wwPDB EM Validation Summary Report (i)

Nov 6, 2022 – 02:24 PM EST

PDB ID	:	6NYM
EMDB ID	:	EMD-0546
Title	:	Helicobacter pylori Vacuolating Cytotoxin A Oligomeric Assembly 2d (OA-2d)
Authors	:	Zhang, K.; Zhang, H.; Li, S.; Au, S.; Chiu, W.
Deposited on	:	2019-02-11
Resolution	:	3.60 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 (i) 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 (i)) were used in the production of this report:

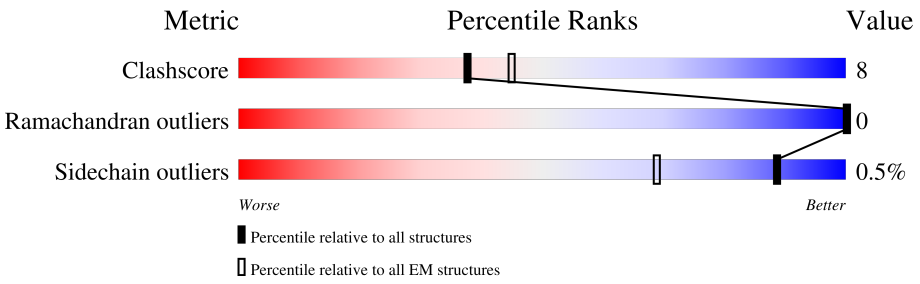
EMDB validation analysis	:	0.0.1.dev43
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
MapQ	:	1.9.9
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.31.2

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 3.60 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 <=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	821	73% 19% 9%
1	B	821	71% 20% 9%
1	C	821	72% 20% 9%
1	D	821	71% 20% 9%
1	E	821	72% 20% 9%
1	F	821	72% 19% 9%
1	G	821	73% 19% 9%
1	H	821	72% 19% 9%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	I	821	72% 19% 9%
1	J	821	72% 20% 9%
1	K	821	72% 19% 9%
1	L	821	72% 19% 9%

2 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 68664 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 Vacuolating cytotoxin autotransporter.

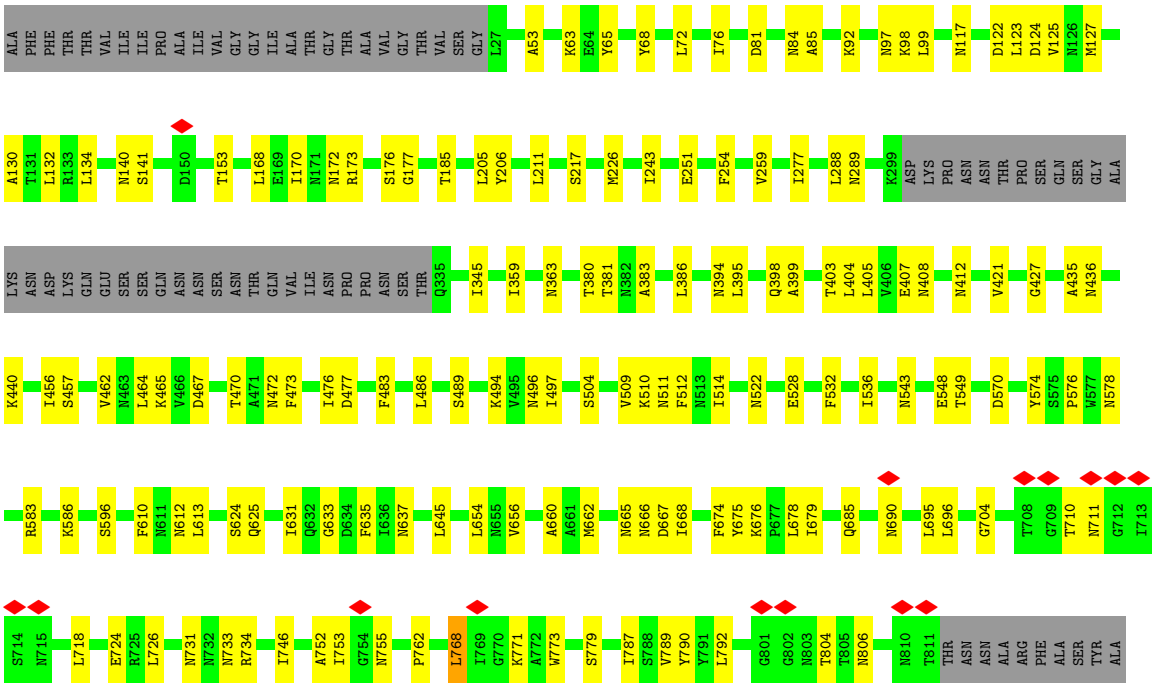
Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
1	A	750	5722	3579	1005	1126	12	0	0
1	B	750	5722	3579	1005	1126	12	0	0
1	C	750	5722	3579	1005	1126	12	0	0
1	D	750	5722	3579	1005	1126	12	0	0
1	E	750	5722	3579	1005	1126	12	0	0
1	F	750	5722	3579	1005	1126	12	0	0
1	G	750	5722	3579	1005	1126	12	0	0
1	H	750	5722	3579	1005	1126	12	0	0
1	I	750	5722	3579	1005	1126	12	0	0
1	J	750	5722	3579	1005	1126	12	0	0
1	K	750	5722	3579	1005	1126	12	0	0
1	L	750	5722	3579	1005	1126	12	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 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.

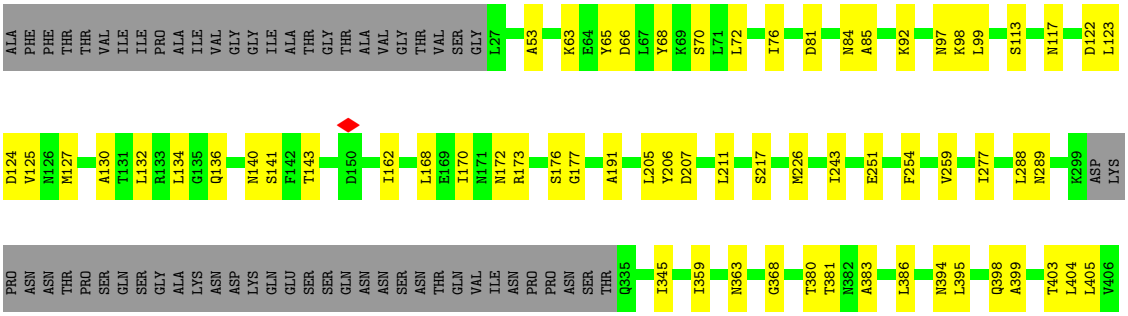
• Molecule 1: Vacuolating cytotoxin autotransporter

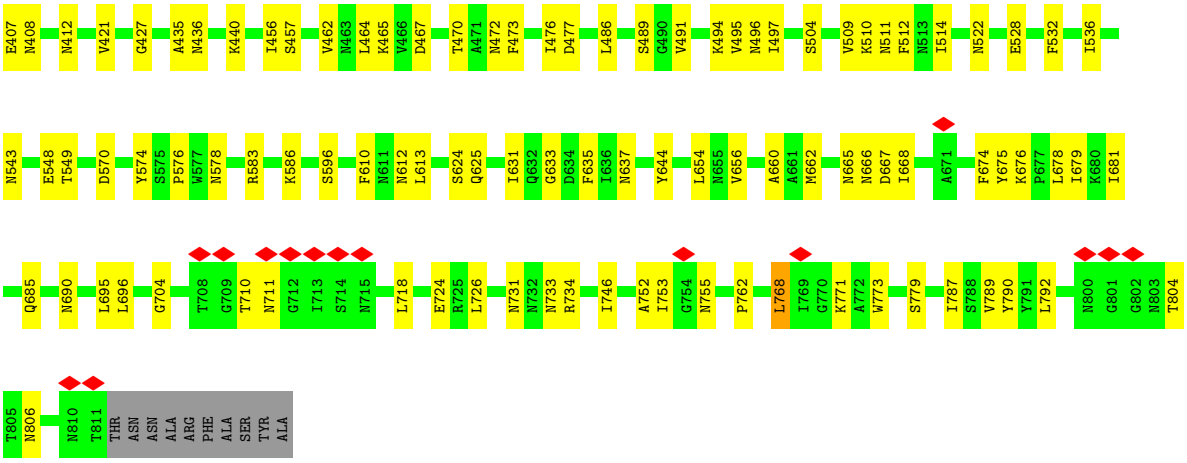
Chain A: 73% 19% 9%



• Molecule 1: Vacuolating cytotoxin autotransporter

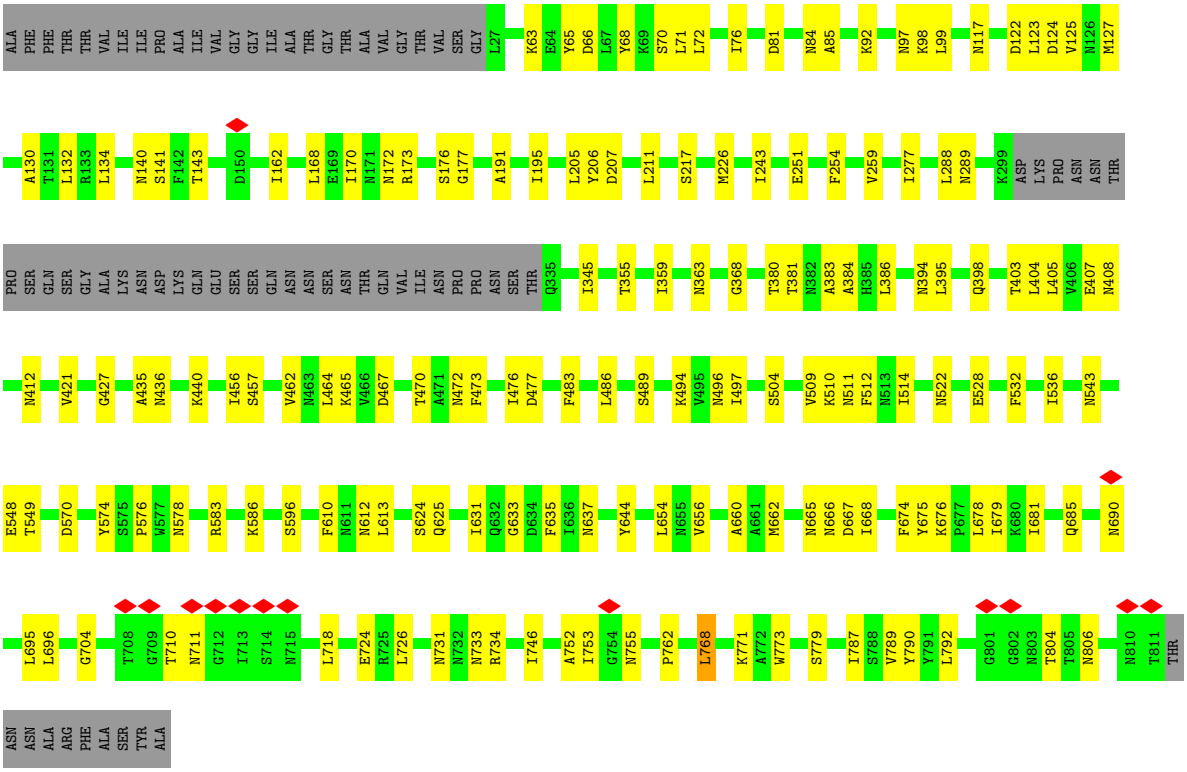
Chain B: 71% 20% 9%





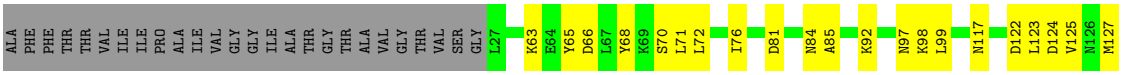
• Molecule 1: Vacuolating cytotoxin autotransporter

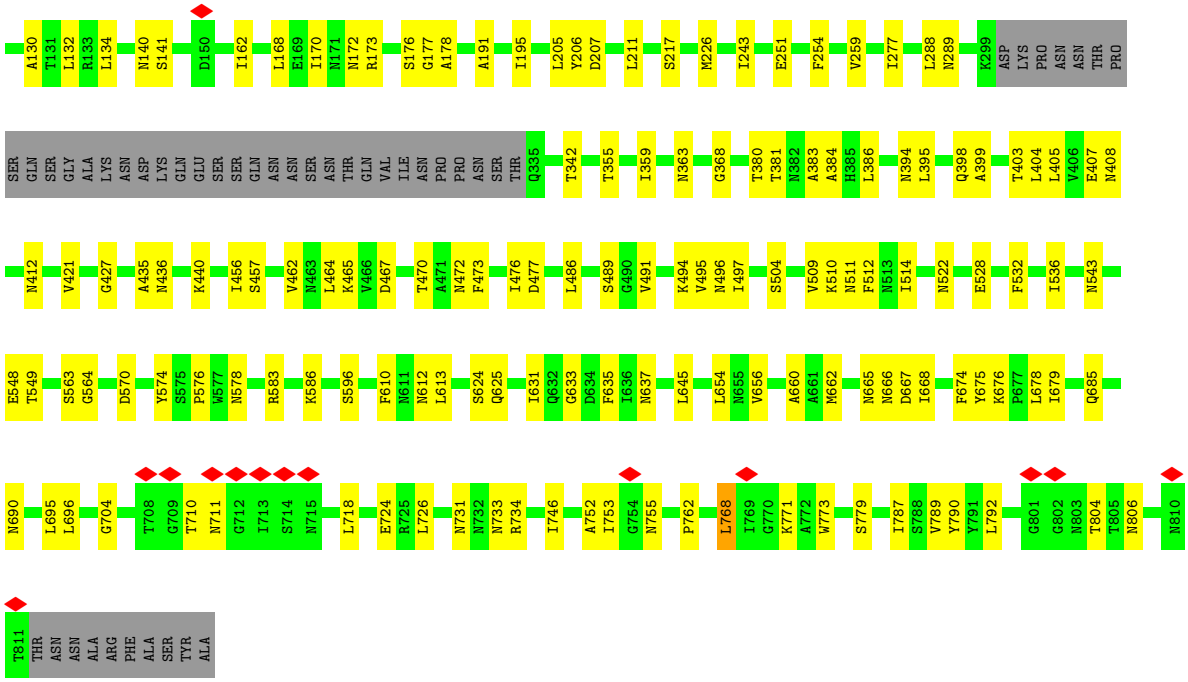
Chain C:



• Molecule 1: Vacuolating cytotoxin autotransporter

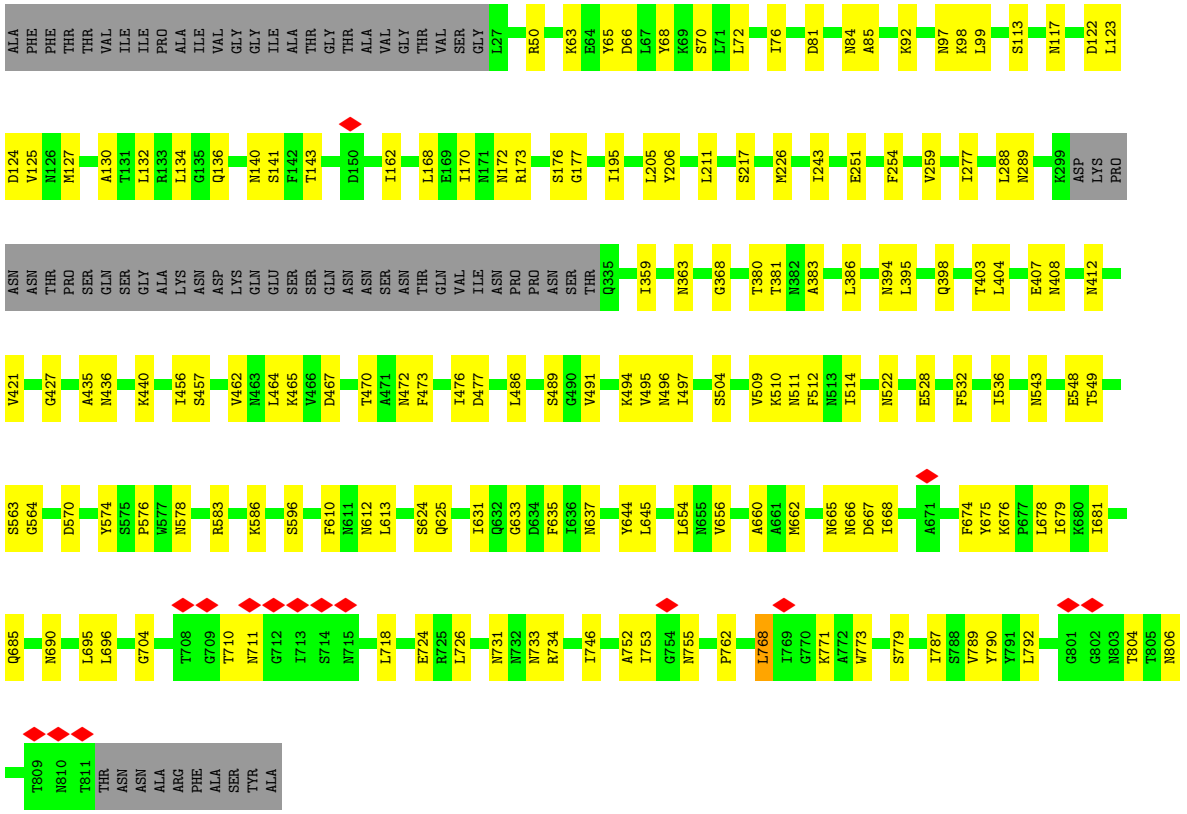
Chain D:





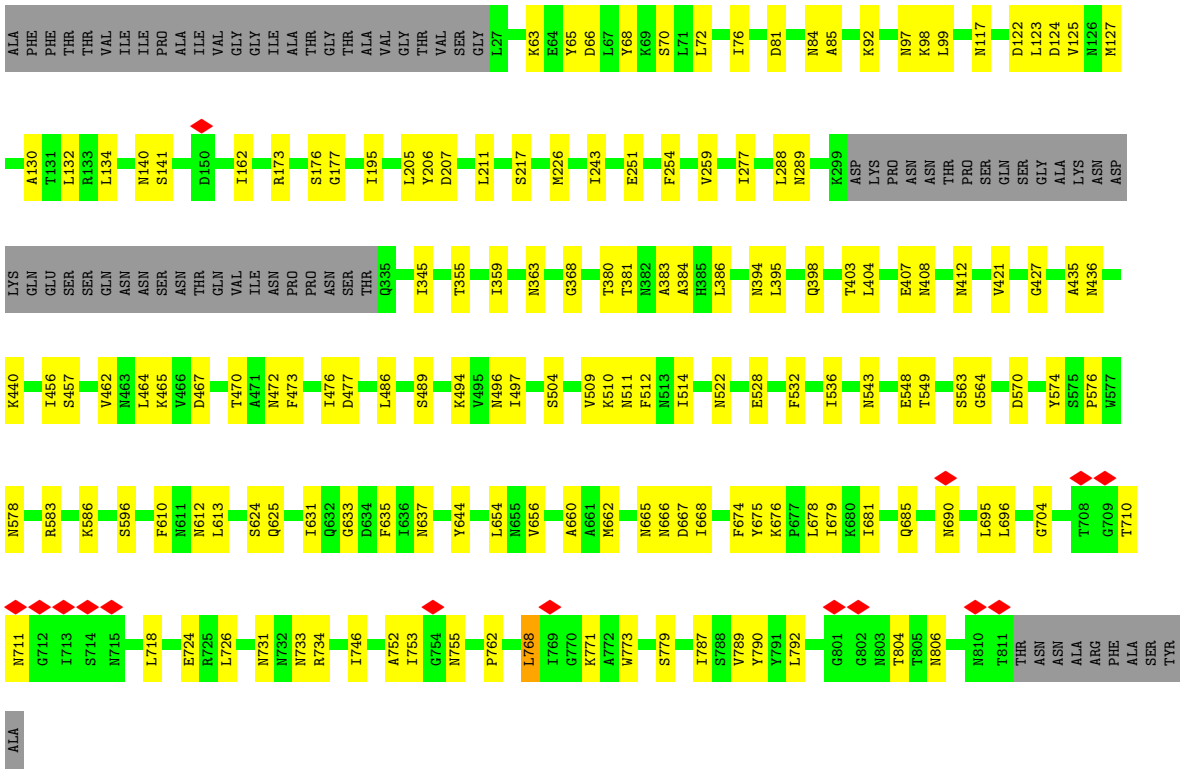
• Molecule 1: Vacuolating cytotoxin autotransporter

Chain E:



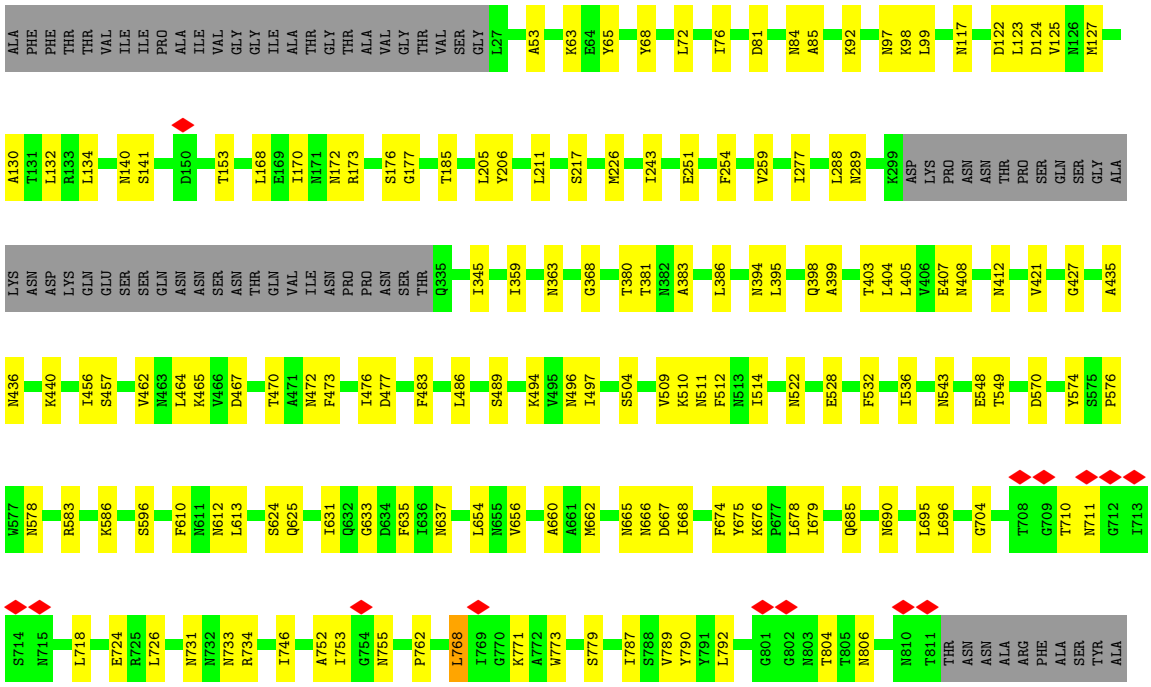
• Molecule 1: Vacuolating cytotoxin autotransporter

Chain F: 72% 19% 9%



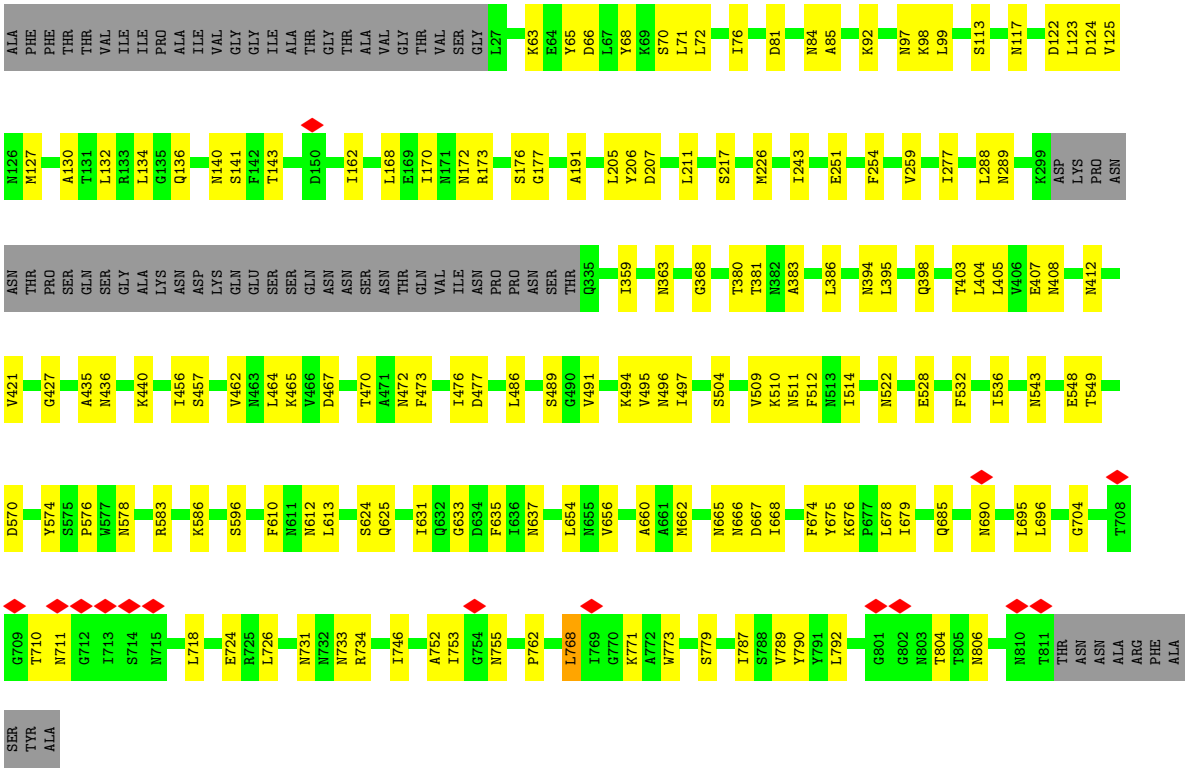
• Molecule 1: Vacuolating cytotoxin autotransporter

Chain G: 73% 19% 9%



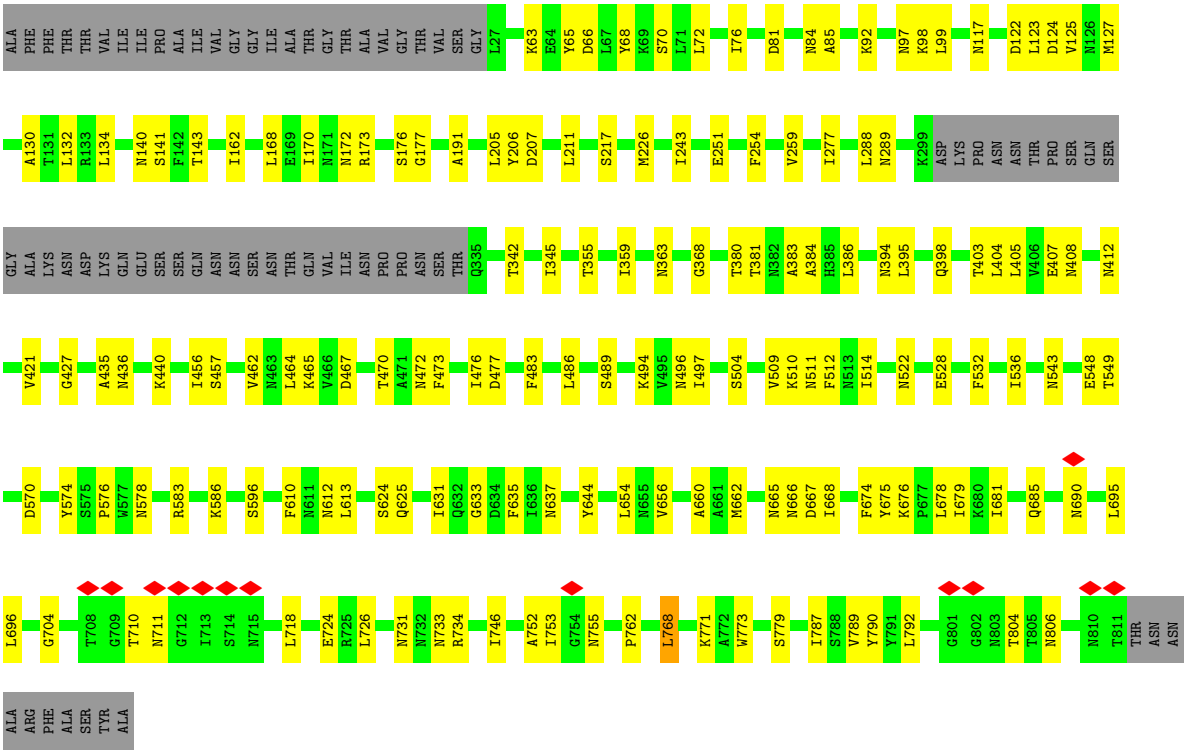
• Molecule 1: Vacuolating cytotoxin autotransporter

Chain H:



- Molecule 1: Vacuolating cytotoxin autotransporter

Chain I:



• Molecule 1: Vacuolating cytotoxin autotransporter

Chain J: 72% 20% 9%

ALA PHE PHE THR THR VAL ILE ILE PRO ALA ILE VAL GLY GLY ILE ALA THR GLY THR ALA VAL GLY THR VAL SER GLY L27 R50 K63 E64 Y65 Y68 L71 L72 I76 D81 N84 A85 K92 N97 K98 L99 N117 D122 L123 D124 V125 N126 M127

A130 T131 L132 R133 L134 N140 S141 T153 I162 N172 R173 S176 G177 A178 T185 A191 I195 L205 Y206 D207 L211 S217 M226 I243 E251 F254 V259 I277 L288 N289 K299 ASP LYS PRO ASN ASN THR PRO SER GLN

SER GLY ALA LYS ASN ASP LYS GLN GLU SER SER GLN ASN ASN SER ASN THR GLN VAL ILE ASN PRO PRO ASN SER THR Q335 T355 I359 N363 G368 T380 T381 N382 A383 A384 H385 L386 N394 L395 Q398 A399 T403 L404 L405 V406 E407 N408 N412 V421

G427 A435 N436 K440 I456 S457 V462 N463 L464 K465 V466 D467 T470 A471 N472 F473 I476 D477 L486 S489 G490 V491 K494 V495 N496 I497 S504 V509 K510 N511 F512 N513 I514 N522 E528 F532 I536 N543 E548 T549 S563

G564 D570 Y574 S575 P576 W577 N578 R583 K586 S596 F610 N611 N612 L613 S624 Q625 I631 Q632 G633 D634 F635 I636 N637 L654 N655 V656 A660 A661 M662 N665 N666 D667 I668 F674 Y675 K676 P677 L678 I679 Q685 N690 L695 L696 G704

T708 G709 T710 N711 G712 I713 S714 N715 L718 E724 R725 L726 N731 N732 N733 R734 I746 A752 I753 G754 N755 P762 L768 I769 G770 K771 A772 W773 S779 I787 S788 V789 Y790 Y791 L792 G801 G802 N803 T804 T805 N806 N810 T811 THR ASN ASN ALA ARG

PHE ALA SER TYR ALA

• Molecule 1: Vacuolating cytotoxin autotransporter

Chain K: 72% 19% 9%

ALA PHE PHE THR THR VAL ILE ILE PRO ALA ILE VAL GLY GLY ILE ALA THR GLY THR ALA VAL GLY THR VAL SER GLY L27 K63 E64 Y65 D66 L67 Y68 K69 S70 L71 L72 I76 D81 N84 A85 K92 N97 K98 L99 S113 N117 D122 L123 D124 V125

N126 M127 A130 T131 L132 R133 L134 G135 Q136 N140 S141 D150 I162 L168 E169 I170 N171 N172 R173 S176 G177 A178 I195 L205 Y206 L211 M226 I243 F254 V259 I277 L288 N289 K299 ASP LYS PRO ASN ASN THR PRO SER GLN SER

GLY ALA LYS ASN ASP LYS GLN GLU SER SER GLN ASN ASN SER ASN THR GLN VAL ILE ASN PRO PRO ASN SER THR Q335 I345 I359 N363 G368 T380 T381 N382 A383 L386 N394 L395 Q398 T403 L404 L405 V406 E407 N408 N412 V421 G427

A435 N436 K440 I456 S457 V462 N463 L464 K465 V466 D467 T470 A471 N472 F473 I476 D477 L486 S489 G490 V491 K494 V495 N496 I497 S504 V509 K510 N511 F512 N513 I514 N522 E528 F532 I536 N543 E548 T549 S563 G564 D570

Y574 S575 P576 W577 N578 R583 K586 S596 F610 N611 N612 L613 S624 Q625 I631 Q632 G633 D634 F635 I636 N637 L645 L654 N655 V656 A660 A661 M662 N665 N666 D667 I668 F674 Y675 K676 P677 L678 I679 Q685 N690 L695 L696 G704

T708 G709 T710 N711 G712 I713 S714 N715 L718 E724 R725 L726 N731 N732 N733 R734 I746 A752 I753 G754 N755 P762 L768 I769 G770 K771 A772 W773 S779 I787 S788 V789 Y790 Y791 L792 N800 G801 G802 N803 T804 T805 N806 N810 T811 THR ASN ASN ALA ARG

PHE ALA SER TYR ALA

- Molecule 1: Vacuolating cytotoxin autotransporter

Chain L: 72% 19% 9%

ALA PHE PHE THR THR VAL ILE ILE PRO ALA ILE VAL GLY GLY ILE ALA THR GLY THR ALA VAL GLY THR VAL SER GLY L27 A53 K63 E64 Y65 D66 L67 Y68 K69 S70 L71 L72 I76 D81 N84 A85 K92 N97 K98 L99 N117 D122 L123 D124 V125

N126 M127 A130 T131 L132 R133 L134 N140 S141 D150 I162 R173 S176 G177 I195 L205 Y206 D207 L211 S217 M226 I243 E251 F254 V259 I277 L288 N289 K299 ASP LYS PRO ASN ASN THR PRO SER GLN SER GLY ALA LYS

ASN ASP LYS GLN GLU SER SER GLN ASN ASN SER ASN THR GLN VAL ILE ASN PRO PRO ASN SER THR Q335 I345 P348 T355 I359 N363 G368 T380 T381 N382 A383 A384 H385 L386 N394 L395 Q398 A399 T403 L404 E407 N408 N412 V421

G427 A435 N436 K440 I456 S457 V462 N463 L464 K465 V466 D467 T470 A471 N472 F473 I476 D477 L486 S489 K494 V495 N496 I497 S504 V509 K510 N511 F512 N513 I514 N522 E528 F532 I536 N543 E548 T549 S563 G564 D570

Y574 S575 P576 W577 N578 R583 K586 S596 F610 N611 N612 L613 S624 Q625 I631 Q632 G633 D634 F635 I636 N637 Y644 L645 L654 N655 V656 A660 A661 M662 N665 N666 D667 I668 F674 Y675 K676 P677 L678 I679 K680 I681 Q685 N690 L695 L696

G704 T708 G709 T710 N711 G712 I713 S714 N715 L718 E724 R725 L726 N731 N732 N733 R734 I746 A752 I753 G754 N755 P762 L768 I769 G770 K771 A772 W773 S779 I787 S788 V789 V790 Y791 L792 N800 G801 G802 N803 T804 T805 N806 N810 T811 THR ASN

ASN ALA ARG PHE ALA SER TYR ALA

4 Experimental information (i)

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, D6	Depositor
Number of particles used	31625	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	PHASE FLIPPING AND AMPLITUDE CORRECTION	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ($e^-/Å^2$)	7	Depositor
Minimum defocus (nm)	1300	Depositor
Maximum defocus (nm)	2500	Depositor
Magnification	130000	Depositor
Image detector	GATAN K2 SUMMIT (4k x 4k)	Depositor
Maximum map value	2.167	Depositor
Minimum map value	-0.896	Depositor
Average map value	0.008	Depositor
Map value standard deviation	0.074	Depositor
Recommended contour level	0.3	Depositor
Map size (Å)	407.03998, 407.03998, 407.03998	wwPDB
Map dimensions	384, 384, 384	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	1.06, 1.06, 1.06	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	0.31	0/5820	0.59	1/7882 (0.0%)
1	B	0.31	0/5820	0.59	1/7882 (0.0%)
1	C	0.31	0/5820	0.59	1/7882 (0.0%)
1	D	0.31	0/5820	0.59	1/7882 (0.0%)
1	E	0.31	0/5820	0.59	1/7882 (0.0%)
1	F	0.31	0/5820	0.59	1/7882 (0.0%)
1	G	0.31	0/5820	0.59	1/7882 (0.0%)
1	H	0.31	0/5820	0.59	1/7882 (0.0%)
1	I	0.31	0/5820	0.59	1/7882 (0.0%)
1	J	0.31	0/5820	0.59	1/7882 (0.0%)
1	K	0.31	0/5820	0.59	1/7882 (0.0%)
1	L	0.31	0/5820	0.59	1/7882 (0.0%)
All	All	0.31	0/69840	0.59	12/94584 (0.0%)

There are no bond length outliers.

The worst 5 of 12 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	G	132	LEU	CA-CB-CG	5.30	127.49	115.30
1	E	132	LEU	CA-CB-CG	5.29	127.48	115.30
1	A	132	LEU	CA-CB-CG	5.29	127.46	115.30
1	H	132	LEU	CA-CB-CG	5.28	127.45	115.30
1	J	132	LEU	CA-CB-CG	5.28	127.45	115.30

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	5722	0	5634	90	0
1	B	5722	0	5634	95	0
1	C	5722	0	5634	95	0
1	D	5722	0	5634	97	0
1	E	5722	0	5634	92	0
1	F	5722	0	5634	88	0
1	G	5722	0	5634	91	0
1	H	5722	0	5634	92	0
1	I	5722	0	5634	94	0
1	J	5722	0	5634	95	0
1	K	5722	0	5634	91	0
1	L	5722	0	5634	92	0
All	All	68664	0	67608	1101	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 8.

The worst 5 of 1101 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:B:654:LEU:H	1:B:696:LEU:HA	1.59	0.67
1:F:654:LEU:H	1:F:696:LEU:HA	1.59	0.67
1:H:654:LEU:H	1:H:696:LEU:HA	1.59	0.67
1:D:654:LEU:H	1:D:696:LEU:HA	1.60	0.67
1:J:654:LEU:H	1:J:696:LEU:HA	1.59	0.67

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	746/821 (91%)	652 (87%)	94 (13%)	0	100	100
1	B	746/821 (91%)	652 (87%)	94 (13%)	0	100	100
1	C	746/821 (91%)	652 (87%)	94 (13%)	0	100	100
1	D	746/821 (91%)	652 (87%)	94 (13%)	0	100	100
1	E	746/821 (91%)	652 (87%)	94 (13%)	0	100	100
1	F	746/821 (91%)	652 (87%)	94 (13%)	0	100	100
1	G	746/821 (91%)	653 (88%)	93 (12%)	0	100	100
1	H	746/821 (91%)	653 (88%)	93 (12%)	0	100	100
1	I	746/821 (91%)	652 (87%)	94 (13%)	0	100	100
1	J	746/821 (91%)	652 (87%)	94 (13%)	0	100	100
1	K	746/821 (91%)	652 (87%)	94 (13%)	0	100	100
1	L	746/821 (91%)	652 (87%)	94 (13%)	0	100	100
All	All	8952/9852 (91%)	7826 (87%)	1126 (13%)	0	100	100

There are no Ramachandran outliers to report.

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	617/674 (92%)	614 (100%)	3 (0%)	88	95
1	B	617/674 (92%)	614 (100%)	3 (0%)	88	95
1	C	617/674 (92%)	614 (100%)	3 (0%)	88	95
1	D	617/674 (92%)	614 (100%)	3 (0%)	88	95
1	E	617/674 (92%)	614 (100%)	3 (0%)	88	95
1	F	617/674 (92%)	614 (100%)	3 (0%)	88	95
1	G	617/674 (92%)	614 (100%)	3 (0%)	88	95
1	H	617/674 (92%)	614 (100%)	3 (0%)	88	95
1	I	617/674 (92%)	614 (100%)	3 (0%)	88	95

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	J	617/674 (92%)	614 (100%)	3 (0%)	88	95
1	K	617/674 (92%)	614 (100%)	3 (0%)	88	95
1	L	617/674 (92%)	614 (100%)	3 (0%)	88	95
All	All	7404/8088 (92%)	7368 (100%)	36 (0%)	89	95

5 of 36 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	J	665	ASN
1	L	768	LEU
1	J	768	LEU
1	K	768	LEU
1	E	613	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 96 such sidechains are listed below:

Mol	Chain	Res	Type
1	H	387	ASN
1	I	810	ASN
1	H	665	ASN
1	I	387	ASN
1	J	665	ASN

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.

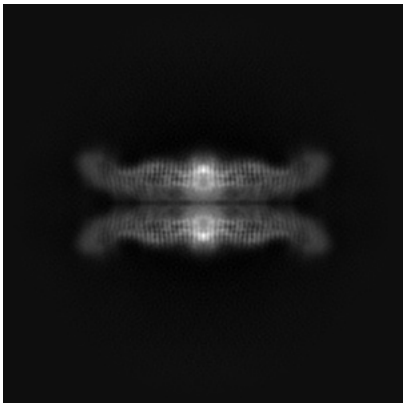
6 Map visualisation (i)

This section contains visualisations of the EMDB entry EMD-0546. These allow visual inspection of the internal detail of the map and identification of artifacts.

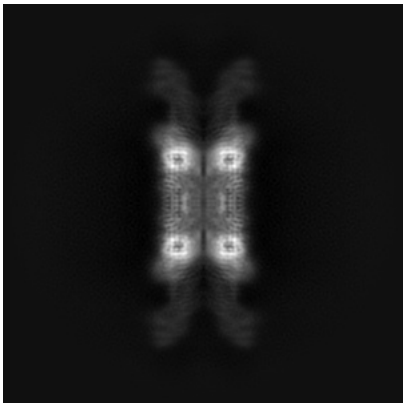
No raw map or half-maps were deposited for this entry and therefore no images, graphs, etc. pertaining to the raw map can be shown.

6.1 Orthogonal projections (i)

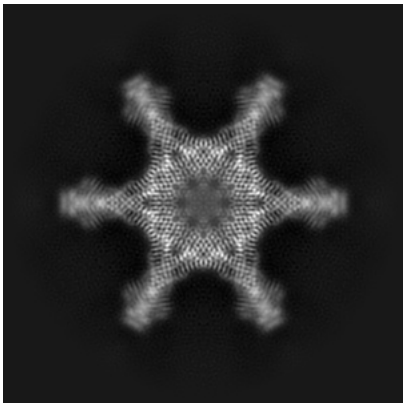
6.1.1 Primary map



X



Y

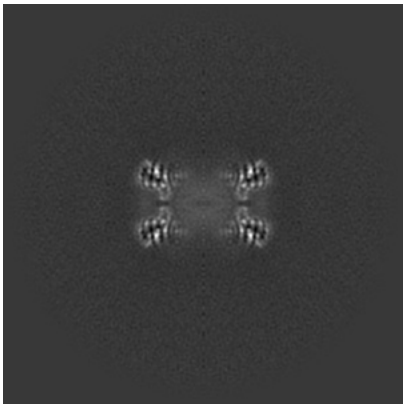


Z

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

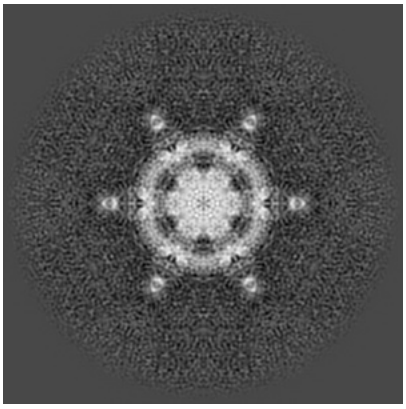
6.2 Central slices (i)

6.2.1 Primary map



X Index: 192

Y Index: 192

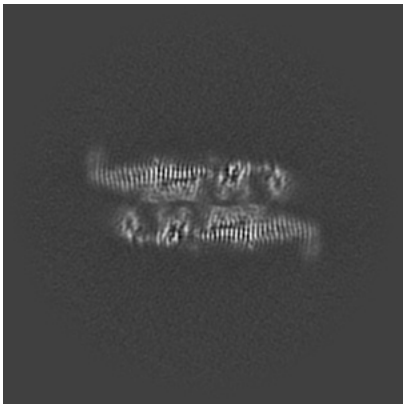


Z Index: 192

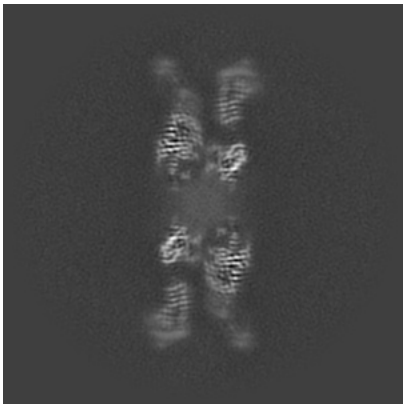
The images above show central slices of the map in three orthogonal directions.

6.3 Largest variance slices (i)

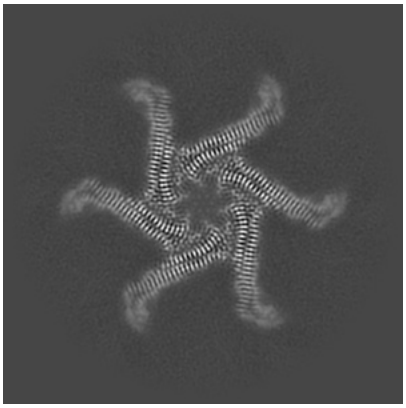
6.3.1 Primary map



X Index: 156



Y Index: 201



Z Index: 162

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

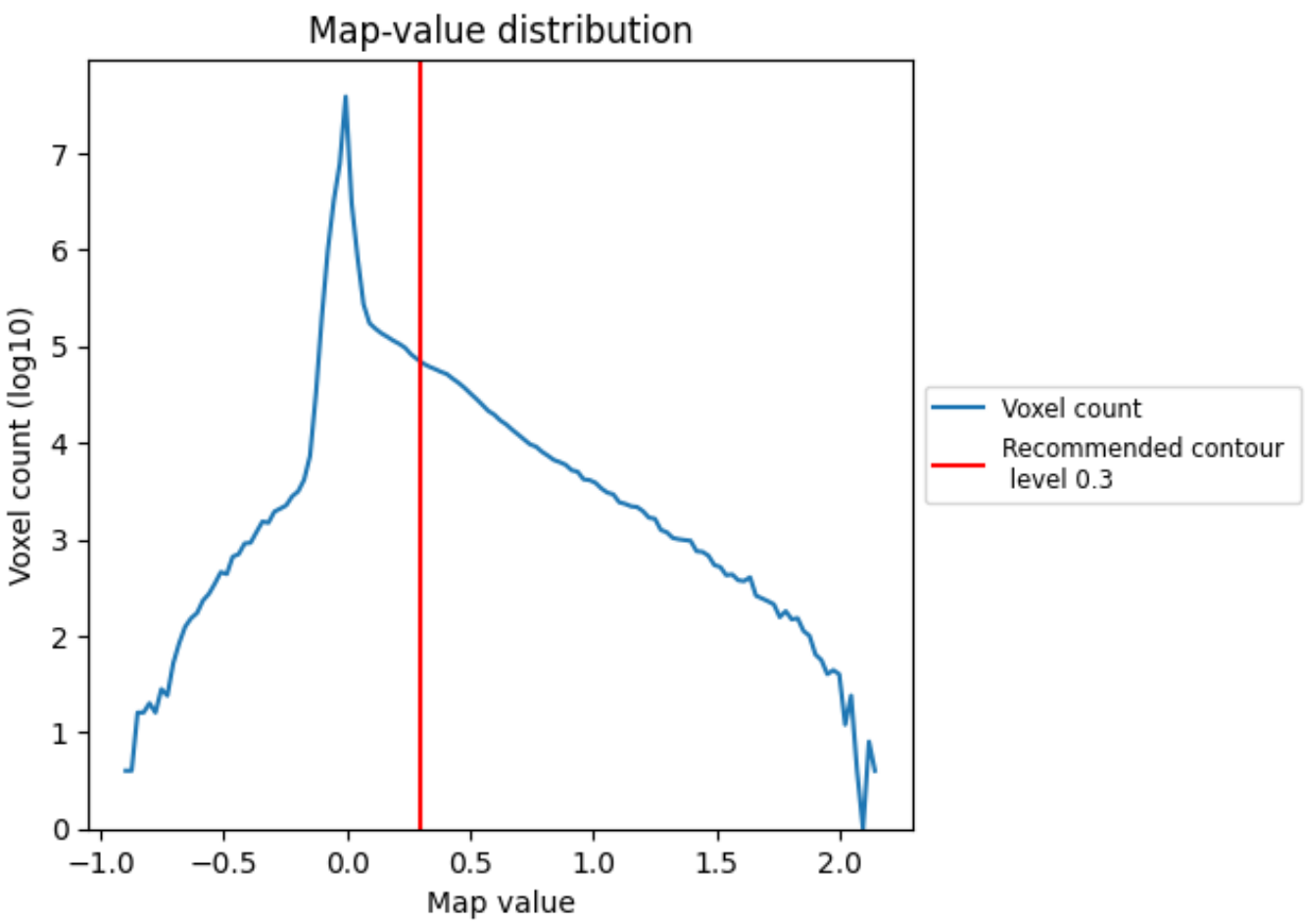
6.5 Mask visualisation (i)

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

7 Map analysis (i)

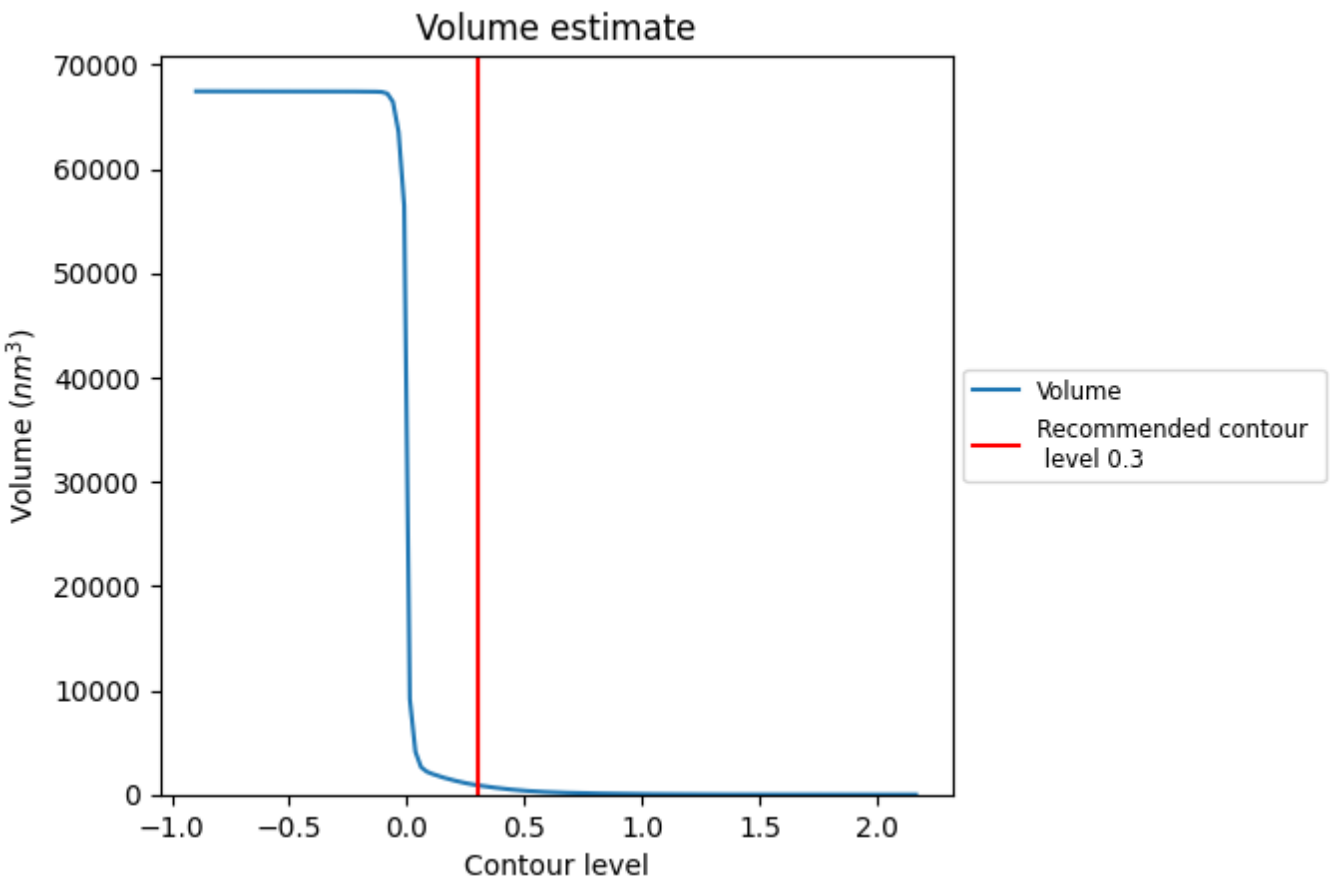
This section contains the results of statistical analysis of the map.

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.

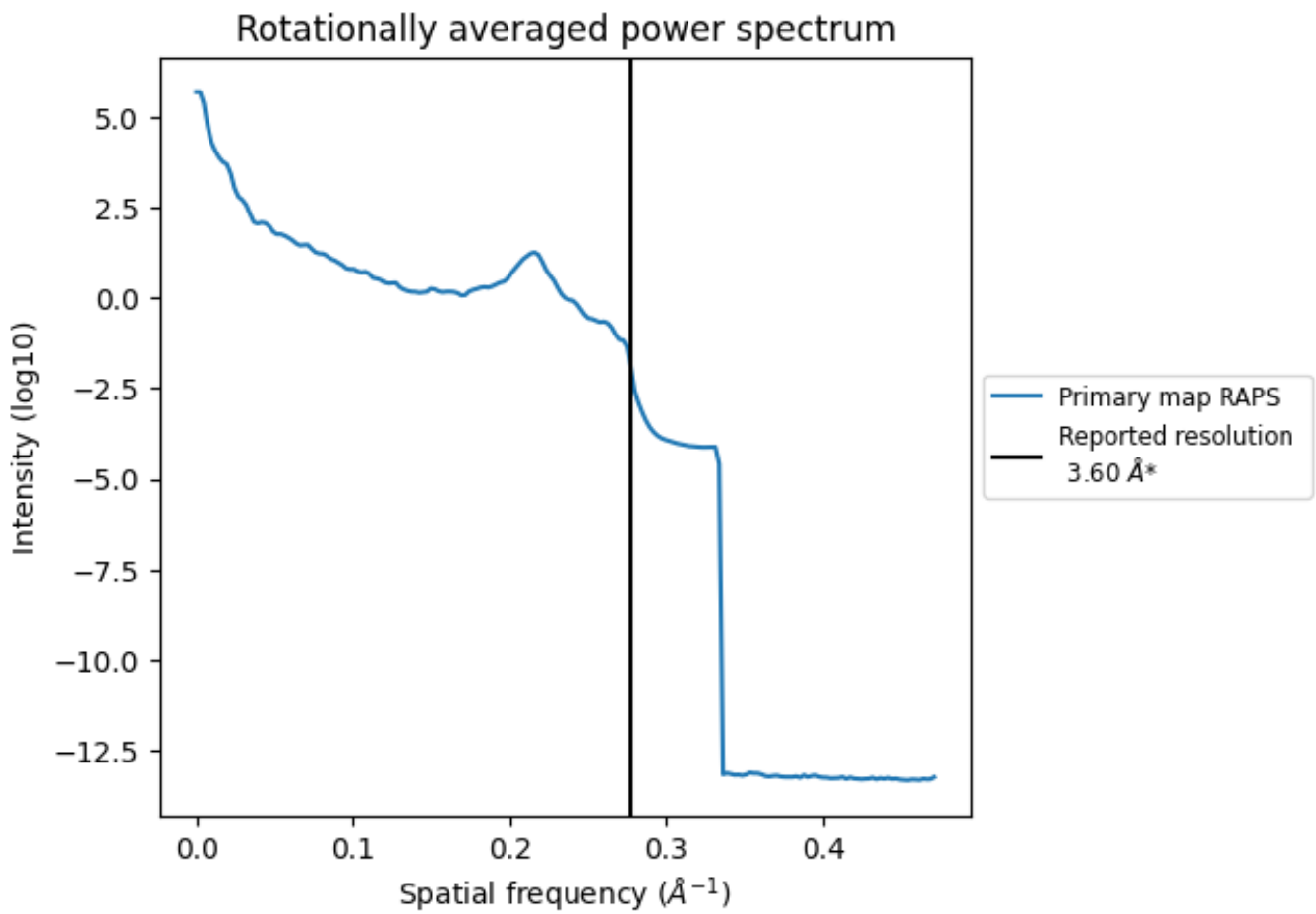
7.2 Volume estimate (i)



The volume at the recommended contour level is 895 nm^3; this corresponds to an approximate mass of 809 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 0.278 Å^{-1}

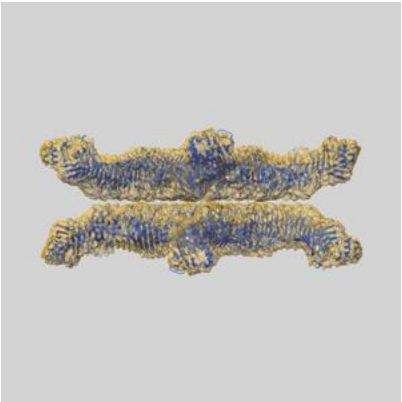
8 Fourier-Shell correlation (i)

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

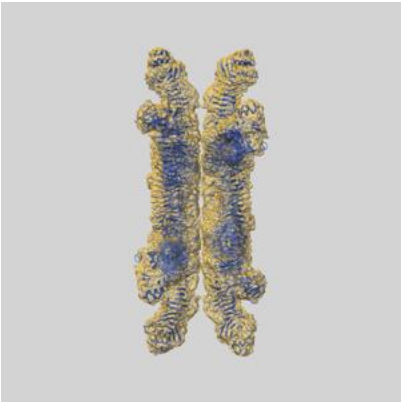
9 Map-model fit (i)

This section contains information regarding the fit between EMDB map EMD-0546 and PDB model 6NYM. Per-residue inclusion information can be found in section 3 on page 5.

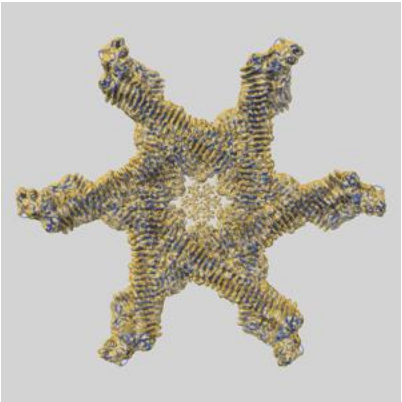
9.1 Map-model overlay (i)



X



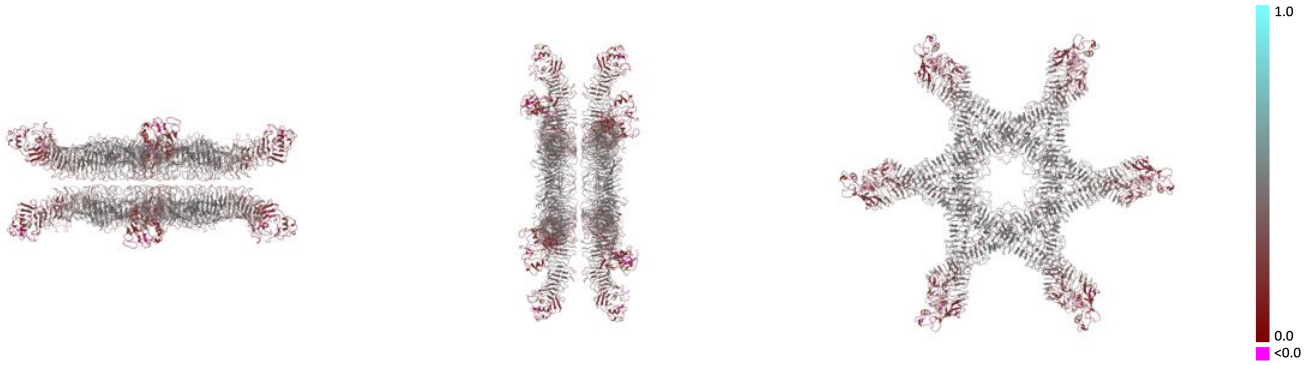
Y



Z

The images above show the 3D surface view of the map at the recommended contour level 0.3 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.

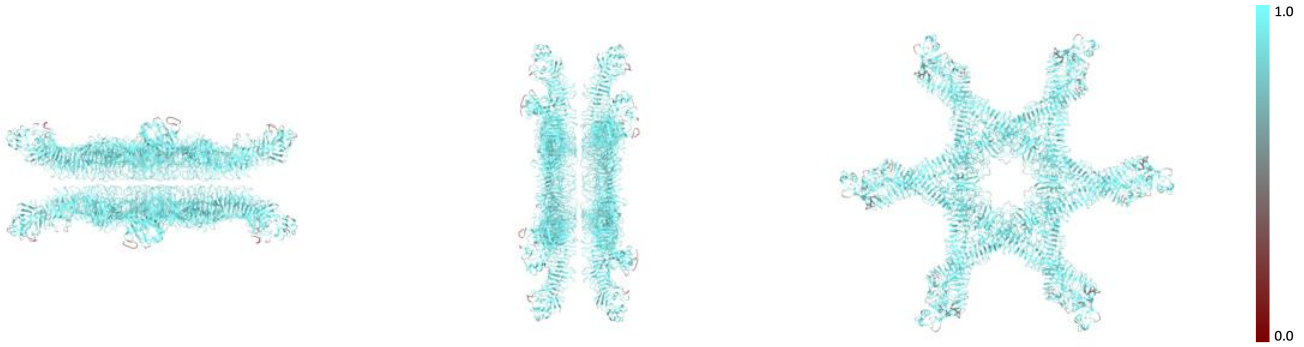
9.2 Q-score mapped to coordinate model (i)



X Y Z

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.

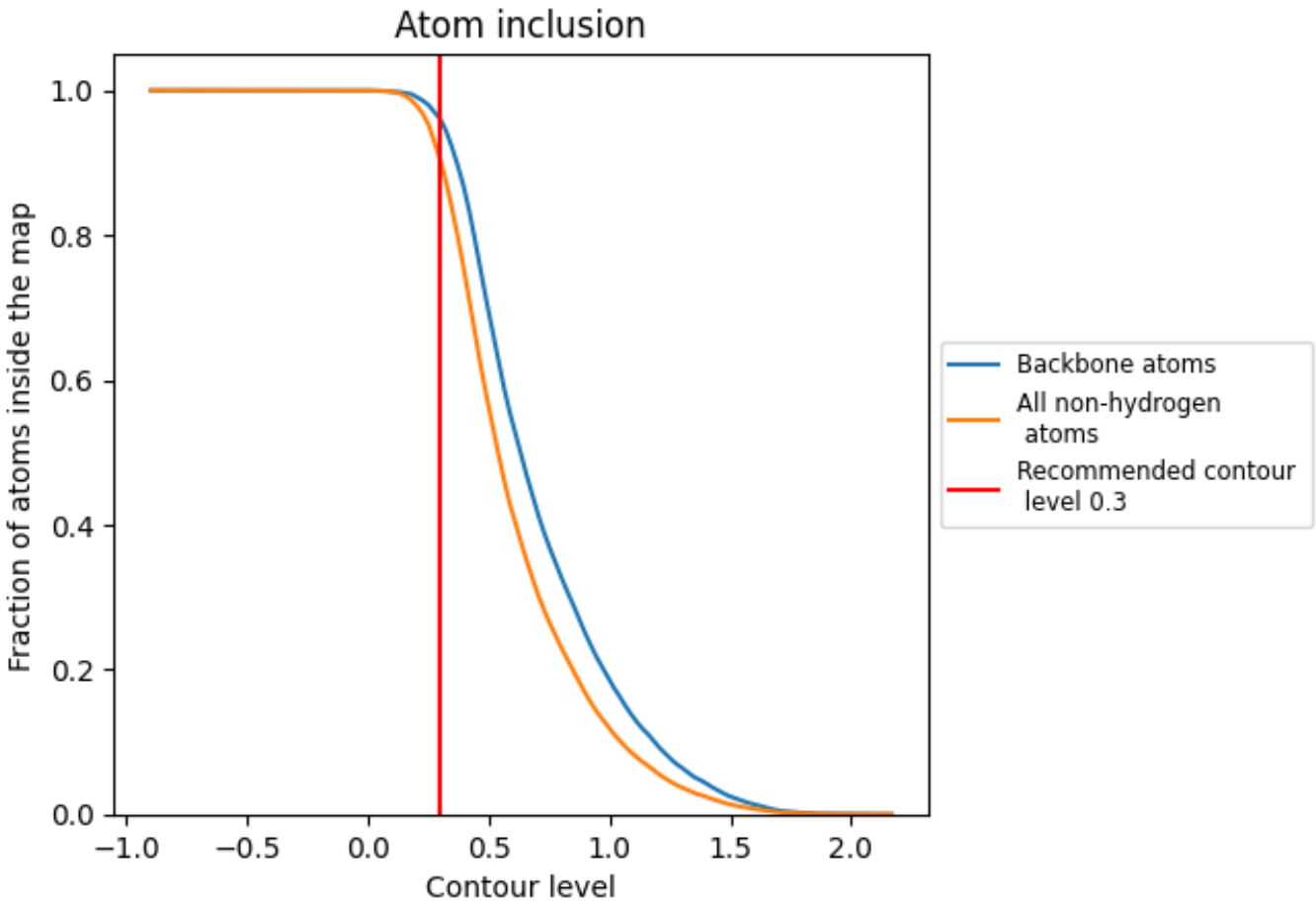
9.3 Atom inclusion mapped to coordinate model (i)



X Y Z

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 (0.3).

9.4 Atom inclusion (i)



At the recommended contour level, 96% of all backbone atoms, 90% of all non-hydrogen atoms, are inside the map.

9.5 Map-model fit summary (i)

The table lists the average atom inclusion at the recommended contour level (0.3) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	0.9047	0.3620
A	0.9045	0.3650
B	0.9013	0.3620
C	0.9059	0.3600
D	0.9063	0.3630
E	0.9036	0.3610
F	0.9061	0.3640
G	0.9033	0.3630
H	0.9031	0.3600
I	0.9066	0.3630
J	0.9068	0.3630
K	0.9026	0.3610
L	0.9061	0.3630