wwPDB EM Validation Summary Report

Nov 20, 2022 – 06:08 PM JST

PDB ID	:	7CG3
EMDB ID	:	EMD-30349
Title	:	Staggered ring conformation of CtHsp104 (Hsp104 from Chaetomium Thermophilum)
Authors	:	Inoue, Y.; Hanazono, Y.; Noi, K.; Kawamoto, A.; Kimatsuka, M.; Harada, R.; Takeda, K.; Iwamasa, N.; Shibata, K.; Noguchi, K.; Shigeta, Y.; Namba, K.; Ogura, T.; Miki, K.; Shinohara, K.; Yohda, M.
Deposited on	:	2020-06-30
Resolution	:	5.10 Å(reported)
Based on initial model	:	5ZUI

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:

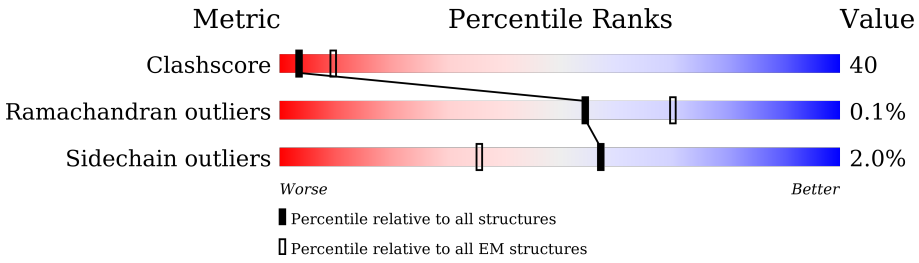
EMDB validation analysis	:	0.0.1.dev43
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
MapQ	:	1.9.9
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.31.3

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 5.10 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	764	8% 26% 48% • 26%
1	B	764	5% 23% 50% • 26%
1	C	764	11% 30% 58% • 10%
1	D	764	5% 33% 55% • 10%
1	E	764	10% 32% 56% • 10%
1	F	764	17% 31% 43% • 26%

2 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 29520 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 Heat shock protein 104.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	F	568	Total	C	N	O	S	0	0
			4441	2793	803	826	19		
1	A	568	Total	C	N	O	S	0	0
			4441	2793	803	826	19		
1	B	568	Total	C	N	O	S	0	0
			4441	2793	803	826	19		
1	C	686	Total	C	N	O	S	0	0
			5399	3377	988	1013	21		
1	D	686	Total	C	N	O	S	0	0
			5399	3377	988	1013	21		
1	E	686	Total	C	N	O	S	0	0
			5399	3377	988	1013	21		

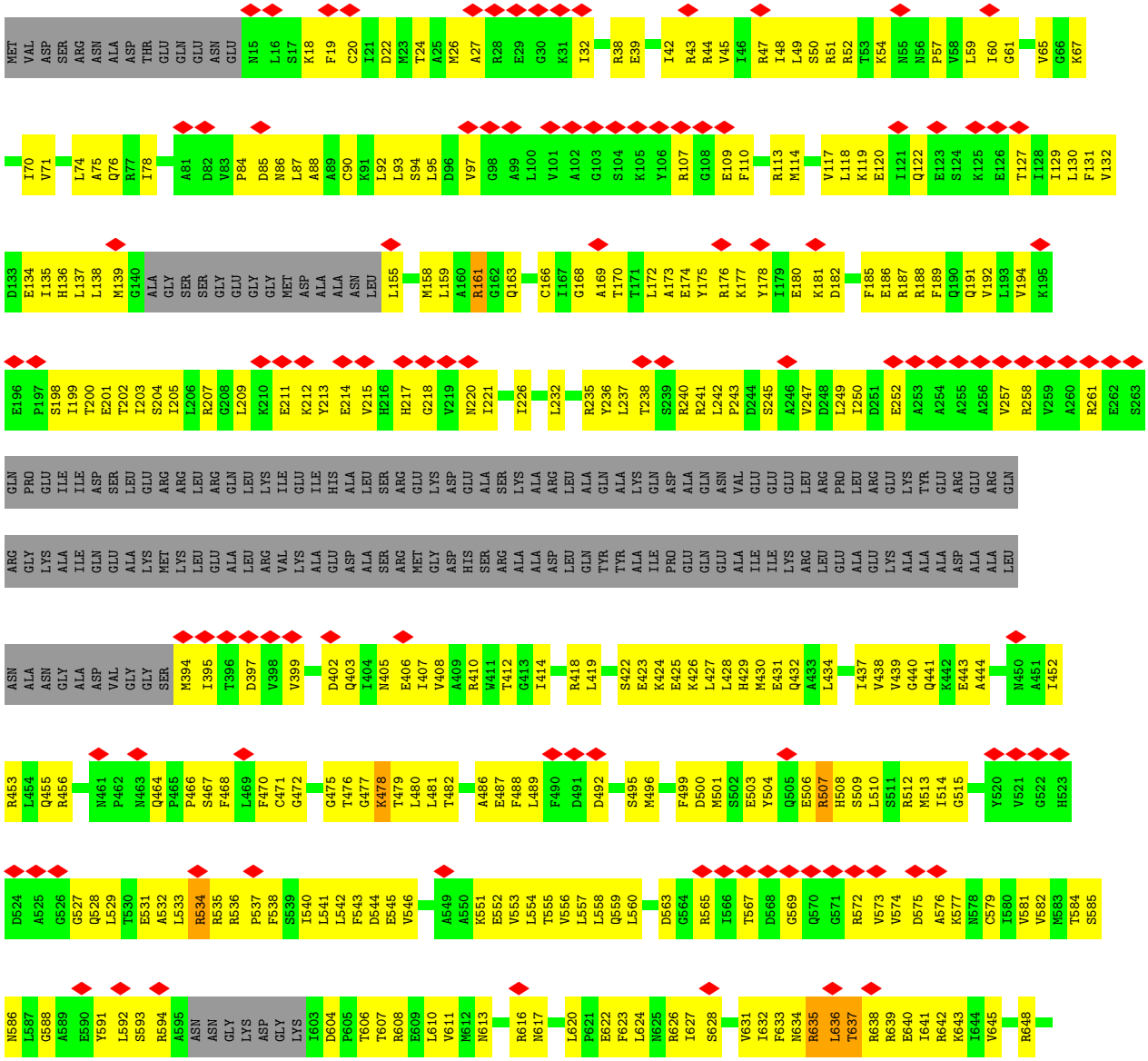
There are 6 discrepancies between the modelled and reference sequences:

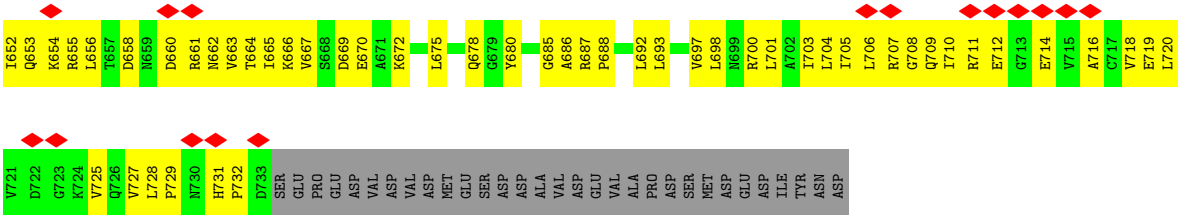
Chain	Residue	Modelled	Actual	Comment	Reference
F	1	MET	-	initiating methionine	UNP A0A2Z6G185
A	1	MET	-	initiating methionine	UNP A0A2Z6G185
B	1	MET	-	initiating methionine	UNP A0A2Z6G185
C	1	MET	-	initiating methionine	UNP A0A2Z6G185
D	1	MET	-	initiating methionine	UNP A0A2Z6G185
E	1	MET	-	initiating methionine	UNP A0A2Z6G185

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.

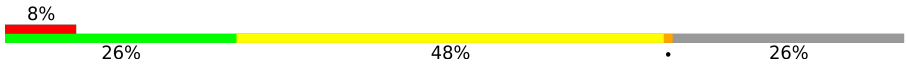
- Molecule 1: Heat shock protein 104

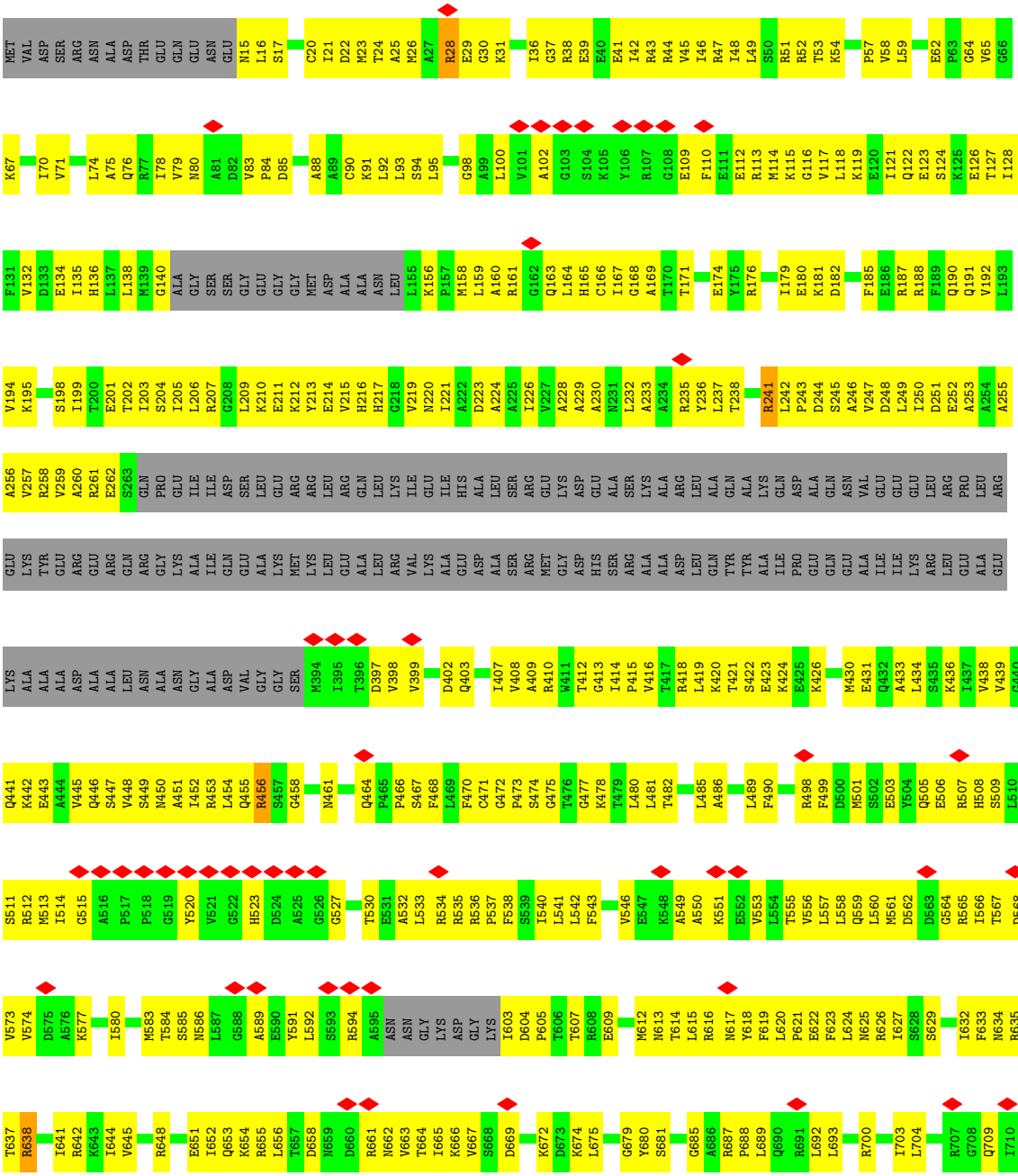


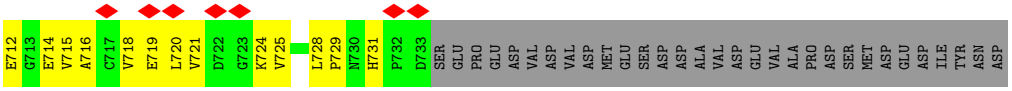


• Molecule 1: Heat shock protein 104

Chain A:

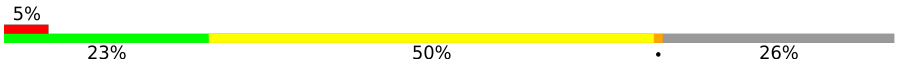


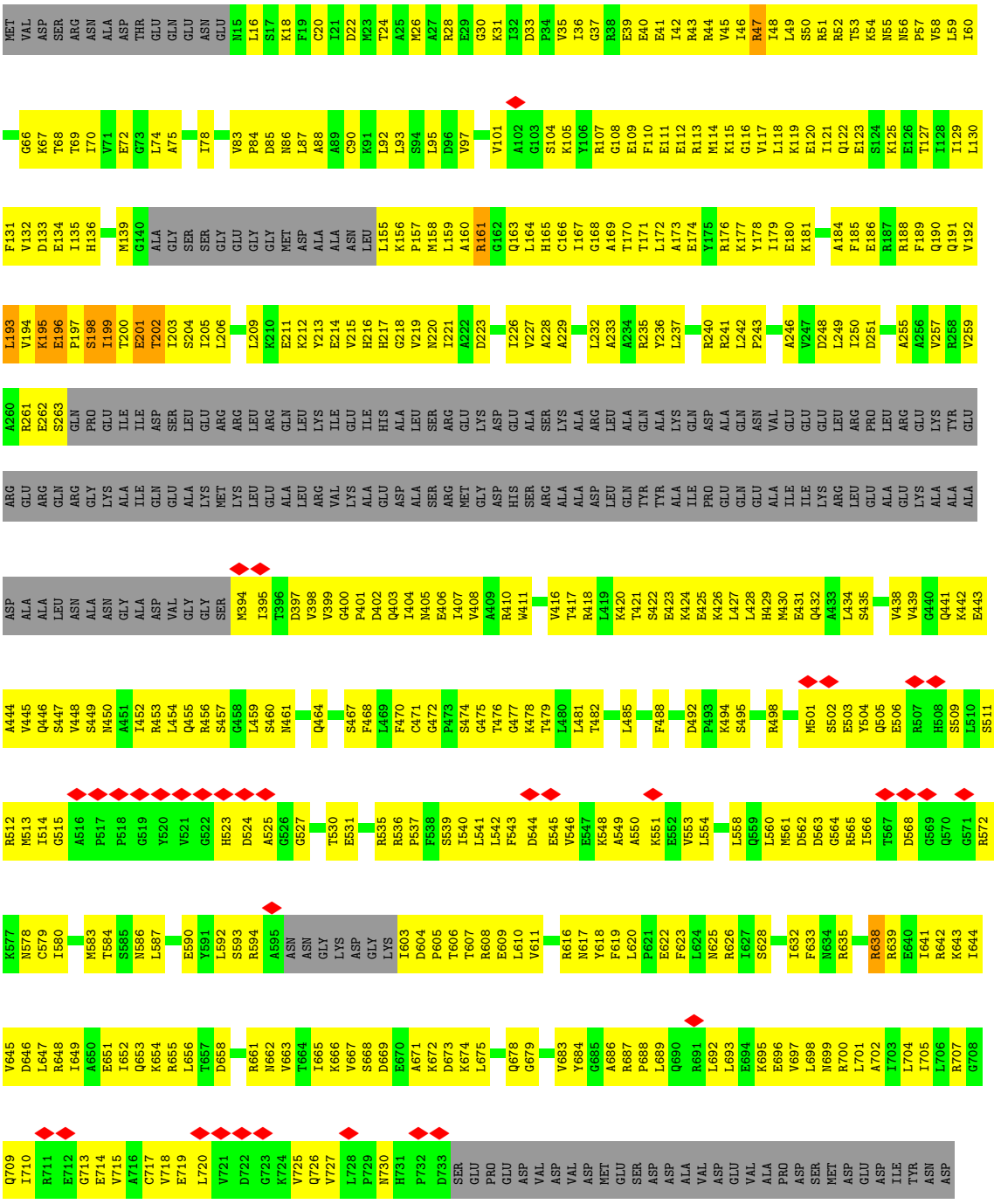




• Molecule 1: Heat shock protein 104

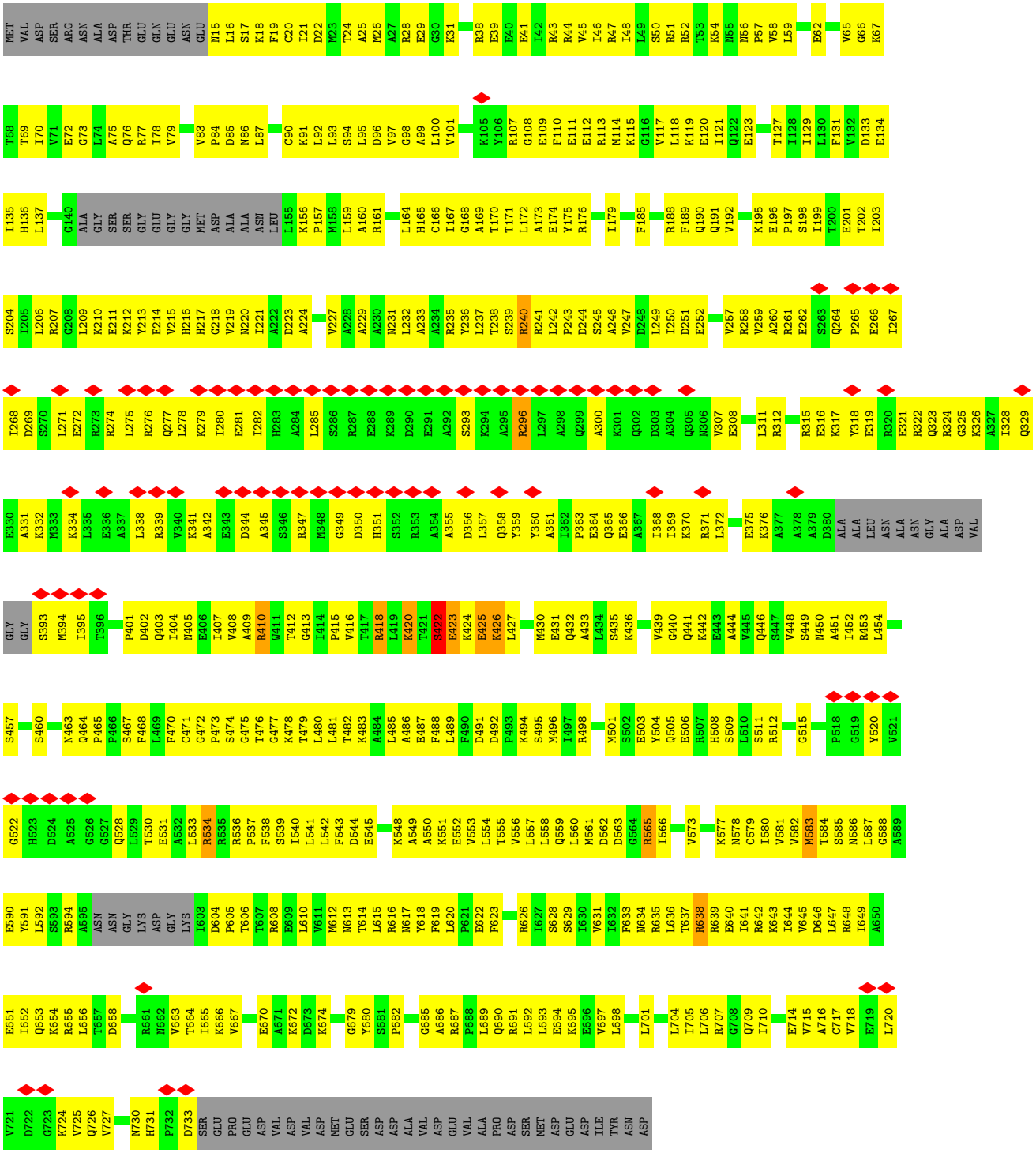
Chain B:





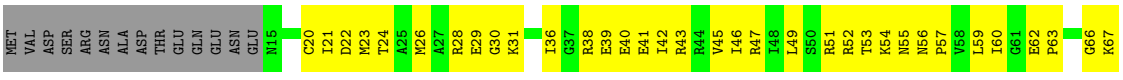
• Molecule 1: Heat shock protein 104

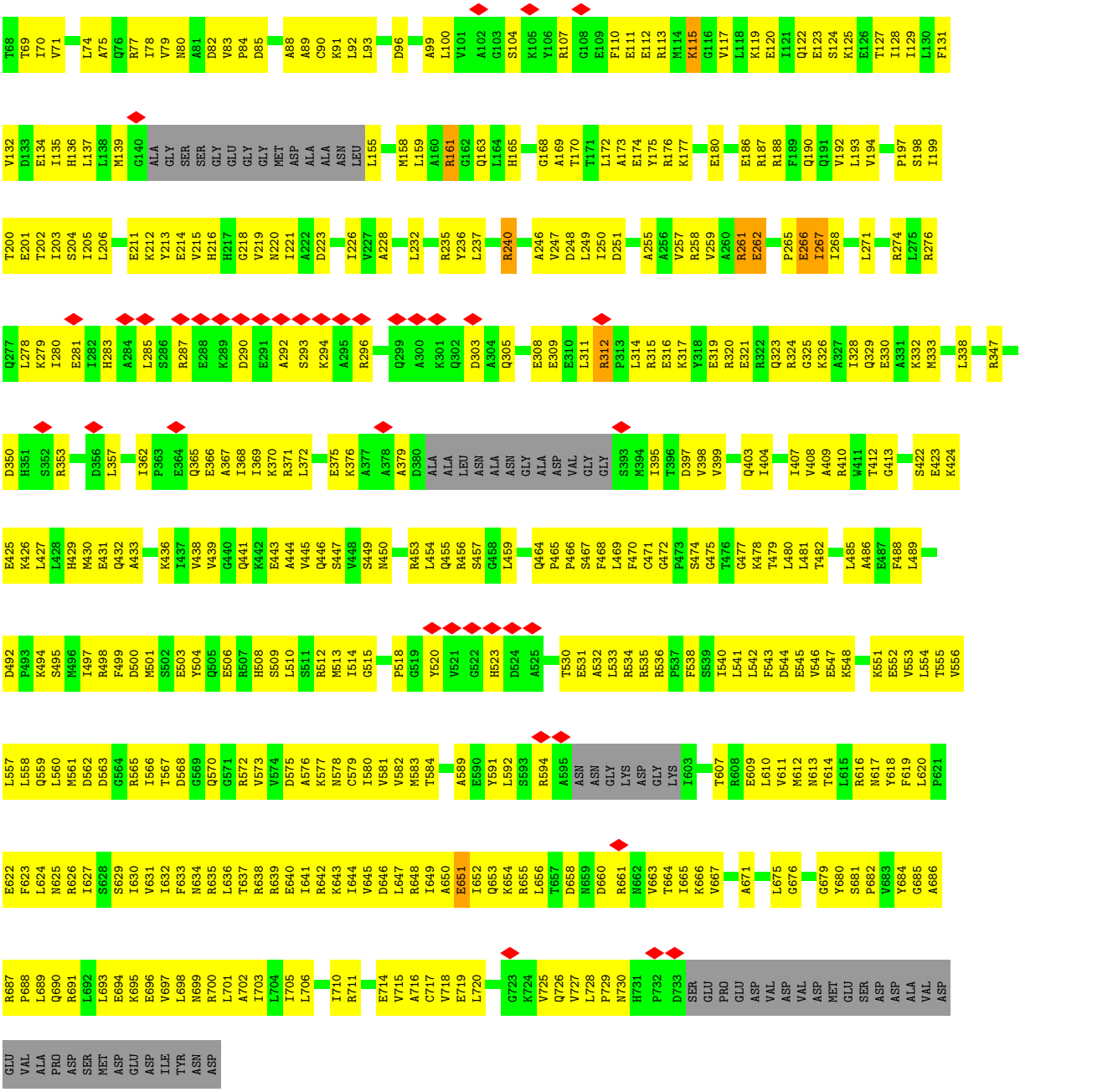
Chain C:



• Molecule 1: Heat shock protein 104

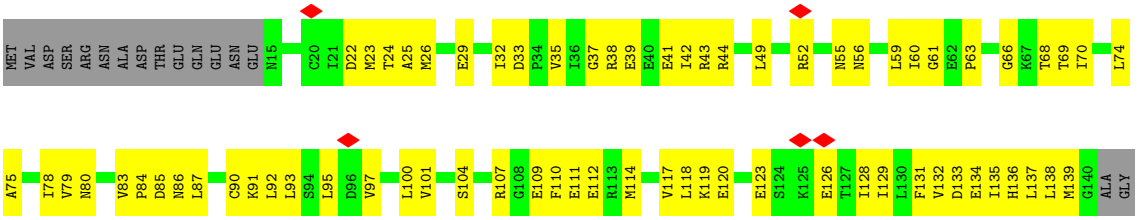
Chain D:

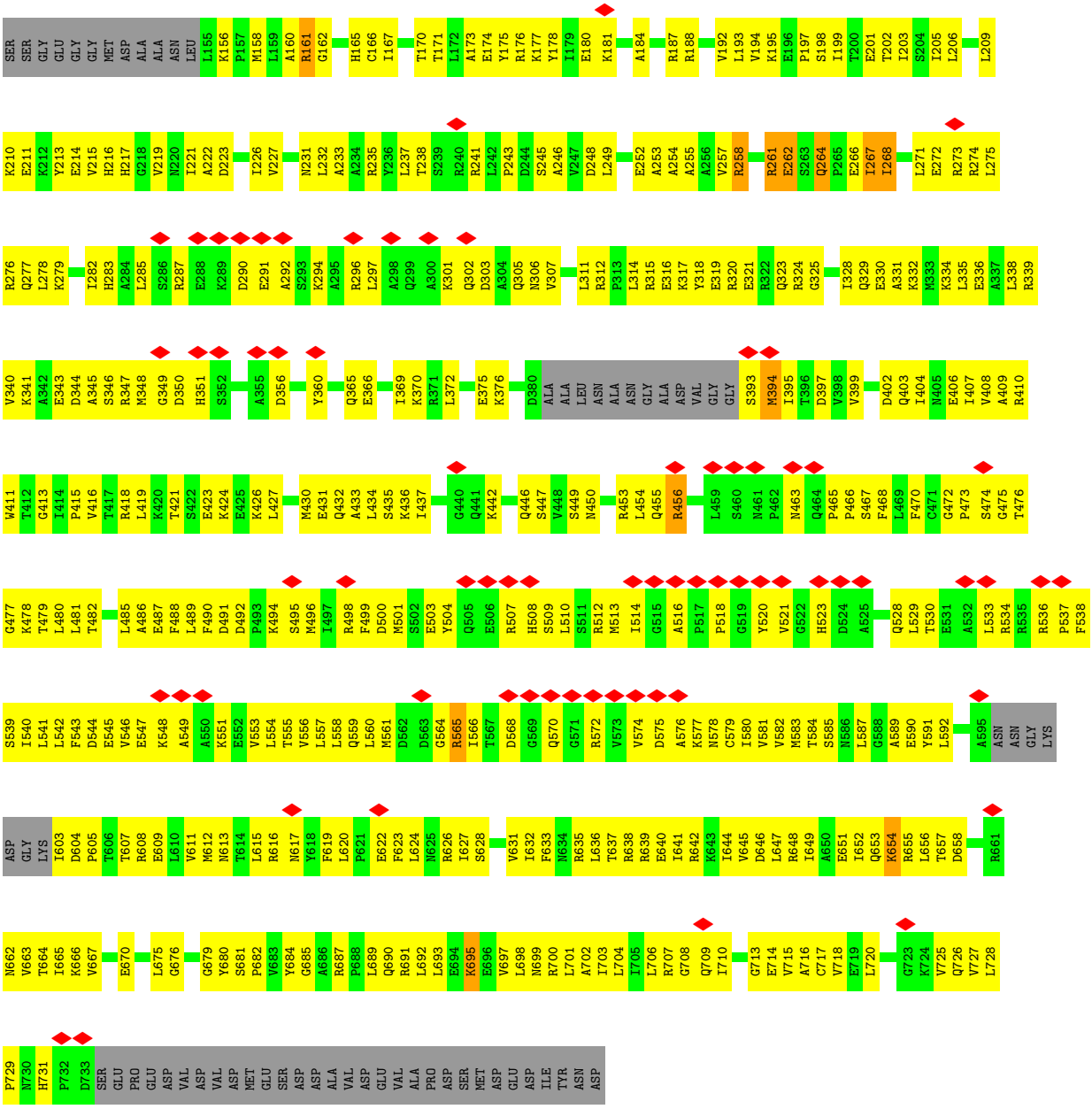




• Molecule 1: Heat shock protein 104

Chain E:




SER SER GLY GLU GLY GLY MET ASP ALA ALA ASN LEU L155 K156 P157 M158 L159 A160 R161 G162 H165 C166 I167 T170 T171 L172 A173 E174 Y175 R176 K177 Y178 I179 E180 K181 A184 R187 R188 V192 L193 V194 K195 E196 P197 S198 I199 T200 E201 T202 I203 S204 I205 L206 L209
K210 E211 K212 Y213 E214 V215 H216 H217 G218 V219 N220 I221 A222 D223 I226 V227 N231 L232 A233 A234 R235 Y236 L237 T238 S239 R240 R241 L242 P243 D244 S245 A246 Y247 D248 L249 E252 A253 A254 A255 A256 V257 R258 R261 E262 S263 Q264 P265 E266 I267 I268 L271 E272 R273 R274 L275
R276 Q277 L278 K279 I282 H283 A284 L285 S286 R287 E288 K289 D290 E291 A292 S293 K294 A295 R296 L297 A298 Q299 A300 K301 Q302 D303 A304 Q305 N306 V307 L311 R312 P313 L314 R315 E316 K317 Y318 E319 R320 E321 R322 Q323 R324 G325 I328 Q329 E330 A331 K332 M333 K334 L335 E336 A337 L338 R339
V340 K341 A342 E343 D344 A345 S346 R347 M348 G349 D350 H351 S352 A355 D356 Y360 Q365 E366 I369 K370 R371 L372 E375 K376 D380 ALA ALA LEU ASN ALA ASN GLY ALA ASP VAL GLY GLY S393 M394 I395 T396 D397 V398 V399 D402 Q403 I404 N405 E406 I407 V408 A409 R410
W411 T412 G413 I414 P415 V416 T417 R418 L419 K420 T421 S422 E423 K424 E425 K426 L427 M430 E431 Q432 A433 L434 S435 K436 I437 G440 Q441 K442 Q446 S447 V448 S449 N450 R453 L454 Q455 R456 L459 S460 N461 P462 N463 Q464 P465 P466 S467 F468 L469 F470 C471 G472 P473 S474 G475 T476
G477 K478 T479 L480 L481 T482 L485 A486 E487 F488 L489 F490 D491 D492 P493 K494 S495 M496 I497 R498 F499 D500 M501 S502 E503 Y504 Q505 E506 R507 H508 S509 L510 S511 R512 M513 I514 G515 A516 P517 P518 G519 Y520 V521 G522 H523 D524 A525 Q528 L529 T530 E531 A532 L533 R534 R535 R536 P537 F538
S539 I540 L541 L542 F543 D544 E545 V546 E547 K548 A549 A550 K551 E552 V553 L554 T555 V556 L557 L558 Q559 L560 M561 D562 D563 G564 R565 I566 T567 D568 G569 Q570 G571 R572 V573 V574 D575 A576 K577 N578 C579 I580 V581 V582 M583 T584 S585 N586 L587 G588 A589 E590 Y591 L592 A595 ASN ASN GLY LYS
ASP GLY LYS I603 D604 P605 T606 T607 R608 E609 L610 V611 M612 N613 T614 L615 R616 N617 Y618 F619 L620 P621 E622 F623 L624 M625 R626 I627 S628 V631 I632 F633 M634 R635 L636 T637 R638 R639 E640 I641 R642 K643 I644 V645 D646 L647 R648 I649 A650 E651 I652 Q653 K654 R655 L656 T657 D658 R661
N662 V663 T664 I665 K666 V667 E670 L675 G676 G679 Y680 S681 P682 V683 Y684 G685 A686 R687 P688 L689 Q690 R691 L692 L693 E694 K695 E696 V697 L698 N699 R700 L701 A702 I703 L704 I705 L706 R707 G708 Q709 I710 G713 E714 V715 A716 C717 V718 E719 L720 G723 K724 V725 Q726 V727 L728
P729 N730 H731 P732 D733 SER GLU PRO GLU ASP VAL ASP VAL ASP MET GLU SER ASP ASP ALA VAL ASP GLU VAL ALA PRO ASP SER MET ASP GLU ASP ILE TYR ASN ASP

4 Experimental information (i)

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, C1	Depositor
Number of particles used	180636	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)	1000	Depositor
Maximum defocus (nm)	3000	Depositor
Magnification	75000	Depositor
Image detector	FEI FALCON III (4k x 4k)	Depositor
Maximum map value	0.056	Depositor
Minimum map value	-0.021	Depositor
Average map value	0.001	Depositor
Map value standard deviation	0.003	Depositor
Recommended contour level	0.0122	Depositor
Map size (Å)	253.44, 253.44, 253.44	wwPDB
Map dimensions	288, 288, 288	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	0.88, 0.88, 0.88	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	0.29	0/4496	0.46	0/6062
1	B	0.32	0/4496	0.48	0/6062
1	C	0.33	0/5462	0.47	0/7351
1	D	0.32	0/5462	0.48	0/7351
1	E	0.29	0/5462	0.45	0/7351
1	F	0.29	0/4496	0.48	0/6062
All	All	0.31	0/29874	0.47	0/40239

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
1	A	0	1

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	241	ARG	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	4441	0	4613	356	0
1	B	4441	0	4613	411	0
1	C	5399	0	5590	485	0
1	D	5399	0	5590	420	0
1	E	5399	0	5588	468	0
1	F	4441	0	4613	330	0
All	All	29520	0	30607	2412	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 40.

The worst 5 of 2412 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:C:426:LYS:HB2	1:C:488:PHE:CZ	1.49	1.47
1:C:426:LYS:NZ	1:C:489:LEU:HA	1.31	1.40
1:C:426:LYS:HE2	1:C:488:PHE:CE2	1.60	1.36
1:E:217:HIS:HA	1:E:258:ARG:NH2	1.42	1.32
1:C:426:LYS:CE	1:C:488:PHE:CE2	2.15	1.28

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	560/764 (73%)	508 (91%)	52 (9%)	0	100	100
1	B	560/764 (73%)	503 (90%)	56 (10%)	1 (0%)	47	81
1	C	678/764 (89%)	626 (92%)	50 (7%)	2 (0%)	41	76
1	D	678/764 (89%)	611 (90%)	67 (10%)	0	100	100
1	E	678/764 (89%)	628 (93%)	50 (7%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
1	F	560/764 (73%)	513 (92%)	47 (8%)	0	100 100
All	All	3714/4584 (81%)	3389 (91%)	322 (9%)	3 (0%)	54 85

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	423	GLU
1	B	476	THR
1	C	422	SER

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	484/640 (76%)	480 (99%)	4 (1%)	81 89
1	B	484/640 (76%)	471 (97%)	13 (3%)	44 65
1	C	581/640 (91%)	567 (98%)	14 (2%)	49 69
1	D	581/640 (91%)	570 (98%)	11 (2%)	57 75
1	E	581/640 (91%)	565 (97%)	16 (3%)	43 64
1	F	484/640 (76%)	477 (99%)	7 (1%)	67 81
All	All	3195/3840 (83%)	3130 (98%)	65 (2%)	57 73

5 of 65 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	E	268	ILE
1	E	394	MET
1	C	195	LYS
1	B	638	ARG
1	E	456	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 47 such sidechains are listed below:

Mol	Chain	Res	Type
1	D	217	HIS
1	E	217	HIS
1	D	302	GLN
1	D	528	GLN
1	E	299	GLN

5.3.3 RNA

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates

There are no monosaccharides in this entry.

5.6 Ligand geometry

There are no ligands in this entry.

5.7 Other polymers

There are no such residues in this entry.

5.8 Polymer linkage issues

There are no chain breaks in this entry.

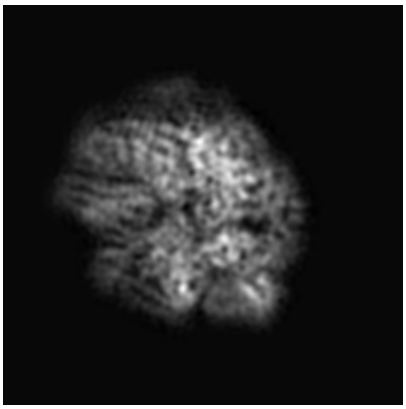
6 Map visualisation

This section contains visualisations of the EMDB entry EMD-30349. These allow visual inspection of the internal detail of the map and identification of artifacts.

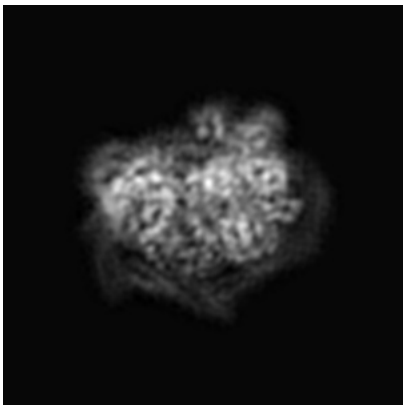
Images derived from a raw map, generated by summing the deposited half-maps, are presented below the corresponding image components of the primary map to allow further visual inspection and comparison with those of the primary map.

6.1 Orthogonal projections

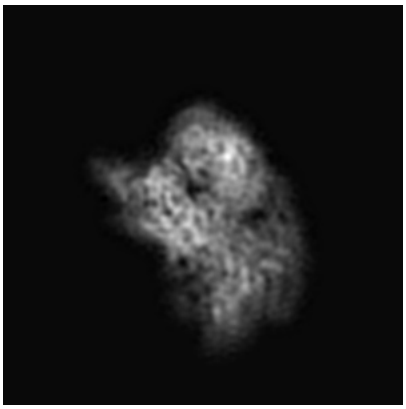
6.1.1 Primary map



X

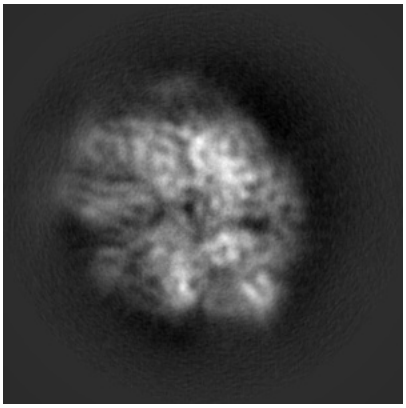


Y

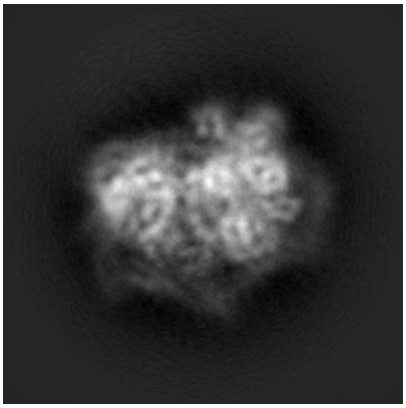


Z

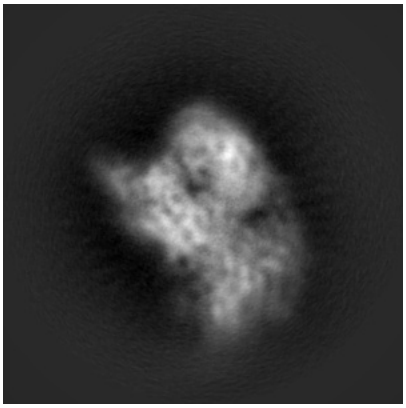
6.1.2 Raw map



X



Y

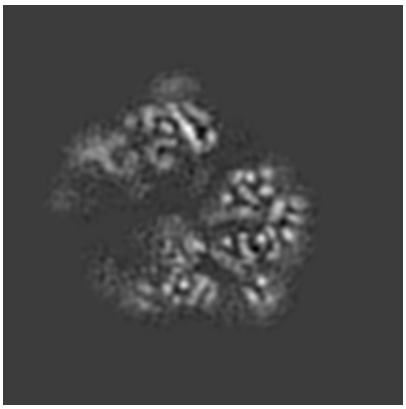


Z

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

6.2 Central slices (i)

6.2.1 Primary map

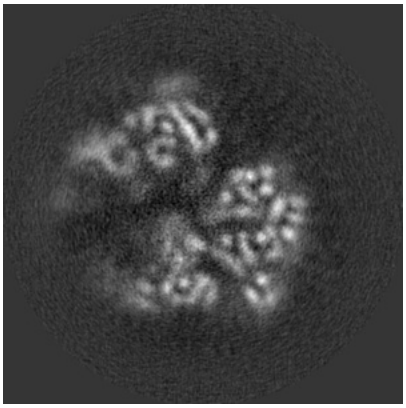


X Index: 144

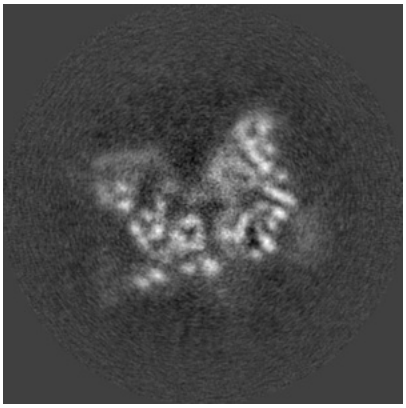
Y Index: 144

Z Index: 144

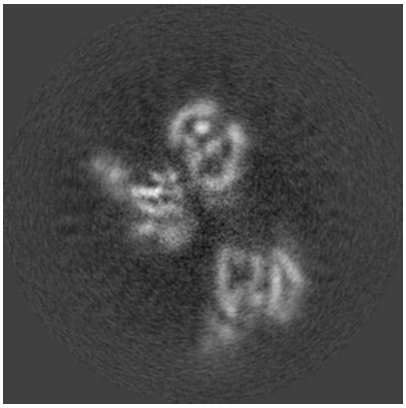
6.2.2 Raw map



X Index: 144



Y Index: 144



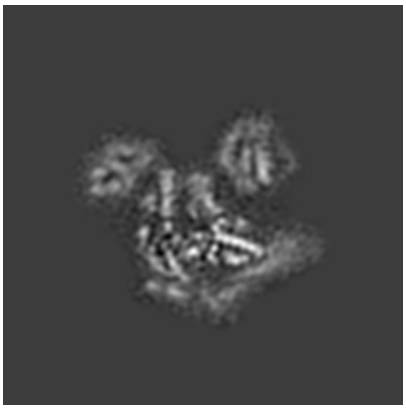
Z Index: 144

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

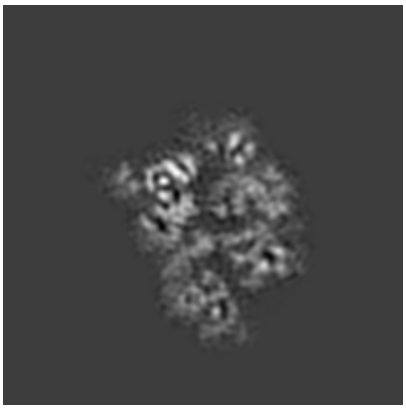
6.3 Largest variance slices (i)

6.3.1 Primary map

X Index: 163

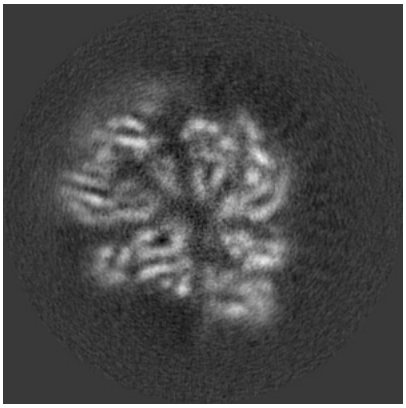


Y Index: 160

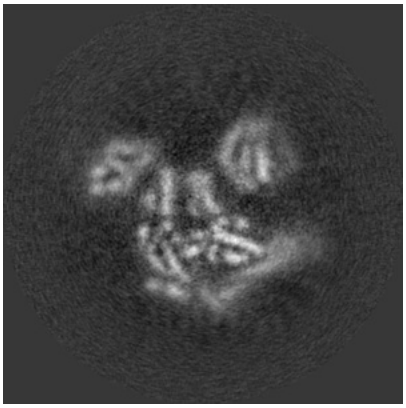


Z Index: 173

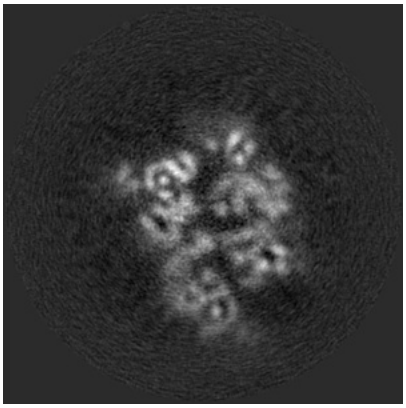
6.3.2 Raw map



X Index: 163



Y Index: 160

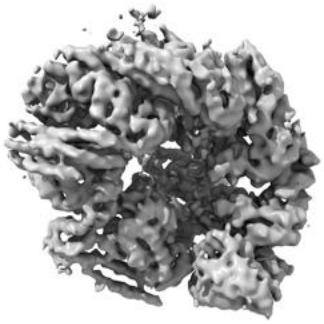


Z Index: 173

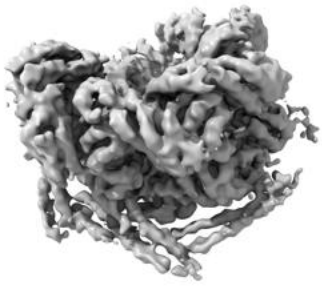
The images above show the largest variance slices of the map in three orthogonal directions.

6.4 Orthogonal surface views (i)

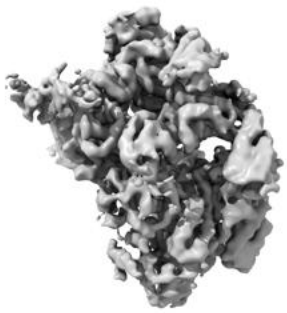
6.4.1 Primary map



X



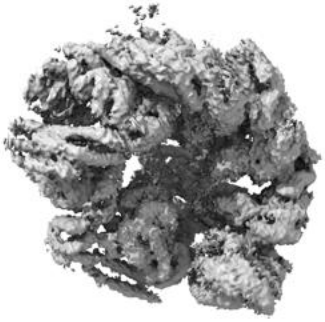
Y



Z

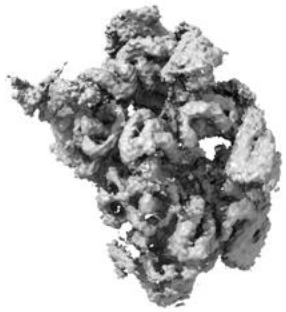
The images above show the 3D surface view of the map at the recommended contour level 0.0122. These images, in conjunction with the slice images, may facilitate assessment of whether an appropriate contour level has been provided.

6.4.2 Raw map



X

Y



Z

These images show the 3D surface of the raw map. The raw map's contour level was selected so that its surface encloses the same volume as the primary map does at its recommended contour level.

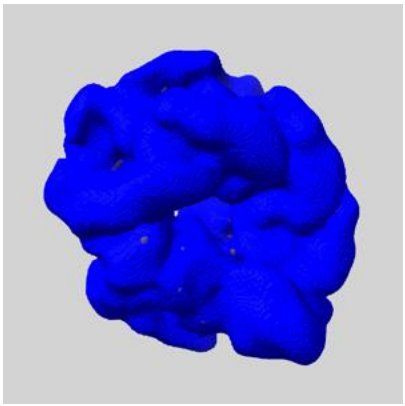
6.5 Mask visualisation (i)

This section shows the 3D surface view of the primary map at 50% transparency overlaid with the specified mask at 0% transparency

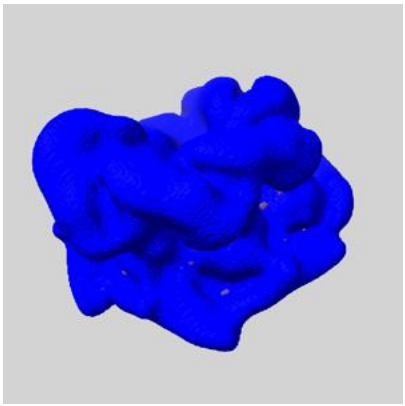
A mask typically either:

- Encompasses the whole structure
- Separates out a domain, a functional unit, a monomer or an area of interest from a larger structure

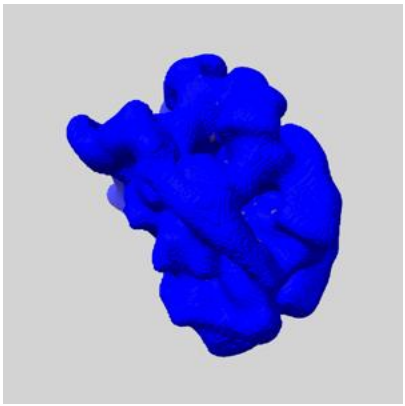
6.5.1 emd_30349_msk_1.map (i)



X



Y

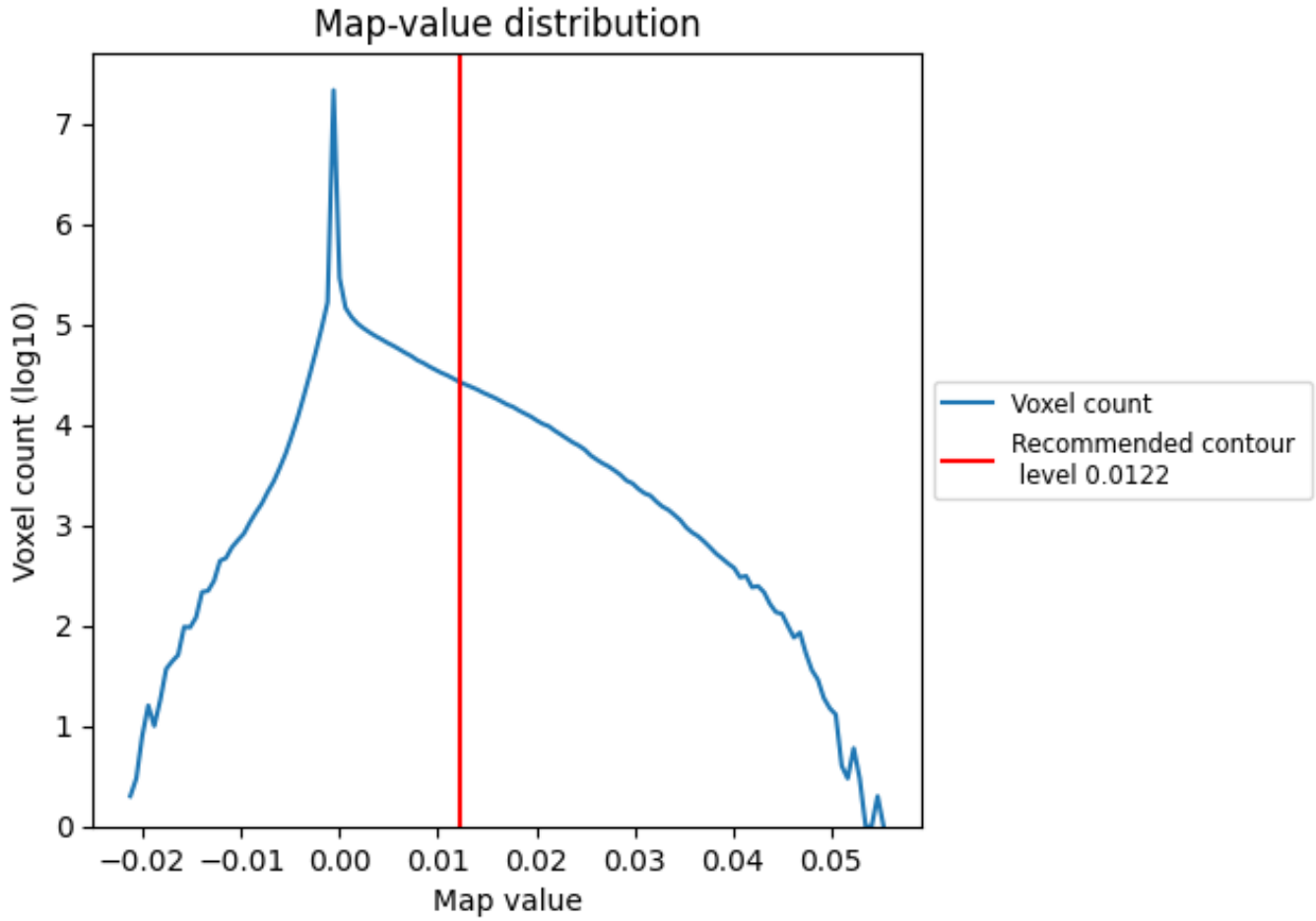


Z

7 Map analysis

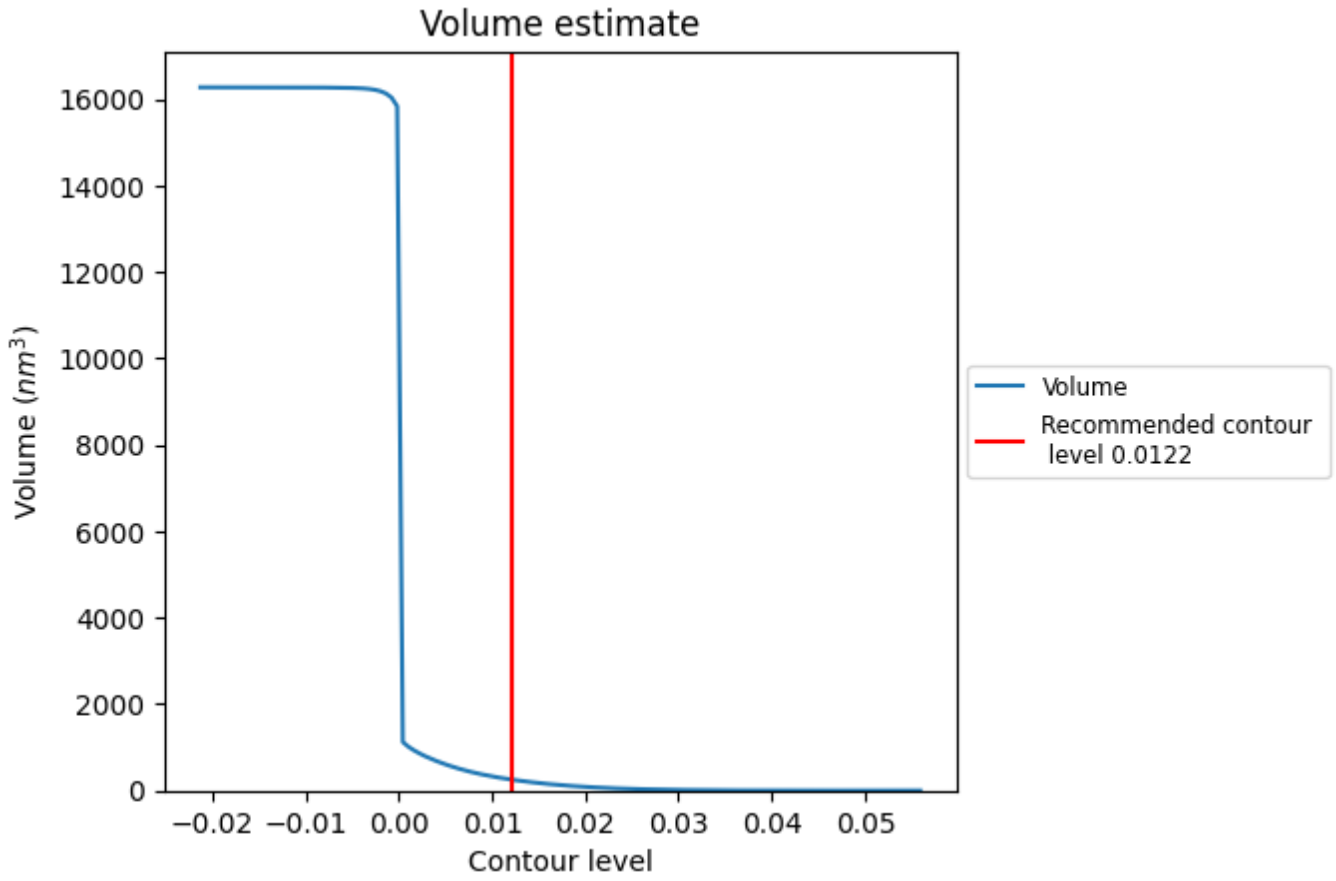
This section contains the results of statistical analysis of the map.

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.

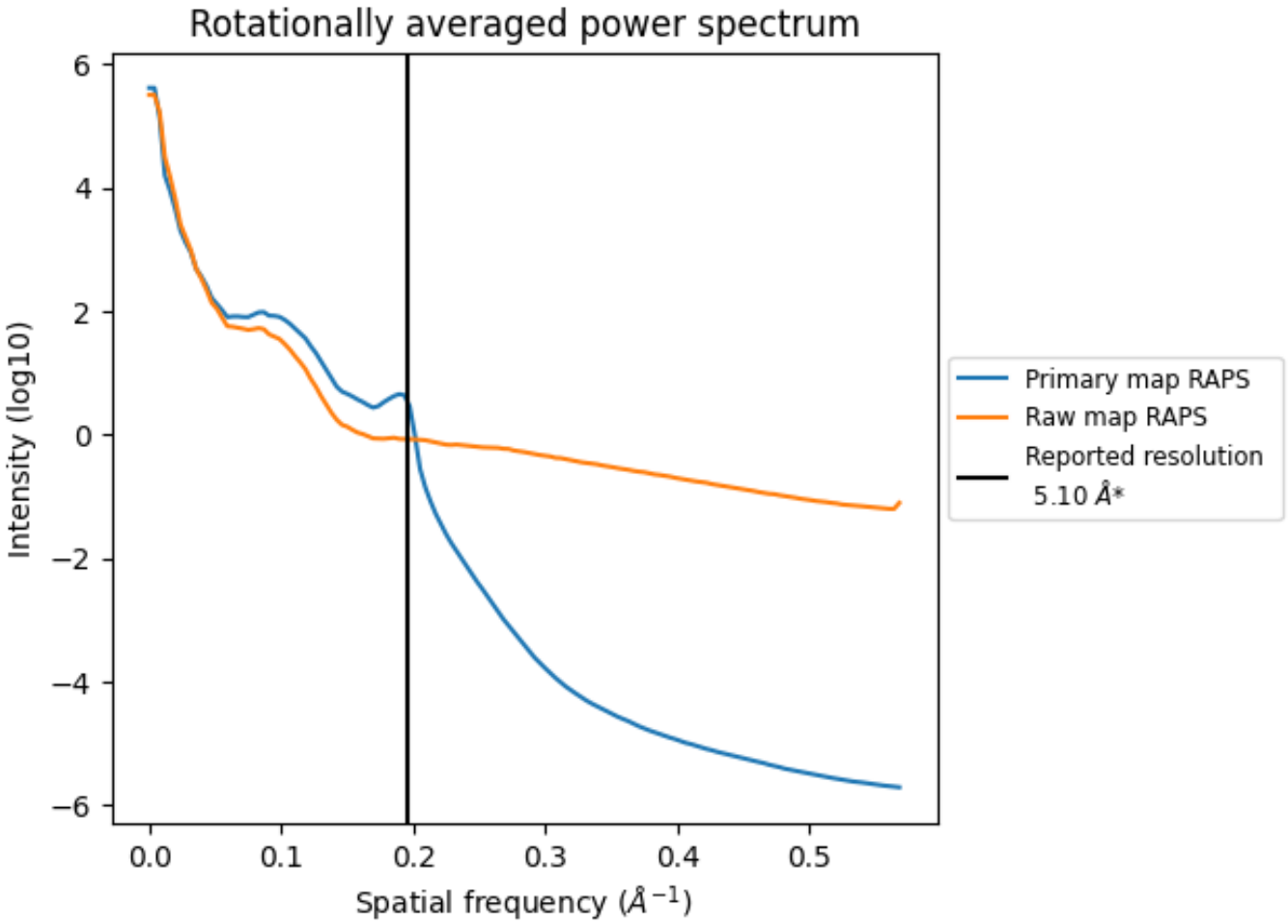
7.2 Volume estimate (i)



The volume at the recommended contour level is 251 nm^3; this corresponds to an approximate mass of 226 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)

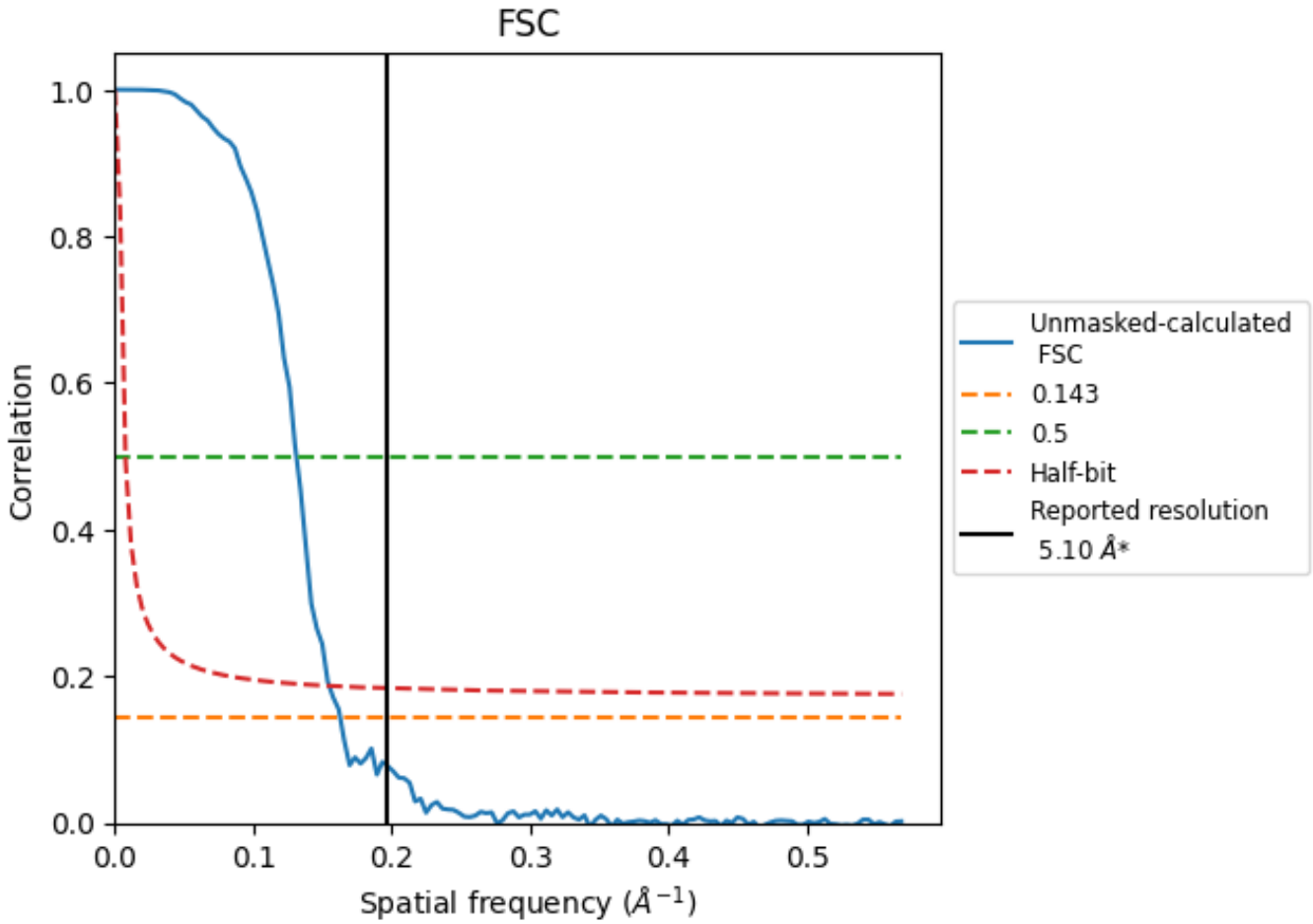


*Reported resolution corresponds to spatial frequency of 0.196 Å^{-1}

8 Fourier-Shell correlation

Fourier-Shell Correlation (FSC) is the most commonly used method to estimate the resolution of single-particle and subtomogram-averaged maps. The shape of the curve depends on the imposed symmetry, mask and whether or not the two 3D reconstructions used were processed from a common reference. The reported resolution is shown as a black line. A curve is displayed for the half-bit criterion in addition to lines showing the 0.143 gold standard cut-off and 0.5 cut-off.

8.1 FSC



*Reported resolution corresponds to spatial frequency of 0.196 Å^{-1}

8.2 Resolution estimates (i)

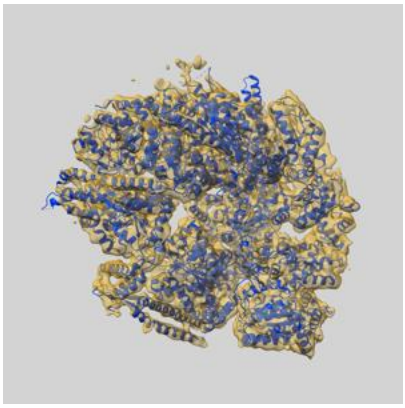
Resolution estimate (Å)	Estimation criterion (FSC cut-off)		
	0.143	0.5	Half-bit
Reported by author	5.10	-	-
Author-provided FSC curve	-	-	-
Unmasked-calculated*	6.14	7.62	6.46

*Resolution estimate based on FSC curve calculated by comparison of deposited half-maps. The value from deposited half-maps intersecting FSC 0.143 CUT-OFF 6.14 differs from the reported value 5.1 by more than 10 %

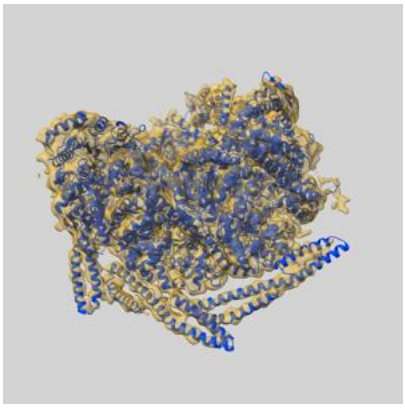
9 Map-model fit

This section contains information regarding the fit between EMDB map EMD-30349 and PDB model 7CG3. Per-residue inclusion information can be found in section 3 on page 4.

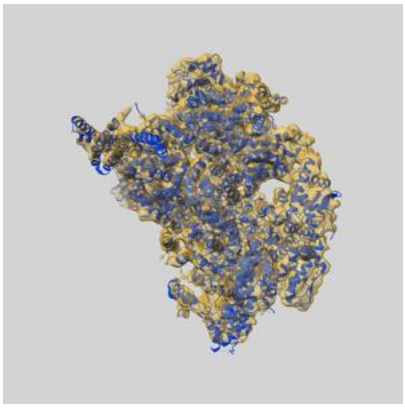
9.1 Map-model overlay



X



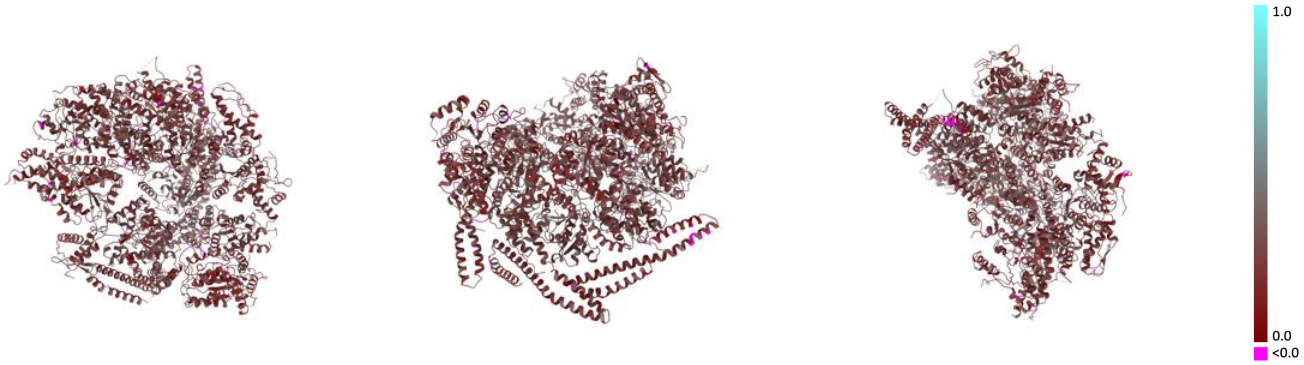
Y



Z

The images above show the 3D surface view of the map at the recommended contour level 0.0122 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.

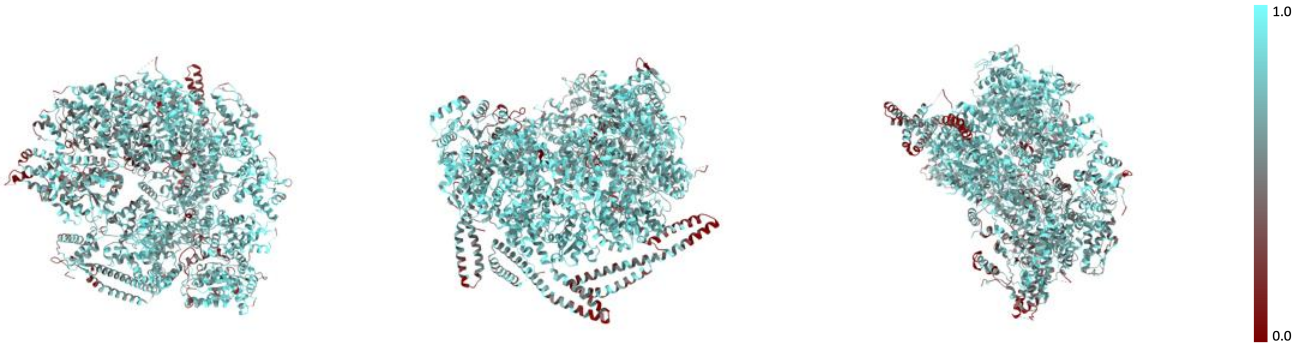
9.2 Q-score mapped to coordinate model (i)



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.

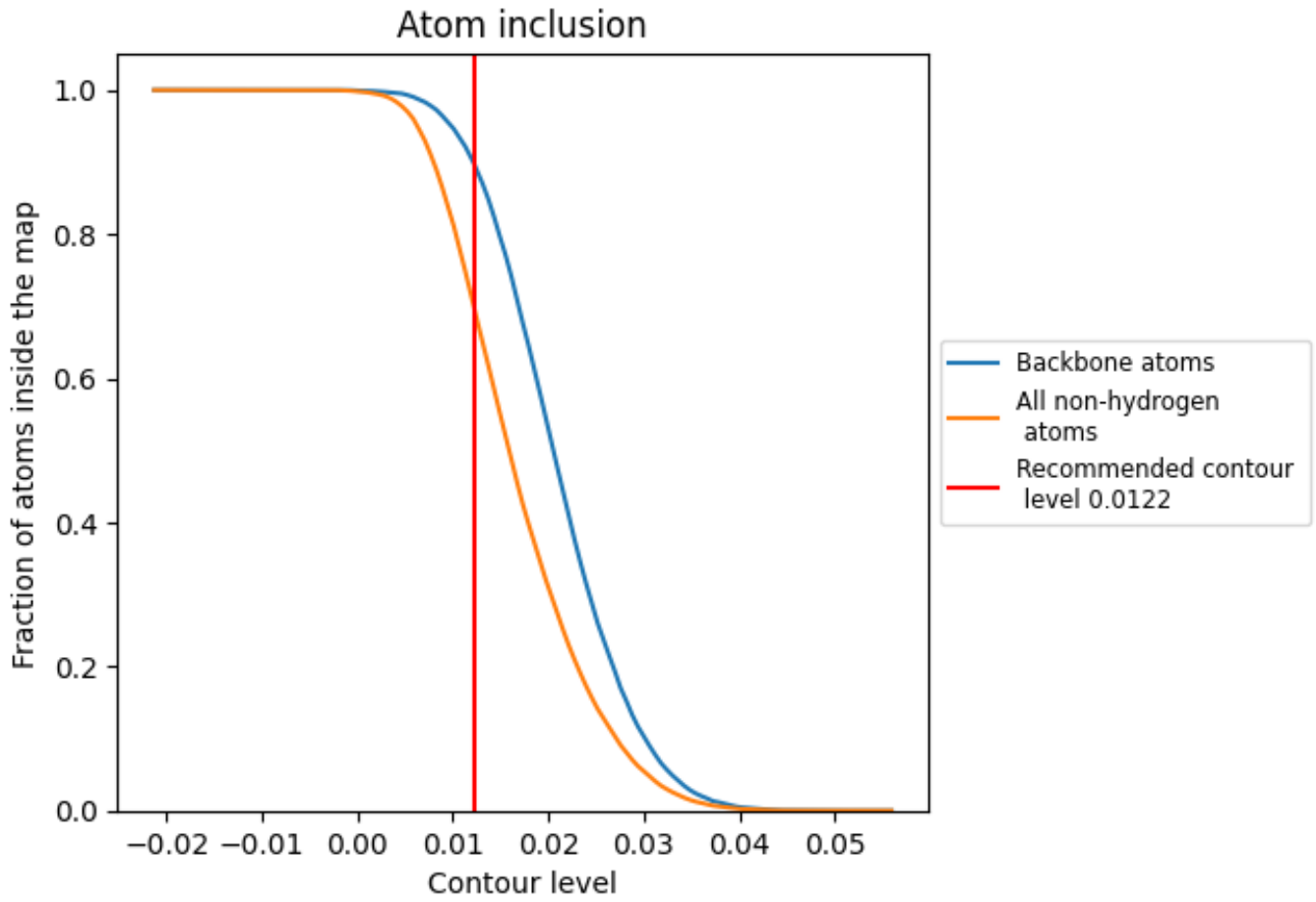
9.3 Atom inclusion mapped to coordinate model (i)



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 (0.0122).

9.4 Atom inclusion (i)



At the recommended contour level, 90% of all backbone atoms, 70% 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 (0.0122) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	0.6984	0.2640
A	0.6684	0.2680
B	0.7369	0.2730
C	0.7266	0.2650
D	0.7580	0.2750
E	0.6966	0.2490
F	0.5857	0.2570