



wwPDB EM Validation Summary Report ⓘ

Nov 29, 2022 – 03:42 AM EST

PDB ID : 7UT4
EMDB ID : EMD-26754
Title : Gea2 closed/closed conformation (composite structure)
Authors : Muccini, A.; Fromme, J.C.
Deposited on : 2022-04-26
Resolution : 3.90 Å (reported)

This is a wwPDB EM Validation Summary Report for a publicly released PDB entry.

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

A user guide is available at

<https://www.wwpdb.org/validation/2017/EMValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

The types of validation reports are described at

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

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

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

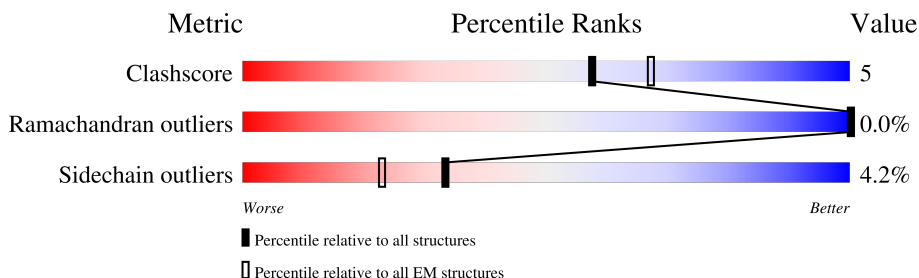
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

ELECTRON MICROSCOPY

The reported resolution of this entry is 3.90 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	EM structures (#Entries)
Clashscore	158937	4297
Ramachandran outliers	154571	4023
Sidechain outliers	154315	3826

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

Mol	Chain	Length	Quality of chain
1	A	1459	
1	B	1459	

2 Entry composition

There is only 1 type of molecule in this entry. The entry contains 39740 atoms, of which 20092 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.

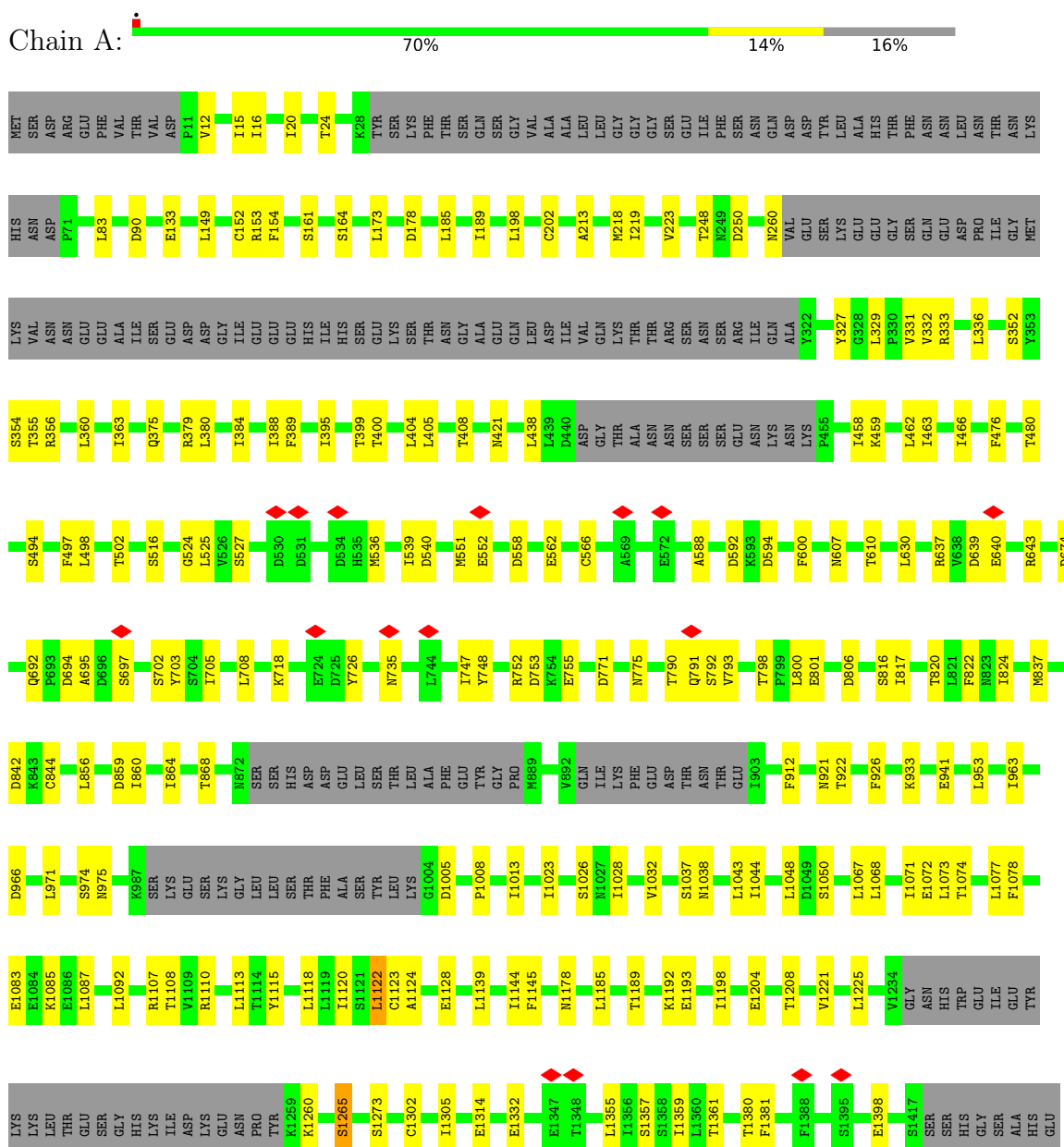
- Molecule 1 is a protein called GEA2 isoform 1.

Mol	Chain	Residues	Atoms					AltConf	Trace	
			Total	C	H	N	O			S
1	A	1224	19870	6306	10046	1603	1878	37	0	0
1	B	1224	19870	6306	10046	1603	1878	37	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: GEA2 isoform 1



ASN THR ASP
THR PRO GLU ARG GLU SER ASN ASP VAL
GLN THR SER ASP
GLU ILE ASP
ALA THR
ALA PRO
ILE ILE
ASP ASP
ASN ASP
ASN THR
THR ASP

• Molecule 1: GEA2 isoform 1

Chain B: 69% 14% 16%

MET SER ASP
ASP P71
L83
L86
L89
D90
N91
I92
K124
I15
I20
T24
K28
TYR
SER
LYS
PHE
THR
SER
GLN
SER
LEU
SER
GLY
VAL
VAL
ALA
ALA
LEU
LEU
GLY
GLY
GLY
SER
GLU
ILE
PHE
SER
ASN
ASN
ASN
ASN
PHE
THR
HIS
ALA
TYR
LEU
LEU
LEU
ALA
HIS
THR
PHE
ASN
ASN
ASN
ASN
LEU
LEU
THR
ASN
HIS

ASN ASP
P71
L83
L86
L89
D90
N91
I92
K124
I15
I20
T24
K28
TYR
SER
LYS
PHE
THR
SER
GLN
SER
LEU
SER
GLY
VAL
VAL
ALA
ALA
LEU
LEU
GLY
GLY
GLY
SER
GLU
ILE
PHE
SER
ASN
ASN
ASN
ASN
PHE
THR
HIS
ALA
TYR
LEU
LEU
LEU
ALA
HIS
THR
PHE
ASN
ASN
ASN
ASN
LEU
LEU
THR
ASN
HIS

T222
S246
D250
D255
VAL
GLU
SER
LYS
GLU
GLY
GLY
SER
SER
GLN
GLU
ASP
PRO
ILE
ILE
GLY
MET
LYS
VAL
ASN
ASN
ASN
ASN
GLU
GLU
GLU
ALA
ALA
ILE
SER
GLU
SER
ASP
GLY
ILE
GLU
GLU
GLU
HIS
HIS
HIS
SER
SER
SER
LEU
VAL
GLN

LYS THR
THR ARG
ARG THR
ASN SER
SER ARG
ILE ARG
ALA Y322
Y327
Y335
L342
E348
S352
Y353
L367
F381
L506
P507
I384
S385
I388
F389
I395
T399
I418
L419
Q426
R433
L438
L439
D440
ASP
GLY
THR
ALA
ALA
ASN
ASN
SER
SER
SER
SER
GLU
ASN
LYS

ASN LYS
P455
I463
T471
R472
F476
T480
D485
C486
N487
L488
D489
R490
S494
F497
T502
L506
P507
E508
S509
S516
S527
S531
D534
D538
I539
D540
E542
M551
E552
N571
D592
M602
R605
M606
N607
L650
HIS

S634
D639
E640
L645
L646
F665
D686
D694
A695
D696
S697
S702
Q716
H720
F723
E724
N729
Y743
D753
Q791
L802
F805
D806
S816
T820
L821
F822
R823
I824
L841
D842
K843
C844
L856
L860
R872
SER
SER
HIS

ASP ASP
LEU LEU
THR THR
LEU LEU
PHE PHE
TYR TYR
GLY GLY
PRO PRO
K689
P890
L891
V892
ILE ILE
LYS LYS
PHE PHE
GLU GLU
ASP ASP
THR THR
ASN ASN
THR THR
I903
F912
Q919
L920
N921
T922
F926
N932
K933
S939
K940
E941
L942
W943
I946
V947
L951
L957
I963
F964
I983
N986

K987
SER LYS
GLU GLU
SER SER
LYS LYS
GLY GLY
LEU LEU
SER SER
THR THR
PHE PHE
ALA ALA
SER SER
TYR TYR
LEU LEU
G1004
T1009
E1010
E1011
E1012
I1013
C1022
I1023
N1027
I1028
V1032
N1038
I1039
D1042
L1043
I1044
S1050
R1061
E1064
L942
W943
I1070
L1073
L1074
L1077
L1087
R1107
T1108
V1109
R1110

L1113
S1121
L1122
C1123
A1124
L1139
I1144
F1145
F1149
F1150
A1151
T1152
L1162
M1178
F1181
Q1315
I1206
V1221
L1224
L1225
I1228
S1229
A1233
GLY
HIS
TRP
GLU
ILE
TYR
LYS
LYS
LEU
THR
SER
GLY
HIS
LYS
ILE
ASP
LYS
GLU
ASN
PRO
TYR
K1259

E1263
L1266
L1271
D1281
E1286
A1289
I1290
Q1297
C1302
K1303
T1312
L1313
E1314
Q1315
T1316
T1325
M1328
E1333
S1357
S1358
I1359
L1360
T1361
Y1367
L1368
H1369
Y1370
G1374
K1375
T1376
S1377
F1381
F1388
I1405
K1408
I1411
S1417
SER
SER
HIS

HIS GLY
SER SER
ALA ALA
HIS HIS
GLU GLU
GLN GLN
THR THR
PRO PRO
GLU GLU
SER SER
ASN ASN
ASP ASP
VAL VAL
GLU GLU
ILE ILE
GLU GLU
ALA ALA
THR THR
ALA ALA
PRO PRO
ILE ILE
ASP ASP
ASP ASP
ASN ASN
THR THR
ASP ASP
ASP ASP
ASN ASN
LYS LYS
PRO PRO
LYS LYS
LEU LEU
SER SER
ASP ASP
VAL VAL
GLU GLU
ASP ASP

4 Experimental information

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	83665	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	NONE	Depositor
Microscope	FEI TALOS ARCTICA	Depositor
Voltage (kV)	200	Depositor
Electron dose ($e^-/\text{\AA}^2$)	1.0	Depositor
Minimum defocus (nm)	1000	Depositor
Maximum defocus (nm)	2000	Depositor
Magnification	Not provided	
Image detector	GATAN K3 (6k x 4k)	Depositor
Maximum map value	125.253	Depositor
Minimum map value	-109.998	Depositor
Average map value	0.002	Depositor
Map value standard deviation	1.500	Depositor
Recommended contour level	5.85	Depositor
Map size (\AA)	499.2, 499.2, 499.2	wwPDB
Map dimensions	300, 300, 300	wwPDB
Map angles ($^\circ$)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (\AA)	1.664, 1.664, 1.664	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	0/9981	0.61	1/13478 (0.0%)
1	B	0.31	0/9981	0.62	1/13478 (0.0%)
All	All	0.31	0/19962	0.61	2/26956 (0.0%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	842	ASP	CB-CG-OD1	9.60	126.94	118.30
1	B	842	ASP	CB-CG-OD2	8.67	126.11	118.30

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	9824	10046	10046	96	0
1	B	9824	10046	10046	105	0
All	All	19648	20092	20092	200	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 5.

The worst 5 of 200 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:953:LEU:HD11	1:A:1023:ILE:HD13	1.64	0.79
1:A:1032:VAL:O	1:A:1038:ASN:ND2	2.24	0.71
1:A:498:LEU:O	1:A:502:THR:OG1	2.09	0.71
1:A:705:ILE:HD12	1:A:747:ILE:HD11	1.74	0.70
1:B:1032:VAL:O	1:B:1038:ASN:ND2	2.23	0.70

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	1208/1459 (83%)	1150 (95%)	58 (5%)	0	100	100
1	B	1208/1459 (83%)	1143 (95%)	64 (5%)	1 (0%)	51	84
All	All	2416/2918 (83%)	2293 (95%)	122 (5%)	1 (0%)	100	100

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	1233	ALA

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	1137/1346 (84%)	1093 (96%)	44 (4%)	32	59

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	B	1137/1346 (84%)	1085 (95%)	52 (5%)	27	55
All	All	2274/2692 (84%)	2178 (96%)	96 (4%)	33	57

5 of 96 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	488	LEU
1	B	724	GLU
1	B	508	GLU
1	B	630	LEU
1	B	802	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

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.

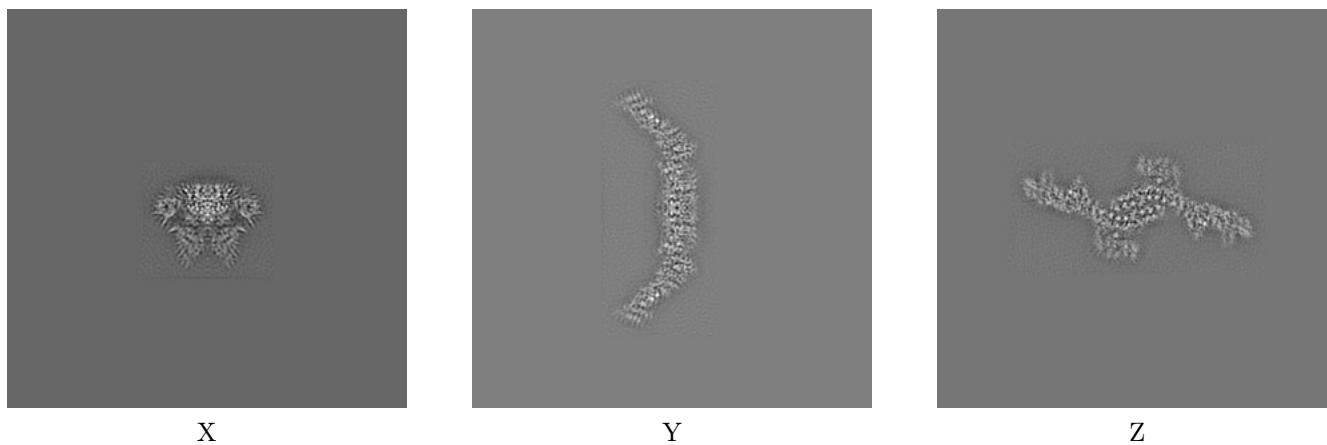
6 Map visualisation [i](#)

This section contains visualisations of the EMDB entry EMD-26754. 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](#)

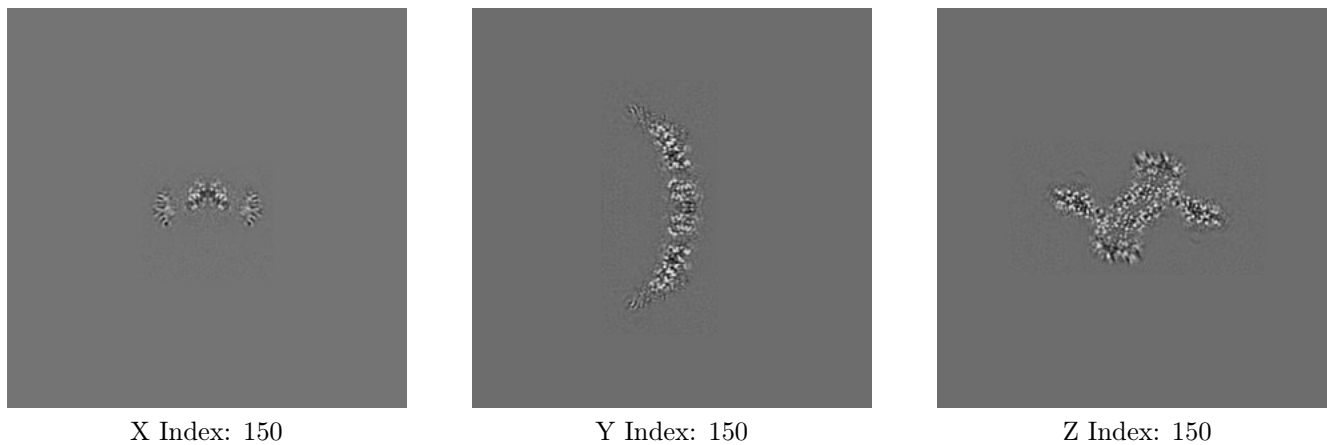
6.1.1 Primary map



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

6.2 Central slices [i](#)

6.2.1 Primary map



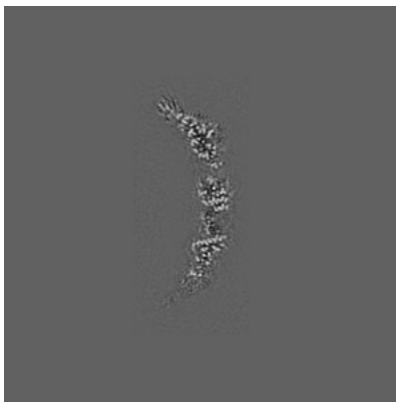
The images above show central slices of the map in three orthogonal directions.

6.3 Largest variance slices [i](#)

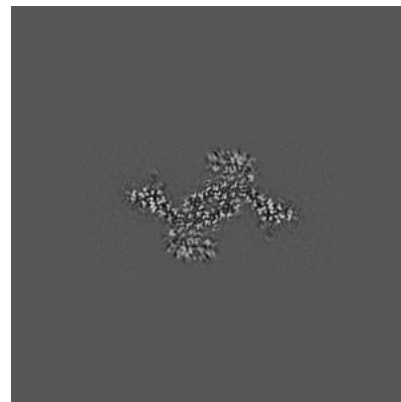
6.3.1 Primary map



X Index: 147



Y Index: 147

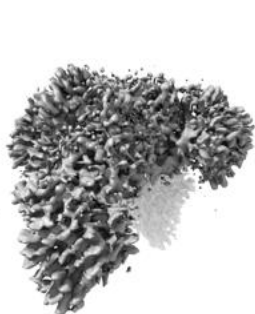


Z Index: 152

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

6.4 Orthogonal surface views [i](#)

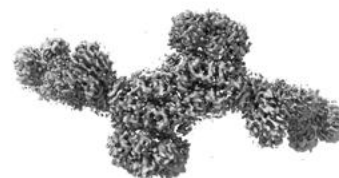
6.4.1 Primary map



X



Y



Z

The images above show the 3D surface view of the map at the recommended contour level 5.85. These images, in conjunction with the slice images, may facilitate assessment of whether an appropriate contour level has been provided.

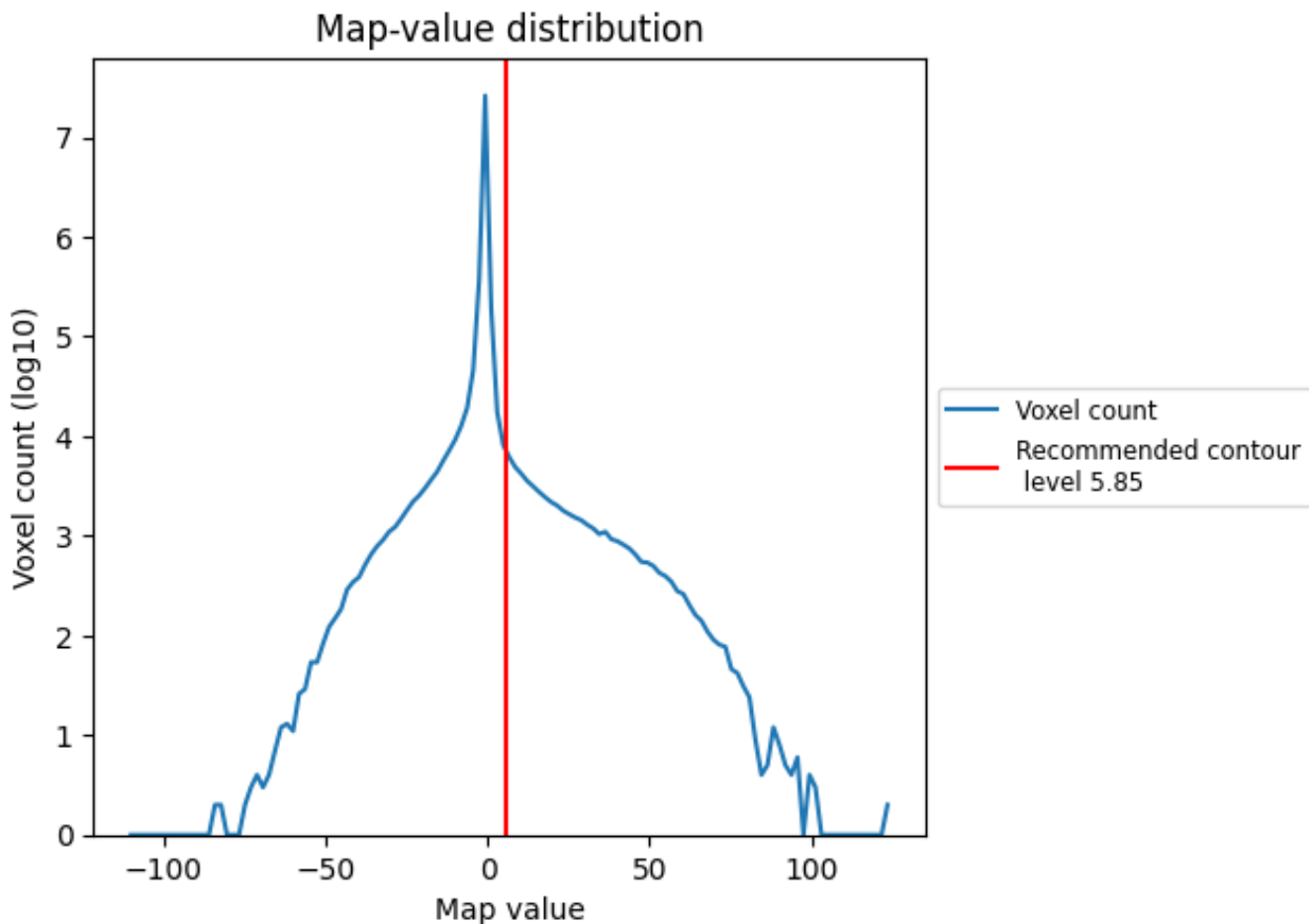
6.5 Mask visualisation

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

7 Map analysis [i](#)

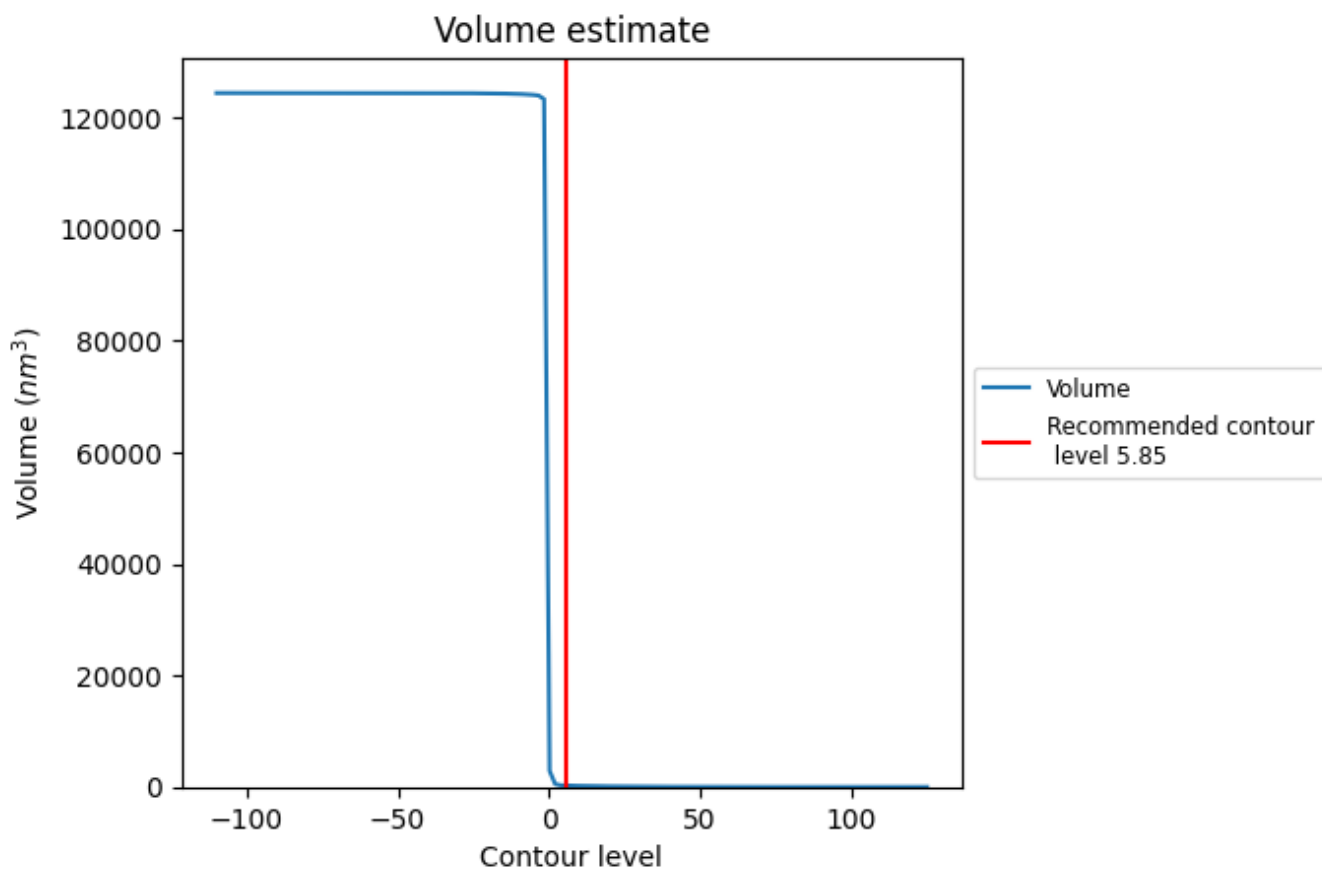
This section contains the results of statistical analysis of the map.

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.

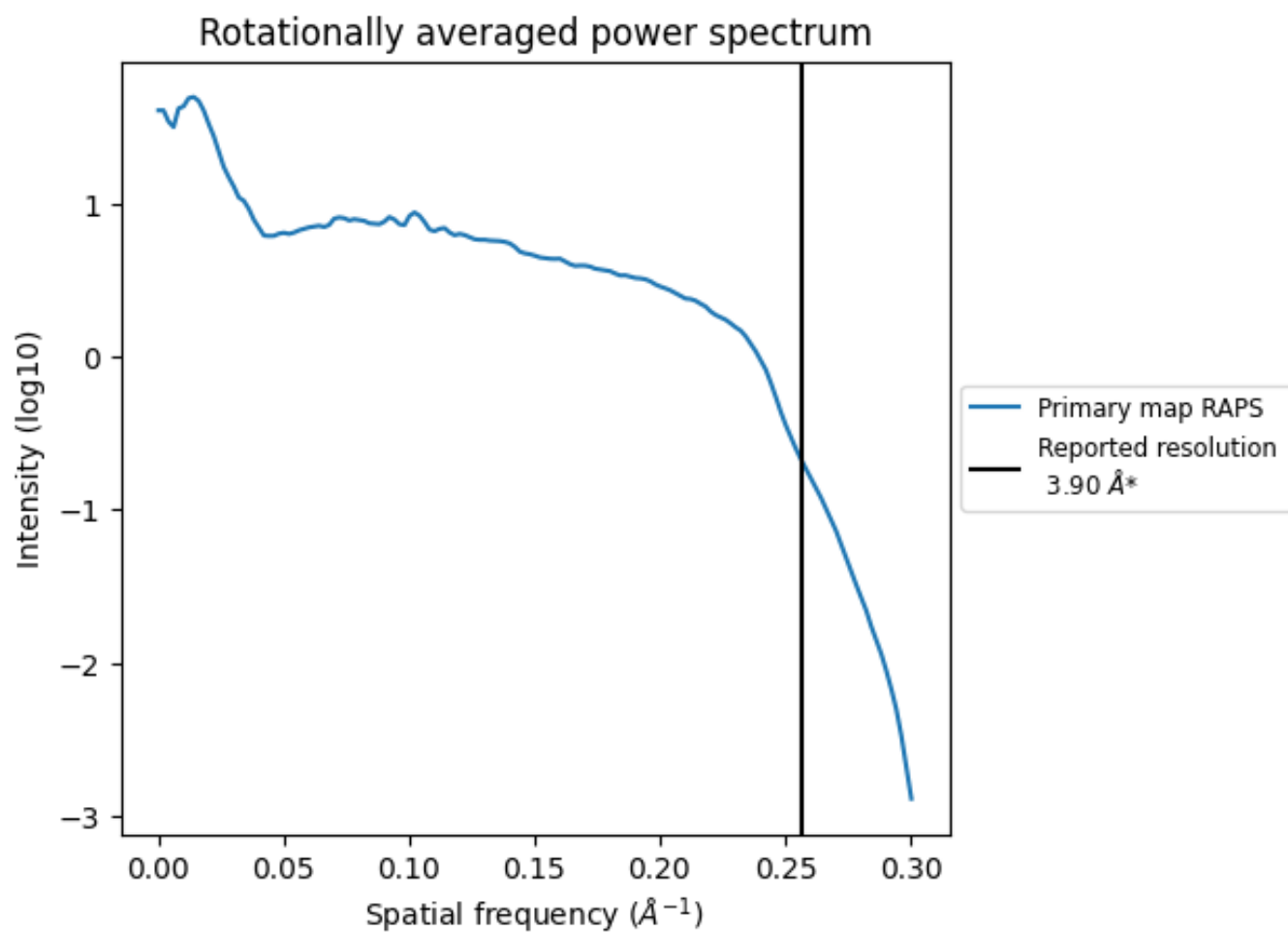
7.2 Volume estimate [i](#)



The volume at the recommended contour level is 250 nm³; 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 0.256 Å⁻¹

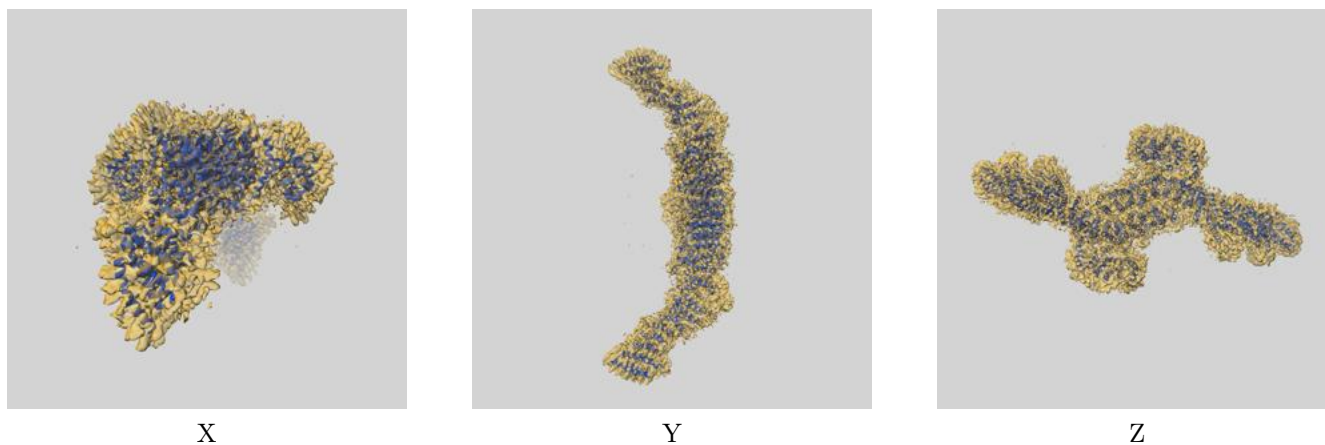
8 Fourier-Shell correlation

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

9 Map-model fit [i](#)

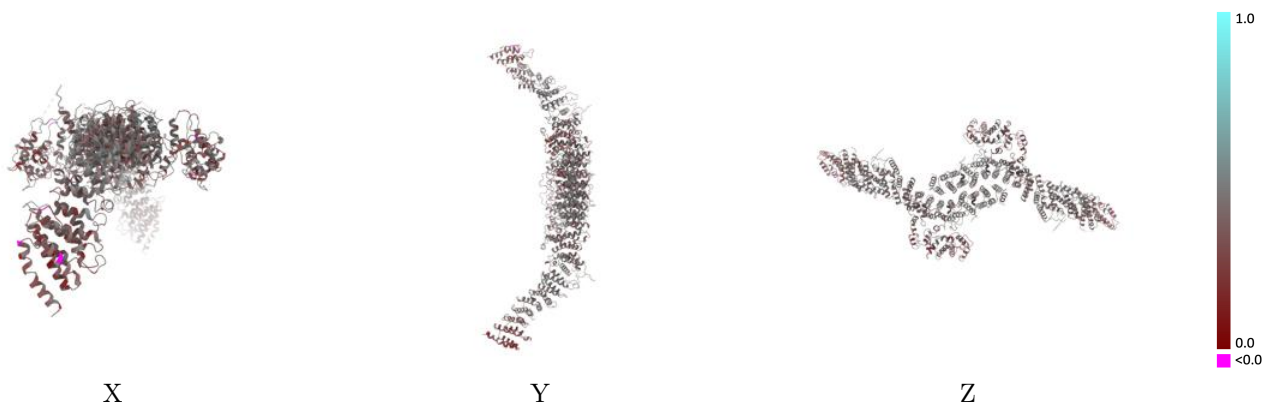
This section contains information regarding the fit between EMDB map EMD-26754 and PDB model 7UT4. Per-residue inclusion information can be found in section [3](#) on page [4](#).

9.1 Map-model overlay [i](#)



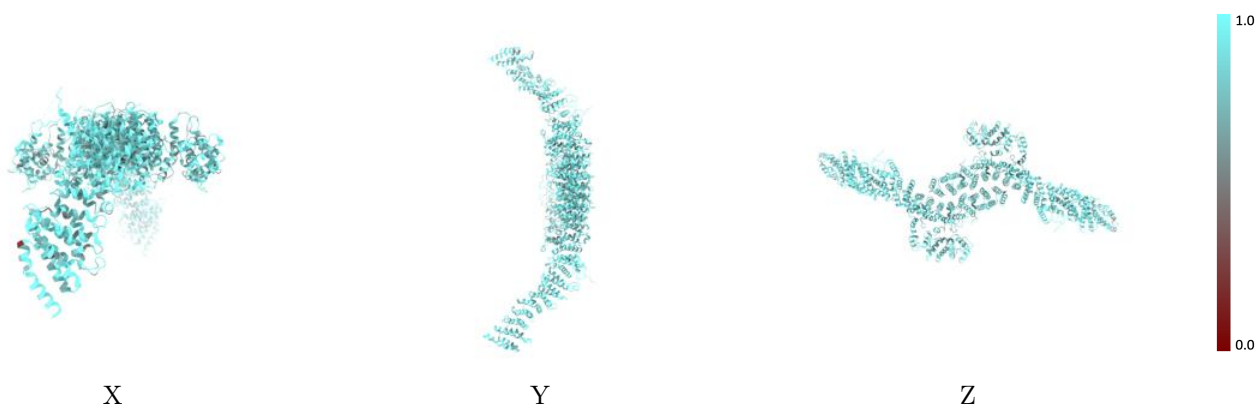
The images above show the 3D surface view of the map at the recommended contour level 5.85 at 50% transparency in yellow overlaid with a ribbon representation of the model coloured in blue. These images allow for the visual assessment of the quality of fit between the atomic model and the map.

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



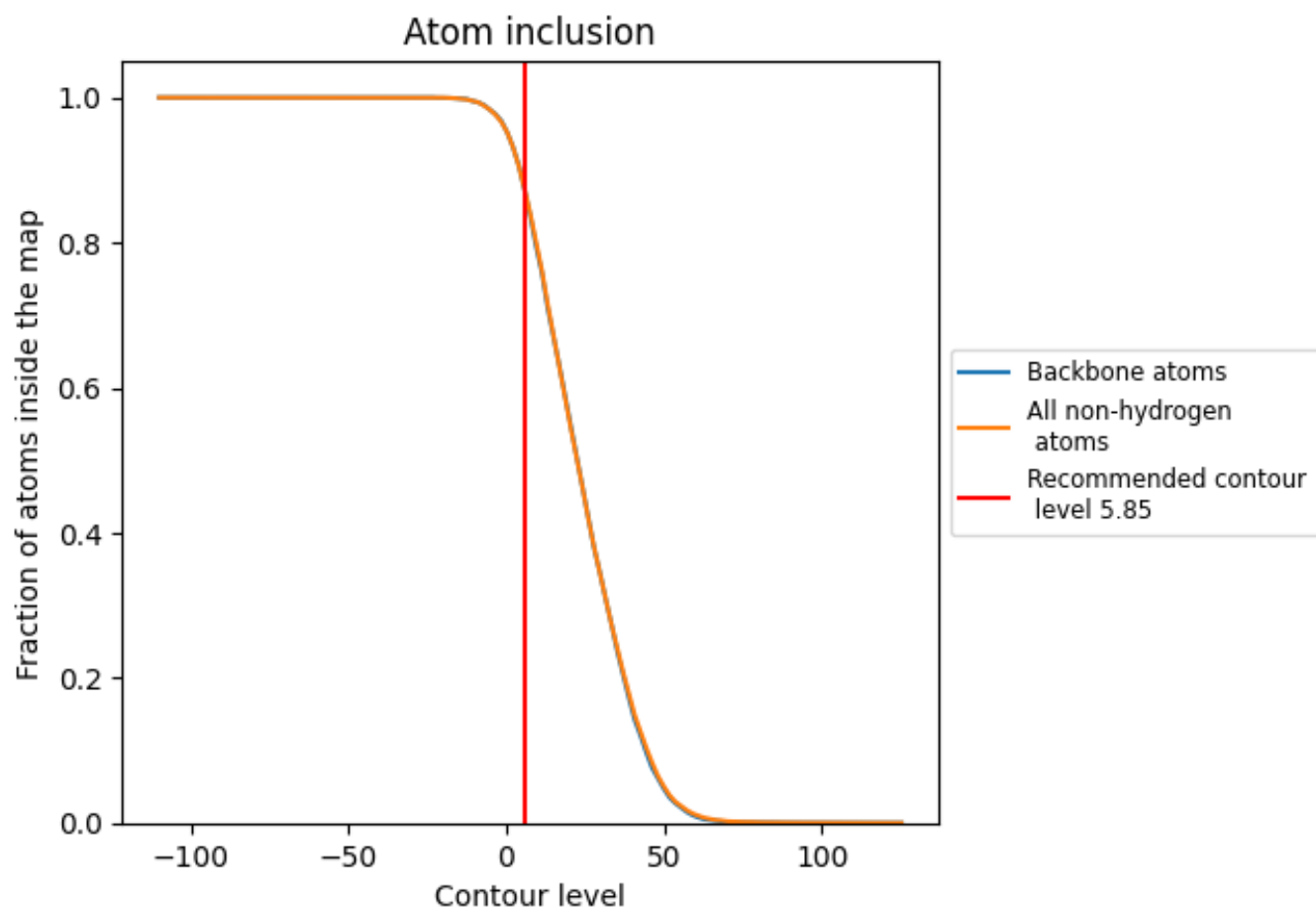
The images above show the model with each residue coloured according its Q-score. This shows their resolvability in the map with higher Q-score values reflecting better resolvability. Please note: Q-score is calculating the resolvability of atoms, and thus high values are only expected at resolutions 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 (5.85).







9.4 Atom inclusion [i](#)



At the recommended contour level, 87% of all backbone atoms, 87% 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 (5.85) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	 0.8727	 0.4120
A	 0.8726	 0.4120
B	 0.8759	 0.4120

