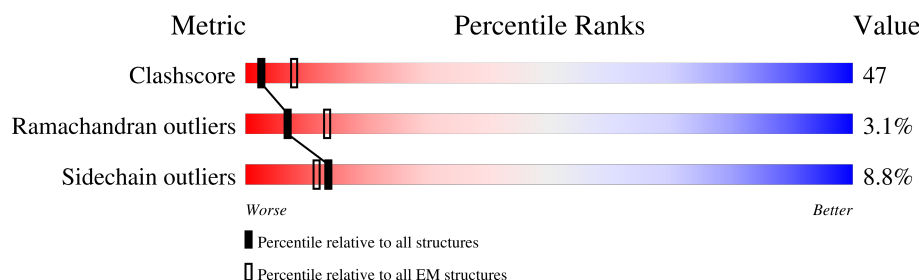
1 Overall quality at a glance

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	158937	4297
Ramachandran outliers	154571	4023
Sidechain outliers	154315	3826

The table below summarises the geometric issues observed across the polymeric chains and their fit to the map. The red, orange, yellow and green segments of the bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions $\leq 5\%$. The upper red bar (where present) indicates the fraction of residues that have poor fit to the EM map (all-atom inclusion $< 40\%$). The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	272	
1	B	272	
1	C	272	
1	D	272	
1	E	272	
1	F	272	
2	U	659	
2	V	659	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	W	659	<div><div><div>7%</div><div>33%</div><div>50%</div><div>9%</div><div>8%</div></div></div>
2	X	659	<div><div><div>7%</div><div>34%</div><div>49%</div><div>9%</div><div>8%</div></div></div>
2	Y	659	<div><div><div>7%</div><div>34%</div><div>50%</div><div>9%</div><div>8%</div></div></div>
2	Z	659	<div><div><div>7%</div><div>33%</div><div>50%</div><div>9%</div><div>8%</div></div></div>

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...

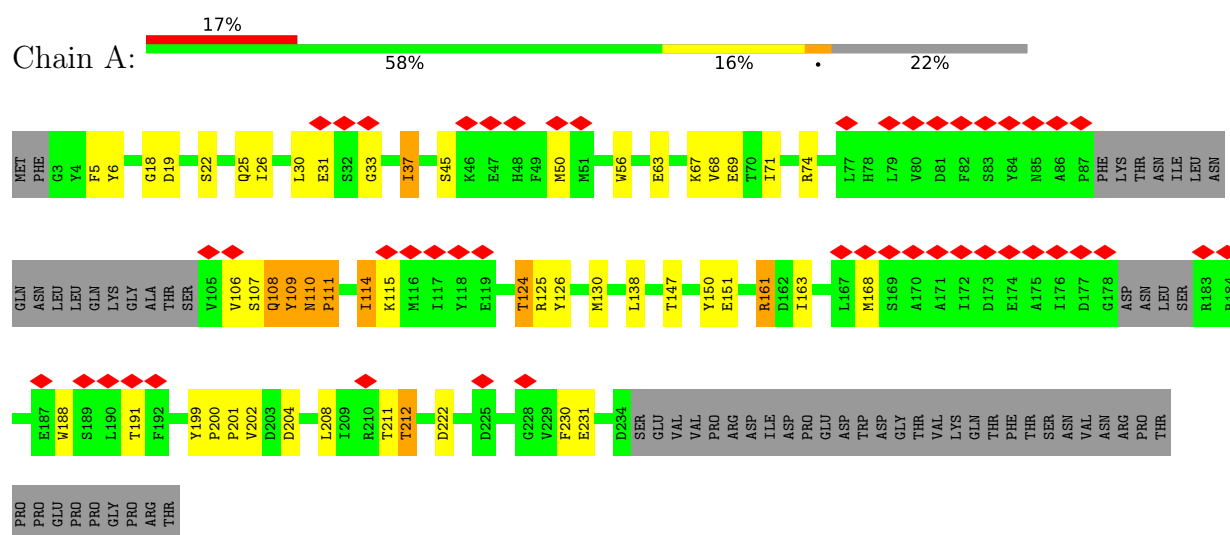
Continued from previous page...

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

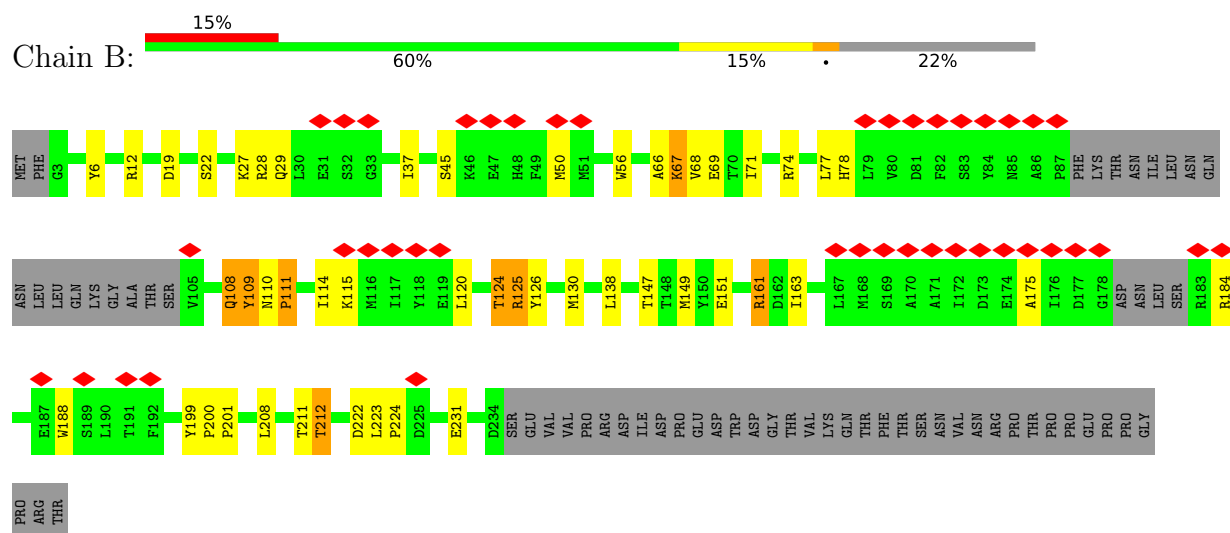
3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and atom inclusion in map density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red diamond above a residue indicates a poor fit to the EM map for this residue (all-atom inclusion < 40%). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

• Molecule 1: Tail connector protein Gp15



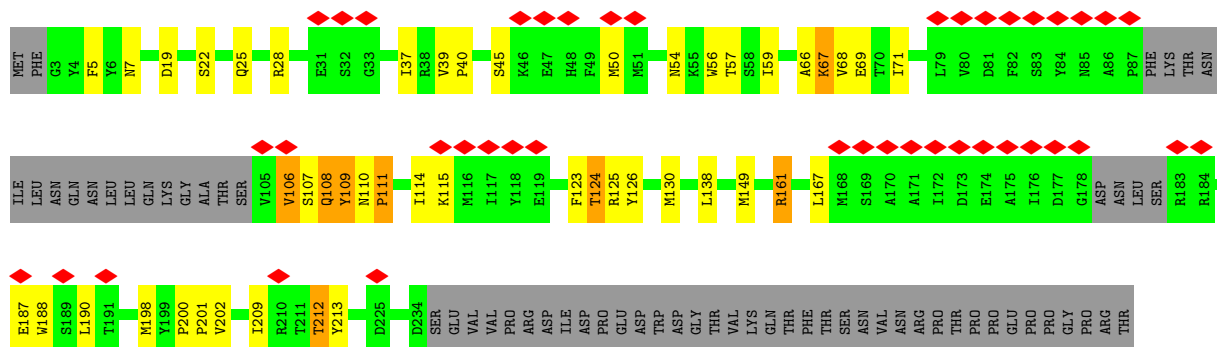
• Molecule 1: Tail connector protein Gp15



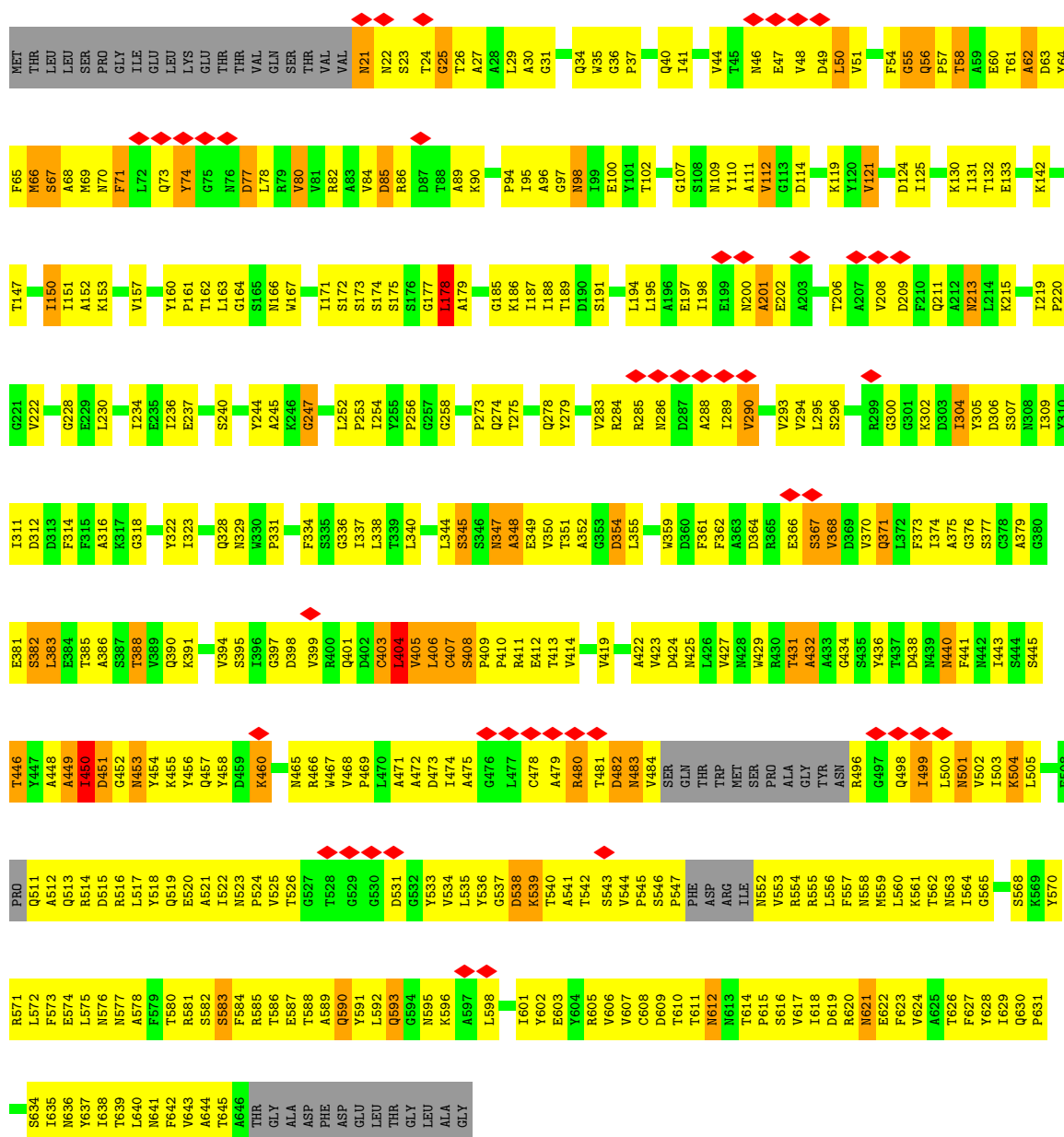
• Molecule 1: Tail connector protein Gp15



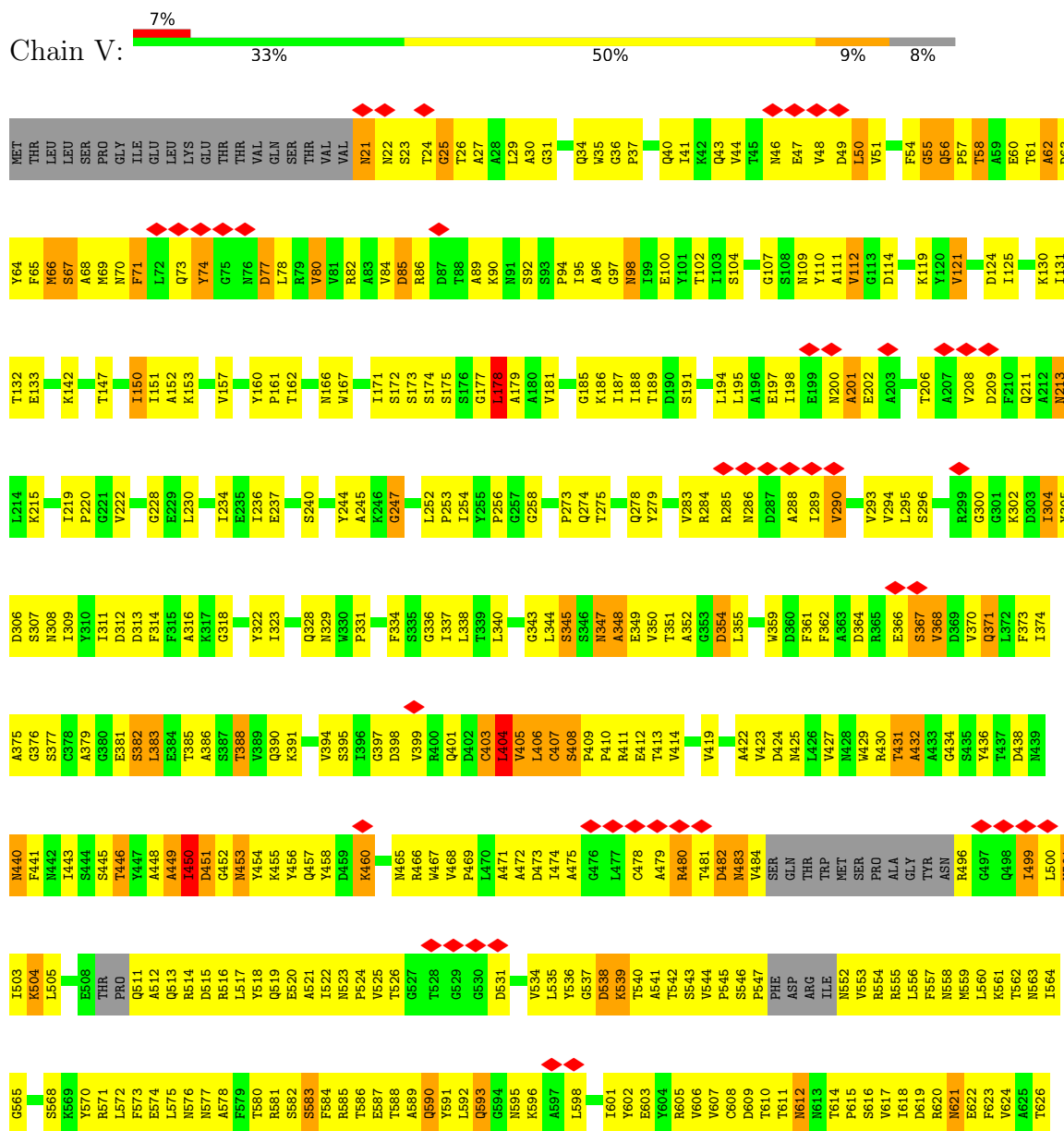




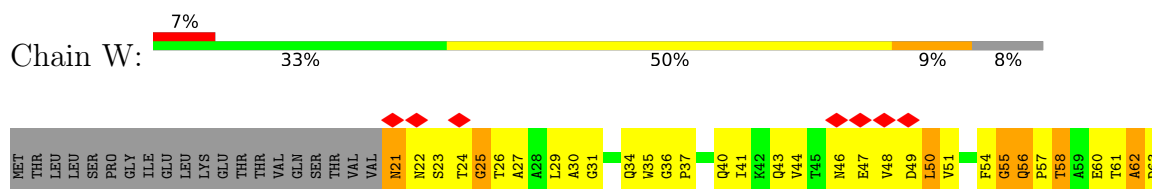
• Molecule 2: Tail sheath protein Gp18

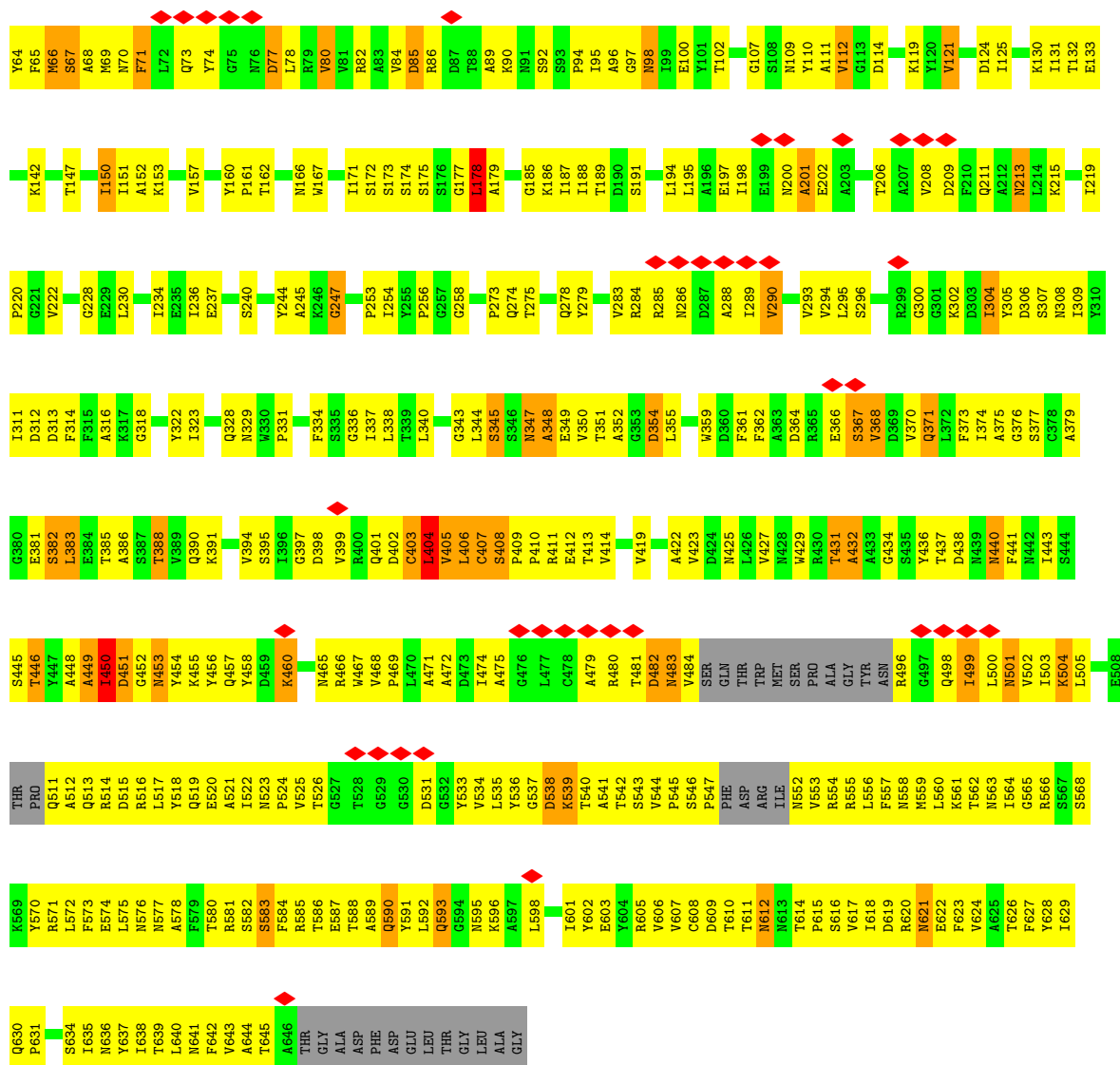


• Molecule 2: Tail sheath protein Gp18

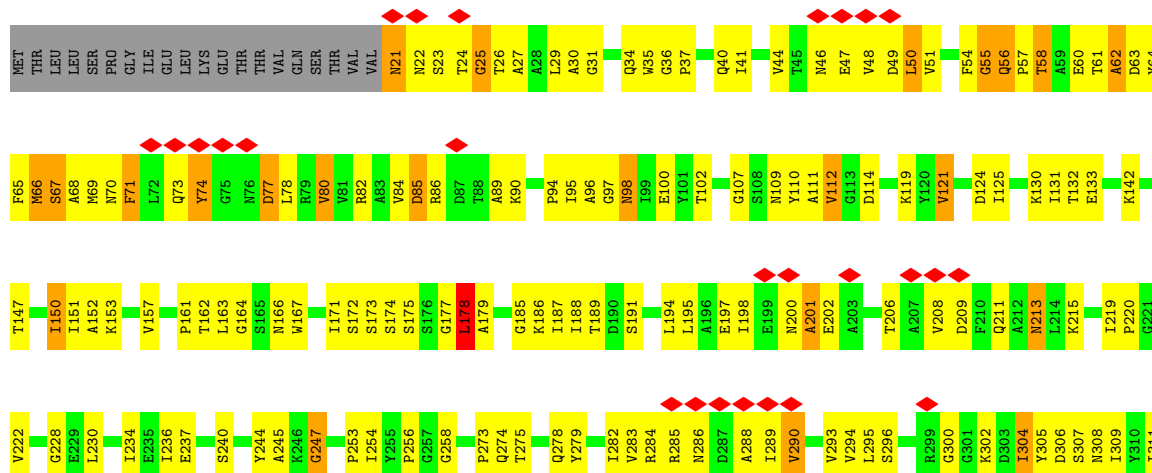


• Molecule 2: Tail sheath protein Gp18

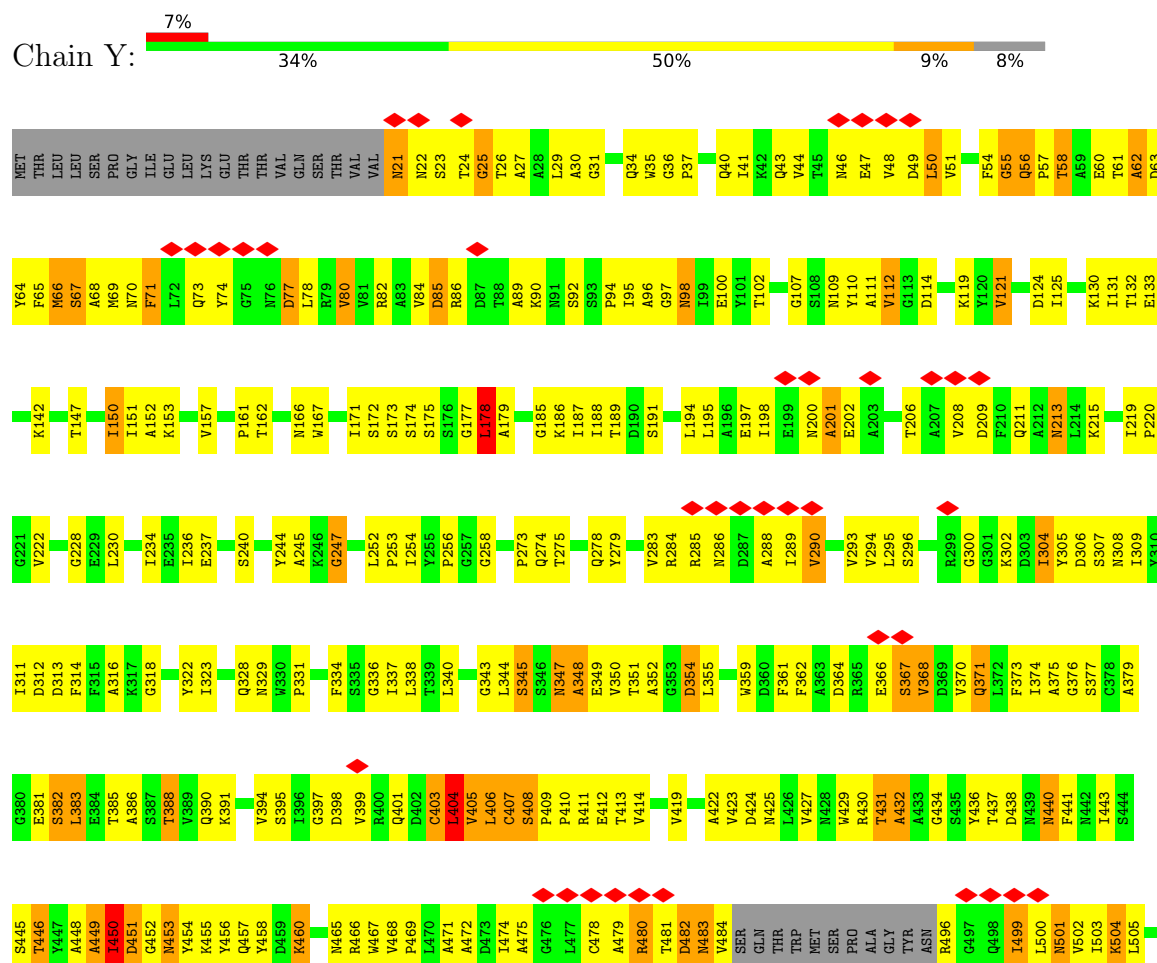


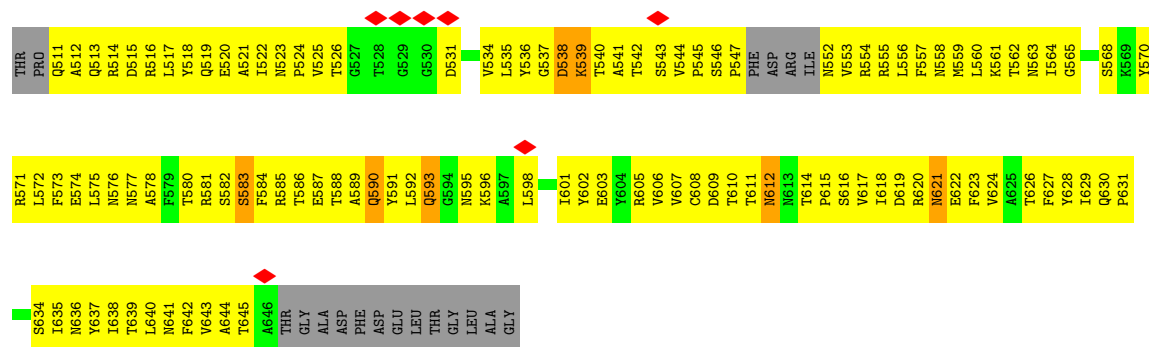


• Molecule 2: Tail sheath protein Gp18

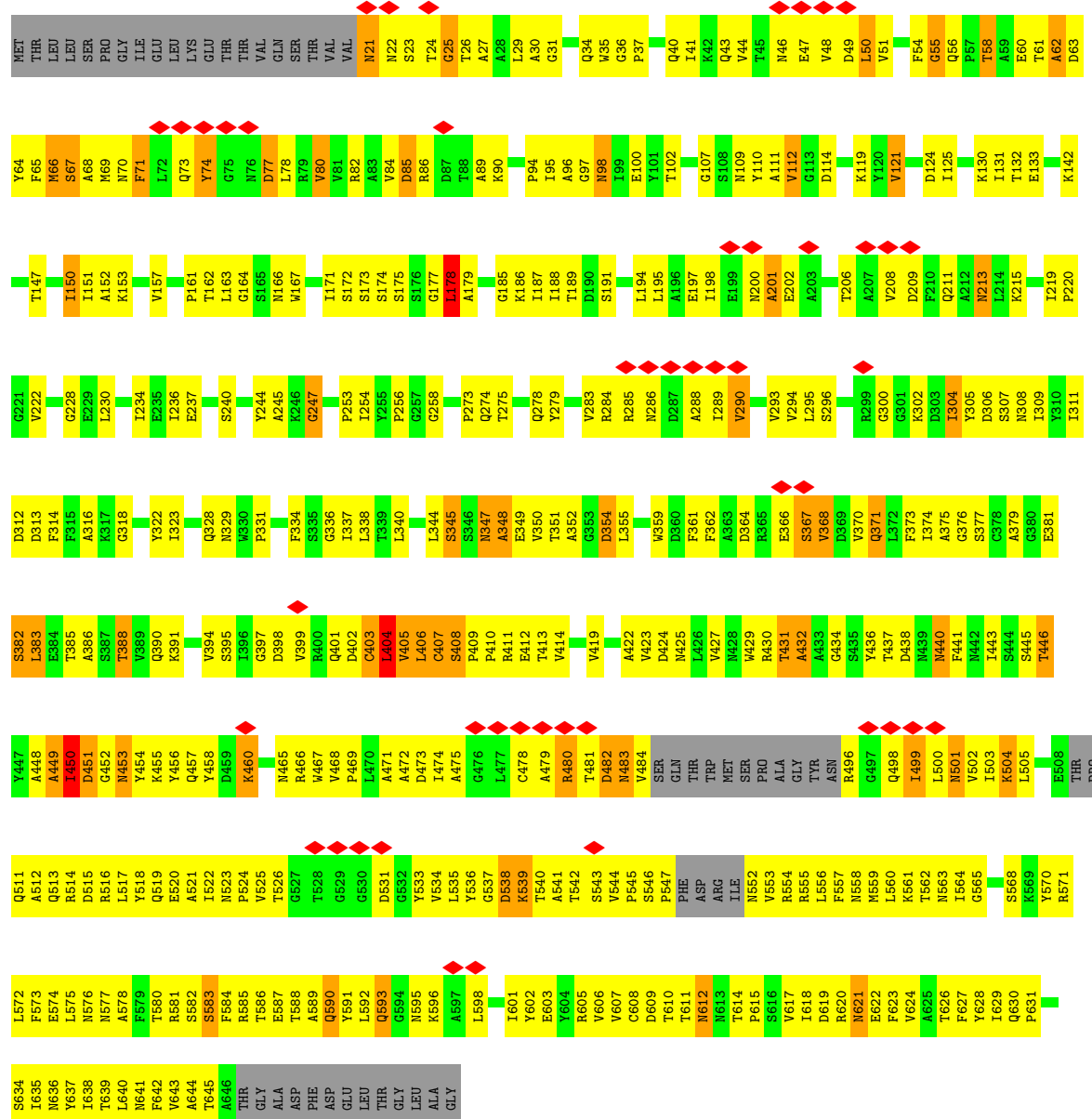


- Molecule 2: Tail sheath protein Gp18





• Molecule 2: Tail sheath protein Gp18



4 Experimental information

Property	Value	Source
EM 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
Maximum map value	14.635	Depositor
Minimum map value	-11.223	Depositor
Average map value	0.000	Depositor
Map value standard deviation	1.000	Depositor
Recommended contour level	1.5	Depositor
Map size (\AA)	1006.8096, 1006.8096, 1006.8096	wwPDB
Map dimensions	256, 256, 256	wwPDB
Map angles ($^\circ$)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (\AA)	3.93285, 3.93285, 3.93285	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 >5	RMSZ	# Z >5
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%)	10	46
1	B	205/272 (75%)	184 (90%)	17 (8%)	4 (2%)	7	38
1	C	205/272 (75%)	184 (90%)	17 (8%)	4 (2%)	7	38
1	D	205/272 (75%)	184 (90%)	17 (8%)	4 (2%)	7	38
1	E	205/272 (75%)	183 (89%)	18 (9%)	4 (2%)	7	38
1	F	205/272 (75%)	184 (90%)	18 (9%)	3 (2%)	10	46
2	U	601/659 (91%)	479 (80%)	101 (17%)	21 (4%)	3	25
2	V	601/659 (91%)	479 (80%)	101 (17%)	21 (4%)	3	25
2	W	601/659 (91%)	479 (80%)	101 (17%)	21 (4%)	3	25
2	X	601/659 (91%)	479 (80%)	101 (17%)	21 (4%)	3	25
2	Y	601/659 (91%)	481 (80%)	99 (16%)	21 (4%)	3	25
2	Z	601/659 (91%)	479 (80%)	101 (17%)	21 (4%)	3	25
All	All	4836/5586 (87%)	3979 (82%)	709 (15%)	148 (3%)	7	27

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%)	11	34
1	B	192/250 (77%)	181 (94%)	11 (6%)	20	45
1	C	192/250 (77%)	176 (92%)	16 (8%)	11	34
1	D	192/250 (77%)	180 (94%)	12 (6%)	18	43
1	E	192/250 (77%)	180 (94%)	12 (6%)	18	43
1	F	192/250 (77%)	179 (93%)	13 (7%)	16	41
2	U	494/536 (92%)	448 (91%)	46 (9%)	9	28
2	V	494/536 (92%)	447 (90%)	47 (10%)	8	27
2	W	494/536 (92%)	447 (90%)	47 (10%)	8	27
2	X	494/536 (92%)	448 (91%)	46 (9%)	9	28
2	Y	494/536 (92%)	446 (90%)	48 (10%)	8	27
2	Z	494/536 (92%)	446 (90%)	48 (10%)	8	27
All	All	4116/4716 (87%)	3754 (91%)	362 (9%)	13	31

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

Mol	Chain	Res	Type
2	X	150	ILE
2	Y	347	ASN
2	X	290	VAL
2	X	499	ILE
2	Y	437	THR

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

Mol	Chain	Res	Type
2	X	457	GLN
2	Y	390	GLN
2	X	513	GLN
2	X	641	ASN
2	Y	511	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 monosaccharides in this entry.

5.6 Ligand geometry [i](#)

There are no ligands in this entry.

5.7 Other polymers [i](#)

There are no such residues in this entry.

5.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

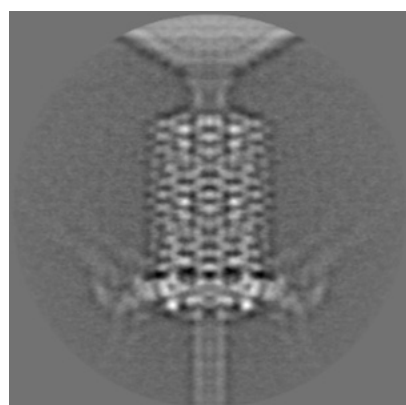
6 Map visualisation [i](#)

This section contains visualisations of the EMDB entry EMD-1086. These allow visual inspection of the internal detail of the map and identification of artifacts.

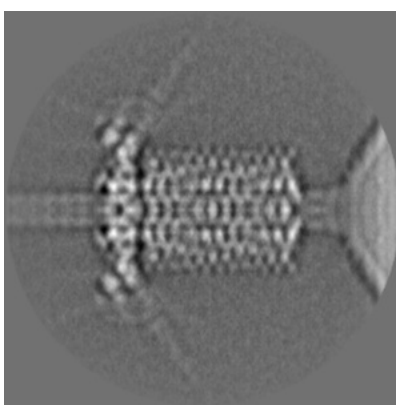
No raw map or half-maps were deposited for this entry and therefore no images, graphs, etc. pertaining to the raw map can be shown.

6.1 Orthogonal projections [i](#)

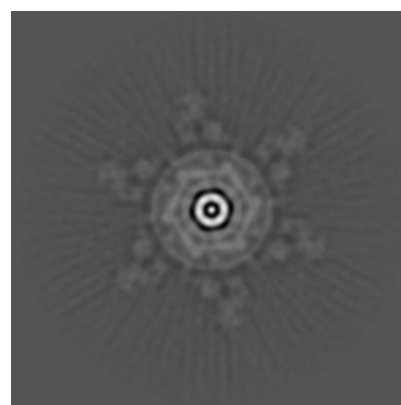
6.1.1 Primary map



X



Y

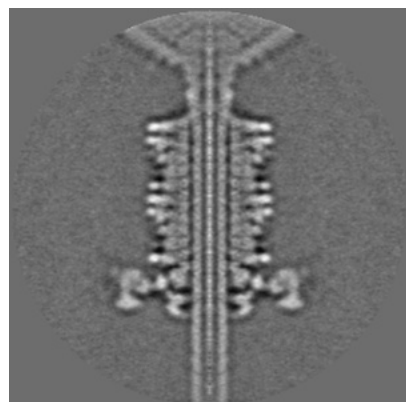


Z

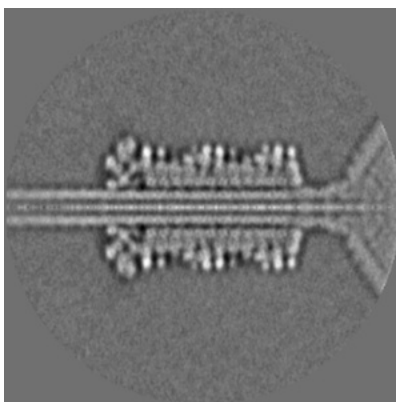
The images above show the map projected in three orthogonal directions.

6.2 Central slices [i](#)

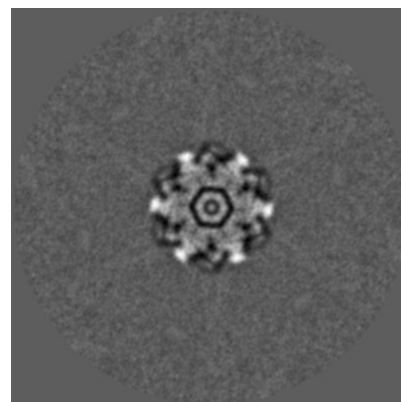
6.2.1 Primary map



X Index: 128



Y Index: 128

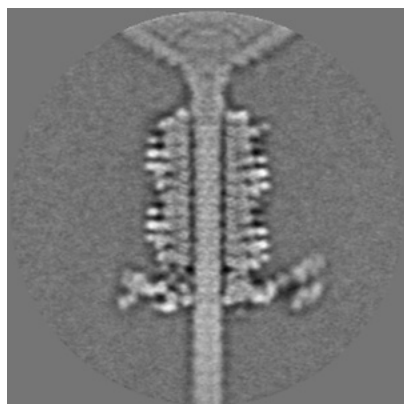


Z Index: 128

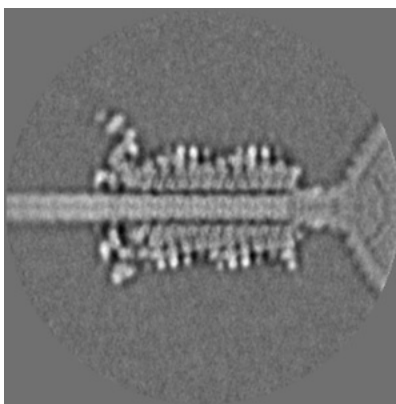
The images above show central slices of the map in three orthogonal directions.

6.3 Largest variance slices [i](#)

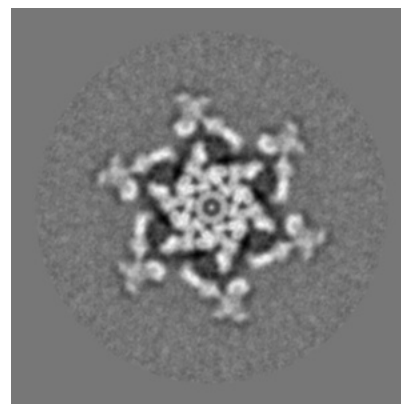
6.3.1 Primary map



X Index: 121



Y Index: 121



Z Index: 70

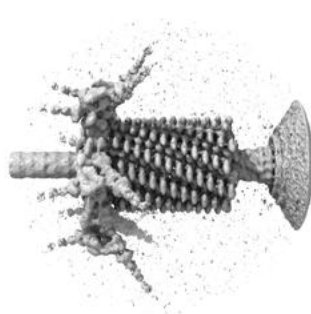
The images above show the largest variance slices of the map in three orthogonal directions.

6.4 Orthogonal surface views [i](#)

6.4.1 Primary map



X



Y



Z

The images above show the 3D surface view of the map at the recommended contour level 1.5. These images, in conjunction with the slice images, may facilitate assessment of whether an appropriate contour level has been provided.

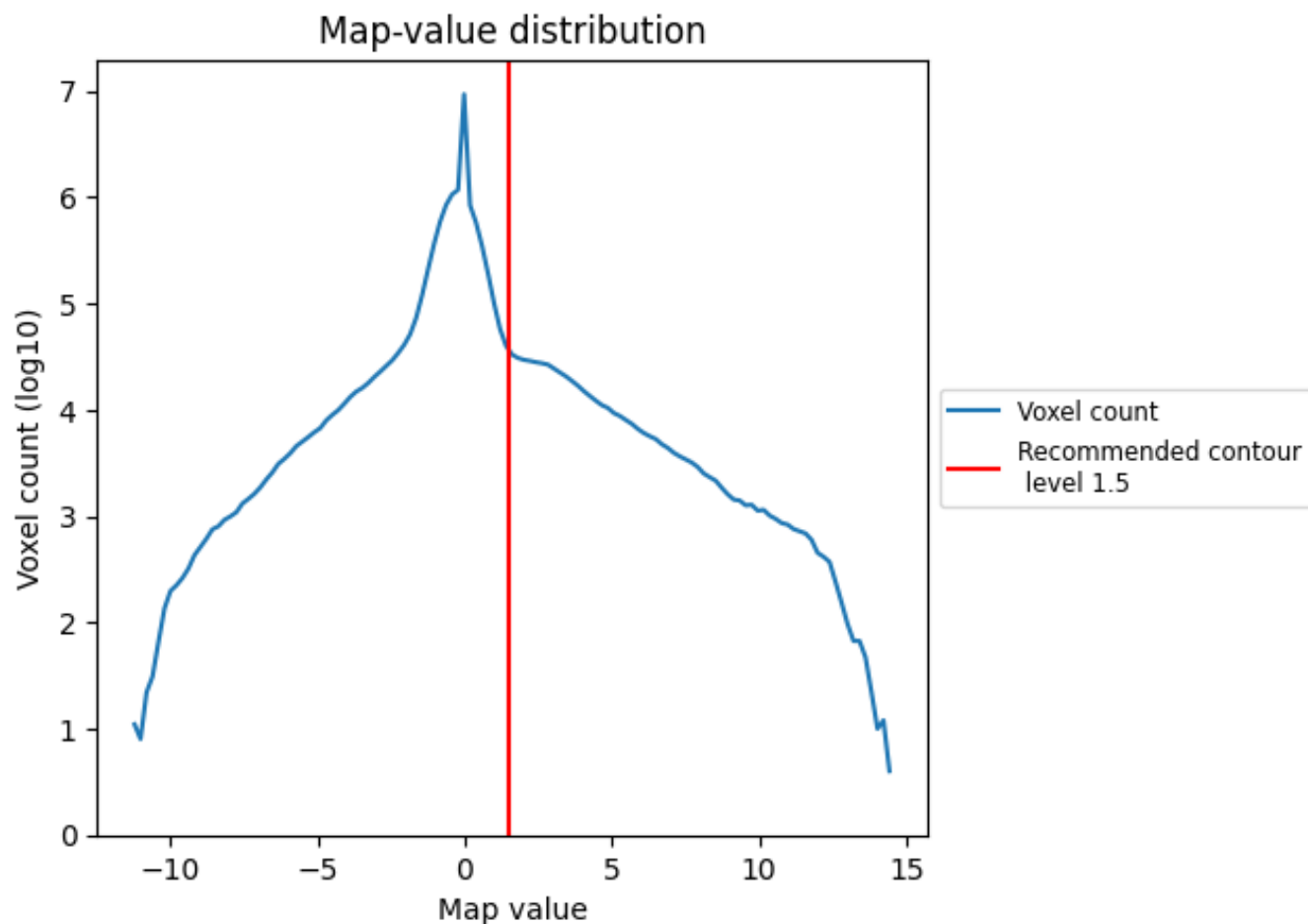
6.5 Mask visualisation

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

7 Map analysis [i](#)

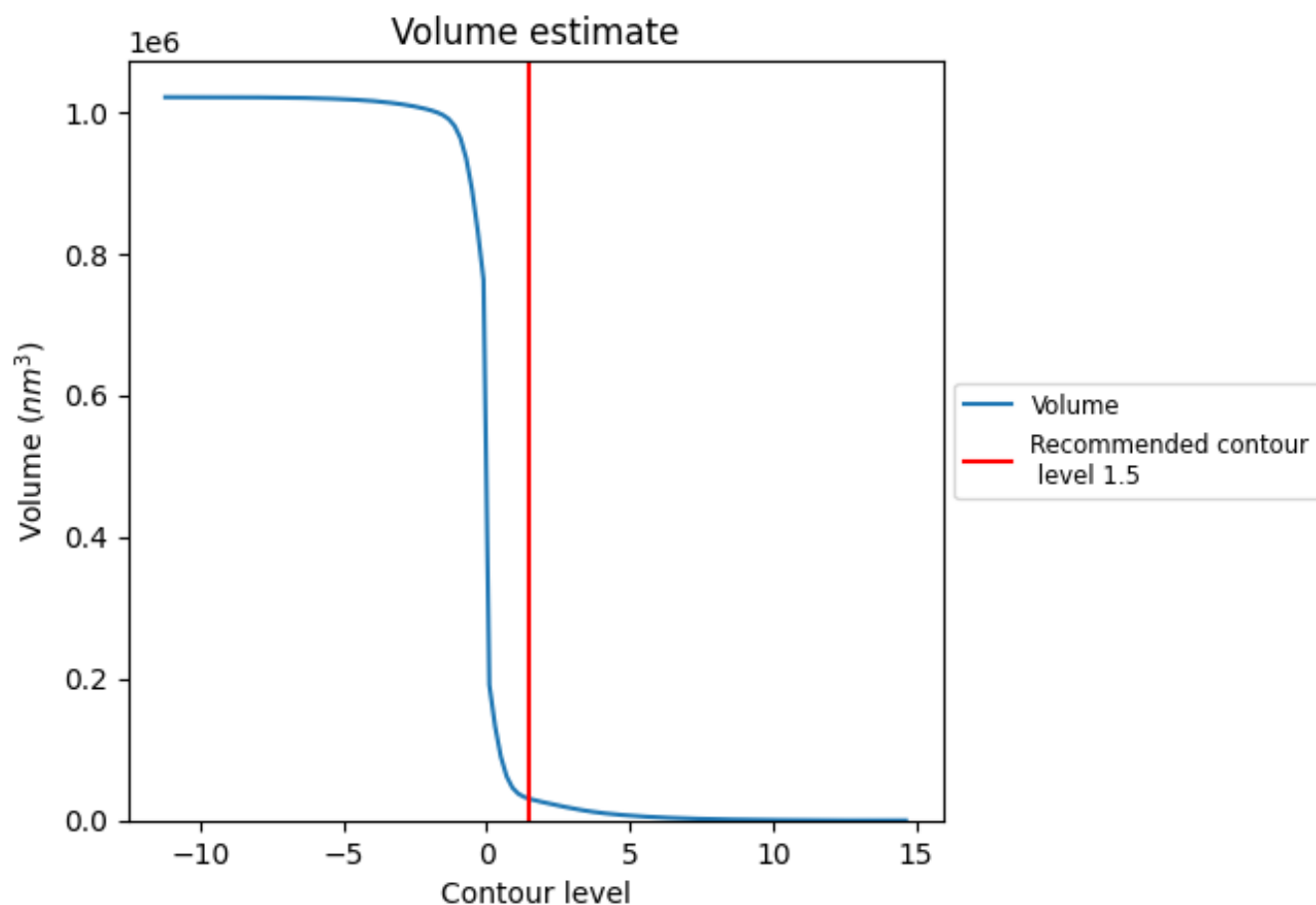
This section contains the results of statistical analysis of the map.

7.1 Map-value distribution [i](#)



The map-value distribution is plotted in 128 intervals along the x-axis. The y-axis is logarithmic. A spike in this graph at zero usually indicates that the volume has been masked.

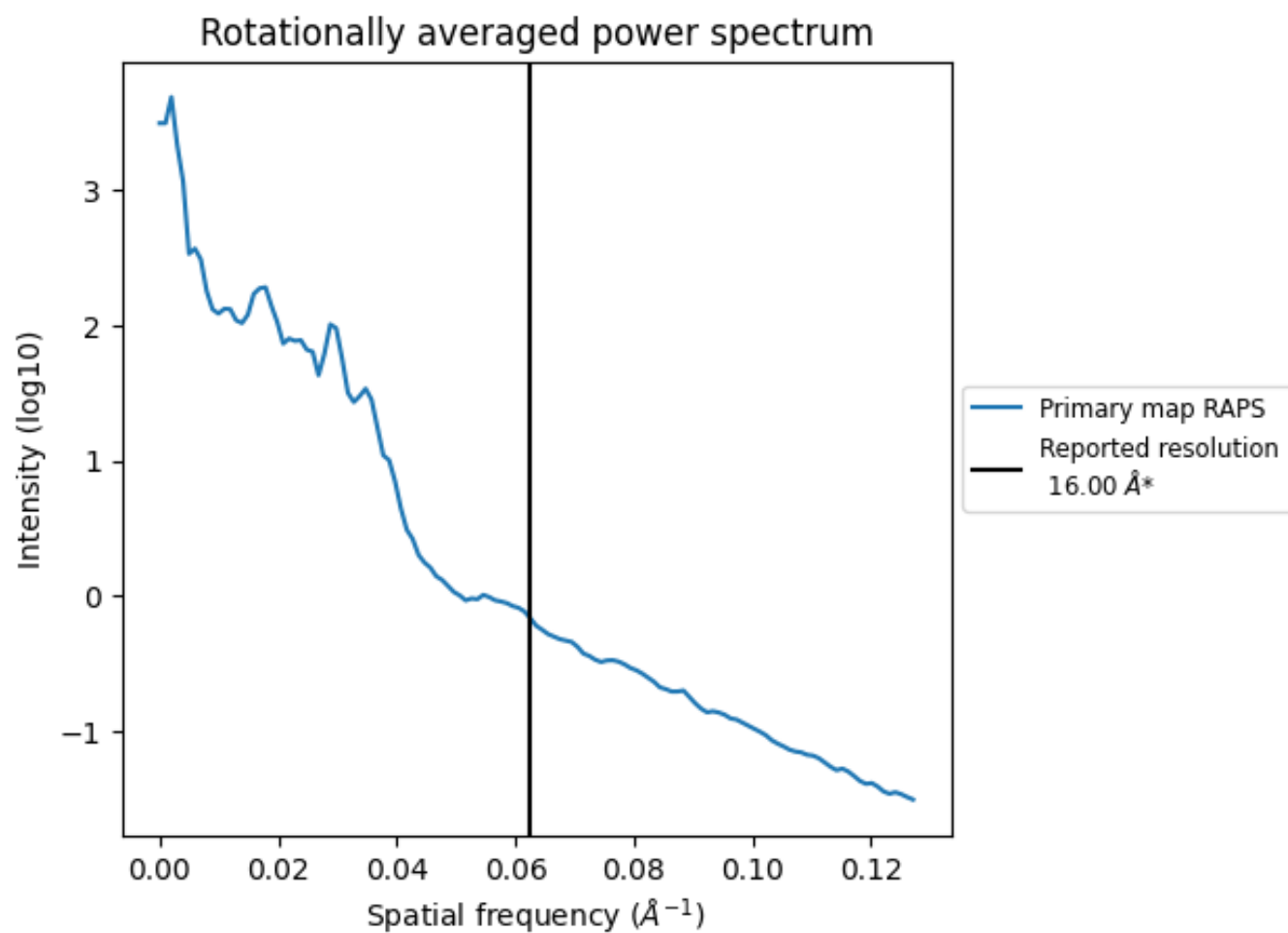
7.2 Volume estimate [i](#)



The volume at the recommended contour level is 30363 nm³; this corresponds to an approximate mass of 27428 kDa.

The volume estimate graph shows how the enclosed volume varies with the contour level. The recommended contour level is shown as a vertical line and the intersection between the line and the curve gives the volume of the enclosed surface at the given level.

7.3 Rotationally averaged power spectrum ⓘ



*Reported resolution corresponds to spatial frequency of 0.062 Å⁻¹

8 Fourier-Shell correlation

This section was not generated. No FSC curve or half-maps provided.

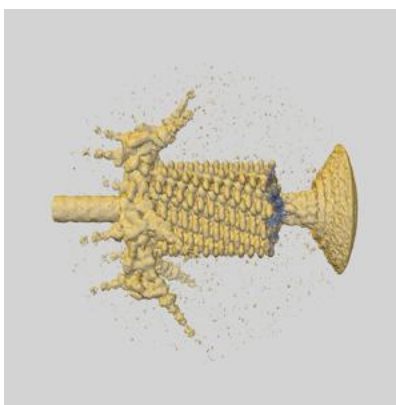
9 Map-model fit [i](#)

This section contains information regarding the fit between EMDB map EMD-1086 and PDB model 3J2N. Per-residue inclusion information can be found in section [3](#) on page [7](#).

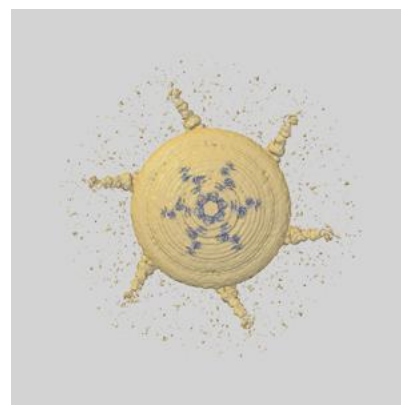
9.1 Map-model overlay [i](#)



X



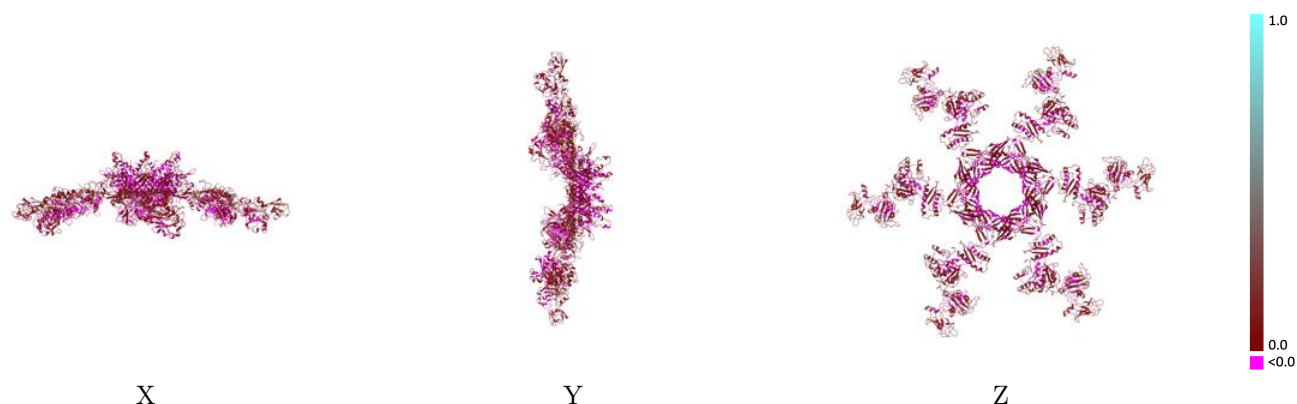
Y



Z

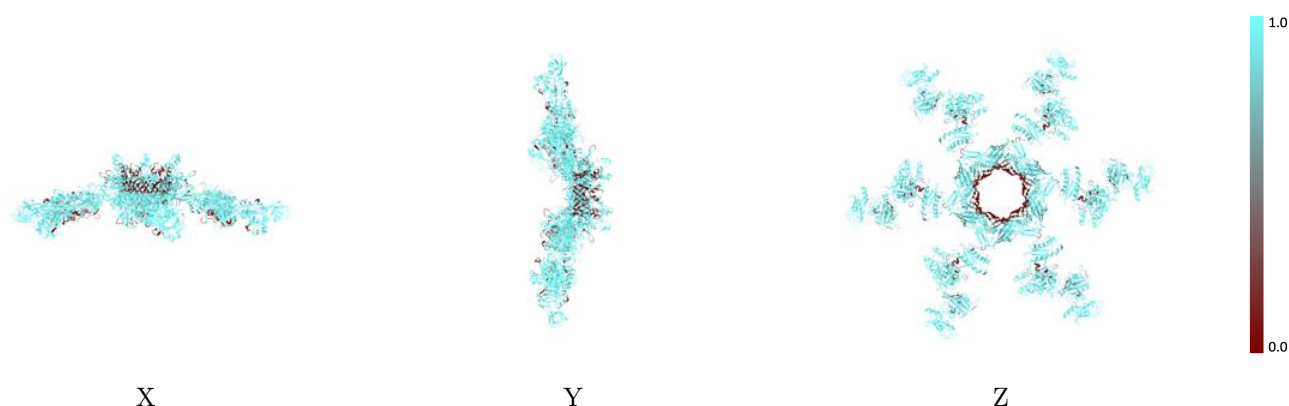
The images above show the 3D surface view of the map at the recommended contour level 1.5 at 50% transparency in yellow overlaid with a ribbon representation of the model coloured in blue. These images allow for the visual assessment of the quality of fit between the atomic model and the map.

9.2 Q-score mapped to coordinate model [i](#)



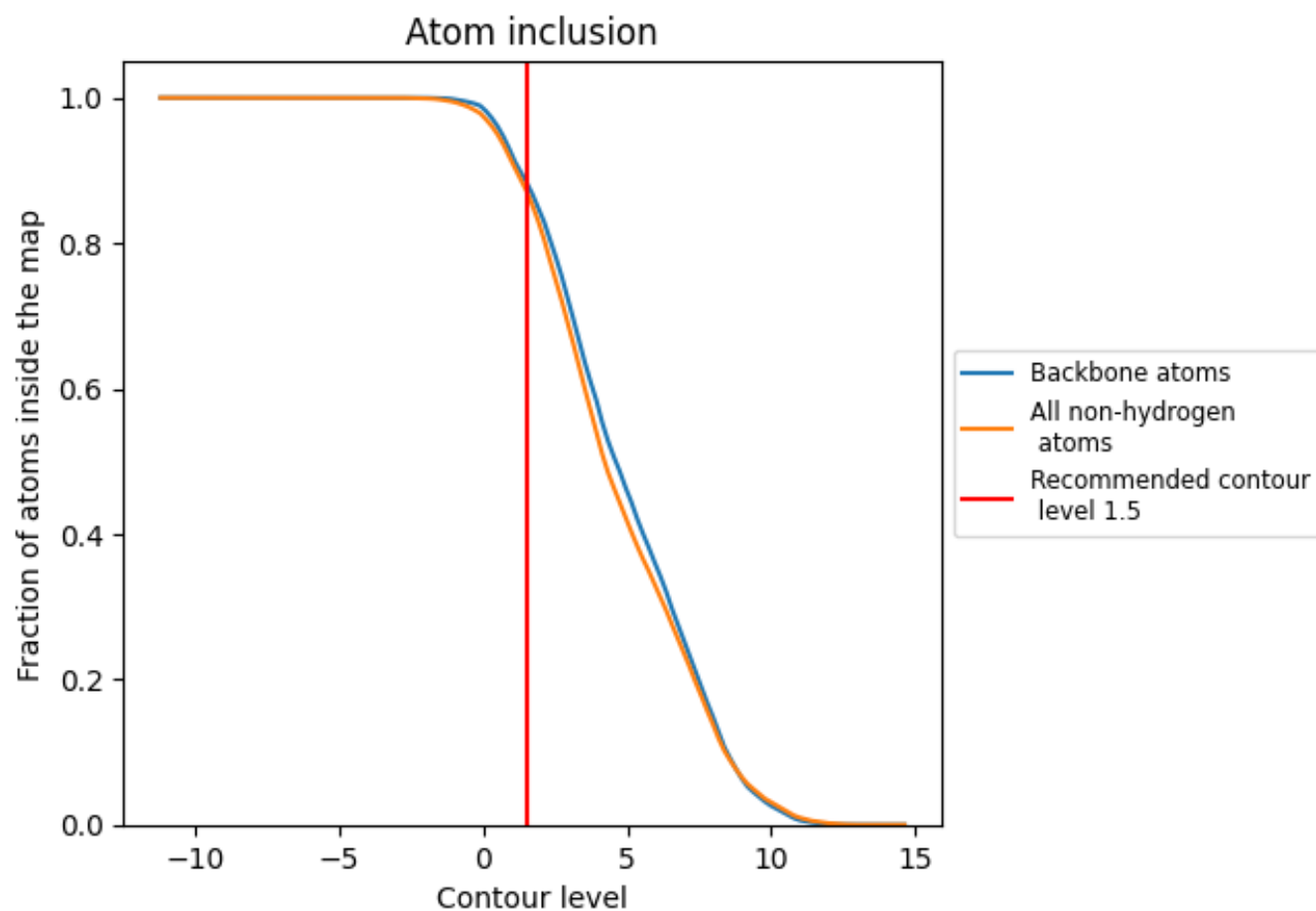
The images above show the model with each residue coloured according to its Q-score. This shows their resolvability in the map with higher Q-score values reflecting better resolvability. Please note: Q-score is calculating the resolvability of atoms, and thus high values are only expected at resolutions at which atoms can be resolved. Low Q-score values may therefore be expected for many entries.

9.3 Atom inclusion mapped to coordinate model [i](#)



The images above show the model with each residue coloured according to its atom inclusion. This shows to what extent they are inside the map at the recommended contour level (1.5).

9.4 Atom inclusion [i](#)



At the recommended contour level, 88% of all backbone atoms, 87% of all non-hydrogen atoms, are inside the map.

9.5 Map-model fit summary ⓘ

The table lists the average atom inclusion at the recommended contour level (1.5) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	<div></div> 0.8716	<div></div> 0.0570
A	<div></div> 0.7760	<div></div> 0.0330
B	<div></div> 0.7765	<div></div> 0.0360
C	<div></div> 0.7765	<div></div> 0.0400
D	<div></div> 0.7777	<div></div> 0.0350
E	<div></div> 0.7806	<div></div> 0.0310
F	<div></div> 0.7842	<div></div> 0.0310
U	<div></div> 0.9070	<div></div> 0.0660
V	<div></div> 0.9063	<div></div> 0.0660
W	<div></div> 0.9066	<div></div> 0.0660
X	<div></div> 0.9052	<div></div> 0.0650
Y	<div></div> 0.9070	<div></div> 0.0650
Z	<div></div> 0.9063	<div></div> 0.0650

