



## wwPDB EM Validation Summary Report ⓘ

Oct 8, 2022 – 09:39 am BST

PDB ID : 7ZU0  
EMDB ID : EMD-14964  
Title : HOPS tethering complex from yeast  
Authors : Shvarev, D.; Schoppe, J.; Koenig, C.; Perz, A.; Fuellbrunn, N.; Kiontke, S.;  
Langemeyer, L.; Janulienė, D.; Schnelle, K.; Kuemmel, D.; Froehlich, F.;  
Moeller, A.; Ungermann, C.  
Deposited on : 2022-05-11  
Resolution : 4.40 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](mailto: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) : 2.31.2

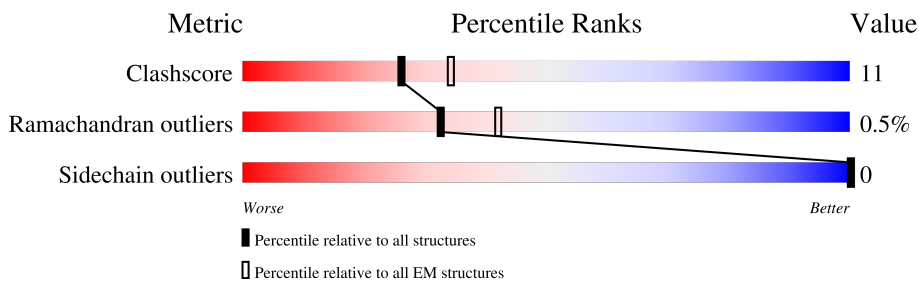
# 1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

*ELECTRON MICROSCOPY*

The reported resolution of this entry is 4.40 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  $\geq 3$ , 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	1029	29% (Upper Red) 29% (Red), 68% (Green), 25% (Yellow), 7% (Grey)
2	B	798	32% (Upper Red) 32% (Red), 66% (Green), 25% (Yellow), 8% (Grey)
3	C	918	7% (Upper Red) 7% (Red), 72% (Green), 19% (Yellow), 8% (Grey)
4	D	691	35% (Upper Red) 35% (Red), 64% (Green), 22% (Yellow), 14% (Grey)
5	E	1049	9% (Upper Red) 9% (Red), 29% (Green), 67% (Grey)
6	F	1016	8% (Upper Red) 8% (Red), 8% (Green), 89% (Grey)

## 2 Entry composition [i](#)

There are 6 unique types of molecules in this entry. The entry contains 25763 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 E3 ubiquitin-protein ligase PEP5.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
1	A	954	6914	4395	1185	1315	19	0	0

- Molecule 2 is a protein called Vacuolar protein sorting-associated protein 16.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
2	B	733	5981	3862	980	1110	29	0	0

- Molecule 3 is a protein called Vacuolar membrane protein PEP3.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
3	C	842	5549	3476	989	1063	21	0	0

- Molecule 4 is a protein called Vacuolar protein sorting-associated protein 33.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
4	D	597	4749	3036	789	904	20	0	0

- Molecule 5 is a protein called Vacuolar morphogenesis protein 6.

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
5	E	343	1704	1018	343	343	0	0

- Molecule 6 is a protein called Vacuolar protein sorting-associated protein 41.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
6	F	112	866	549	144	170	3	0	0

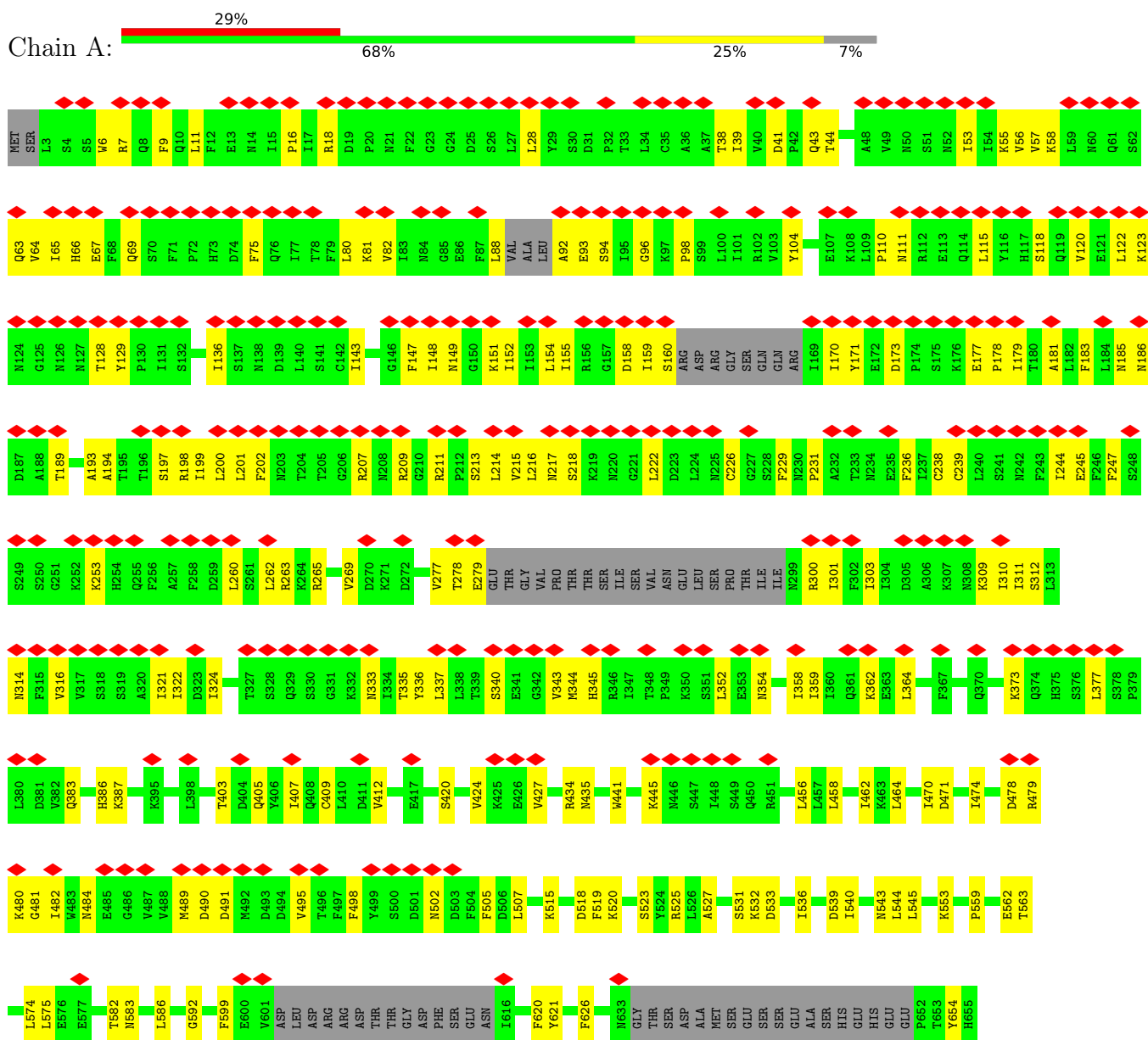
There are 24 discrepancies between the modelled and reference sequences:

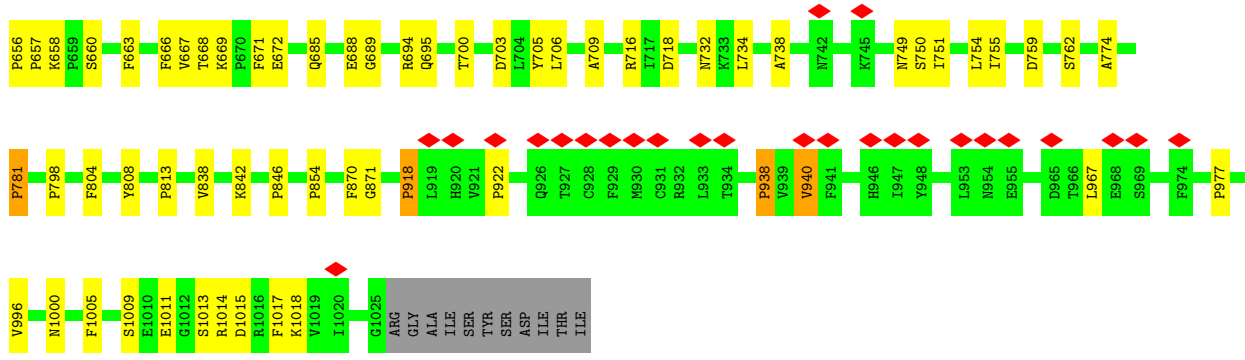
Chain	Residue	Modelled	Actual	Comment	Reference
F	993	ASP	-	expression tag	UNP P38959
F	994	TYR	-	expression tag	UNP P38959
F	995	LYS	-	expression tag	UNP P38959
F	996	ASP	-	expression tag	UNP P38959
F	997	ASP	-	expression tag	UNP P38959
F	998	ASP	-	expression tag	UNP P38959
F	999	ASP	-	expression tag	UNP P38959
F	1000	LYS	-	expression tag	UNP P38959
F	1001	ASP	-	expression tag	UNP P38959
F	1002	TYR	-	expression tag	UNP P38959
F	1003	LYS	-	expression tag	UNP P38959
F	1004	ASP	-	expression tag	UNP P38959
F	1005	ASP	-	expression tag	UNP P38959
F	1006	ASP	-	expression tag	UNP P38959
F	1007	ASP	-	expression tag	UNP P38959
F	1008	LYS	-	expression tag	UNP P38959
F	1009	ASP	-	expression tag	UNP P38959
F	1010	TYR	-	expression tag	UNP P38959
F	1011	LYS	-	expression tag	UNP P38959
F	1012	ASP	-	expression tag	UNP P38959
F	1013	ASP	-	expression tag	UNP P38959
F	1014	ASP	-	expression tag	UNP P38959
F	1015	ASP	-	expression tag	UNP P38959
F	1016	LYS	-	expression tag	UNP P38959

### 3 Residue-property plots

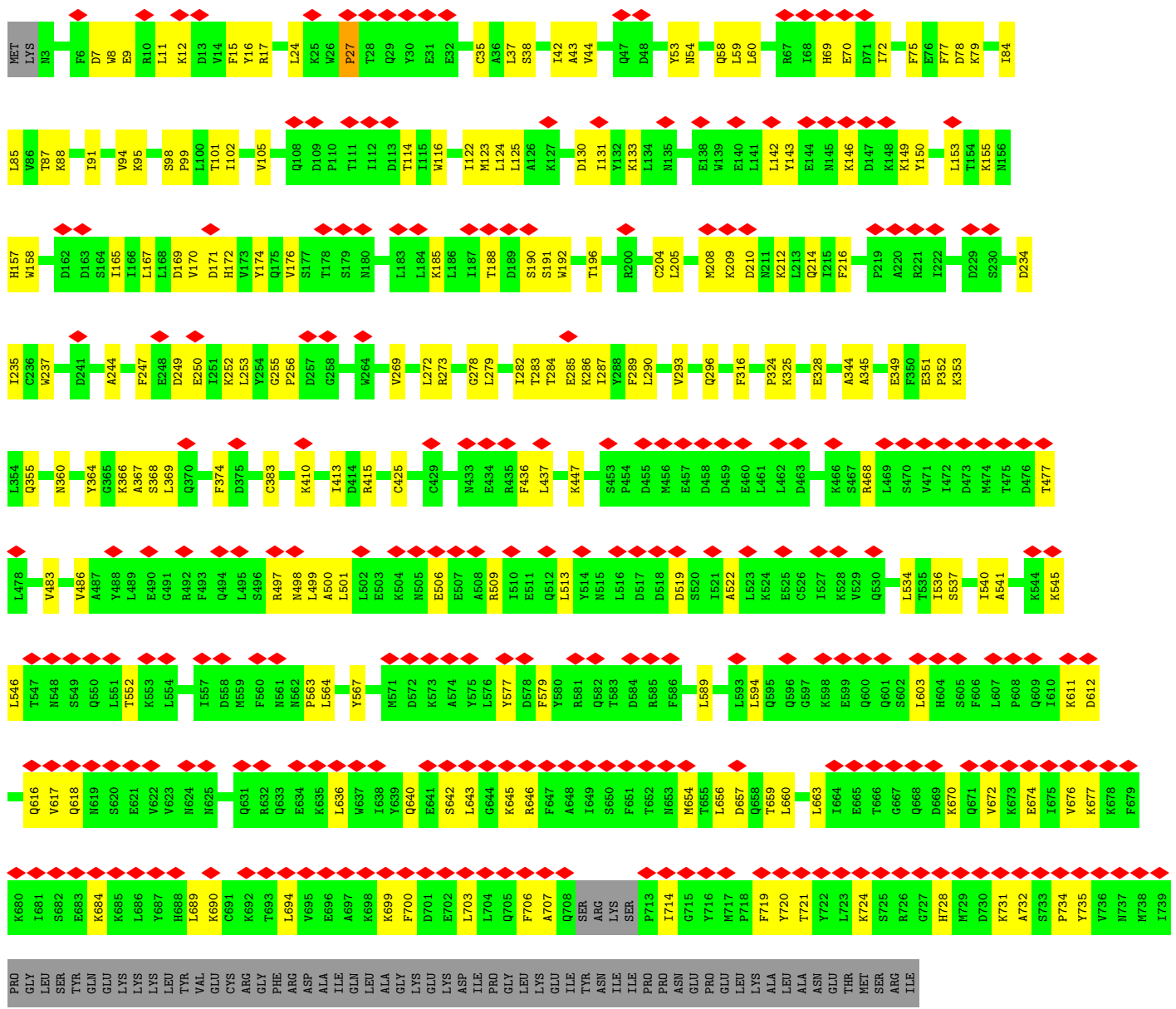
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and atom inclusion 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: E3 ubiquitin-protein ligase PEP5

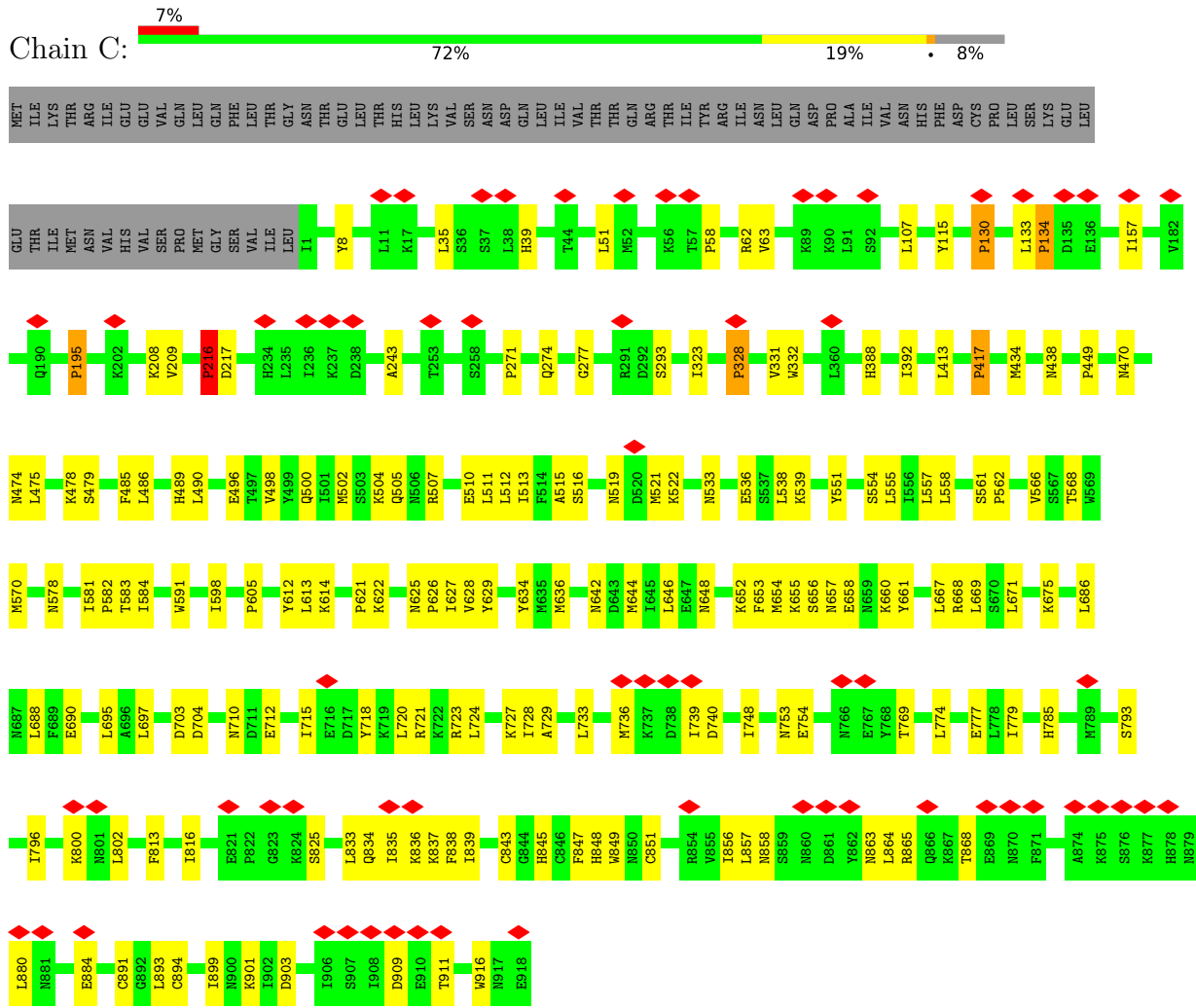




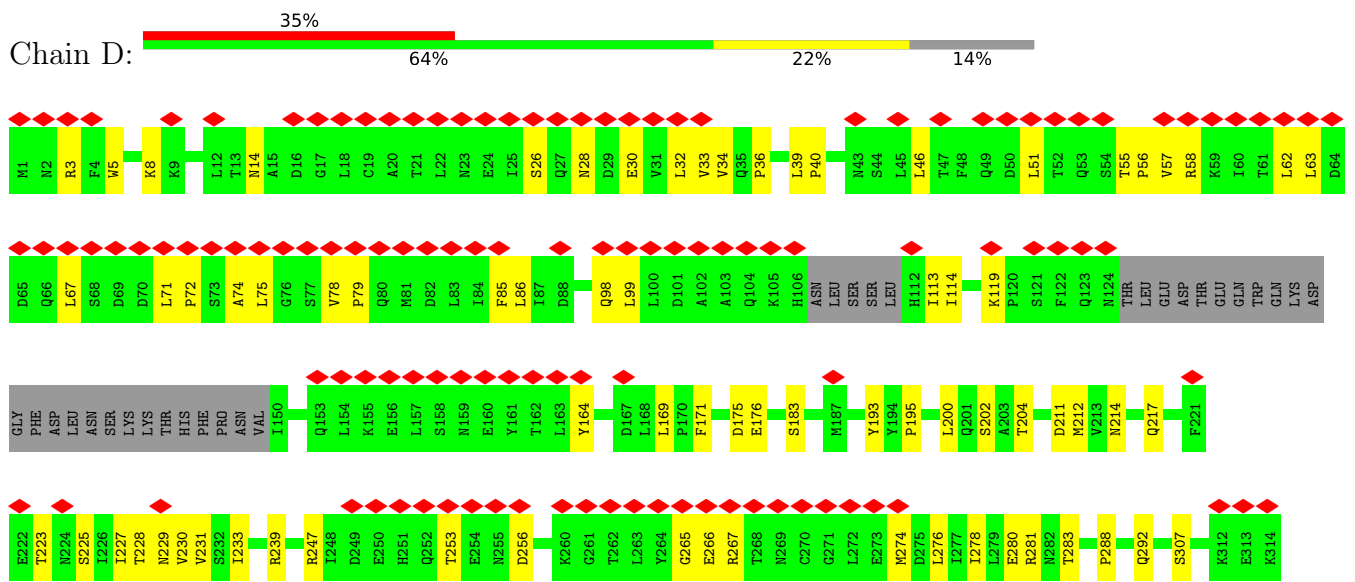
• Molecule 2: Vacuolar protein sorting-associated protein 16

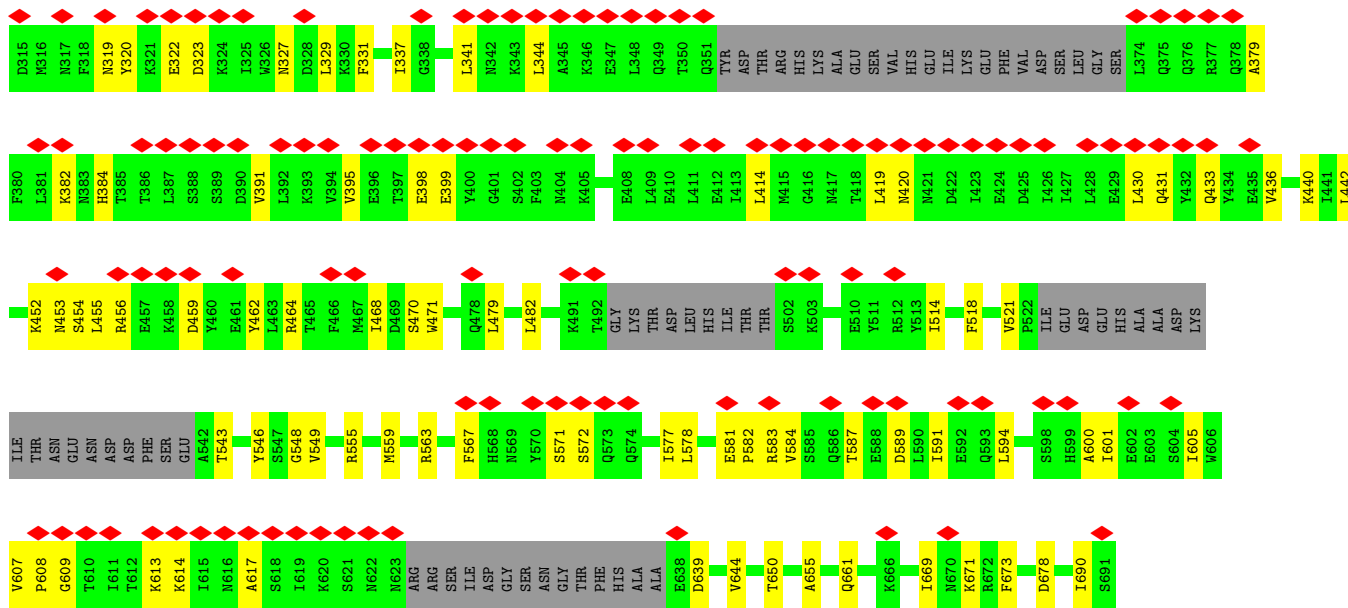


• Molecule 3: Vacuolar membrane protein PEP3

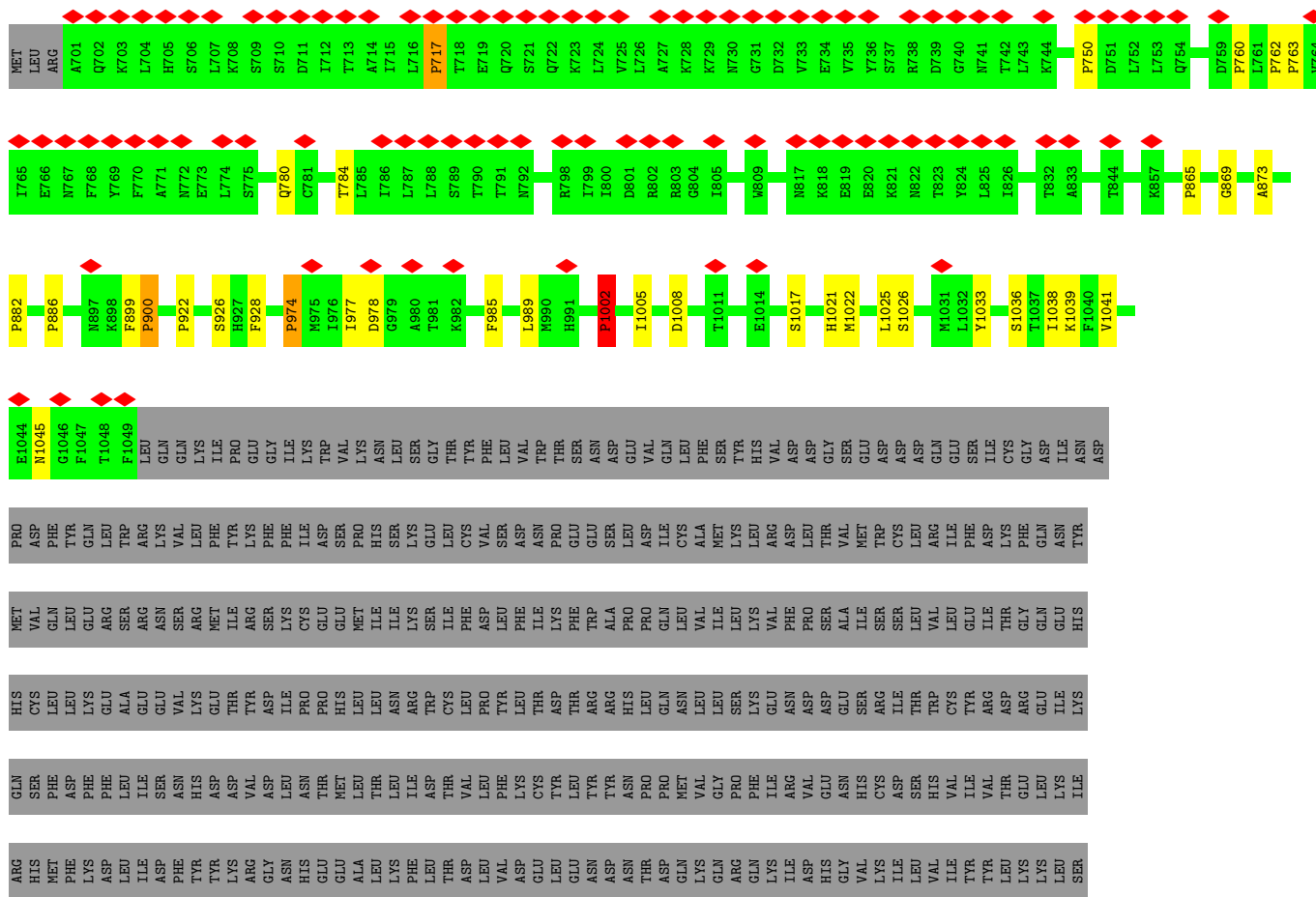


• Molecule 4: Vacuolar protein sorting-associated protein 33





• Molecule 5: Vacuolar morphogenesis protein 6







LEU	PHE	CYS	TYR	LEU	ILE	LYS	PHE	LEU	ASP	THR	ASP	ASP	SER	PHE	MET	ILE	SER	PRO	TYR	GLU	ASN	GLN	LEU	ILE	GLU	LEU	TYR	SER	GLU	TYR	P866	Q867	L868	F872	N873	R874	N875	W876	G877	D878	S882	E883	L884	W885	E886	F887	W888	I889	N890	Y891	S892	L893	D894	K895	P896	N897	F898	T899	K900	A901
LYS	ASP	ASP	ASP	ASP	LYS	ASP	ASP	ASP	ASP	ASP	ASP	ASP	LYS	TYR	TYR	ASP	D922	D923	L924	Q925	I926	D927	R928	L929	K934	Y937	Q938	E939	N940	S941	L942	R947	D948	N949	I950	R969	F975	Q976	W977	ASP	GLU	SER	ASP	ILE	GLU	ILE	ASN	ASP	ASP	LEU	ASN	GLY	VAL	LEU	LEU	ASP	TYR			

## 4 Experimental information

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	244661	Depositor
Resolution determination method	OTHER	Depositor
CTF correction method	NONE	Depositor
Microscope	TFS GLACIOS	Depositor
Voltage (kV)	200	Depositor
Electron dose ( $e^-/\text{\AA}^2$ )	50	Depositor
Minimum defocus (nm)	800	Depositor
Maximum defocus (nm)	2800	Depositor
Magnification	Not provided	
Image detector	FEI FALCON IV (4k x 4k)	Depositor
Maximum map value	1.893	Depositor
Minimum map value	-0.818	Depositor
Average map value	-0.001	Depositor
Map value standard deviation	0.027	Depositor
Recommended contour level	0.35	Depositor
Map size ( $\text{\AA}$ )	620.92804, 620.92804, 620.92804	wwPDB
Map dimensions	672, 672, 672	wwPDB
Map angles ( $^\circ$ )	90.0, 90.0, 90.0	wwPDB
Pixel spacing ( $\text{\AA}$ )	0.9240001, 0.9240001, 0.9240001	Depositor

## 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	0.26	0/7025	0.50	9/9555 (0.1%)
2	B	0.28	0/6101	0.49	1/8240 (0.0%)
3	C	0.26	0/5601	0.49	9/7644 (0.1%)
4	D	0.26	0/4827	0.47	2/6533 (0.0%)
5	E	0.44	0/1702	0.65	12/2373 (0.5%)
6	F	0.28	0/875	0.52	1/1180 (0.1%)
All	All	0.28	0/26131	0.50	34/35525 (0.1%)

There are no bond length outliers.

The worst 5 of 34 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	E	1002	PRO	N-CA-CB	6.80	111.46	103.30
3	C	134	PRO	N-CA-CB	6.75	111.40	103.30
3	C	216	PRO	N-CA-CB	6.63	111.26	103.30
3	C	195	PRO	N-CA-CB	6.58	111.20	103.30
5	E	900	PRO	N-CA-CB	6.50	111.10	103.30

There are no chirality outliers.

There are no planarity outliers.

### 5.2 Too-close contacts [i](#)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry-related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	6914	0	6244	166	0
2	B	5981	0	6031	148	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	C	5549	0	4384	129	0
4	D	4749	0	4710	96	0
5	E	1704	0	731	20	0
6	F	866	0	830	23	0
All	All	25763	0	22930	546	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 11.

The worst 5 of 546 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)
1:A:479:ARG:HH22	1:A:490:ASP:HA	1.47	0.79
2:B:153:LEU:HB2	2:B:157:HIS:CE1	2.19	0.78
3:C:834:GLN:HA	3:C:838:PHE:HZ	1.48	0.77
2:B:447:LYS:HG2	2:B:486:VAL:HG21	1.67	0.76
1:A:750:SER:HB2	3:C:627:ILE:HD13	1.65	0.76

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	940/1029 (91%)	840 (89%)	97 (10%)	3 (0%)	41 76
2	B	729/798 (91%)	646 (89%)	82 (11%)	1 (0%)	51 85
3	C	820/918 (89%)	751 (92%)	61 (7%)	8 (1%)	15 54
4	D	583/691 (84%)	548 (94%)	35 (6%)	0	100 100
5	E	339/1049 (32%)	320 (94%)	14 (4%)	5 (2%)	10 46
6	F	108/1016 (11%)	96 (89%)	12 (11%)	0	100 100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
All	All	3519/5501 (64%)	3201 (91%)	301 (9%)	17 (0%)	32 68

5 of 17 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	938	PRO
2	B	27	PRO
3	C	130	PRO
3	C	134	PRO
3	C	195	PRO

### 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 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	652/952 (68%)	652 (100%)	0	100 100
2	B	666/730 (91%)	666 (100%)	0	100 100
3	C	396/873 (45%)	396 (100%)	0	100 100
4	D	526/640 (82%)	526 (100%)	0	100 100
6	F	92/938 (10%)	92 (100%)	0	100 100
All	All	2332/4133 (56%)	2332 (100%)	0	100 100

There are no protein residues with a non-rotameric sidechain to report.

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	333	ASN
2	B	494	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](#)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
3	C	10
5	E	1
6	F	1
1	A	1

The worst 5 of 13 chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	C	115:TYR	C	121:TRP	N	13.77
1	E	991:HIS	C	998:ASN	N	13.23
1	C	23:GLN	C	27:LEU	N	12.99
1	C	140:PHE	C	150:GLU	N	12.97
1	C	11:LEU	C	17:LYS	N	10.46

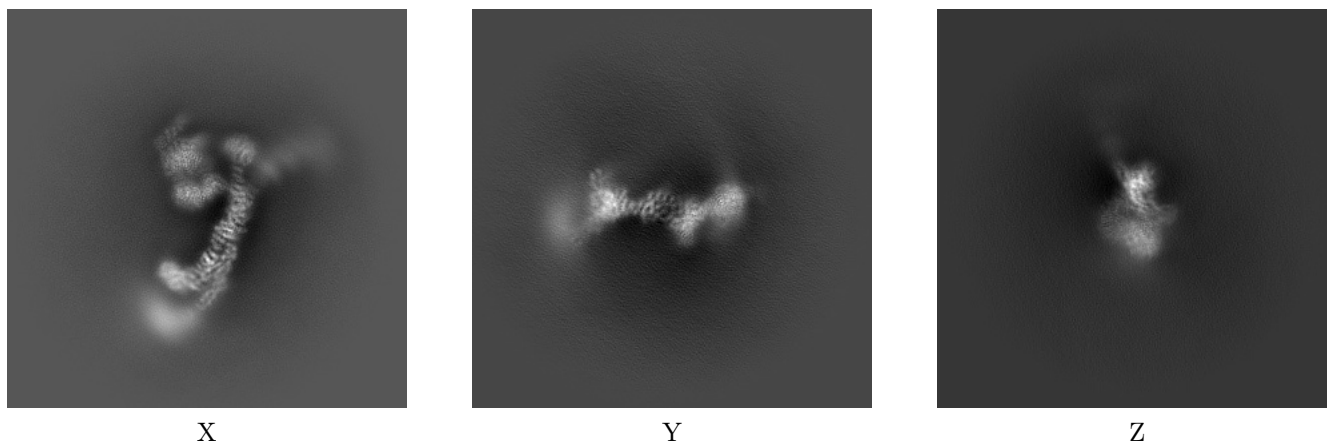
## 6 Map visualisation [i](#)

This section contains visualisations of the EMDB entry EMD-14964. 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](#)

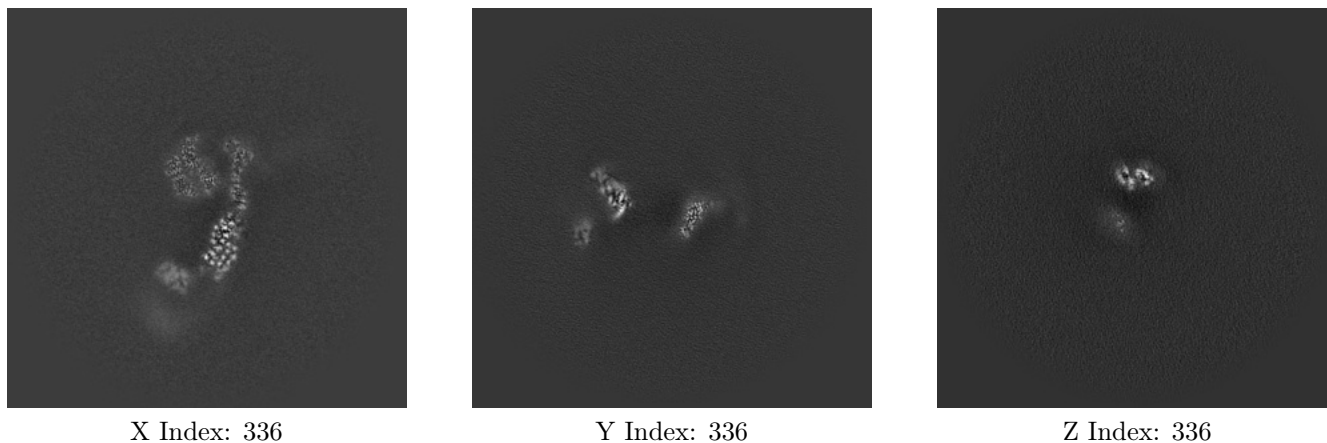
#### 6.1.1 Primary map



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

### 6.2 Central slices [i](#)

#### 6.2.1 Primary map





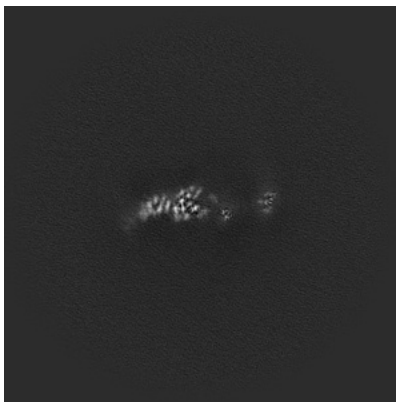
The images above show central slices of the map in three orthogonal directions.

## 6.3 Largest variance slices [i](#)

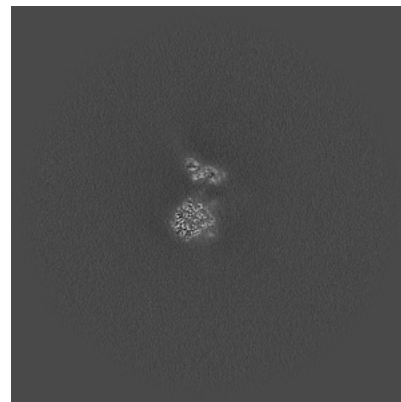
### 6.3.1 Primary map



X Index: 342



Y Index: 372



Z Index: 358

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 0.35. These images, in conjunction with the slice images, may facilitate assessment of whether an appropriate contour level has been provided.

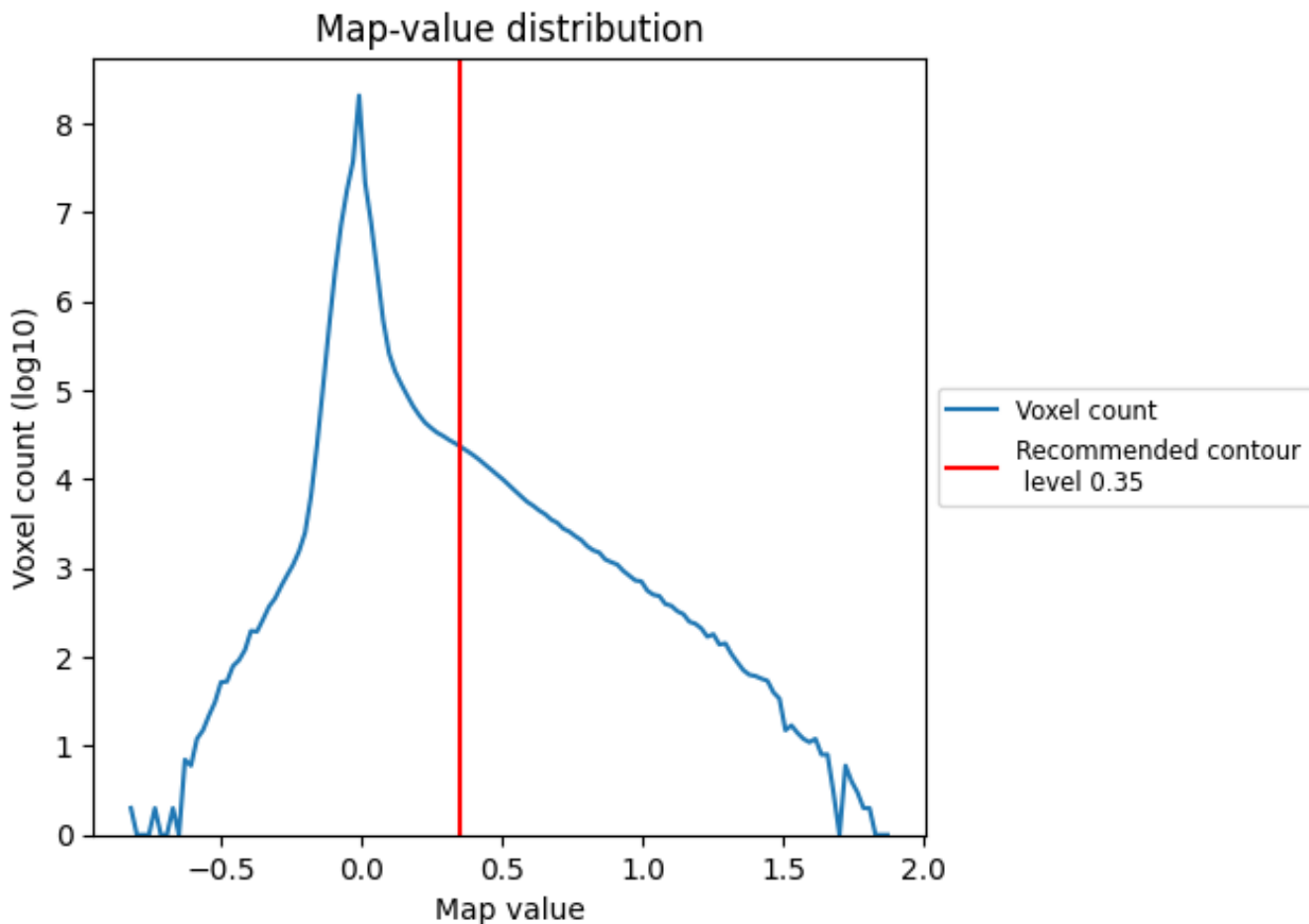
## 6.5 Mask visualisation

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

## 7 Map analysis [i](#)

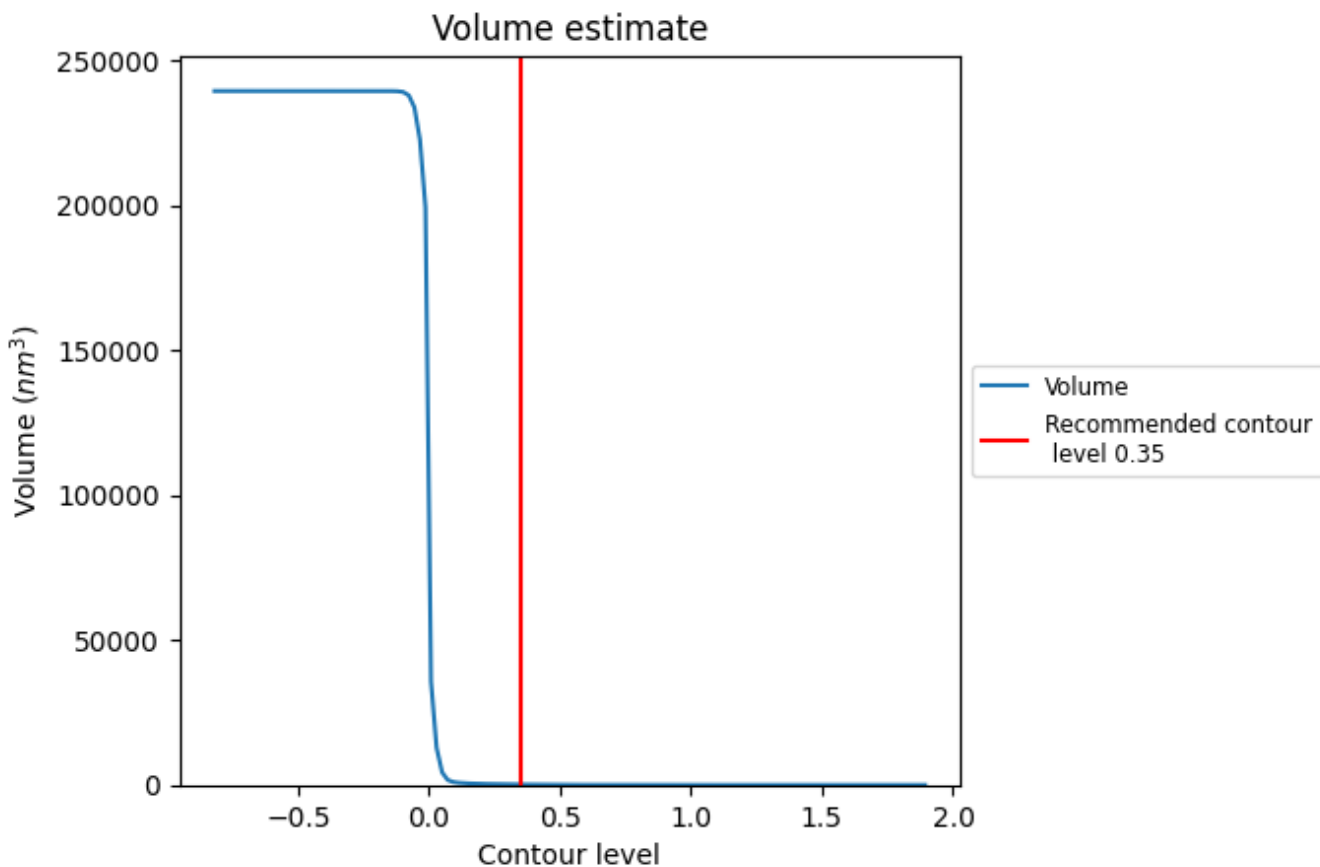
This section contains the results of statistical analysis of the map.

### 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.

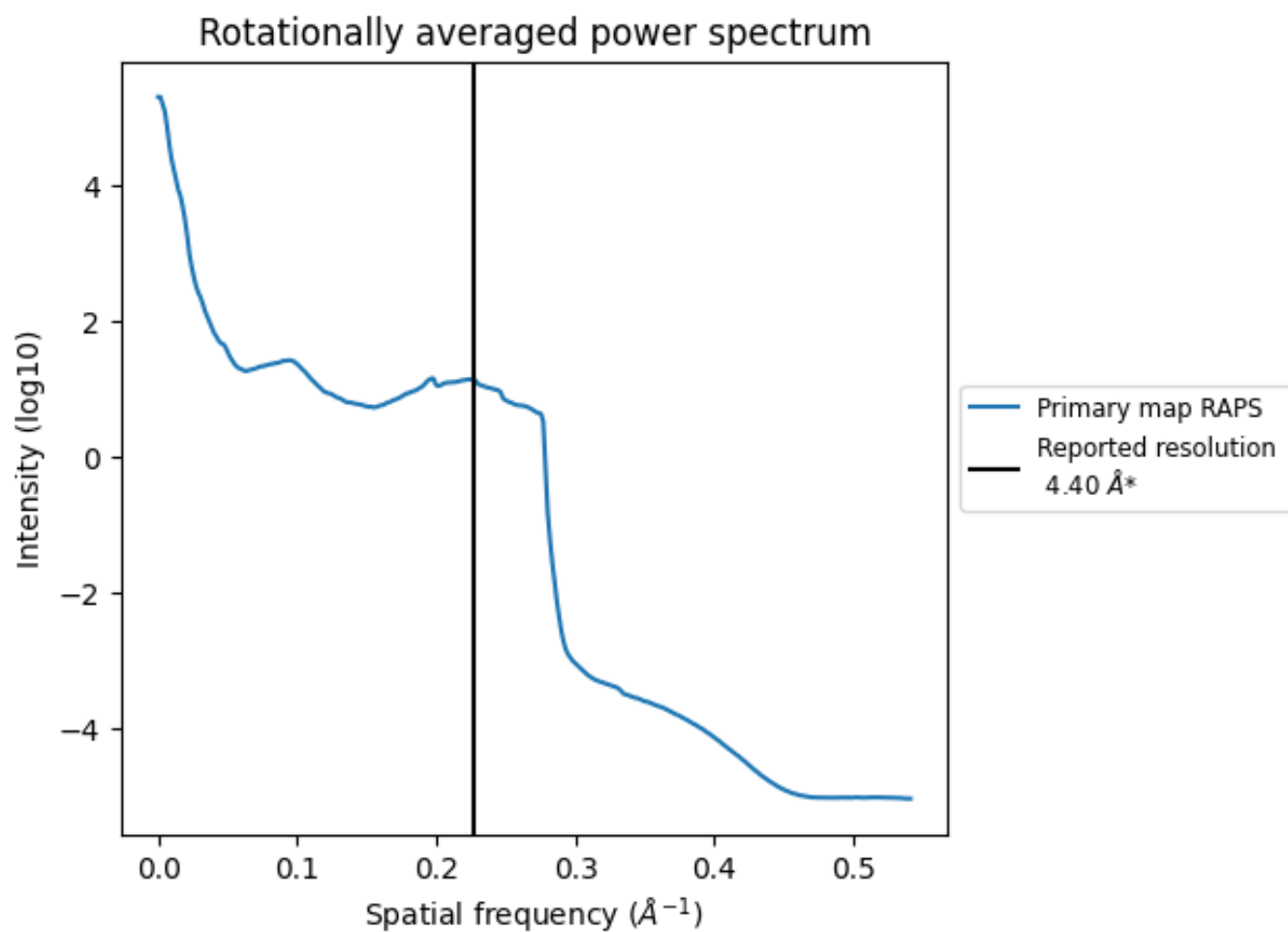
## 7.2 Volume estimate [\(i\)](#)



The volume at the recommended contour level is 164 nm<sup>3</sup>; this corresponds to an approximate mass of 149 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 [i](#)



\*Reported resolution corresponds to spatial frequency of  $0.227 \text{\AA}^{-1}$

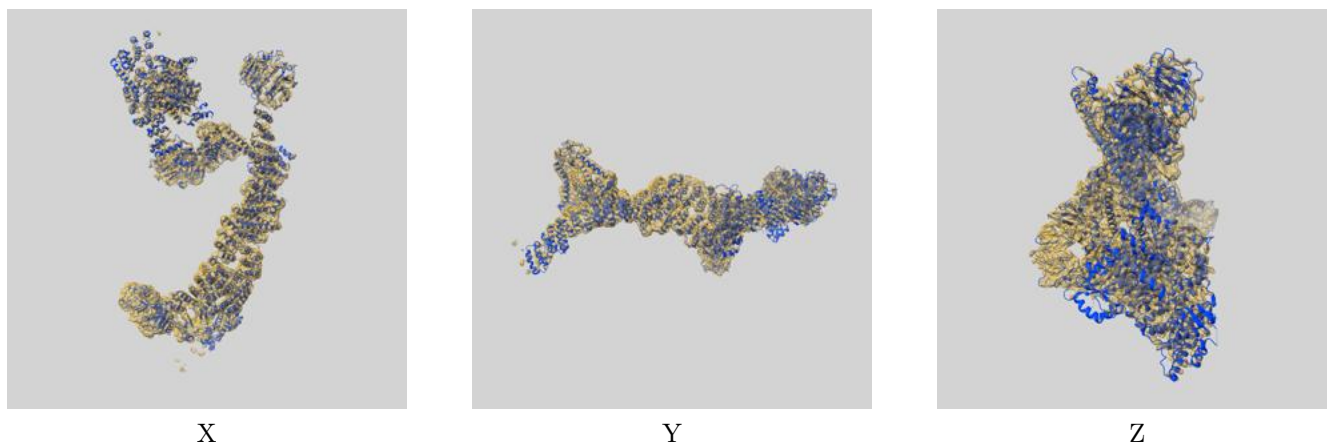
## 8 Fourier-Shell correlation

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

## 9 Map-model fit [i](#)

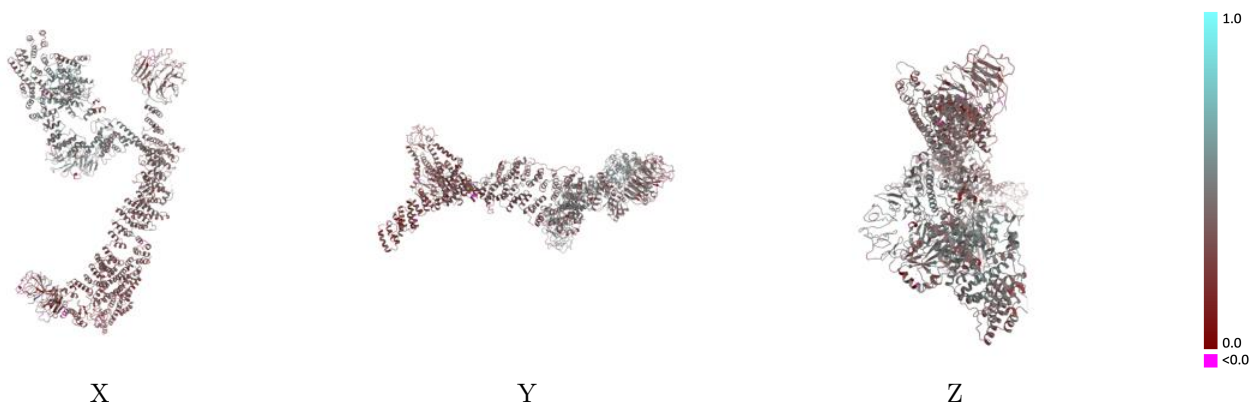
This section contains information regarding the fit between EMDB map EMD-14964 and PDB model 7ZU0. Per-residue inclusion information can be found in section 3 on page 5.

### 9.1 Map-model overlay [i](#)



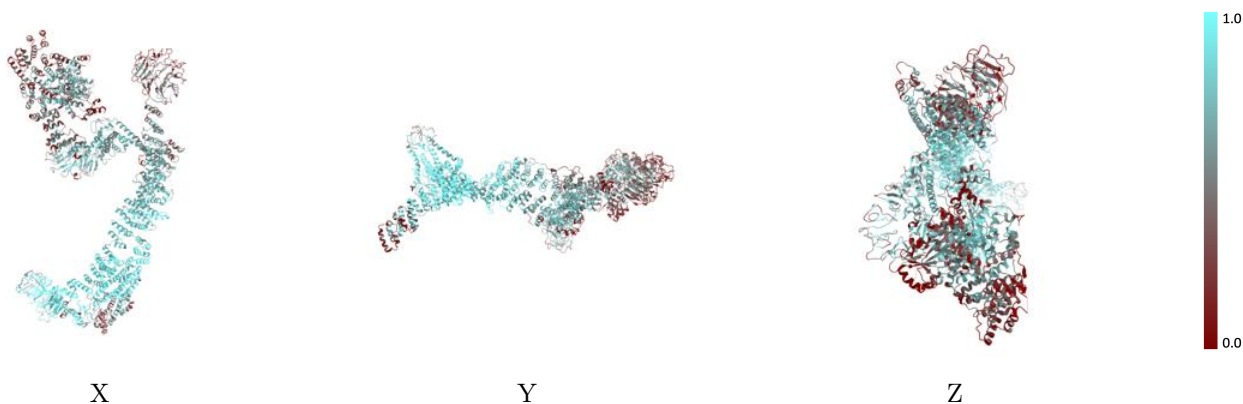
The images above show the 3D surface view of the map at the recommended contour level 0.35 at 50% transparency in yellow overlaid with a ribbon representation of the model coloured in blue. These images allow for the visual assessment of the quality of fit between the atomic model and the map.

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



The images above show the model with each residue coloured according 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.

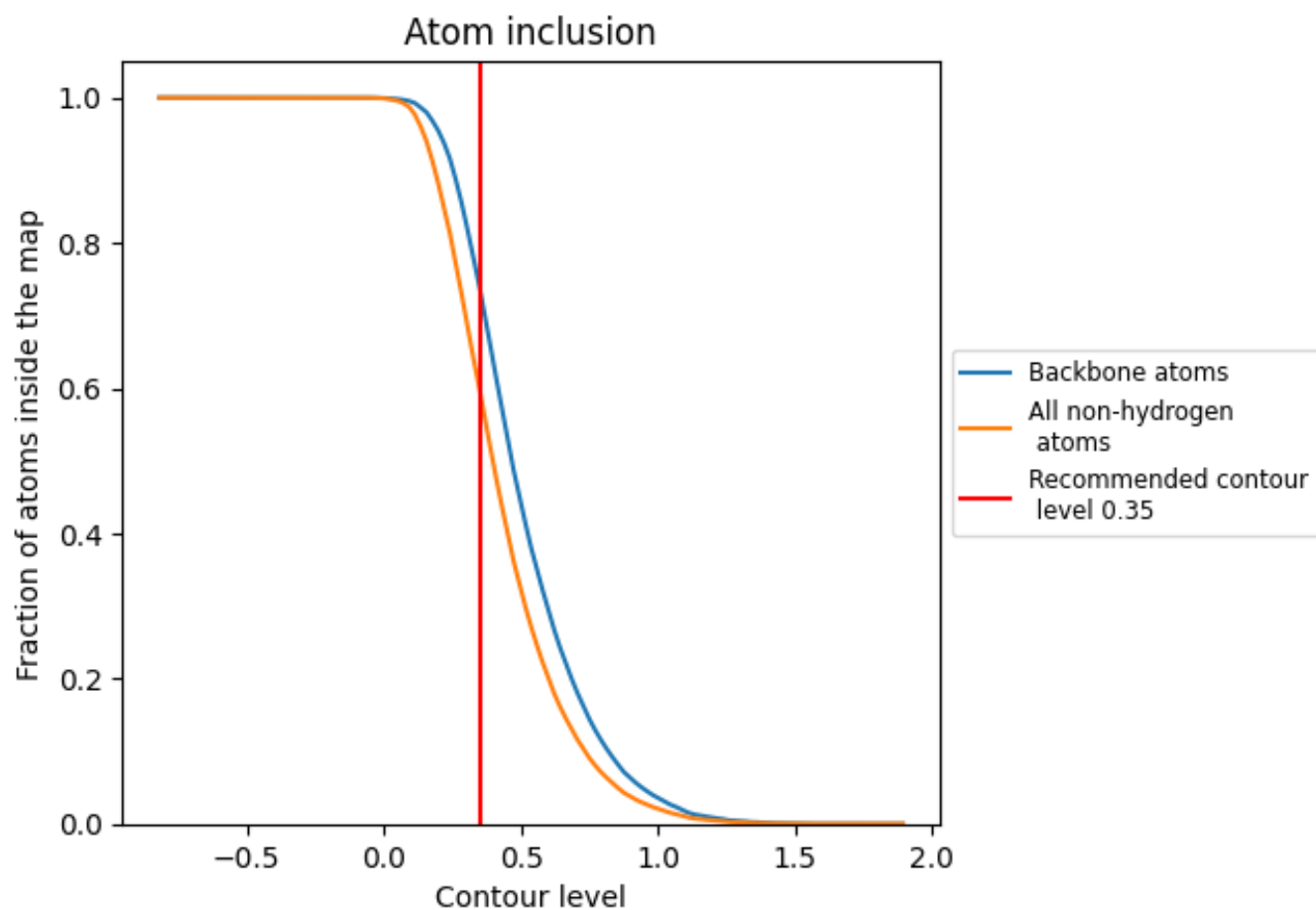
## 9.3 Atom inclusion mapped to coordinate model [i](#)



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 (0.35).

















## 9.4 Atom inclusion [i](#)



At the recommended contour level, 74% of all backbone atoms, 60% 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 (0.35) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	 0.5978	 0.3930
A	 0.5654	 0.3590
B	 0.5315	 0.4490
C	 0.8004	 0.3510
D	 0.4708	 0.4620
E	 0.6743	 0.2820
F	 0.5496	 0.4010

