wwPDB EM Validation Summary Report (i)

Nov 12, 2022 – 11:48 PM EST

PDB ID	:	6U5Z
EMDB ID	:	EMD-20659
Title	:	Cryo-EM structure of E. coli LonA S679A
Authors	:	Botos, I.; Lountos, G.T.; Weimin, W.; Wlodawer, A.
Deposited on	:	2019-08-28
Resolution	:	3.50 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:

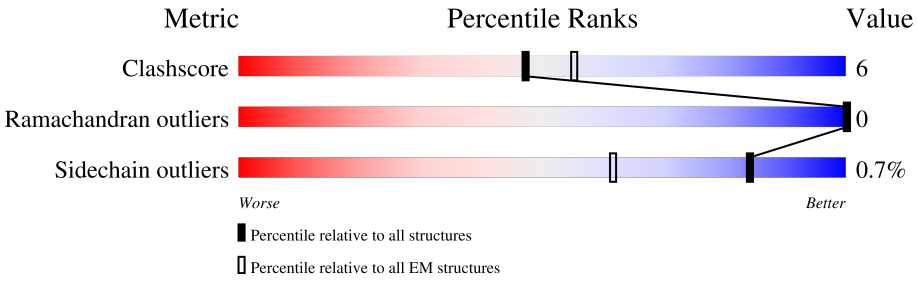
EMDB validation analysis	:	0.0.1.dev43
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
MapQ	:	1.9.9
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.31.2

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 3.50 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	784	39% 56% 11% 33%
1	B	784	9% 56% 11% 33%
1	C	784	56% 11% 33%
1	D	784	56% 11% 33%
1	E	784	11% 55% 12% 33%
1	F	784	57% 56% 11% 33%

2 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 24606 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 Lon protease.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	529	Total	C	N	O	S	0	0
			4101	2577	730	773	21		
1	B	529	Total	C	N	O	S	0	0
			4101	2577	730	773	21		
1	C	529	Total	C	N	O	S	0	0
			4101	2577	730	773	21		
1	D	529	Total	C	N	O	S	0	0
			4101	2577	730	773	21		
1	E	529	Total	C	N	O	S	0	0
			4101	2577	730	773	21		
1	F	529	Total	C	N	O	S	0	0
			4101	2577	730	773	21		

There are 6 discrepancies between the modelled and reference sequences:

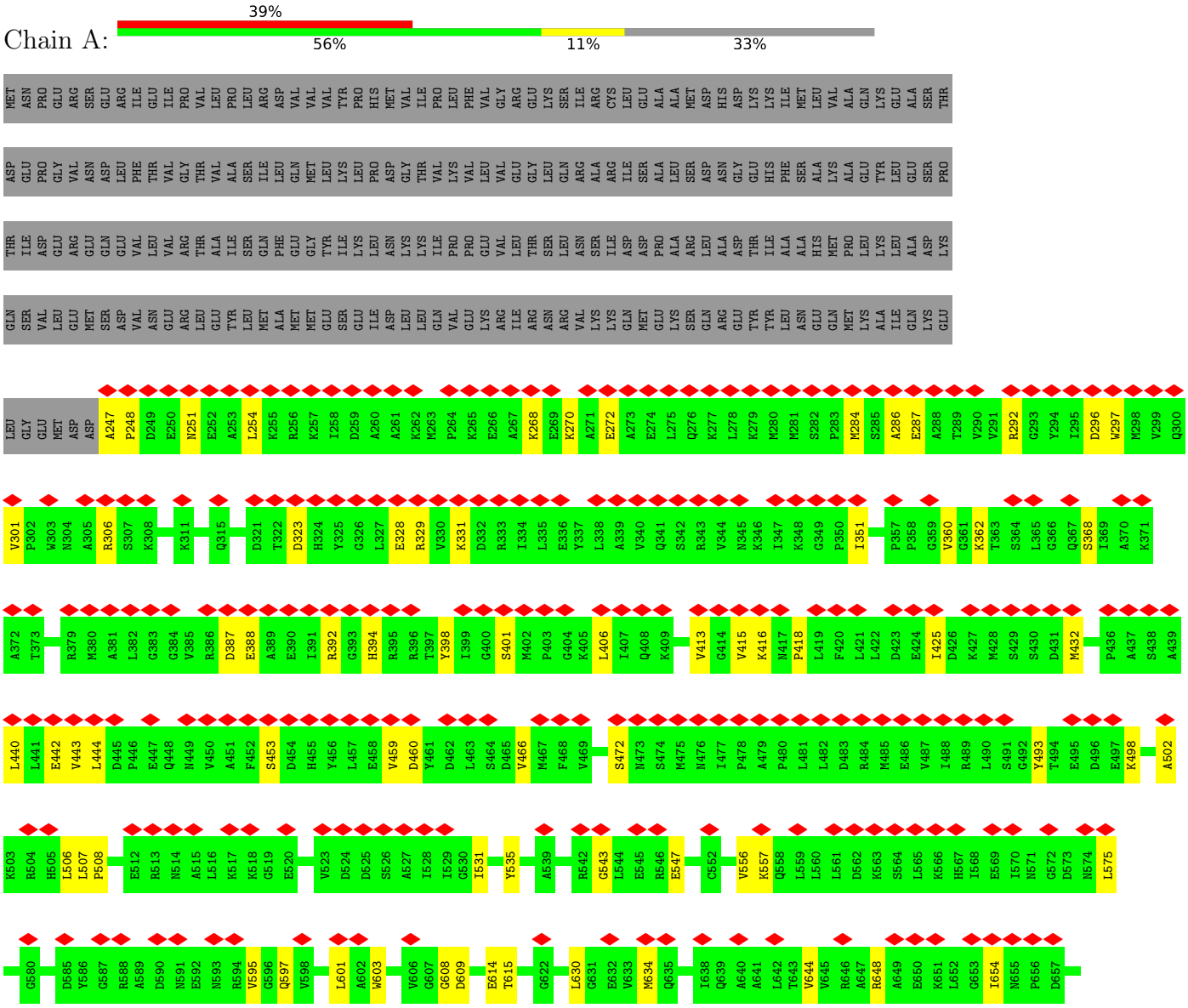
Chain	Residue	Modelled	Actual	Comment	Reference
A	679	ALA	SER	engineered mutation	UNP C3TLS2
B	679	ALA	SER	engineered mutation	UNP C3TLS2
C	679	ALA	SER	engineered mutation	UNP C3TLS2
D	679	ALA	SER	engineered mutation	UNP C3TLS2
E	679	ALA	SER	engineered mutation	UNP C3TLS2
F	679	ALA	SER	engineered mutation	UNP C3TLS2

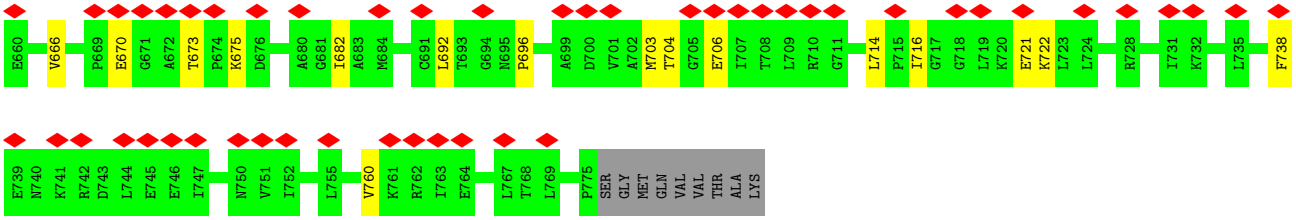
3 Residue-property plots (i)

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: Lon protease

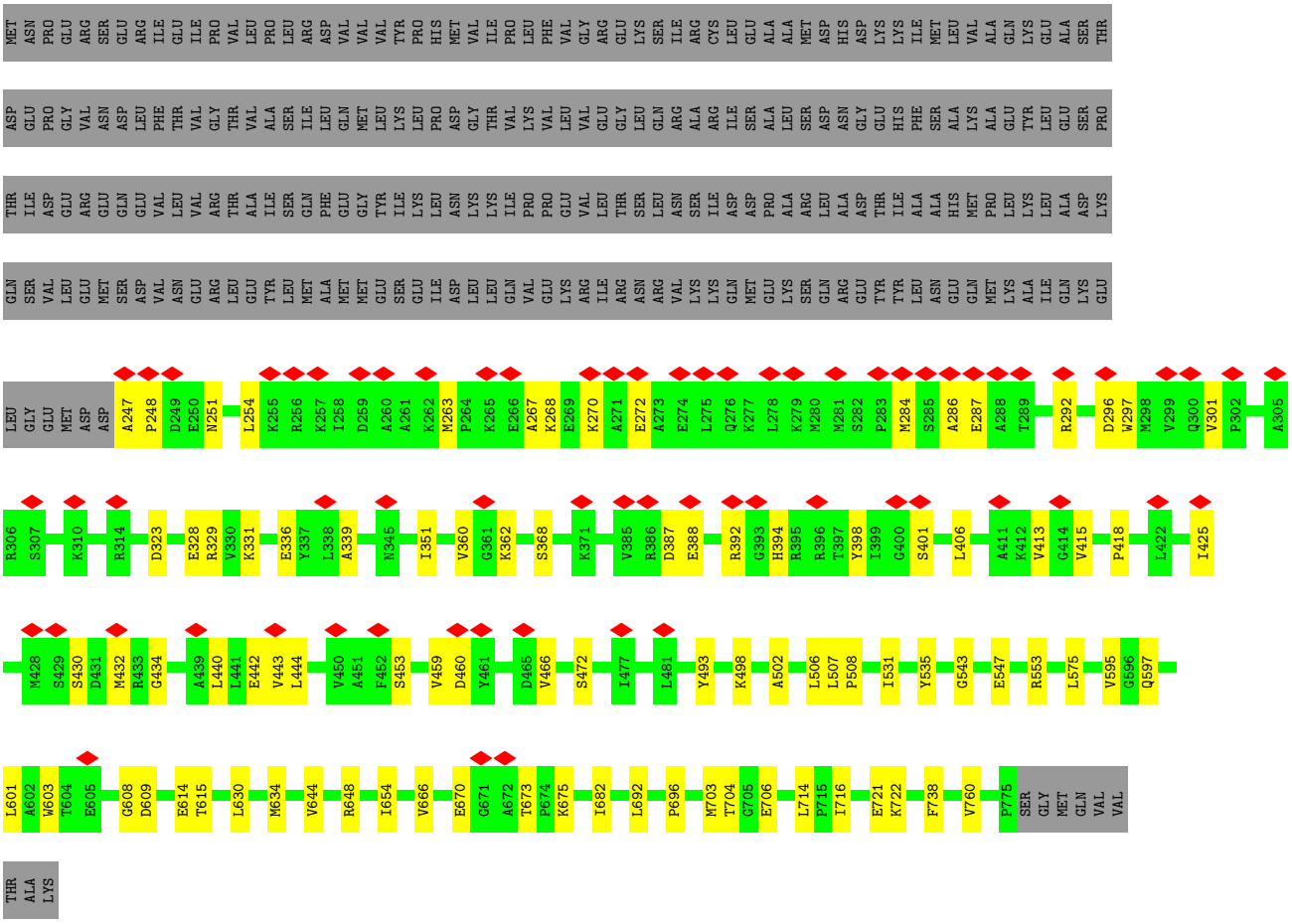
Chain A:





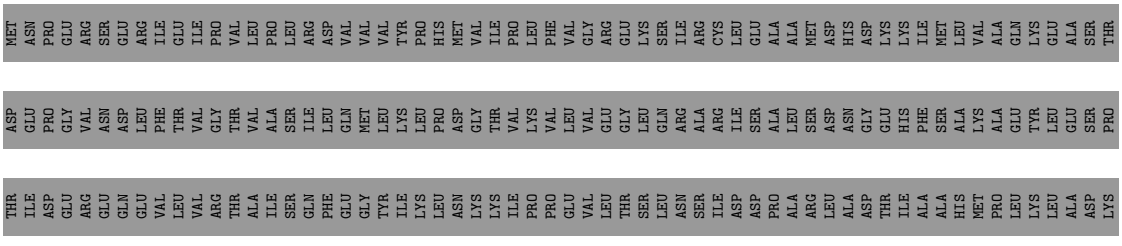
• Molecule 1: Lon protease

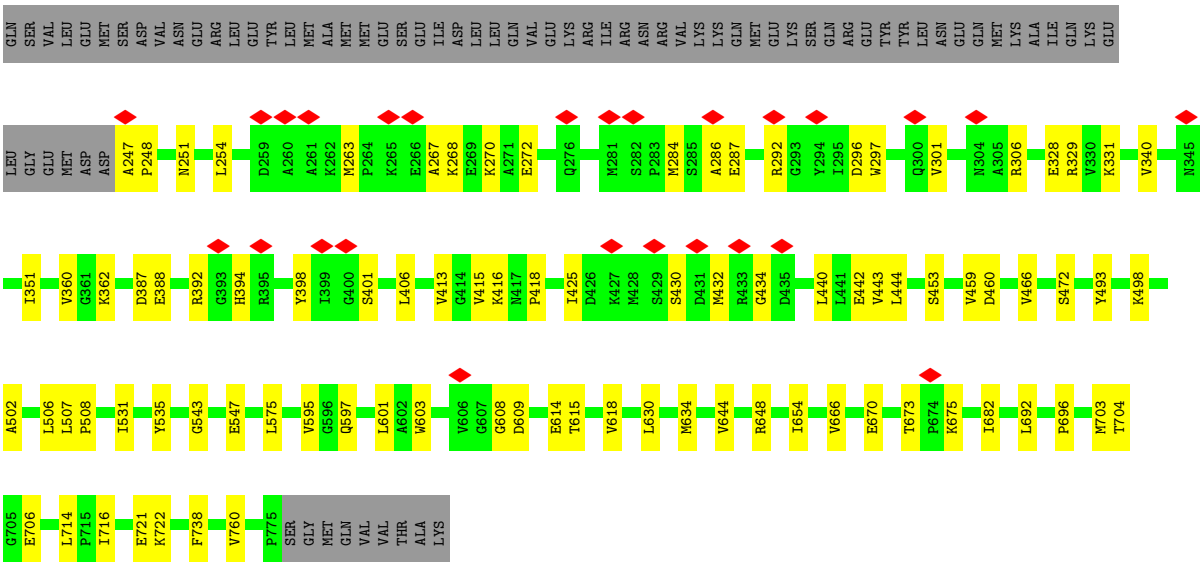
Chain B:



• Molecule 1: Lon protease

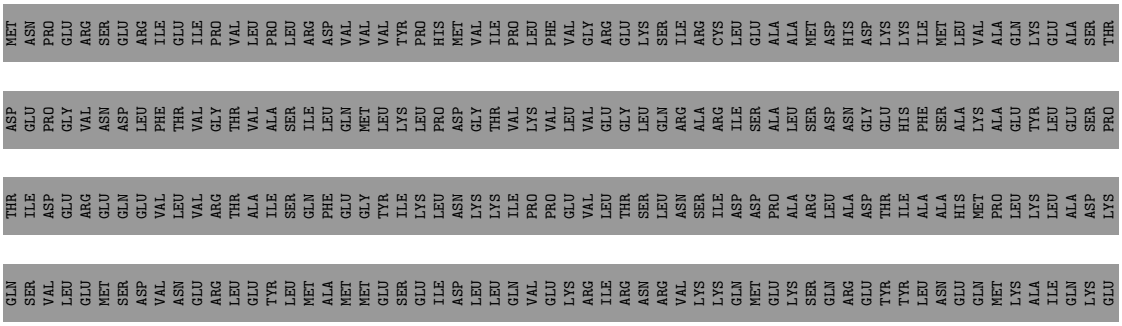
Chain C:





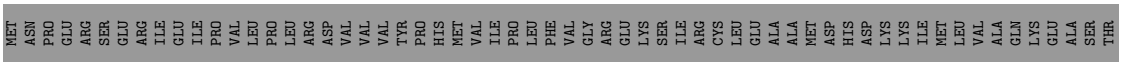
• Molecule 1: Lon protease

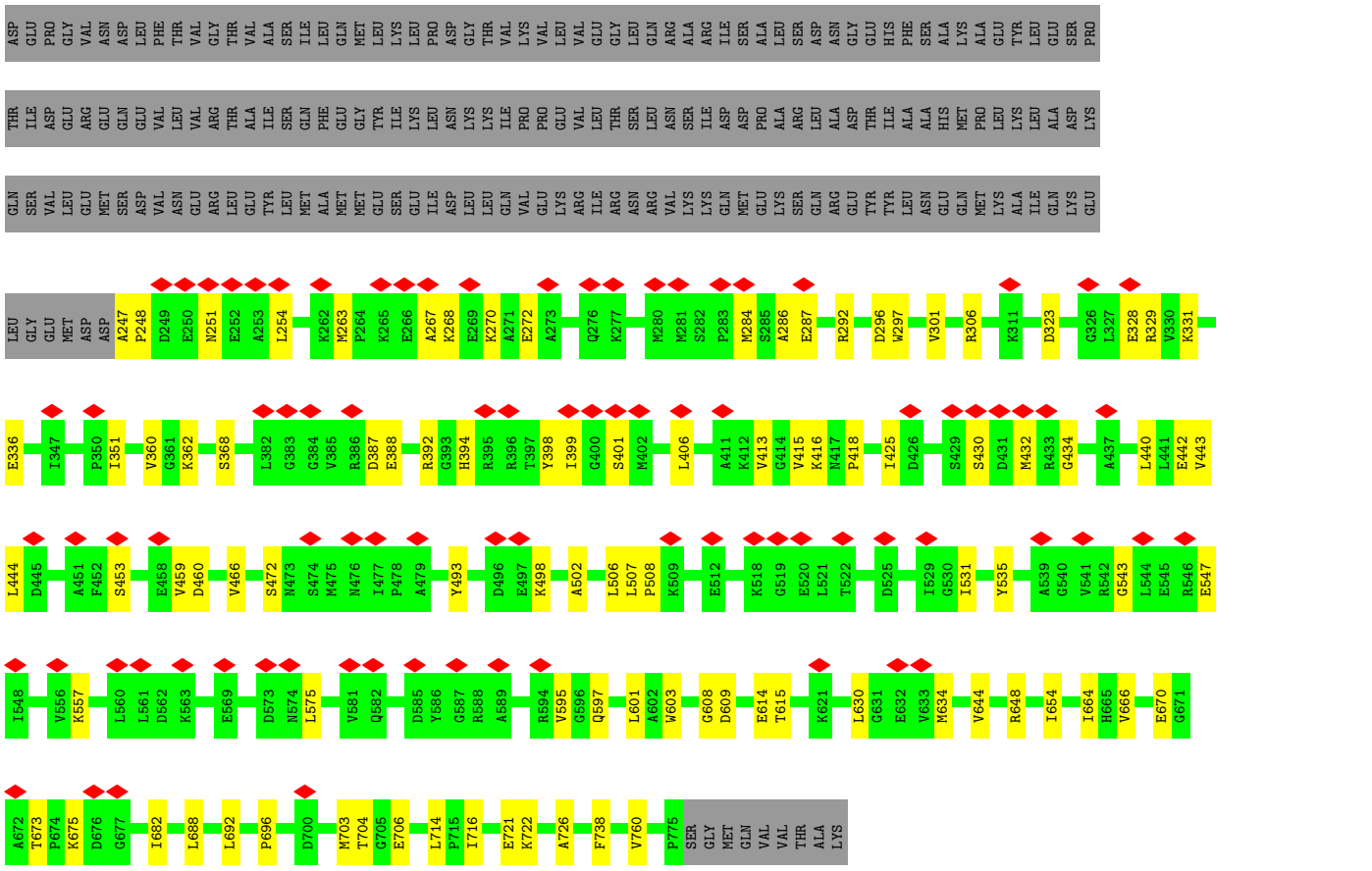
Chain D:



• Molecule 1: Lon protease

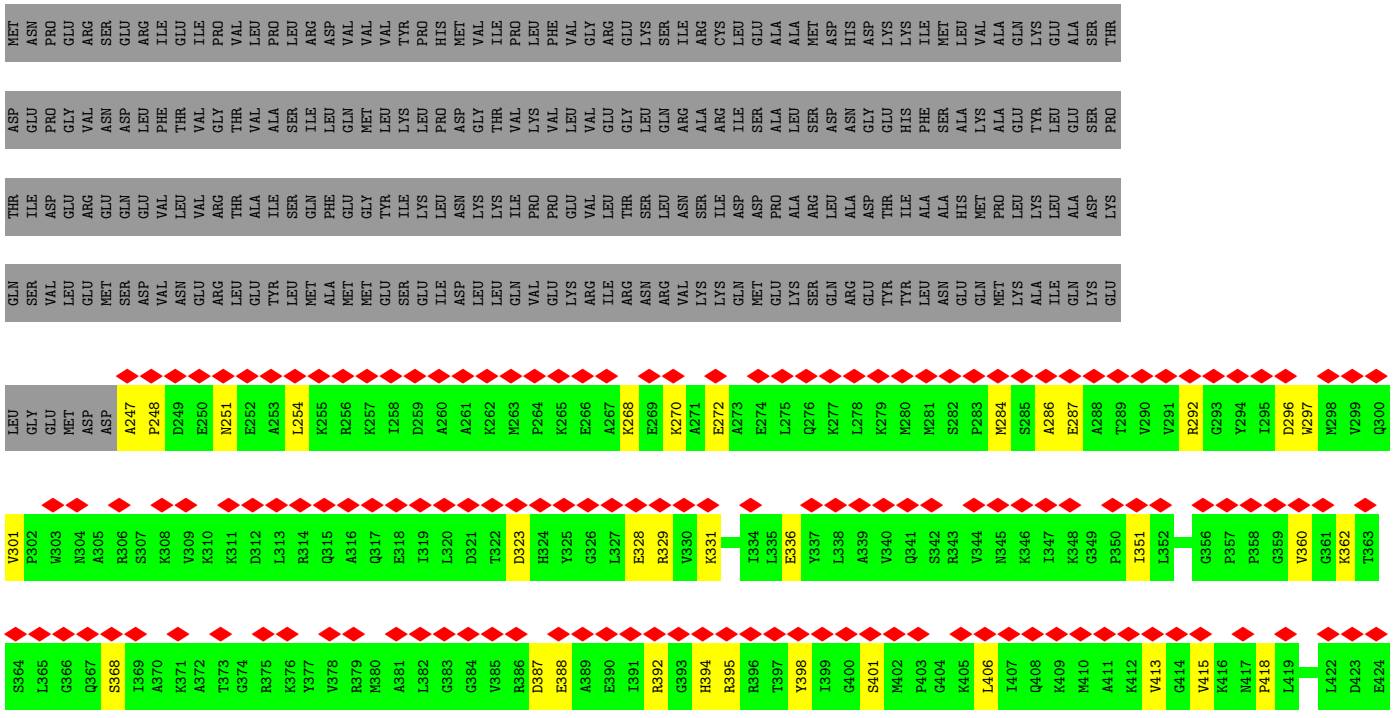
Chain E:

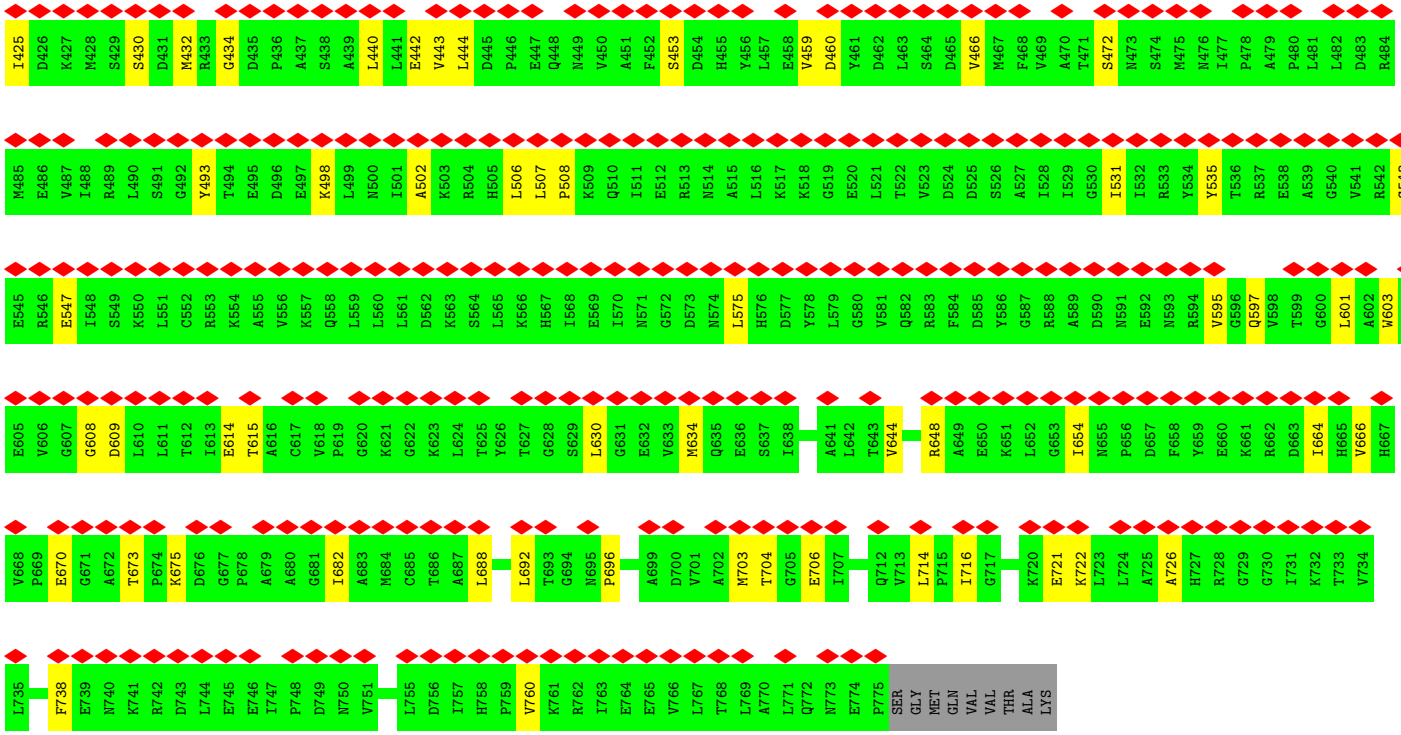




• Molecule 1: Lon protease

Chain F:




I425 D426 K427 M428 S429 S430 D431 M432 R433 G434 D435 P436 A437 S438 A439 L440 L441 E442 V443 L444 D445 P446 E447 Q448 N449 V450 A451 F452 S453 D454 H455 Y456 L457 E458 V459 D460 Y461 D462 L463 S464 D465 V466 M467 F468 V469 A470 T471 S472 N473 S474 M475 N476 I477 P478 A479 P480 L481 L482 D483 R484
M485 E486 V487 I488 R489 L490 S491 G492 Y493 T494 E495 D496 E497 K498 L499 N500 I501 A502 K503 R504 H505 L506 L507 P508 K509 Q510 I511 E512 R513 N514 A515 L516 K517 K518 G519 E520 L521 T522 V523 D524 D525 S526 A527 I528 I529 G530 I531 I532 R533 Y534 Y535 T536 R537 E538 A539 G540 V541 R542 G543 L544
E545 R546 E547 I548 S549 K550 L551 C552 R553 K554 A555 V556 K557 Q558 L559 L560 L561 D562 K563 S564 L565 K566 H567 I568 E569 I570 N571 G572 D573 N574 L575 H576 D577 Y578 L579 G580 V581 Q582 R583 F584 D585 Y586 G587 R588 A589 D590 N591 E592 N593 R594 V595 G596 Q597 V598 T599 G600 L601 A602 W603 T604
E605 V606 G607 G608 D609 L610 L611 T612 I613 E614 T615 A616 C617 V618 P619 G620 K621 G622 K623 L624 T625 Y626 T627 G628 S629 L630 G631 E632 V633 M634 Q635 E636 S637 I638 A641 L642 T643 V644 R648 A649 E650 K651 L652 G653 I654 N655 P656 D657 F658 Y659 E660 K661 R662 D663 I664 H665 V666 H667
V668 P669 E670 G671 A672 T673 P674 K675 D676 G677 P678 A679 A680 G681 I682 A683 M684 C685 T686 A687 L688 L692 T693 G694 N695 P696 A699 D700 V701 A702 M703 T704 G705 E706 I707 Q712 V713 L714 P715 I716 G717 K720 E721 K722 L723 L724 A725 A726 H727 R728 G729 G730 I731 K732 T733 V734
L735 F738 E739 N740 K741 R742 D743 L744 E745 E746 I747 P748 D749 N750 V751 L755 D756 I757 H758 P759 V760 K761 R762 I763 E764 E765 V766 L767 T768 L769 A770 L771 Q772 N773 E774 P775 SER GLY MET GLN VAL VAL THR ALA LYS

4 Experimental information (i)

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, C1	Depositor
Number of particles used	274765	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 ($e^-/Å^2$)	50	Depositor
Minimum defocus (nm)	Not provided	
Maximum defocus (nm)	Not provided	
Magnification	Not provided	
Image detector	GATAN K2 SUMMIT (4k x 4k)	Depositor
Maximum map value	0.099	Depositor
Minimum map value	-0.063	Depositor
Average map value	0.000	Depositor
Map value standard deviation	0.002	Depositor
Recommended contour level	0.0048	Depositor
Map size (Å)	271.36, 271.36, 271.36	wwPDB
Map dimensions	320, 320, 320	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	0.84799993, 0.84799993, 0.84799993	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	0.24	0/4161	0.42	0/5614
1	B	0.24	0/4161	0.42	0/5614
1	C	0.24	0/4161	0.42	0/5614
1	D	0.24	0/4161	0.42	0/5614
1	E	0.24	0/4161	0.42	0/5614
1	F	0.24	0/4161	0.42	0/5614
All	All	0.24	0/24966	0.42	0/33684

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

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	4101	0	4229	54	0
1	B	4101	0	4229	55	0
1	C	4101	0	4229	55	0
1	D	4101	0	4229	52	0
1	E	4101	0	4229	58	0
1	F	4101	0	4229	54	0
All	All	24606	0	25374	318	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 6.

The worst 5 of 318 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:E:557:LYS:HD3	1:F:336:GLU:HG2	1.61	0.81
1:A:670:GLU:O	1:A:673:THR:HG23	1.83	0.79
1:E:670:GLU:O	1:E:673:THR:HG23	1.83	0.78
1:F:670:GLU:O	1:F:673:THR:HG23	1.83	0.78
1:B:670:GLU:O	1:B:673:THR:HG23	1.83	0.78

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	527/784 (67%)	502 (95%)	25 (5%)	0	100	100
1	B	527/784 (67%)	502 (95%)	25 (5%)	0	100	100
1	C	527/784 (67%)	502 (95%)	25 (5%)	0	100	100
1	D	527/784 (67%)	502 (95%)	25 (5%)	0	100	100
1	E	527/784 (67%)	502 (95%)	25 (5%)	0	100	100
1	F	527/784 (67%)	502 (95%)	25 (5%)	0	100	100
All	All	3162/4704 (67%)	3012 (95%)	150 (5%)	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	442/670 (66%)	439 (99%)	3 (1%)	84 93
1	B	442/670 (66%)	439 (99%)	3 (1%)	84 93
1	C	442/670 (66%)	439 (99%)	3 (1%)	84 93
1	D	442/670 (66%)	439 (99%)	3 (1%)	84 93
1	E	442/670 (66%)	439 (99%)	3 (1%)	84 93
1	F	442/670 (66%)	439 (99%)	3 (1%)	84 93
All	All	2652/4020 (66%)	2634 (99%)	18 (1%)	84 93

5 of 18 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	E	609	ASP
1	F	609	ASP
1	F	329	ARG
1	C	609	ASP
1	E	329	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 6 such sidechains are listed below:

Mol	Chain	Res	Type
1	D	510	GLN
1	E	510	GLN
1	F	510	GLN
1	B	510	GLN
1	A	510	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.

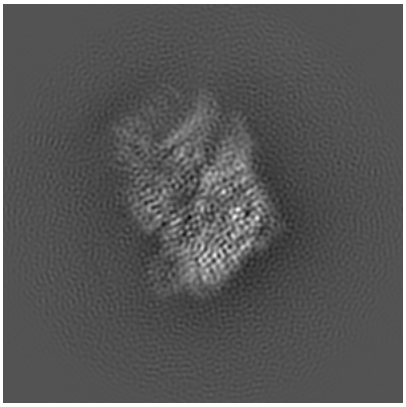
6 Map visualisation

This section contains visualisations of the EMDB entry EMD-20659. These allow visual inspection of the internal detail of the map and identification of artifacts.

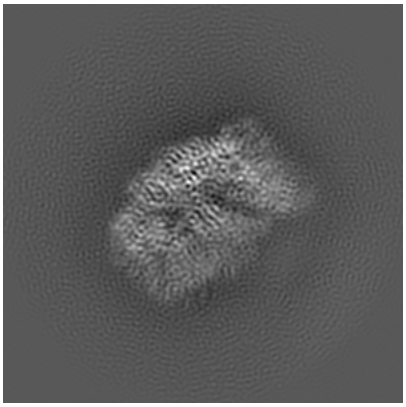
No raw map or half-maps were deposited for this entry and therefore no images, graphs, etc. pertaining to the raw map can be shown.

6.1 Orthogonal projections

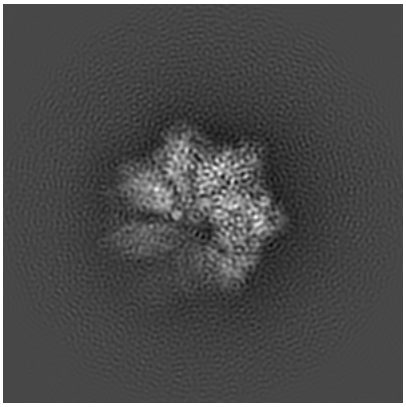
6.1.1 Primary map



X



Y

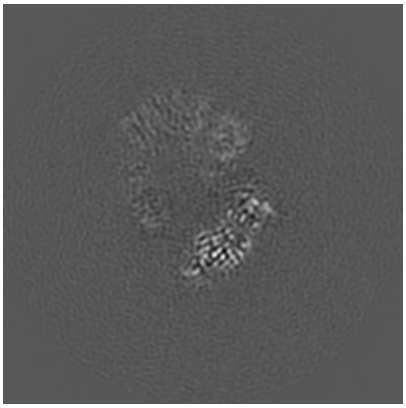


Z

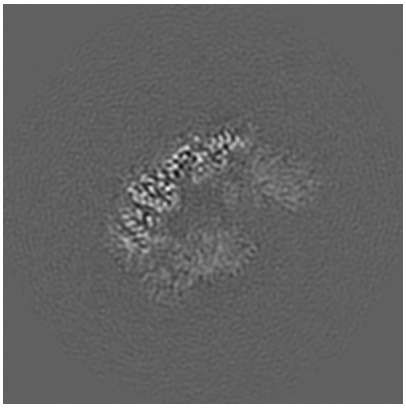
The images above show the map projected in three orthogonal directions.

6.2 Central slices

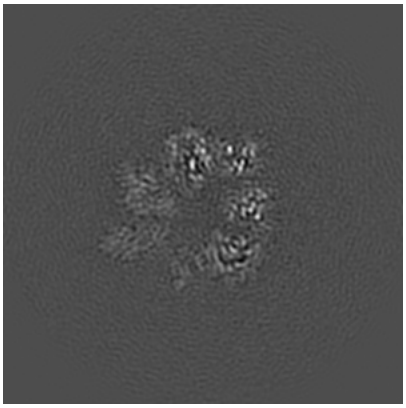
6.2.1 Primary map



X Index: 160



Y Index: 160

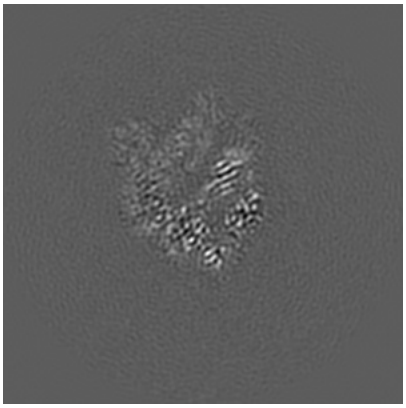


Z Index: 160

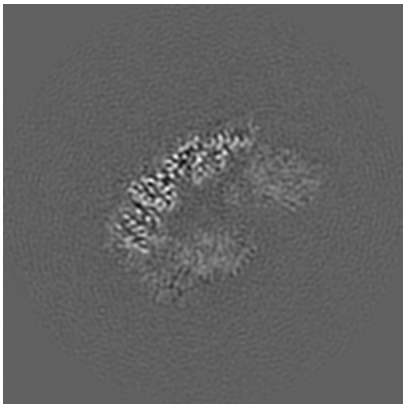
The images above show central slices of the map in three orthogonal directions.

6.3 Largest variance slices (i)

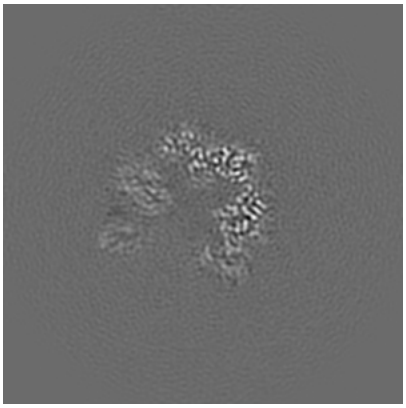
6.3.1 Primary map



X Index: 184



Y Index: 159



Z Index: 147

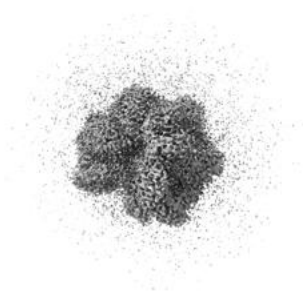
The images above show the largest variance slices of the map in three orthogonal directions.

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 0.0048. These images, in conjunction with the slice images, may facilitate assessment of whether an appropriate contour level has been provided.

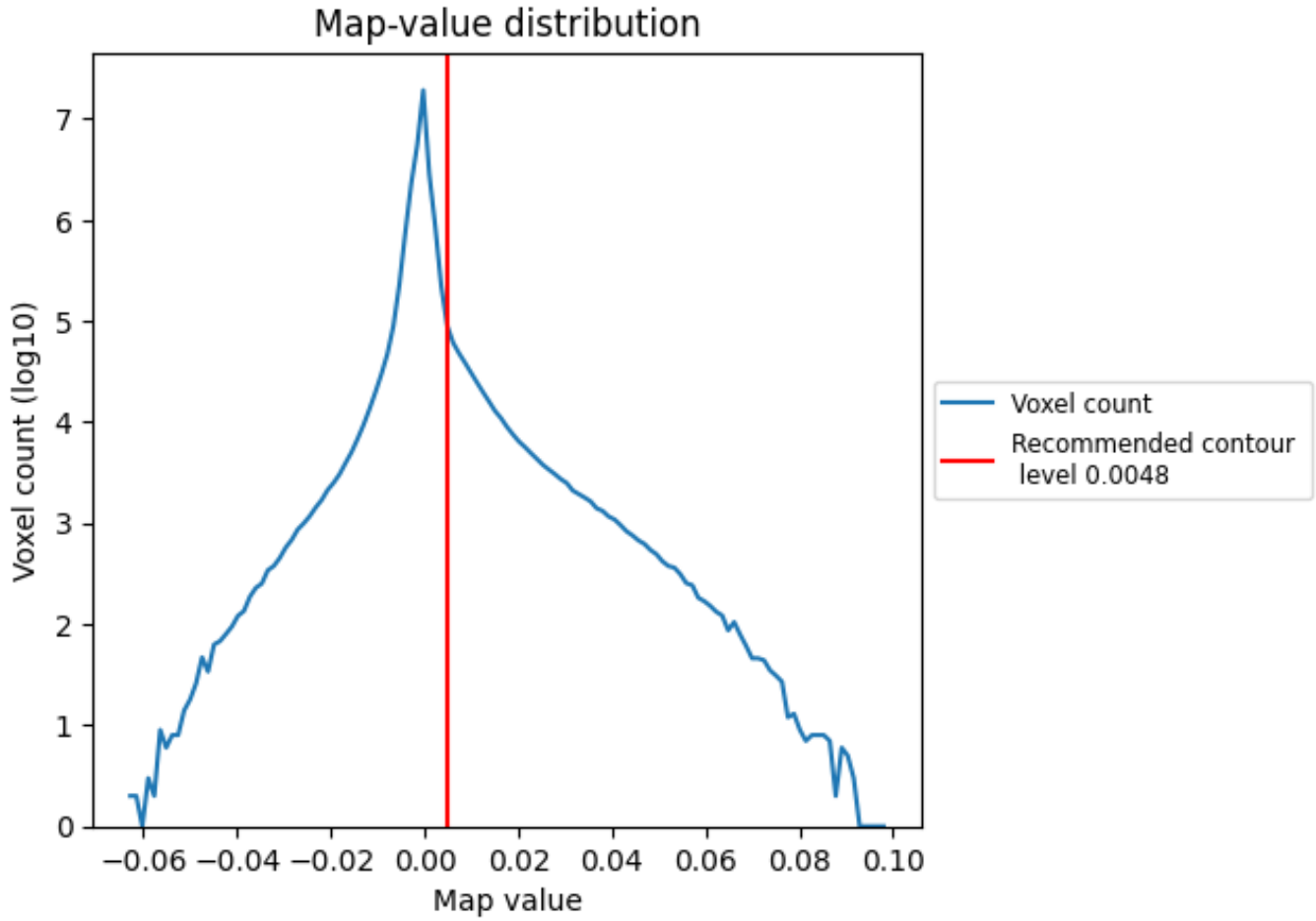
6.5 Mask visualisation (i)

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

7 Map analysis

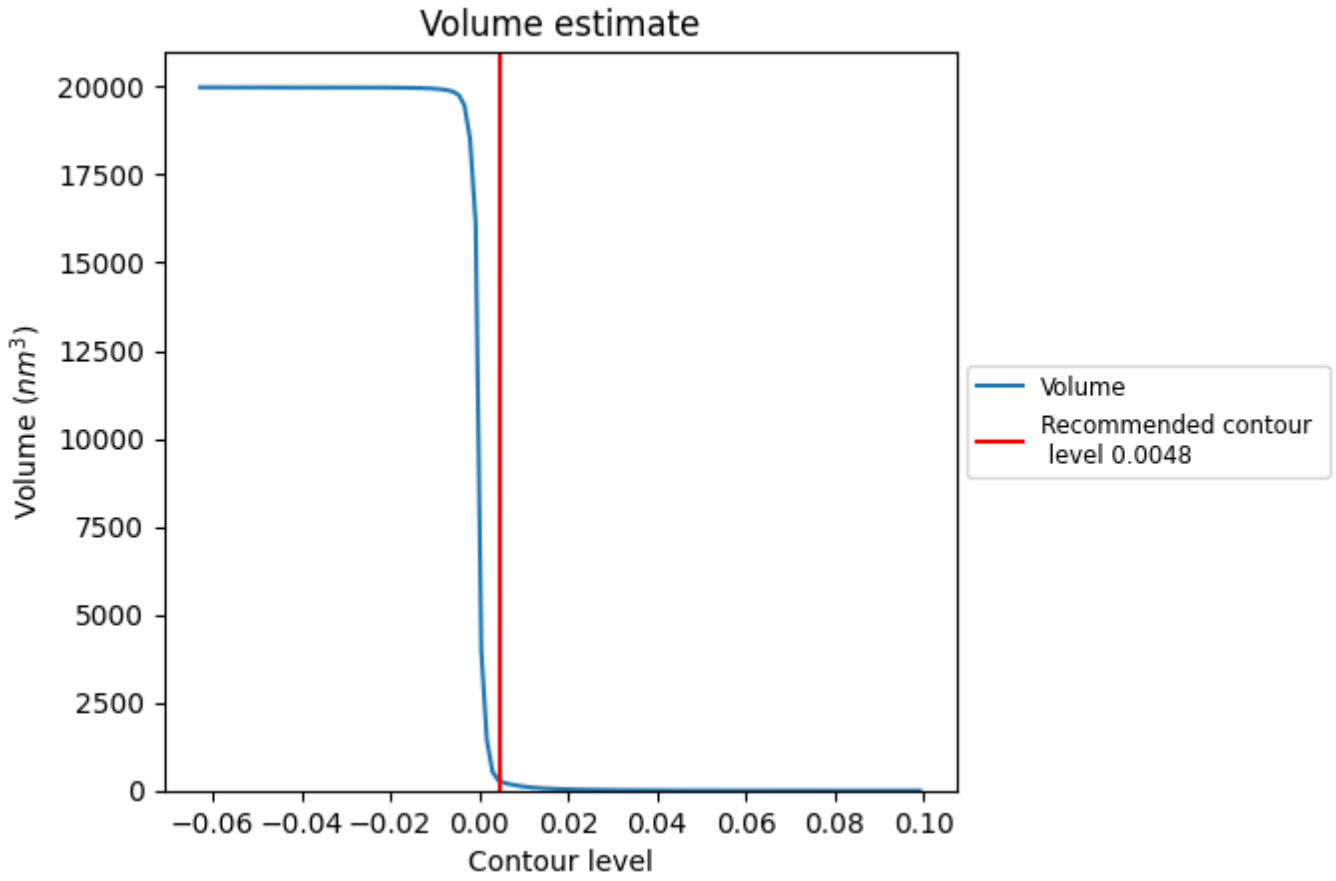
This section contains the results of statistical analysis of the map.

7.1 Map-value distribution



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.

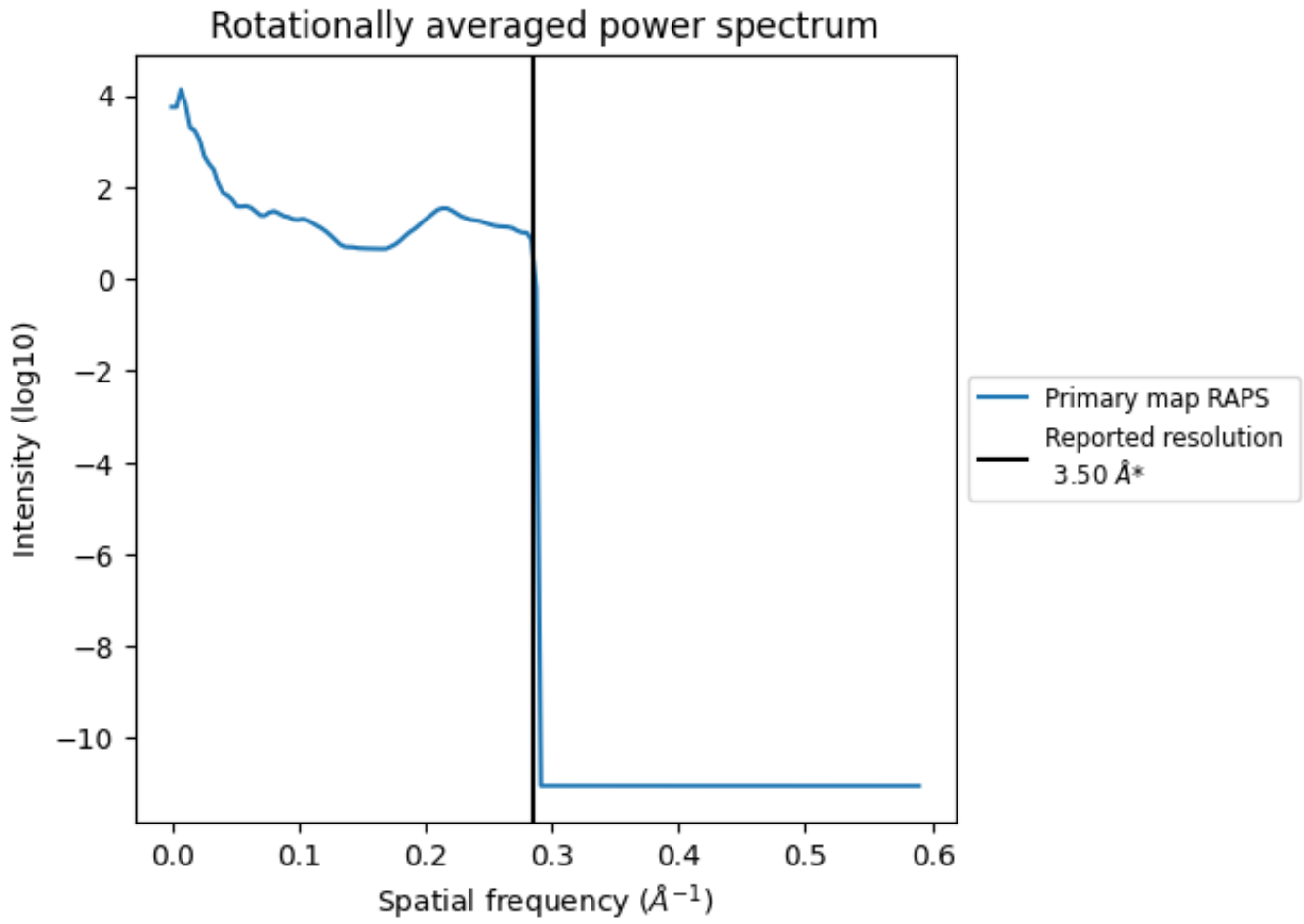
7.2 Volume estimate (i)



The volume at the recommended contour level is 269 nm^3; this corresponds to an approximate mass of 243 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 0.286 Å^{-1}

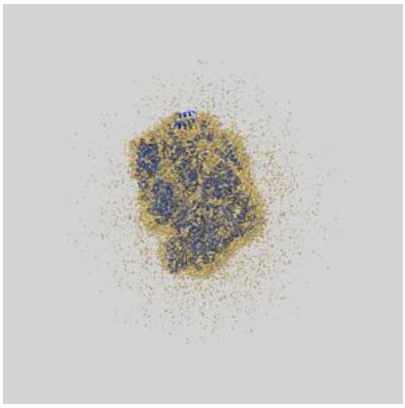
8 Fourier-Shell correlation

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

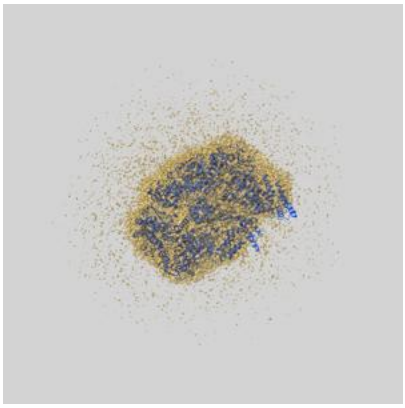
9 Map-model fit

This section contains information regarding the fit between EMDB map EMD-20659 and PDB model 6U5Z. Per-residue inclusion information can be found in section 3 on page 4.

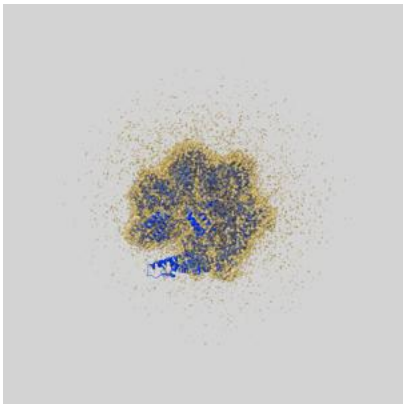
9.1 Map-model overlay



X



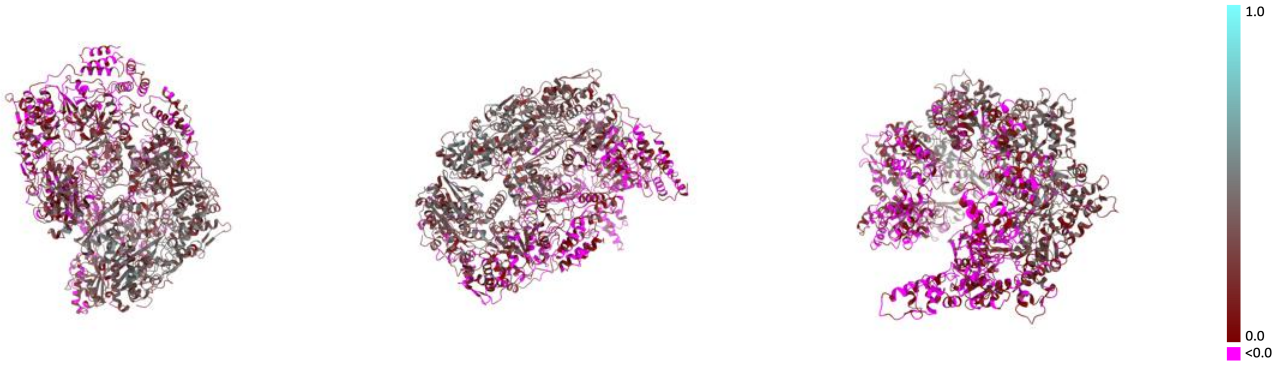
Y



Z

The images above show the 3D surface view of the map at the recommended contour level 0.0048 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.

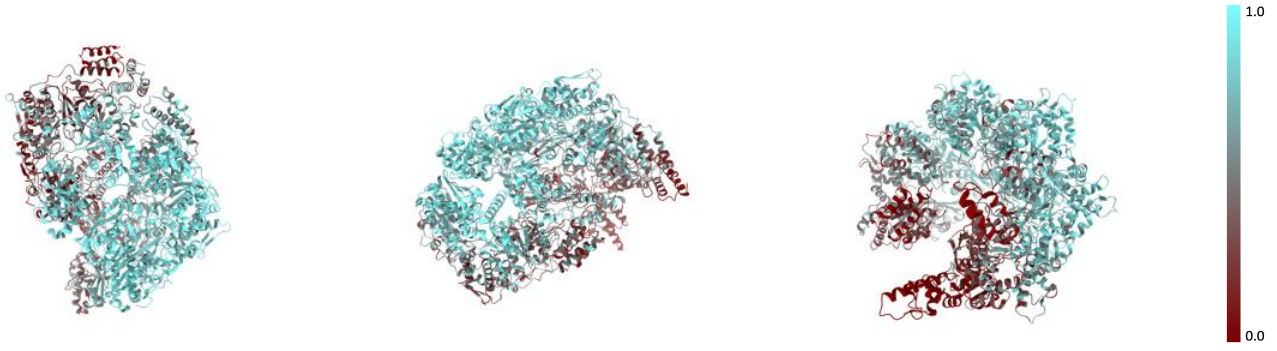
9.2 Q-score mapped to coordinate model (i)



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.

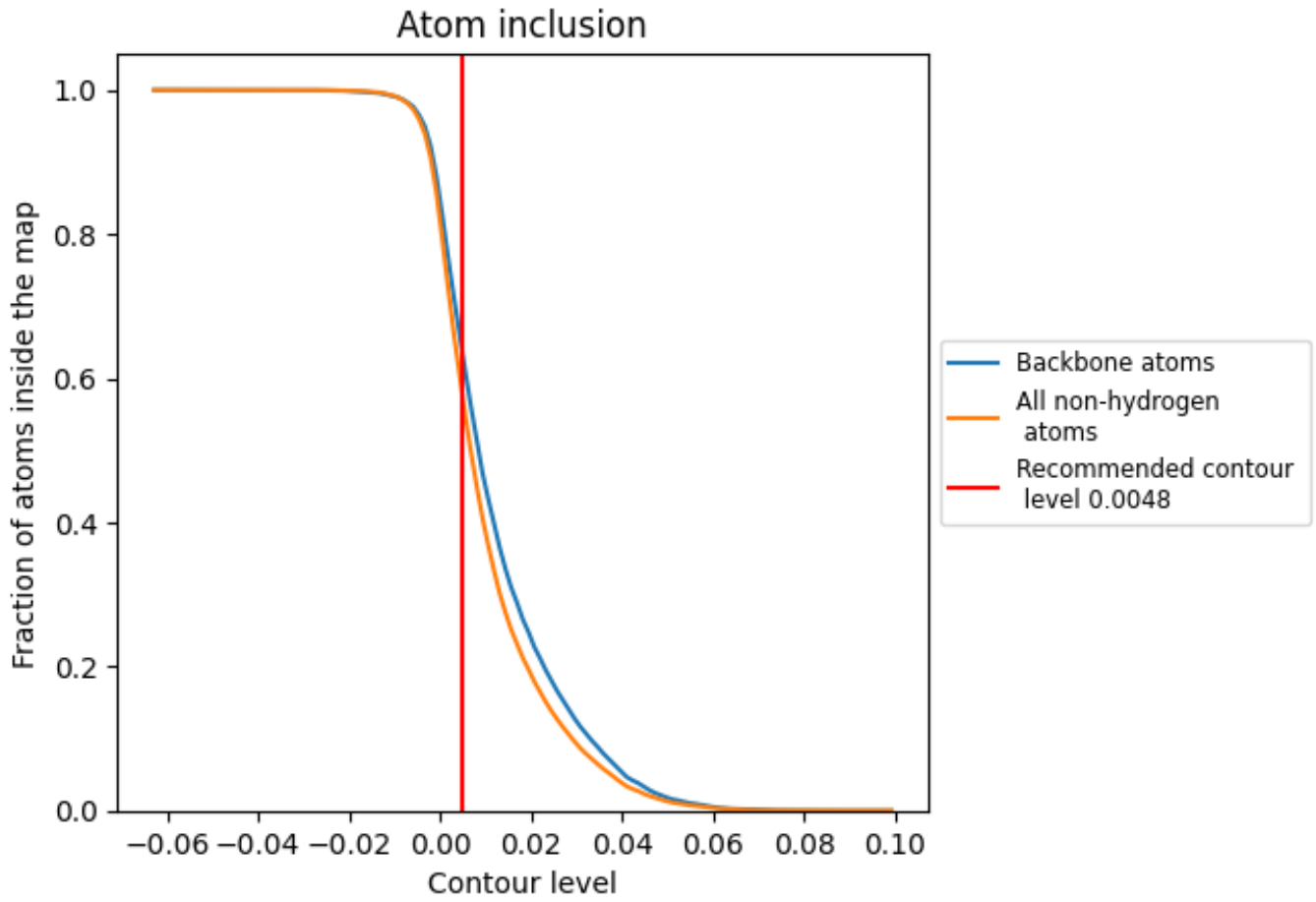
9.3 Atom inclusion mapped to coordinate model (i)



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 (0.0048).

9.4 Atom inclusion (i)



At the recommended contour level, 64% of all backbone atoms, 58% 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 (0.0048) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	0.5831	0.1830
A	0.3623	0.0810
B	0.7303	0.2250
C	0.7891	0.2970
D	0.7916	0.2900
E	0.6472	0.1790
F	0.1780	0.0250